Cálculo en una variable

Luis Navas

Última actualización: 2 de marzo de 2019

Índice

El texto en este color indica un hipervínculo mientras que en este otro color es simplemente texto resaltado **igual que este otro** y también este. El color de un recuadro indica el tipo de contenido:

- definición
- resultado
- atención
- pregunta

Símbolos

El alfabeto griego

minúscula	variante	mayúscula	nombre	pronunciación
α			alpha	alfa
β			beta	beta
γ		Γ	gamma	gama
δ		Δ	delta	delta
ε	ϵ		epsilon	épsilon
φ	ϕ	Φ	phi	fi
η			eta	eta
θ	ϑ	Θ	theta	cita
ι			iota	llota
κ	$\varkappa$		kappa	capa
λ		Λ	lambda	lambda
μ			mi	mu
ν			ni	nu
o			omicron	ómicron
π		Π	pi	pi
ρ	ϱ		rho	ro
σ	ς	Σ	sigma	sigma
τ			tau	tau
υ		Υ	ípsilon	úpsilon
ξ		Ξ	xi	ksi
χ			chi	chi ó ji
ψ		Ψ	psi	psi
ω		Ω	omega	oméga
ζ			zeta	zeta

El alfabeto griego, reordenado para indicar mejor la correspondencia con el latino. Sólo se indica la mayúscula cuando es distinta al alfabeto latino. Algunas minúsculas tienen variantes.

• **Ejemplo:** Usos comunes (no son obligatorios, sólo habituales)

- θ, φ para un ángulo. También α, β, γ.
- ρ para un radio, también para densidad.
- δ, ε para una cantidad pequeña pero positiva.
- ω para una frecuencia angular.
- μ para una media.

Símbolos lógicos y de conjuntos

símbolo	significado
$\exists$	existe
$\nexists$	no existe
$\exists!$	existe un único
$\forall$	para todo
$\in$	pertenece, elemento de, en
$\notin$	no pertenece
$\varnothing$	conjunto vacío
$\implies$	implica
$\impliedby$	implicado por
$\iff$	equivale, si y sólo si
$:$	tal que
$\mid$	tal que (variante)
$\therefore$	por tanto
$\because$	porque
$\vee$	ó (inclusivo: $p \vee q$: de p, q, uno o los dos son ciertos)
$\wedge$	y ($p \wedge q$: p y q son ciertos. También se usa una coma: p, q)
$\sim$	negación ($\sim p$: lo contrario de p)
$\neg$	negación (variante)
$\cap$	intersección de conjuntos
$\cup$	unión de conjuntos
A^c	complementario de A (elementos que no pertenecen á A)
∞	infinito

• **Ejemplo:** "El cuadrado de un número real siempre es positivo o nulo"

$$x^2 \geq 0 \quad \forall x \in \mathbb{R}$$

• **Ejemplo:** "Todo número real positivo tiene una única raíz cuadrada positiva:"

$$\forall\, a \in \mathbb{R} : a > 0 \;\; \exists!\; b \in \mathbb{R} : b > 0 \wedge b^2 = a$$

• **Ejemplo:** "El cuadrado de un real no es negativo, por tanto no hay ningún número real cuyo cuadrado sea -1"

$$x^2 \geq 0 \quad \forall x \in \mathbb{R} \quad \therefore \quad \nexists\; x \in \mathbb{R} : x^2 = -1$$

• **Ejemplo:** Suma de desigualdades:

$$(a < b) \wedge (c < d) \implies a + c < b + d$$

1. Preliminares

1.1. Aritmética básica

Repasamos los distintos conjuntos de números que se utilizan.

- $\mathbb{N}$ – los números **naturales** o "de contar"

$$1, 2, 3, \ldots$$

  a los cuales a veces se añade el 0.

- $\mathbb{Z}$ – los números **enteros**

$$\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots$$

  es decir, los naturales, junto con el 0 y los números **negativos.**

- $\mathbb{Q}$ – los números **racionales,** es decir, las **fracciones** o **cocientes** de enteros:

$$r = \frac{a}{b}, \quad a, b \in \mathbb{Z}, b \neq 0.$$

  Habitualmente se suponen escritos en **forma reducida,** es decir, con denominador b positivo, y a, b **sin factores comunes.**

- $\mathbb{R}$ – los números **reales** son **límites de sucesiones** de números racionales. Se pueden representar por un signo y su **expansión decimal**

$$x = \pm a_1 a_2 \ldots a_k, d_1 d_2 d_3 \ldots, \quad a_i, d_j \in \{0, 1, 2, \ldots, 9\}$$

  (o en cualquier base entera $b \geq 2$) donde posiblemente hacen falta infinitos dígitos a la derecha del separador. Geométricamente, representan los puntos de una **recta**, dividida en dos partes por el 0, estando a la derecha los positivos y a la izquierda los negativos.

- $\mathbb{C}$ – los números **complejos** son combinaciones de dos números reales con el **imaginario** i que por definición cumple $i^2 = -1$:

$$z = x + iy, \quad x, y \in \mathbb{R}, \quad i^2 = -1.$$

  Geométricamente, representan los puntos de un **plano** (x, y) con eje de abscisas el **eje real** de números reales $z = x$ y eje de ordenadas el **eje imaginario** de números **puramente imaginarios** $z = iy$.

  Si $z = x + iy \in \mathbb{C}$, con $x, y \in \mathbb{R}$, decimos que x es la **parte real** de z, escrita $x = \operatorname{Re} z$, y que y es la **parte imaginaria** de z, escrita $y = \operatorname{Im} z$.

Se suponen conocidas las reglas de la aritmética, que enumeramos a continuación.

- **Asociatividad de la suma** $(a+b)+c = a+(b+c)$
- **Conmutatividad de la suma** $a+b = b+a$
- **Elemento neutro de la suma** $a+0 = 0+a = a$
- **Asociatividad del producto** $(a \cdot b) \cdot c = a \cdot (b \cdot c)$
- **Conmutatividad del producto** $a \cdot b = b \cdot a$
- **Elemento neutro del producto** $a \cdot 1 = 1 \cdot a = a$
- **Ley distributiva** $(a+b) \cdot c = a \cdot c + b \cdot c$ y $a \cdot (b+c) \cdot c = a \cdot b + a \cdot c$
- **Absorción** $0 \cdot a = a \cdot 0 = 0$
- **Negativo, opuesto o inverso aditivo** $a + (-a) = (-a) + a = 0$
- **Recíproco o inverso multiplicativo** $a \cdot \frac{1}{a} = \frac{1}{a} \cdot a = 1$ cuando $a \neq 0$.
- **Dividir significa multiplicar por el recíproco** $\frac{a}{b} \overset{\text{def}}{=} a \cdot \frac{1}{b}$
- Se dice que "la división por cero no está definida" o "no es un número." El enunciado preciso sería

  0 no tiene recíproco: ningún número multiplicado por 0 da 1.

- **Reglas de signos** $(+) \cdot (+) = (+)$, $(-) \cdot (+) = (+) \cdot (-) = (-)$, $(-) \cdot (-) = (+)$
- **Cancelación aditiva** $a + x = a + y \implies x = y$.
- **Cancelación multiplicativa** $ax = ay \implies x = y$ cuando $a \neq 0$.

1.2. El orden en los reales

Dados dos números reales x, y, recordar que se dice que x **es menor o igual** que y, escrito $x \leq y$, si $y - x \geq 0$. Se dice que x es **(estrictamente) menor** que y si $y - x > 0$.

Propiedades del orden

- $a < b \iff a + c < b + c$
- $a < b,\ c < d \implies a + c < b + d$
- $a < b \iff ac < bc$, si $c > 0$
- $a < b \iff ac > bc$, si $c < 0$
- $0 < a < b \implies \dfrac{1}{a} > \dfrac{1}{b} > 0$
- $a < b < 0 \implies 0 > \dfrac{1}{a} > \dfrac{1}{b}$

El **valor absoluto** de un número real se define por

$$|x| = \begin{cases} x & \text{si } x \geq 0 \\ -x & \text{si } x < 0 \end{cases}$$

Así pues, $|1| = 1, |-1| = 1$. Tomar el valor absoluto equivale a convertir los signos negativos en positivos, teniendo también en cuenta que $|0| = 0$. Se tiene pues

- $|-x| = |x|$
- Si $\delta > 0$, entonces $|x| < \delta \iff -\delta < x < \delta$.

La **distancia** entre dos números reales a, b se define por

$$d(a, b) = |b - a|$$

1.3. El orden de operaciones

Muchos resultados importantes comparten una forma común que puede expresarse como el hecho de conmutar dos operaciones, es decir, que el resultado sea el mismo sin importar el orden en el que se realizan dichas operaciones. Por ejemplo, consideremos las dos operaciones

(1) Operación "S:" sumar dos números.

(2) Operación "M_c:" multiplicar por un número c dado.

Empezamos con dos números, a, b. Si primero sumamos y luego multiplicamos, obtenemos

$$a, b \xrightarrow{S} a + b \xrightarrow{M_c} (a + b) \cdot c$$

Dado que cuando simbolizamos la aplicación de una función f a una variable x, escribimos $f(x)$ y no $x(f)$, este orden de las operaciones se simboliza por M_cS, ya que S se aplica primero. Si primero multiplicamos y luego sumamos, obtenemos en cambio

$$a, b \xrightarrow{M_c} ac, bc \xrightarrow{S} ac + bc$$

y se simboliza por SM_c ya que ahora primero se aplica M_c.

Resulta que en este ejemplo no importa el orden en el cual se realizan las operaciones. Dan el mismo resultado, lo cual expresa la Ley Distributiva

$$(a+b)\cdot c = a\cdot c + b\cdot c$$

cualesquiera que sean los números a, b, c. Decimos que las dos operaciones conmutan, y se puede simbolizar por la ecuación

$$SM_c = M_cS.$$

Este tipo de relación entre dos factores que se yuxtaponen, donde "el orden de los factores no altera el producto" se llama conmutatividad.

Se dice que dos operaciones $\mathcal{A}$ y $\mathcal{B}$ **conmutan** si siempre se obtiene el mismo resultado al aplicar una seguida de otra, sin importar el orden en el que se realizan. Se simboliza en general por

$$\mathcal{AB} = \mathcal{BA}.$$

Algunas reglas básicas de la Aritmética pueden expresarse también como relaciones de conmutación. Por ejemplo, si a, b son positivos y x es cualquier número real,

$$\boxed{(ab)^x = a^x b^x.}$$

Es decir, "elevar á x" conmuta con multiplicar números positivos.

La conmutatividad como relación simbólica nos es familiar puesto que la suma y el producto de números tienen esta propiedad:

$$a+b=b+a, \qquad a\cdot b = b\cdot a.$$

Sin embargo, **la conmutatividad es la excepción, no la regla.**

Si ahora M es la operación de multiplicar dos números, y S_c la de sumar un número c dado, tenemos por un lado

$$MS_c : a, b \xrightarrow{S_c} a+c, b+c \xrightarrow{M} (a+c)(b+c)$$

y por otro

$$S_cM : a, b \xrightarrow{M} ab \xrightarrow{S_c} ab+c$$

que no dan el mismo resultado, pues en general $(a+c)(b+c) \neq ab+c$. Entonces, M y S_c no conmutan.

Ejemplos de operaciones que no conmutan

- Sumar y elevar al cuadrado: $(a+b)^2 \neq a^2 + b^2$
- Sumar y tomar recíprocos: $\dfrac{1}{a+b} \neq \dfrac{1}{a} + \dfrac{1}{b}$
- Sumar y tomar la raíz cuadrada: $\sqrt{a+b} \neq \sqrt{a} + \sqrt{b}$
- Sumar y tomar logaritmos: $\log(a+b) \neq \log(a) + \log(b)$
- Sumar y tomar el seno: $\sin(a+b) \neq \sin(a) + \sin(b)$
- Multiplicar y exponenciar: $e^{ab} \neq e^a e^b$
- Multiplicar y tomar logaritmos: $\log(ab) \neq \log(a) \cdot \log(b)$
- Multiplicar y tomar el seno: $\sin(ab) \neq \sin(a) \cdot \sin(b)$

Es importante entender que para poder asegurar que dos operaciones conmutan, es preciso saber que al cambiar su orden dan el mismo resultado cualesquiera que sean los valores iniciales sobre los que se operan. No basta con verificarlo para unos valores concretos, hace falta saberlo para todos. En cambio, si se encuentran algunos valores concretos para los cuales el aplicar las operaciones en distinto orden da resultados distintos, entonces podemos asegurar que las operaciones no conmutan.

Ante la duda de si dos operaciones conmutan, se puede probar a **sustituir valores concretos** de las variables y ver si dan el mismo resultado numéricamente. Por ejemplo,

$$\frac{1}{1+1} \neq \frac{1}{1} + \frac{1}{1}, \quad \sqrt{1+1} \neq \sqrt{1} + \sqrt{1}, \quad (1+1)^2 \neq 1^2 + 1^2.$$

Hay una oferta sobre un artículo que queremos comprar. Nos ofrecen dos posibilidades: primero aplicar el descuento y luego el IVA, o al revés, primero el IVA y después el descuento. ¿Cuál es la mejor opción?

Como es poco frecuente que dos operaciones conmuten, cuando sí ocurre suele ser la señal de una propiedad importante que da lugar a un resultado básico y útil, tanto teórico como práctico. Por ejemplo, la propiedad de ser una función lineal f entre dos espacios vectoriales se puede expresar como la propiedad de conmutar tanto con la operación S de suma de dos vectores como con el producto M_c de un vector por un número real fijo c,

$$f(a+b) = f(a) + f(b), \quad f(cx) = c\, f(x).$$

1.4. Intervalos

1.4.1. Intervalos acotados

Un **intervalo acotado** es un subconjunto de $\mathbb{R}$ correspondiente al segmento delimitado por dos números reales a, b, considerados también como puntos de la recta real. Según estos números se incorporen o no como miembros del conjunto, cambia la terminología para describir el tipo de intervalo.

Sean a, b dos números reales con $a < b$. El **intervalo abierto** entre a y b es el conjunto

$$(a, b) = \{x : a < x < b\}$$

El **intervalo cerrado** entre a y b es el conjunto

$$[a, b] = \{x : a \leq x \leq b\}$$

El **intervalo semiabierto a la derecha** entre a y b es el conjunto

$$[a, b) = \{x : a \leq x < b\}$$

El **intervalo semiabierto a la izquierda** entre a y b es el conjunto

$$(a, b] = \{x : a < x \leq b\}$$

A los números o puntos a, b se les llama los **extremos** del intervalo.

intervalo compacto es sinónimo de intervalo cerrado y acotado.

Como puede verse en las definiciones, lo que distingue “abierto” de “cerrado” es si el extremo pertenece o no al intervalo.

En el caso de ser $a > b$, no hay números x con $a \leq x \leq b$ y por tanto estas definiciones en principio son vacías. Sin embargo, es habitual entender que en ese caso uno se refiere a los intervalos **con los extremos cambiados de orden,** es decir, con b, a en vez de a, b, aún siendo menos habitual hablar de “el intervalo entre 1000000 y 1” en vez de “el intervalo entre 1 y 1000000.”

Cuando $a = b$, todos los intervalos anteriores son vacíos excepto el cerrado, que se reduce a un punto, $[a, a] = \{a\}$. En este caso también se dice que el intervalo ha **degenerado a un punto.**

1.4.2. Intervalos no acotados

Sea c un número real. El intervalo **abierto, infinito a la derecha** con extremo en c es

$$(c, \infty) = \{x : c < x\}$$

y el intervalo **abierto, infinito a la izquierda** con extremo en c es

$$(-\infty, c) = \{x : x < c\}.$$

De manera similar, el intervalo **cerrado, infinito a la derecha** con extremo en c es

$$[c, \infty) = \{x : c \leq x\}$$

y el intervalo **cerrado, infinito a la izquierda** con extremo en c es

$$(-\infty, c] = \{x : x \leq c\}.$$

También se pueden llamar **semiabiertos** a estos dos últimos. Finalmente, la recta real completa es lo mismo que el intervalo con dos extremos infinitos

$$(-\infty, \infty) \stackrel{\text{def}}{=} \mathbb{R}.$$

Por razones que no vamos a abordar aquí, éste se considera tanto abierto como cerrado.

1.4.3. Entornos

Sea a un número real y δ un real **positivo**. El **entorno** de **centro** a y **radio** δ es el intervalo abierto acotado

$$(a - \delta, a + \delta) = \{x : a - \delta < x < a + \delta\}.$$

Es fácil comprobar que esto es lo mismo que el conjunto descrito por los x que cumplen la siguiente desigualdad con el valor absoluto:

$$\boxed{x \in (a - \delta, a + \delta) \iff |x - a| < \delta}$$

es decir, son los reales x cuya distancia al centro a es menor que el radio δ, lo cual justifica esta terminología geométrica.

La frase "en un entorno de a" se usará a menudo en lo que sigue. Significa simplemente que existe algún $\delta > 0$ tal que $|x - a| < \delta$. La intención es que δ sea "pequeño," aunque esto es una noción relativa y no rigurosa. Para esta idea se empleará la **notación de aproximación**

$$\boxed{x \approx a}$$

que se lee como "x en un entorno de a" o también, como sinónimos, "x cerca de a," "x aproximadamente igual á a" y frases similares.

Como veremos en el apartado dedicado a los límties, a menudo es es conveniente excluir al propio punto a de un entorno. Esto da lugar a la siguiente definición.

Un **entorno reducido** del punto a es un conjunto de puntos de la forma

$$0 < |x - a| < \delta$$

o sea, se tiene o bien $a - \delta < x < a$ ó $a < x < a + \delta$ pero siempre con $x \neq a$. Es decir, un entorno reducido de a es un entorno de a al que se le ha quitado su propio centro a.

Simbolizaremos la pertenencia a un entorno reducido de a por la notación de aproximación

$$\boxed{x \approx a^*}$$

También es útil en ocasiones considerar las dos "mitades" de un entorno reducido dadas por la condición de ser mayor que a (estar a su derecha) o menor que a (estar a su izquierda). Estas condiciones corresponden a las desigualdades

$$\boxed{a - \delta < x < a, \text{ simbolizada por } x \approx a^+}$$

y

$$\boxed{a < x < a + \delta, \text{ simbolizada por } x \approx a^-}$$

respectivamente. Se dirá que x "se aproxima á a por la derecha/izquierda." También, por abuso de la terminología, diremos de vez en cuando "entorno (reducido) de $a^\pm$."

Entornos generales

En las definiciones anteriores, usamos para entornos de un punto solo intervalos centrados en ese punto. Esto, aunque es habitual, no es necesario. Por entorno de un punto se puede también entender cualquier intervalo abierto que contiene al punto.

La razón que los dos puntos de vista sean equivalentes es que, dado un intervalo abierto $I = (a, b)$ que contiene al punto c, existe $\delta > 0$ tal que $c \in (c - \delta, c + \delta) \subseteq I$. Puede comprobarse que basta con tomar $\delta < \text{mín}(c - a, b - c)$.

Es decir, un entorno descentrado de un punto siempre se puede encoger a uno centrado.

1.5. Completitud de los reales

Propiedad de los intervalos encajados

Sea $I_n = [a_n, b_n], n \in \mathbb{N}$ *una sucesión de intervalos* ***encajados,*** *es decir, tales que*

$$I_1 \supseteq I_2 \supseteq I_3 \supseteq \cdots \supseteq I_n \supseteq \cdots$$

y tales que la longitud de I_n *tiende á* $0 : \lim_n (b_n - a_n) = 0$*. Entonces existe un único punto* c *común a todos los intervalos. Es decir,*

$$\bigcap_{n=1}^{\infty} I_n = \{c\}, \quad \textit{siendo de hecho} \quad c = \lim_n a_n = \lim_n b_n.$$

Intuitivamente, esta propiedad viene a expresar que los reales $\mathbb{R}$ son **un continuo** de puntos, es decir, que $\mathbb{R}$ “no tiene agujeros.” Esto también se llama **completitud.**

Un subconjunto $A \subseteq \mathbb{R}$ **acotado superiormente** si existe $M \in \mathbb{R}$ tal que $x \leq M$ para todo $x \in A$. Decimos que tal M es una **cota superior.**

Decimos que $A \subseteq \mathbb{R}$ está **acotado inferiormente** si existe $m \in \mathbb{R}$ tal que $m \leq x$ para todo $x \in A$. Un tal m es una **cota inferior.**

$A \subseteq \mathbb{R}$ es **acotado** si está acotado tanto superior como inferiormente.

Existencia de ínfimos y supremos

Todo subconjunto $A \subseteq \mathbb{R}$ *acotado superiormente tiene una* ***mínima cota superior*** *ó* ***supremo*** $\beta = \sup A$*. Es decir,* β *es una cota superior de* A *y, si* M *es cualquier otra cota superior de* A*, se tiene* $\beta \leq M$*. Si* A *no está acotado superiormente, ponemos* $\sup A = \infty$*.*

Si $A \subseteq \mathbb{R}$ *está acotado inferiormente, tiene una* ***máxima cota inferior*** *o* ***ínfimo,*** *denotada por* $\inf A$*. Si* A *no está acotado inferiormente, ponemos* $\inf A = -\infty$*.*

Ínfimo y **supremo** no son lo mismo que **mínimo** y **máximo.** Por ejemplo, para un intervalo abierto y acotado $I = (a, b)$, se tiene $a = \inf I$ y $b = \sup I$, pero I no tiene elemento mínimo ni elemento máximo, pues dado cualquier $c \in I$, siempre existe un número más pequeño, $c' \in I$ con $c' < c$. Si es cierto que el ínfimo y supremo son **límites** de elementos del conjunto.

La propiedad de existencia del supremo y del ínfimo no la comparte por ejemplo el conjunto de números racionales $\mathbb{Q}$. El conjunto $A = \{x \in \mathbb{Q} : x^2 < 2\}$ está acotado superiormente (por ejemplo, por $\frac{3}{2}$), pero no tiene una mínima cota superior. Esta debería ser un número cuyo cuadrado es 2, pero no existe tal número racional.

Existencia de límites de sucesiones crecientes

Toda sucesión $\{a_n\}$ *creciente y acotada de números reales tiene límite; es decir, existe* $a \in \mathbb{R}$ *tal que* $\lim_n a_n = a$. *Si la sucesión es creciente pero no acotada,* $\lim_n a_n = \infty$.

Resulta que estas tres propiedades, la de los intervalos encajados, la existencia de ínfimos y supremos, y la existencia de límites de sucesiones crecientes, son equivalentes.

Puede verse, por ejemplo, que si $I_n = [a_n, b_n]$, $n \in \mathbb{N}$, es una sucesión de intervalos encajados, entonces $c = \sup\{a_n : n \in \mathbb{N}\} = \inf\{b_n : n \in \mathbb{N}\}$, y además, la sucesión a_n es creciente con límite c, y b_n es decreciente con el mismo límite.

1.6. Funciones

Una **función** consta de tres datos:

- Un conjunto D, llamado **dominio** de la función.
- Un conjunto R, llamado **recorrido** de la función.
- Una **correspondencia** que a cada miembro del dominio, $x \in D$, le asigna un **único** valor de la imagen, $y \in R$. La función se denota por una **fórmula** f, o más detalladamente, $y = f(x)$.

Escribimos, para denotar una función,

$$f : D \to R \quad \text{ó también} \quad D \xrightarrow{f} R$$

indicando todas las componentes de la definición.

Técnicamente hablando, la función consta de estos tres datos, y si varía alguno de ellos, se considera que hemos cambiado de función, aunque en la práctica se le suele dar más importancia a la fórmula que al dominio o recorrido, llegando incluso a obviarlos o a darlos como sobreentendidos.

Por ejemplo, la fórmula “elevar al cuadrado” asigna a un número real x el número real $x^2 = x \cdot x$. Hablamos pues de $f(x) = x^2$. Para tener una función, podemos tomar como dominio $D = \mathbb{R}$ y como recorrido $R = [0, \infty)$, dado que un cuadrado nunca es negativo.

El dominio y el recorrido de una función no son únicos. Se puede considerar una misma fórmula sobre distintos dominios y recorridos.

Dada una función $f : D \to R$, y un subconjunto D' del dominio D, se obtiene una nueva función $f : D' \to R$ aplicando la fórmula sólo a los $x \in D'$. Esta nueva función se llama la **restricción** de f al subdominio D'.

Por ejemplo, para la fórmula “elevar al cuadrado” también podíamos tomar como recorrido todos los reales, $R = \mathbb{R}$. Asimismo, podíamos restringirla al dominio dado por el intervalo $D = [0, 1]$. En este último caso, también valdría como recorrido el mismo intervalo: $R = D = [0, 1]$.

Hay que tener un cierto cuidado, sobre todo con el dominio. Por ejemplo, la fórmula “tomar la raíz cuadrada de un número real,” denotada $f(x) = \sqrt{x}$, sólo se puede aplicar a un número real no negativo. No tiene sentido hablar de la raíz cuadrada de -1 en los números reales, pues por raíz cuadrada de a se entiende un número x que elevado al cuadrado da a, y ningún número real x elevado al cuadrado da -1. En otras palabras, la ecuación

$$x^2 = -1$$

no tiene ninguna solución real x. Precisamente, para poder hablar de $\sqrt{-1}$ se amplían los números reales a los números complejos.

Sea $f : D \to R$ una función. La **imagen directa** (o simplemente imagen) de f es el conjunto

$$f(D) \stackrel{\text{def}}{=} \{f(x) : x \in D\}$$

formado por todos los valores que f asume sobre su dominio.

Siempre se puede tomar como recorrido de una función $f : D \to R$, su imagen $f(D)$. La imagen es el recorrido más pequeño de la función si fijamos el dominio D.

1.6.1. Algunos tipos de funciones

Hay algunas propiedades generales que pueden tener las funciones que son útiles tener en cuenta.

Decimos que una función $f : D \to R$ es **inyectiva** si a valores distintos de la variable corresponden valores distintos de la función. Es decir, si

$$a \neq b \implies f(a) \neq f(b)$$

La inyectividad resalta que una función consta no sólo de la fórmula sino también el dominio. Por ejemplo, $f(x) = x^2$ no es inyectiva sobre $D = \mathbb{R}$ ya que $(-x)^2 = x^2$ para todo x, pero sí es inyectiva sobre $D = [0, \infty)$.

Una función $f : D \to R$ es **sobreyectiva** o **epiyectiva** si para cualquier $y \in R$ hay al menos un $x \in D$ tal que $y = f(x)$. Es decir, si cada miembro del recorrido proviene de algún miembro del dominio al aplicar f.

Esta propiedad depende del recorrido. Por ejemplo, $f(x) = x^2 : \mathbb{R} \to [0, \infty)$ es epiyectiva, pero si consideramos que el recorrido es $\mathbb{R}$, entonces $f(x) = x^2 : \mathbb{R} \to \mathbb{R}$ no es epiyectiva, ya que ningún número negativo es el cuadrado de un número real.

Una función $f : D \to R$ es **biyectiva** si es inyectiva y epiyectiva.

Por ejemplo, $f(x) = x^2 : [0, \infty) \to [0, \infty)$ es biyectiva, así como $f(x) = x^2 : [0, 1] \to [0, 1]$.

Sea D un conjunto. La función $f : D \to D$ que a cada $x \in D$ le asigna ese mismo x, se llama la **función identidad sobre D.**

La función identidad es un ejemplo sencillo de función biyectiva.

1.6.2. Composición de funciones

Componer dos funciones significa aplicar una seguida de otra. Esto no siempre es posible, debido a que posiblemente no sean compatibles los dominios y recorridos. Para resaltar el problema, escribiremos dominios y recorridos con un subíndice que indica la fórmula a la que están asociados. Sean

$$f : D_f \to R_f, \quad g : D_g \to R_g.$$

- Empezamos con un elemento $a \in D_f$.
- Aplicamos f a este elemento.
- Obtenemos el elemento $b = f(a) \in R_f$.
- Para poder aplicar ahora g, necesitamos que $b \in D_g$.
- Si efectivamente $b \in D_g$, entonces obtenemos $c = g(f(a)) \in R_g$.

Si estos pasos se pueden realizar cualquiera que sea el elemento $a \in D_f$ con el cual empezamos, decimos que f y g se pueden componer, y escribimos

$$\boxed{(g \circ f)(a) = g(f(a))}$$

llamada la **composición** de g con f.

Observar que el orden de aplicación de las funciones va de dentro hacia fuera, es decir, se dice “g compuesta con f” y se escribe $g \circ f$ aunque primero se aplica f y luego g, ya que la función se escribe a la izquierda de la variable.

Razonando de esta manera, vemos que:

Para que tenga sentido la composición de g con f basta con que **el recorrido de f esté contenido en el dominio de g:**

$$R_f \subseteq D_g.$$

En este caso se obtiene una función

$$g \circ f : D_f \to R_g.$$

A veces, si dos funciones no se pueden componer, cambiando los dominios y/o recorridos hace posible su composición.

• **Ejemplo:** Sean

$$f(x) = x - 1 : \mathbb{R} \to \mathbb{R}.$$

$$g(x) = \sqrt{x} : [0, \infty) \to [0, \infty),$$

El recorrido de f es $\mathbb{R}$, que no está contenido en el dominio $[0, \infty)$ de g. Sin embargo, cambiando el dominio de f á $D = [1, \infty)$, y su recorrido a la imagen $f(D) = [0, \infty)$, sí tiene sentido la composición

$$(g \circ f)(x) = \sqrt{x - 1} : [1, \infty) \to [0, \infty).$$

En sentido opuesto, con las funciones originales, el recorrido $[0, \infty)$ de g está contenido en el dominio $\mathbb{R}$ de f, luego no hay ningún problema en componer

$$(f \circ g)(x) = \sqrt{x} - 1 : [0, \infty) \to \mathbb{R}$$

Este ejemplo muestra que

No siempre se pueden definir $g \circ f$ y $f \circ g$ simultáneamente y, aunque esto sea posible, en general no son la misma función. Suele ocurrir que $f \circ g \neq g \circ f$. En otras palabras, **la composición no es conmutativa.**

La composición de funciones es **asociativa**:

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

cuando estén definidas todas estas composiciones. La función identidad $i(x) = x$ actúa como **elemento neutro**:

$$f \circ i = i \circ f = f$$

Siendo más rigurosos, si $f : D \to R$, entonces cuando hablamos de $f \circ i$ nos referimos a la función identidad sobre D, $i : D \to D$ y cuando ponemos $i \circ f$ nos referimos a la función identidad sobre R, $i : R \to R$.

1.6.3. Inversión de funciones

Consideremos las funciones

$$f(x) = \sqrt{x} : [0, \infty) \to [0, \infty), \qquad g(x) = x^2 : \mathbb{R} \to [0, \infty).$$

El recorrido $[0, \infty)$ de f está contenido en el dominio $\mathbb{R}$ de g, y se tiene

$$(g \circ f)(x) = (\sqrt{x})^2 = x : [0, \infty) \to [0, \infty).$$

Si consideramos estas mismas funciones, pero en el orden opuesto, es decir, $f \circ g$ en vez de $g \circ f$, vemos que también tiene sentido esta composición, pues el recorrido $[0, \infty)$ de g es igual al dominio de f. Ahora se tiene

$$(f \circ g)(x) = \sqrt{x^2} = |x| : \mathbb{R} \to [0, \infty)$$

pues por definición, la función “raíz cuadrada” se refiere a la raíz cuadrada positiva. Un número real positivo a tiene dos raíces cuadradas, una positiva y otra negativa, pero se utiliza la convención que al decir “la” raíz cuadrada, escrita $\sqrt{a}$, nos referimos a la positiva. La otra es entonces $-\sqrt{a}$.

Comprobar que efectivamente $\sqrt{x^2} = |x|$ para todo número real x.

Si en el ejemplo anterior restringimos $g(x) = x^2$ al dominio $[0, \infty)$, entonces siguen teniendo sentido las composiciones $f \circ g$ y $g \circ f$ y ahora sí son la misma función, de hecho, la función identidad: $(\sqrt{x})^2 = \sqrt{x^2} = x$ si $x \geq 0$. Esta situación es de gran importancia y nos lleva a la siguiente definición.

Una función $f : D \to R$ es **invertible** si existe otra función que “va al revés” $g : R \to D$ tal que

$$(g \circ f)(x) = x : D \to D, \qquad (f \circ g)(x) = x : R \to R.$$

En este caso, la función g está determinada de manera única por f, y se denomina la **inversa** de f, escrita

$$g = f^{-1}.$$

Decimos que f y g son **inversas** la una de la otra.

Una función $f : D \to R$ es invertible si y sólo si es biyectiva. En este caso, a un $y \in R$ le corresponde un único $x \in D$ tal que $y = f(x)$, y la inversa es la función definida por $x = f^{-1}(y)$.

Por ejemplo, la función $f(x) = x^2 : [0, \infty) \to [0, \infty)$ es invertible, y su inversa es la función raíz cuadrada: $f^{-1}(x) = \sqrt{x}$.

Otro par importante de funciones inversas son la exponencial y el logaritmo (natural o “neperiano”)

$$e^x : \mathbb{R} \to (0, \infty), \qquad \log : (0, \infty) \to \mathbb{R}.$$

Asimismo, se utilizan también frecuentemente las inversas de las funciones trigonométricas.

1.6.4. Interpretación geométrica de algunas composiciones

La **traslación por** k es la función $\tau(x) = x + k$.

Sean $f(x)$ una función con dominio y recorrido en los números reales. Si consideramos su composición con $\tau(x) = x + k$, según se componga a la derecha o a la izquierda, obtendremos

$$(\tau \circ f)(x) = f(x) + k, \qquad (f \circ \tau)(x) = f(x + k),$$

que no serán en general la misma función, pero que tienen una interpretación sencilla en términos de la **gráfica** de f.

La gráfica de $f(x) + k$ se obtiene a partir de la gráfica de $f(x)$ moviéndola

$$|k| \textbf{ unidades} \begin{cases} \textbf{hacia arriba} & \textbf{si } k > 0 \\ \textbf{hacia abajo} & \textbf{si } k < 0 \end{cases}$$

La gráfica de $f(x + k)$ se obtiene a partir de la gráfica de $f(x)$ moviéndola

$$|k| \textbf{ unidades} \begin{cases} \textbf{hacia la izquierda} & \textbf{si } k > 0 \\ \textbf{hacia la derecha} & \textbf{si } k < 0 \end{cases}$$

Hay funciones, como por ejemplo $f(x) = \sin x$ ó $f(x) = \cos x$, para las cuales existe un k tal que $f(x + k) = f(x)$. En el caso de sin y cos, vale $k = 2\pi$. Geométricamente, esto significa que al mover la gráfica $|k|$ unidades a la izquierda, se superpone exactamente sobre la gráfica original. Las funciones con esta propiedad se llaman **funciones periódicas** y el número k se llama un **periodo** de la función.

La **multiplicación** u **homotecia** de razón k es la función $\mu(x) = k \cdot x$.

Si $k > 0$, la gráfica de $k \cdot f(x)$ se obtiene a partir de la gráfica de $f(x)$ **multiplicando la distancia al eje x por k unidades.**

Si $k > 0$, la gráfica de $f(kx)$ se obtiene a partir de la gráfica de $f(x)$ **dividiendo la distancia al eje y por k unidades.**

Para entender lo que ocurre cuando $k < 0$, basta con aplicar lo anterior y entender lo qué ocurre para el caso $k = -1$.

La gráfica de $-f(x)$ se obtiene a partir de la gráfica de $f(x)$ **reflejándola en el eje x.**

La gráfica de $f(-x)$ se obtiene a partir de la gráfica de $f(x)$ **reflejándola en el eje y.**

2. Límites y Continuidad

2.1. La definición de límite.

La idea de límite forma la base de la rama de las Matemáticas conocida como Análisis, y que se inicia con el Cálculo. Está estrechamente ligada a la idea de **aproximación.**

Empezamos con un punto $a \in \mathbb{R}$ y una función de variable real $f(x)$, definida en algún entorno de a. Buscamos dar un sentido riguroso, lógico, a la idea expresada en la siguiente frase:

"Existe un número ℓ tal que, a medida que la variable x se acerca al punto a, el valor $f(x)$ de la función se acerca al valor ℓ."

La idea de entorno es lo que se necesita para precisar esta idea. Se divide en varias partes para desentrañar la lógica detrás de la definición que daremos:

(1) Nuestro **objetivo** es que el valor $f(x)$ esté cerca de ℓ. Esto se expresa por $f(x) \approx \ell$, o sea, que $f(x)$ está en un entorno de ℓ, que nosotros especificamos dando su radio, tradicionalmente denotado por un número $\varepsilon > 0$. O sea, queremos que se cumpla

$$|f(x) - \ell| < \varepsilon$$

donde nosotros hemos dado un $\varepsilon > 0$ cualquiera. **Cuánto más pequeño ε, más cerca estará $f(x)$ de ℓ.**

(2) El objetivo debe alcanzarse haciendo que la variable x se acerque al punto a. Es decir, eligiendo un entorno $|x - a| < \delta$.

(3) Otra idea implícita en la frase "x se acerca á a" es que de hecho, **no hace falta que x alcance el valor a.** Esto se traduce en que de hecho sólo hace falta elegir un entorno reducido $0 < |x - a| < \delta$ de a. Por tanto, simbólicamente podemos escribir

$$\boxed{x \approx a^* \implies f(x) \approx \ell}$$

usando la notación de aproximación.

En particular, para hablar de "límite de $f(x)$ en a" no hace falta que f esté definida **en el propio punto a.**

Definición de límite de una función

*Sea $f(x)$ una función definida en un entorno reducido de a. Decimos que el **límite de $f(x)$ cuando x tiende á a es ℓ** si, dado cualquier $\varepsilon > 0$, existe $\delta > 0$ tal que cuando $0 < |x - a| < \delta$, se tiene $|f(x) - \ell| < \varepsilon$. En la notación de la lógica simbólica,*

$$\forall\, \varepsilon > 0,\ \exists\, \delta > 0 : 0 < |x - a| < \delta \implies |f(x) - \ell| < \varepsilon.$$

Notación estándar de límite. Se escribe

$$\lim_{x\to a} f(x) = \ell, \quad \text{y también} \quad f(x) \to \ell \quad (x \to a),$$

para denotar que el límite de $f(x)$ cuando x tiende á a es ℓ.

Si no nos interesa especificar el valor concreto del límite, diremos simplemente que **existe el límite** de $f(x)$ cuando x tiende á a.

2.1.1. Límites laterales

En algunas aplicaciones hace falta restringir la dirección de aproximación de la variable x a un punto a, es decir, hacer que $x \to a$ pero siendo $x > a$ ó hacer que $x \to a$ con $x < a$. Estas ideas se expresan simbólicamente por

$$\boxed{x \approx a^+ \implies f(x) \approx \ell} \quad \text{y} \quad \boxed{x \approx a^- \implies f(x) \approx \ell}$$

respectivamente. Las nociones correspondientes de límite sólo requieren una ligera modificación respecto a la dada anteriormente.

Decimos que el **límite de** $f(x)$ **cuando** x **tiende á** a **por la derecha es** ℓ**,** simbolizado por

$$\lim_{x\to a^+} f(x) = \ell$$

si, dado cualquier $\varepsilon > 0$, existe $\delta > 0$ tal que cuando $a < x < a+\delta$, se tiene $|f(x)-\ell| < \varepsilon$.

Decimos que el **límite de** $f(x)$ **cuando** x **tiende á** a **por la izquierda es** ℓ**,** simbolizado por

$$\lim_{x\to a^-} f(x) = \ell$$

si, dado cualquier $\varepsilon > 0$, existe $\delta > 0$ tal que cuando $a-\delta < x < a$, se tiene $|f(x)-\ell| < \varepsilon$.

Relación entre límite y límites laterales

Sea $f(x)$ una función definida en un entorno de a. El límite $\lim_{x\to a} f(x) = \ell$ existe si y sólo si existen los límites laterales $\lim_{x\to a^+} f(x)$, $\lim_{x\to a^-} f(x)$ ***y son iguales:***

$$\lim_{x\to a} f(x) = \ell \iff \lim_{x\to a^+} f(x) = \lim_{x\to a^-} f(x) = \ell.$$

• **Ejemplo:** Con la función valor absoluto, $\lim_{x\to 0^+} |x| = \lim_{x\to 0^+} x = 0$ y $\lim_{x\to 0^-} |x| = \lim_{x\to 0^-}(-x) = 0$, por tanto $\lim_{x\to 0} |x| = 0$.

Notación de valor para límites laterales

A menudo se emplea la notación

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

*Esta notación sugiere pero **no indica el valor en un punto,** sino el del límite lateral.*

2.1.2. Algunos límites sencillos

En general, una función cualquiera no tiene por qué tener límite en un punto dado. Por ello debemos empezar mostrando algunos ejemplos de límites sencillos para ver que no hemos dado una definición vacía de sentido.

Si $f(x)$ es la función **constante** con valor c, entonces tiene límite c en cualquier punto. Es decir

$$\lim_{x \to a} c = c$$

para todo $a \in \mathbb{R}$.

Si f es la **función identidad,** o sea, $f(x) = x$, entonces

$$\lim_{x \to a} x = a$$

para todo $a \in \mathbb{R}$.

Intentar demostrar o razonar los enunciados anteriores.

• **Ejemplo:** La función **signo**

$$\operatorname{sgn}(x) = \begin{cases} \dfrac{x}{|x|} & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases} = \begin{cases} +1 & \text{si } x > 0 \\ \ \ 0 & \text{si } x = 0 \\ -1 & \text{si } x < 0 \end{cases}$$

tiene $\lim_{x \to 0^+} \operatorname{sgn}(x) = +1$, $\lim_{x \to 0^-} \operatorname{sgn}(x) = -1$, luego no tiene límite en 0.

Para obtener el valor de otros límites, el camino más fácil es desarrollar algunas de las propiedades generales de los límites que permiten combinar funciones y, dados los valores de los límites de las funciones originales, deducir el valor del límite de su combinación. El tipo fundamental de combinación consiste en realizar operaciones aritméticas sobre funciones.

2.2. Propiedades fundamentales de los límites

Supongamos que f y g son dos funciones, ambas definidas en algún entorno de a, y tales que existen sus límites cuando $x \to a$, $\lim_{x \to a} f(x)$, $\lim_{x \to a} g(x)$. Entonces tiene sentido hablar de la suma, resta, producto, y cociente (si no estamos dividiendo por 0) de f y de g. Cabe preguntarse si estas nuevas funciones a su vez tienen límite cuando $x \to a$. La respuesta es afirmativa, y de hecho se tiene:

Límites y operaciones aritméticas

- $\lim_{x\to a}(f(x)+g(x)) = \lim_{x\to a} f(x) + \lim_{x\to a} g(x)$
- $\lim_{x\to a}(f(x)-g(x)) = \lim_{x\to a} f(x) - \lim_{x\to a} g(x)$
- $\lim_{x\to a}(f(x)\cdot g(x)) = \lim_{x\to a} f(x) \cdot \lim_{x\to a} g(x)$
- $\lim_{x\to a}(f(x)/g(x)) = \lim_{x\to a} f(x) / \lim_{x\to a} g(x)$ *si* $\lim_{x\to a} g(x) \neq 0$

Un caso particular de la fórmula para productos es el hecho que, como $\lim_{x\to a} k = k$ para cualquier constante k, se tiene

$$\boxed{\lim_{x\to a} kf(x) = k \lim_{x\to a} f(x)}$$

Según las ideas consideradas en la sección 1.3, estas propiedades se pueden resumir de la siguiente manera, diciendo que

las operaciones aritméticas conmutan con la operación de límite en un punto.

Se puede observar cómo el propio lenguaje refleja la conmutatividad. Por ejemplo, la primera propiedad se puede leer como "el límite de la suma es igual a la suma de los límites," la tercera como "el límite del producto es igual al producto de los límites" y las demás se expresan de manera similar. Los mismos resultados son ciertos para los límites laterales.

Límites laterales y operaciones aritméticas

- $\lim_{x\to a^\pm}(f(x)+g(x)) = \lim_{x\to a^\pm} f(x) + \lim_{x\to a^\pm} g(x)$
- $\lim_{x\to a^\pm}(f(x)-g(x)) = \lim_{x\to a^\pm} f(x) - \lim_{x\to a^\pm} g(x)$
- $\lim_{x\to a^\pm}(f(x)\cdot g(x)) = \lim_{x\to a^\pm} f(x) \cdot \lim_{x\to a^\pm} g(x)$
- $\lim_{x\to a^\pm}(f(x)/g(x)) = \lim_{x\to a^\pm} f(x) / \lim_{x\to a^\pm} g(x)$ *si* $\lim_{x\to a^\pm} g(x) \neq 0$

Estos resultados son extremadamente útiles tanto en la teoría como en la práctica, pues permiten **determinar límites más complicados a partir de otros más sencillos.**

• **Ejemplo:** Dado que $\lim_{x\to x_0} k = k$ para constantes k y $\lim_{x\to x_0} x = x_0$, se tiene que

$$\lim_{x\to x_0}(mx+k) = \lim_{x\to x_0} mx + \lim_{x\to x_0} k = m \lim_{x\to x_0} x + k = mx_0 + k,$$

cualesquiera que sean $x_0, m, k \in \mathbb{R}$.

• **Ejemplo:** Consideremos la función definida **a trozos** por

$$f(x) = \begin{cases} 3x+1 & \text{si } x \geq 0 \\ -2x+1 & \text{si } x < 0 \end{cases}$$

Entonces por lo que hemos visto, existen los límites laterales

$$\lim_{x \to 0^+} f(x) = 3 \cdot 0 + 1 = 1, \qquad \lim_{x \to 0^-} f(x) = -2 \cdot 0 + 1 = 1,$$

y, como coinciden, se tiene $\lim_{x \to 0} f(x) = 1$.

La **composición** de funciones no se considera una operación aritmética. Su comportamiento respecto a los límites es algo más complicado y se estudiará más adelante.

Límites y Orden

Sean f, g funciones definidas en un entorno reducido de a, ambas con límite en a.

- *Si $f(x) \leq g(x)$ para $x \approx a^*$, entonces $\lim_{x \to a} f(x) \leq \lim_{x \to a} g(x)$.*
- *Si $\lim_{x \to a} f(x) < \lim_{x \to a} g(x)$, entonces $f(x) < g(x)$ para $x \approx a^*$.*

Los mismos resultados son ciertos también para límites laterales, $x \to a^{\pm}$ reemplazando el entorno reducido $x \approx a^*$ por las mitades respectivas $x \approx a^{\pm}$.

Demostración. La segunda parte es lógicamente equivalente a la primera intercambiando f, g. Para demostrar la segunda, ponemos $h(x) = g(x) - f(x)$. Entonces

$$\ell \overset{\text{def}}{=} \lim_{x \to a} h(x) = \lim_{x \to a} g(x) - \lim_{x \to a} f(x) > 0$$

y queremos demostrar que $h(x) > 0$ para $x \approx a^*$. Por definición de límite, tomando $\varepsilon = \ell/2$, que es positivo, existe $\delta > 0$ tal que

$$a - \delta < x < a + \delta \implies |h(x) - \ell| < \varepsilon = \frac{\ell}{2} \implies -\frac{\ell}{2} < h(x) - \ell \implies \frac{\ell}{2} < h(x),$$

en particular, $x \approx a^* \implies h(x) > 0$ ya que $\ell/2 > 0$. $\square$

Conservación del signo en los límites.

Sea f una función definida en un entorno reducido de a con límite en a.

- *Si $f(x) \geq 0$ para $x \approx a^*$, entonces $\lim_{x \to a} f(x) \geq 0$ (y respectivamente con $\leq$).*
- *Si $\lim_{x \to a} f(x) > 0$, entonces $f(x) > 0$ para $x \approx a^*$ (y respectivamente con $<$).*

Los mismos resultados son ciertos también para límites laterales, $x \to a^{\pm}$ reemplazando el entorno reducido $x \approx a^*$ por las mitades respectivas $x \approx a^{\pm}$.

Demostración. Es el caso particular del anterior con una de las funciones nula. $\square$

No es verdad que si $f(x) < g(x)$ para $x \approx a^*$ y existen los límites de f, g en a, entonces $\lim_{x\to a} f(x) < \lim_{x\to a} g(x)$. Por ejemplo, $0 < x^2$ para $x \approx 0^*$ pero $0 = \lim_{x\to 0} x^2$.

El Principio del Sandwich

Si m, f, M son funciones definidas en un entorno reducido de a tales que

$$m(x) \leq f(x) \leq M(x)$$

y los límites de m, M en a existen y coinciden:

$$\lim_{x\to a} m(x) = \lim_{x\to a} M(x) = \ell,$$

entonces el límite de f en a existe y coincide con los anteriores:

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

El mismo resultado es cierto también con límites laterales y entornos laterales de a.

Demostración. Dado $\varepsilon > 0$, existen $\delta_1, \delta_2 > 0$ tales que

$$\begin{aligned} 0 < |x-a| < \delta_1 &\implies |m(x) - \ell| < \varepsilon \\ 0 < |x-a| < \delta_2 &\implies |M(x) - \ell| < \varepsilon. \end{aligned}$$

Tomando $\delta = \min(\delta_1, \delta_2)$ se tiene entonces $\delta > 0$ con

$$a - \delta < x < a + \delta \implies |m(x) - \ell| < \varepsilon \text{ y } |M(x) - \ell| < \varepsilon$$

$$\implies \ell - \varepsilon < m(x) \leq f(x) \leq M(x) < \ell + \varepsilon \implies |f(x) - \ell| < \varepsilon,$$

por tanto $\lim_{x\to a} f(x) = \ell$. □

• **Ejemplo:** Se tiene

$$\lim_{x\to 0} x \sin\frac{1}{x} = 0, \quad \text{ya que} \quad 0 \leq \left| x \sin\frac{1}{x} \right| \leq |x| \quad \text{y} \quad \lim_{x\to 0} |x| = 0.$$

Límites nulos

$$\lim_{x\to a} f(x) = 0 \iff \lim_{x\to a} |f(x)| = 0$$

y en general

$$\lim_{x\to a} f(x) = \ell \iff \lim_{x\to a} |f(x) - \ell| = 0$$

Demostración. La primera propiedad es un caso particular de la segunda, y ésta es consecuencia directa de la definición: ambos enunciados dicen que, dado $\varepsilon > 0$, existe $\delta > 0$ tal que si $0 < |x-a| < \delta$, entonces $|f(x) - \ell| < \varepsilon$, pues $\big| |y| \big| = |y|$ para todo y. □

2.3. Continuidad

La noción de función continua está en la base del Análisis Matemático. Se trata de expresar el fenómeno siguiente:

"A pequeños cambios en la variable x corresponden pequeños cambios en el valor $f(x)$."

Esta idea se basa en observaciones experimentales de muchos fenómenos reales. Expresa la noción de "cambio gradual" de alguna cantidad observada. Por ejemplo, la variación de la temperatura de un gas como función de la presión o del volumen, la variación de la velocidad como función del tiempo en el movimiento uniformemente acelerado, o el precio de un producto en función de la demanda.

Simbólicamente, la condición que hemos descrito esto se puede expresar en la notación de cercanía o entornos por

$$\boxed{x \approx a \implies f(x) \approx f(a)}$$

e inmediatamente lleva a su formulación rigurosa en términos de límites.

Definición de continuidad en un punto

Una función f es ***continua en el punto*** *a si*

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

es decir, si

- *$f(x)$ está definida para x en un entorno de a,*
- *Existe el límite $\lim_{x \to a} f(x)$,*
- *El límite anterior vale $f(a)$.*

Estamos, una vez más, ante una propiedad que puede expresarse en términos de la conmutación de dos operaciones. Efectivamente, la definición de continuidad en un punto se puede formular diciendo que

f es continua en a si la operación de límite en a conmuta con f, es decir

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

Existe el concepto análogo de **continuidad lateral** enunciado en términos de límites laterales. Se dice que f es

- **continua por la derecha en** a si $f(a^+) \stackrel{\text{def}}{=} \lim_{x\to a^+} f(x) = f(a)$.
- **continua por la izquierda en** a si $f(a^-) \stackrel{\text{def}}{=} \lim_{x\to a^-} f(x) = f(a)$.

Hemos usado la notación de valor para límites laterales. Precisamente ahora, cuando la función es continua lateralmente, sí es un valor, el de f en a.

De las conmutación de los límites con las operaciones algebraicas se deduce inmediatamente el siguiente resultado fundamental.

Conservación de la continuidad bajo las operaciones aritméticas:

Las sumas, restas, productos y composiciones de funciones continuas en un punto a son también continuas en a. El cociente f/g de dos funciones f, g continuas en a es también continua en a si $g(a) \neq 0$.

Decimos que una función f definida en un dominio D es **continua** a secas si lo es en cada punto $a \in D$.

De hecho, tampoco es difícil demostrar que la continuidad se conserva incluso bajo la operación de composición.

La composición de funciones continuas es también continua

Estos resultados son importantes porque permiten **construir funciones continuas** complicadas a partir de sencillas, así como **detectar fácilmente** que una función es continua.

Los mismos resultados son válidos para la continuidad por la derecha y por la izquierda.

Por ejemplo, sabemos que $\lim_{x\to a} c = c$ para cualquier constante c y $\lim_{x\to a} x = a$, es decir, las funciones constantes y la función identidad son continuas. Por tanto, sumando, restando y multiplicando estas funciones, obtendremos funciones continuas. Primero, multiplicando un número natural n de veces, vemos que las potencias naturales $x^n, n \in \mathbb{N}$ son continuas. Ahora, tomando combinaciones lineales de potencias naturales nos dará también funciones continuas. Dichas combinaciones son los polinomios,

$$p(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_n x^n, \quad c_0, c_1, \ldots, c_n \in \mathbb{R}.$$

Por tanto, hemos demostrado fácilmente que

Los polinomios son continuos en $\mathbb{R}$

Si tomamos cocientes de polinomios, obtenemos las funciones racionales

$$r(x) = \frac{p(x)}{q(x)} = \frac{a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n}{b_0 + b_1 x + b_2 x^2 + \cdots + b_m x^m}$$

con lo cual vemos que

Las funciones racionales son continuas en los puntos donde su denominador no es 0

• **Ejemplo:** Consideremos la función racional reducida

$$f(x) = \frac{x}{x^2 - 3x + 2}.$$

La inspección directa o la fórmula para las raíces de una ecuación de segundo grado muestran que el denominador factoriza como

$$x^2 - 3x + 2 = (x - 1)(x - 2),$$

por tanto $f(x)$ es continua en los puntos $a \neq 1, 2$.

En general no hay ninguna fórmula que nos dé las raíces de una ecuación de grado $n \geq 5$, por tanto en la práctica se emplean métodos numéricos para hallarlas.

Sin embargo, el **algoritmo de Euclides** (división con resto) sirve para hallar el máximo común divisor de dos polinomios y por tanto se utiliza para reducir una función racional eliminando factores comunes entre su numerador y su denominador.

Ver por ejemplo http://es.wikipedia.org/wiki/Algoritmo_de_Euclides.

Sabiendo que las funciones sin, cos, exp y log son continuas (ver la sección 2.7.1), es inmediato deducir que funciones complicadas como

$$\exp(\sin(\cos^3(\exp(17x)) + \log^{1000} x))$$

son continuas, siempre y cuando tengan sentido. En este caso, hace falta restringir a $x > 0$ para que el logaritmo esté definido.

Haciendo uso de la continuidad de log y exp, podemos demostrar que

Si f y g son continuas en a y $f(a) > 0$ entonces $f(x)^{g(x)}$ está definida en un entorno de a y es continua en a.

Se trata de observar que entonces $f(x) > 0$ cerca de a y por tanto $f(x)^{g(x)}$ está definida y es positiva cerca de a, y se tiene

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

por tanto $f(x)^{g(x)}$ es continua en a al ser combinación (composiciones y productos) de funciones continuas.

2.4. Límites de composiciones

Supongamos que tenemos dos funciones tales que

$$\lim_{x \to a} f(x) = b \quad \text{y} \quad \lim_{y \to b} g(y) = \ell.$$

Cuando $x \to a$, sustituyendo $y = f(x)$, se tiene $y \to b$ y por tanto $g(y) \to \ell$. Parece que

$$\lim_{x\to a} f(x) = b,\ \lim_{y\to b} g(y) = \ell \implies \lim_{x\to a} g(f(x)) = \ell \quad (?)$$

Sin más condiciones, **no** es verdad, aunque, dicho sea de paso, "suele" ser verdad, en el sentido que, para las funciones "normales," se cumple, y hace falta construir ejemplos un tanto artificiales para dar un caso donde no se cumpla.

Dados a, b, sea $f(x)$ la función constante con valor b y sea $g(x)$ la función definida por

$$g(y) = \begin{cases} 0 & \text{si } y \neq b \\ 1 & \text{si } y = b. \end{cases}$$

Entonces $g(f(x)) = g(b) = 1$, por tanto

$$\lim_{x\to a} f(x) = b, \quad \lim_{y\to b} g(y) = 0, \quad \lim_{x\to a} g(f(x)) = 1,$$

y no se cumple la propiedad de sustitución.

Intentar encontrar otros ejemplos que **no** cumplan la propiedad de sustitución.

La continuidad de g es suficiente para que se cumpla:

Límites de composiciones I.

Sea g, f un par de funciones tales que:

- *$f(x)$ está definida en un entorno reducido de a, y existe*

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

- *$g(y)$ está definida en un entorno de b y es continua en b; es decir,*

$$\lim_{y\to b} g(y) = g(b)$$

Entonces

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

Demostración. Versión abreviada:

$$x \approx a^* \implies f(x) \approx b, \quad y \approx b \implies g(y) \approx g(b)$$

por tanto, haciendo $y = f(x)$, queda

$$x \approx a^* \implies g(f(x)) \approx g(b).$$

Versión larga habitual:

Sea $\varepsilon > 0$. La continuidad de g implica que existe $\delta > 0$ tal que

$$|y - b| < \delta \implies |g(y) - g(b)| < \varepsilon$$

(no necesitamos reducir el entorno pues con $y = b$ también $|g(y) - g(b)| = 0 < \varepsilon$). Ahora, para este grado de proximidad δ, existe $\rho > 0$ tal que

$$0 < |x - a| < \rho \implies |f(x) - b| < \delta$$

(aquí si es un entorno reducido pues no suponemos nada sobre $f(x)$ cuando $x = a$). Entonces

$$0 < |x - a| < \rho \implies |f(x) - b| < \delta \implies |g(f(x)) - g(b)| < \varepsilon$$

(haciendo $y = f(x)$). Por tanto $\lim_{x\to a}(g \circ f)(x) = g(b)$. □

En particular con este resultado se deduce que la composición de continuas es continua pues en este caso podemos poner $b = f(a)$ y concluir que $\lim_{x\to a} g(f(x)) = g(f(a))$.

Otra condición suficiente para "compatibilizar" la composición con los límites es que $f(x)$ no sea igual a su límite b para $x \approx a^*$:

Límites de composiciones II.

Sea g, f un par de funciones tales que:

- *$f(x)$ está definida en un entorno reducido de a, y existe*

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

- *$f(x) \neq b$ para x en un entorno reducido de a.*
- *$g(y)$ está definida en un entorno reducido de b, y existe*

$$\lim_{y\to b} g(y) = \ell$$

Entonces

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

Demostración. Versión corta:

$$x \approx a^* \implies f(x) \approx b,\ f(x) \neq b \quad y \approx b^* \implies g(y) \approx \ell$$

por tanto, haciendo $y = f(x)$, queda

$$x \approx a^* \implies f(x) \approx b^* \implies g(f(x)) \approx \ell.$$

Versión larga habitual:

Sea $\varepsilon > 0$. Existe $\delta > 0$ tal que

$$0 < |y - b| < \delta \implies |g(y) - \ell| < \varepsilon.$$

Ahora, para este grado de proximidad δ, existe $\rho > 0$ tal que

$$0 < |x - a| < \rho \implies |f(x) - b| < \delta.$$

Dada la hipótesis sobre f, podemos tomar ρ suficientemente pequeño de forma que además $f(x) \neq b$ en el entorno reducido $0 < |x - a| < \rho$. Entonces de hecho

$$0 < |x - a| < \rho \implies 0 < |f(x) - b| < \delta \implies |g(f(x)) - \ell| < \varepsilon$$

con lo cual $\lim_{x \to a} g(f(x)) = \ell$. □

El resultado anterior también es válido para límites laterales, debidamente modificado:

Si $\lim_{x \to a^\pm} f(x) = b$ y $f(x) \neq b$ para $x \approx a^\pm$, y $\lim_{y \to b} g(y) = \ell$, entonces también $\lim_{x \to a^\pm} g(f(x)) = \ell$.

Suponer solo un límite lateral $\lim_{y \to b^\pm} g(y) = \ell^\pm$ requiere suponer respectivamente $f(x) > b$ ó $f(x) < b$ para x en un entorno reducido de a (que también puede ser lateral y de signo independiente del usado para g) para poder concluir que $\lim_{x \to a} f(x) = \ell^\pm$.

Comentamos varios corolarios útiles de estos resultados, empezando primero con el caso de la composición a la izquierda con una función continua.

Límites y composición a la izquierda con invertibles

Sean $I, J \subseteq \mathbb{R}$ intervalos abiertos de cualquier tipo, y $\varphi : I \to J$ continua e invertible con inversa continua $\varphi^{-1} : J \to I$. Entonces existe $\lim_{x \to a} f(x)$ y pertenece á I si y sólo si existe $\lim_{x \to a} (\varphi \circ f)(x)$ y pertenece á J, y en tal caso

$$\lim_{x \to a} f(x) = \varphi^{-1}\Big(\lim_{x \to a} (\varphi \circ f)(x)\Big)$$

Demostración. Si existe $\lim_{x \to a} f(x) = \ell$ y $\ell \in I$, la continuidad de φ en ℓ implica que

$$\lim_{x \to a} \varphi(f(x)) = \varphi\Big(\lim_{x \to a} f(x)\Big) = \varphi(\ell) \in J.$$

Aplicando la inversa queda $\ell = \lim_{x \to a} f(x) = \varphi^{-1}(\lim_{x \to a} (\varphi \circ f)(x))$.

Recíprocamente, si existe $\lim_{x \to a} (\varphi \circ f)(x) = \mathcal{L}$ y $\mathcal{L} \in J$, la continuidad de φ^{-1} en $\mathcal{L}$ implica que existe

$$\lim_{x \to a} f(x) = \lim_{x \to a} (\varphi^{-1} \circ \varphi \circ f)(x) = \varphi^{-1}\Big(\lim_{x \to a} (\varphi \circ f)(x)\Big) = \varphi^{-1}(\mathcal{L}) \in I.$$

□

Como caso particular, tenemos otro resultado práctico para calcular límites de potencias.

Límites compuestos con la exponencial y el logaritmo

$\lim_{x\to a} f(x)$ existe si y sólo si existe $\lim_{x\to a} e^{f(x)}$ y es positivo, y en tal caso

$$e^{\lim_{x\to a} f(x)} = \lim_{x\to a} e^{f(x)} \quad \textit{o sea} \quad \lim_{x\to a} f(x) = \log\Big(\lim_{x\to a} e^{f(x)}\Big)$$

$\lim_{x\to a} f(x)$ existe y es positivo si y sólo si existe $\lim_{x\to a} \log f(x)$, y en tal caso

$$\log \lim_{x\to a} f(x) = \lim_{x\to a} \log f(x) \quad \textit{o sea} \quad \lim_{x\to a} f(x) = e^{\lim_{x\to a} \log f(x)}$$

Demostración. Estos son los casos de las funciones $\exp : (-\infty, \infty) \to (0, \infty)$ y $\log : (0, \infty) \to (-\infty, \infty)$, que son inversas mutuas. □

Por ejemplo, se deduce el siguiente resultado.

Límites de potencias

Sean $f(x), g(x)$ funciones definidas en un entorno reducido de a, ambas con límite en a, y con $\lim_{x\to a} f(x) > 0$. Entonces

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

Demostración. Sea $b = \lim_{x\to a} f(x)$ y $c = \lim_{x\to a} g(x)$. Como $b > 0$, sabemos que $f(x) > 0$ en un entorno reducido de a y por tanto $h(x) = f(x)^{g(x)}$ está definida en un entorno reducido de a y $h(x) > 0$, siendo $\log h(x) = g(x) \log f(x)$.

Aplicando el resultado anterior para composición con el logaritmo, sabemos que existe $\lim_{x\to a} \log f(x) = \log b$. Entonces también existe

$$\lim_{x\to a} \log h(x) = \lim_{x\to a} g(x) \log f(x) = \lim_{x\to a} g(x) \cdot \lim_{x\to a} \log f(x) = c \cdot \log b = \log(b^c)$$

y por el mismo resultado, existe

$$\lim_{x\to a} h(x) = \exp(\lim_{x\to a} \log h(x)) = \exp(\log(b^c)) = b^c,$$

que es lo que queríamos demostrar. □

Límites y composición a la derecha con invertibles

Sean $I, J \subseteq \mathbb{R}$ intervalos abiertos de cualquier tipo, y $\varphi : I \to J$ continua e invertible con inversa continua $\varphi^{-1} : J \to I$. Dado un punto $b \in J$,

$$\lim_{y \to b} g(y) = \lim_{x \to \varphi^{-1}(b)} (g \circ \varphi)(x)$$

Es decir, un límite existe si y sólo si existe el otro y en tal caso son iguales.

Demostración. Supongamos que existe $\ell = \lim_{y \to b} g(y)$. Sea $a = \varphi^{-1}(b)$, de modo que $a \in I$. Como φ es continua, $\lim_{x \to a} \varphi(x) = \varphi(a) = b$. Como φ es biyectiva, $\varphi(x) = b \iff x = a$, por tanto $\varphi(x) \neq b$ para $x \approx a^*$. Entonces el segundo resultado sobre límites y composiciones, aplicado al par g, φ, implica que $\lim_{x \to a} g(\varphi(x)) = \ell$.

Recíprocamente, supongamos que existe $\ell = \lim_{x \to a} (g \circ \varphi)(x)$. Como φ^{-1} es continua, $\lim_{y \to b} \varphi^{-1}(y) = \varphi^{-1}(b) = a$. Como φ^{-1} es biyectiva, $\varphi^{-1}(y) \neq a$ para $y \approx b^*$. El mismo resultado citado antes, aplicado al par $g \circ \varphi$, φ^{-1}, implica que $\lim_{y \to b} (g \circ \varphi)(\varphi^{-1}(y)) = \lim_{y \to b} g(y) = \ell$. □

Con estas observaciones obtenemos resultados del siguiente tipo:

Límites y traslación

Cualquier límite en un punto a se puede trasladar a un límite en 0 y en consecuencia, a un punto cualquiera b:

$$\lim_{x \to a} f(x) = \ell \iff \lim_{h \to 0} f(a + h) = \ell \iff \lim_{h \to b} f(a - b + h) = \ell$$

Demostración. Aplicar el resultado anterior a la traslación $\tau(x) = x - a + b$. Es continua y biyectiva, con $\tau^{-1}(y) = y + a - b$ también continua y biyectiva. Se tiene $\tau^{-1}(b) = a$. □

2.5. Límites en infinito

Que una cantidad y se "aproxime a infinito" es sinónimo de que esa cantidad se hace arbitrariamente grande, es decir, que dada cualquier cota $M > 0$, se tendrá eventualmente $y > M$. Esta idea sirve tanto para dar la definición de límite tomado cuando la variable tiende a infinito, como de un límite cuyo valor es infinito. Empezamos describiendo el primer caso.

La frase "x tiende a infinito" se simboliza por $x \to \infty$ o también $x \to +\infty$ si queremos resaltar que el signo es el positivo.

Límite en ∞. Sea $f(x)$ una función definida en algún intervalo de la forma (c, ∞). Para un número $\ell \in \mathbb{R}$, decimos que **el límite de $f(x)$ cuando x tiende a ∞ es ℓ,** simbolizado por

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

si dado $\varepsilon > 0$, existe una cota $M > 0$ tal que

$$x > M \implies |f(x) - \ell| < \varepsilon$$

El límite cuando $x \to -\infty$ se define análogamente. Ahora se quiere expresar que la variable se hace arbitrariamente grande en valor absoluto, pero es negativa. Basta con cambiar el sentido de las desigualdades en la definición.

Límite en $-\infty$. Sea $f(x)$ una función definida en algún intervalo de la forma $(-\infty, c)$. Para un número $\ell \in \mathbb{R}$, decimos que **el límite de $f(x)$ cuando x tiende a $-\infty$ es ℓ,** simbolizado por

$$\lim_{x\to-\infty} f(x) = \ell$$

si dado $\varepsilon > 0$, existe una cota $M < 0$ tal que

$$x < M \implies |f(x) - \ell| < \varepsilon$$

Un límite en $-\infty$ se puede calcular haciendo el cambio de variable $x \to -x$ y calculando el límite resultante en ∞:

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

Al igual que ocurre con los límites en un número real, los límites en $\pm\infty$ también conmutan con las operaciones aritméticas.

Límites en infinito y operaciones aritméticas

- $\lim_{x\to\infty}(f(x) + g(x)) = \lim_{x\to\infty} f(x) + \lim_{x\to\infty} g(x)$
- $\lim_{x\to\infty}(f(x) - g(x)) = \lim_{x\to\infty} f(x) - \lim_{x\to\infty} g(x)$
- $\lim_{x\to\infty}(f(x) \cdot g(x)) = \lim_{x\to\infty} f(x) \cdot \lim_{x\to\infty} g(x)$
- $\lim_{x\to\infty}(f(x)/g(x)) = \lim_{x\to\infty} f(x)/ \lim_{x\to\infty} g(x)$ *si* $\lim_{x\to\infty} g(x) \neq 0$

Los mismos resultados son ciertos con $-\infty$ en vez de ∞.

• **Ejemplo:** Un ejemplo básico de límite en ∞ lo proporciona

$$\boxed{\lim_{x\to\infty} \frac{1}{x} = 0}$$

Se puede pensar que los límites en $\pm\infty$ son por definición límites "laterales," ya que sólo hay una dirección de aproximación a $\pm\infty$.

De hecho, un límite en infinito es equivalente a un límite lateral en un punto finito, por ejemplo el 0

$$\lim_{x\to\pm\infty} f(x) = \ell \iff \lim_{t\to 0^\pm} f\left(\frac{1}{t}\right) = \ell$$

Esto resulta del cambio de variable $t = 1/x$, que transforma una condición del tipo $x > M$ en una condición del tipo $0 < t < \delta$, siendo $\delta = 1/M$.

• **Ejemplo:** Como $\lim_{x\to\infty} \sin x$ no existe, ya que $\sin x$ no se acerca a ningún valor concreto sino que periódicamente repite todo valor entre -1 y 1, tampoco puede existir

$$\lim_{t\to 0^+} \sin\frac{1}{t}.$$

Respecto a lo cambios de variable en general, es incluso más sencillo con límites en infinito, pues la condición de no ser igual al límite se cumple automáticamente. Por ejemplo,

$$\left\{\begin{array}{l}\lim\limits_{x\to a} \varphi(x) = \infty \\ \lim\limits_{y\to\infty} f(y) = \ell\end{array}\right\} \implies \lim_{x\to a} f(\varphi(x)) = \ell.$$

Es decir, automáticamente no puede ser $\varphi(x) = \infty$. Este cambio de variable en un límite de valor infinito también es válido para límites laterales y puntos infinitos a.

Dar un razonamiento riguroso que justifique el enunciado anterior y modificaciones válidas para límites laterales.

• **Ejemplo:** Como

$$\lim_{y\to\infty}\left(1+\frac{r}{y}\right)^y = e^r,$$

se tiene que

$$\lim_{x\to a}\left(1+\frac{r}{\varphi(x)}\right)^{\varphi(x)} = e$$

para cualquier función $\varphi(x)$ tal que $\lim_{x\to a} \varphi(x) = \infty$.

2.6. Límites de valor infinito

En la sección 2.5 se cambiaba un punto real $a \in \mathbb{R}$ por $\pm\infty$ en la definición de límite. Ahora haremos esta modificación pero en lo que respecta al valor ℓ del límite. Se trata de expresar la idea de que los valores de $f(x)$ se hacen arbitrariamente grandes a medida que x se acerca á a.

Límites de valor $\pm\infty$ en un punto real $a \in \mathbb{R}$. Sea $a \in \mathbb{R}$ y $f(x)$ una función definida en un entorno reducido de a. Decimos que **el límite de $f(x)$ cuando x tiende á a es ∞,** simbolizado por

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

si, dada cualquier cota $M > 0$, existe $\delta > 0$ tal que

$$0 < |x-a| < \delta \implies f(x) > M.$$

Decimos que **el límite de $f(x)$ cuando x tiende á a es $-\infty$,** simbolizado por

$$\lim_{x\to a} f(x) = -\infty$$

si, dada cualquier cota $M < 0$, existe $\delta > 0$ tal que

$$0 < |x-a| < \delta \implies f(x) < M.$$

Los conceptos correspondientes de **límites laterales** de valor $\pm\infty$

$$\lim_{x\to a^{\pm}} f(x) = \pm\infty$$

se definen reemplazando la condición $0 < |x-a| < \delta$ por $a < x < a+\delta$ y por $a-\delta < x < a$ respectivamente.

Se tiene que $\lim_{x\to a} f(x) = \infty$ si y sólo si $\lim_{x\to a^+} f(x) = \lim_{x\to a^-} f(x) = \infty$.

• **Ejemplo:** Se tiene

$$\boxed{\lim_{x\to 0^+} \frac{1}{x} = \infty} \qquad \boxed{\lim_{x\to 0^-} \frac{1}{x} = -\infty}$$

y no existe $\lim_{x\to 0} 1/x$ ya que no coinciden. Sin embargo, $\lim_{x\to 0} 1/x^2 = \infty$.

Análogamente, se definen los conceptos de límite y límite lateral con valor $-\infty$.

Habiendo llegado a este punto, podemos juntar ambas modificaciones a la definición de límite y dar la definición de límite en infinito con valor infinito. Ahora se trata de expresar la idea de que los valores de $f(x)$ se hacen arbitrariamente grandes a medida que x se va haciendo también arbitrariamente grande.

Límite de valor infinito en infinito. Sea $a \in \mathbb{R}$ y $f(x)$ una función definida en un entorno reducido de a. Decimos que **el límite de** $f(x)$ **cuando** x **tiende á** ∞ **es** ∞**,** simbolizado por

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

si, dada cualquier cota $M > 0$, existe otra cota $K > 0$ tal que

$$x > K \implies f(x) > M$$

Las restantes combinaciones de signos,

$$\lim_{x\to-\infty} f(x) = \infty, \quad \lim_{x\to\infty} f(x) = -\infty, \quad \lim_{x\to-\infty} f(x) = -\infty$$

se definen análogamente cambiando el sentido de las desigualdades y las cotas.

A diferencia de lo que ocurre con valores numéricos de un límite, ahora aparecen dificultades al expresar propiedades de conmutación de límites con valores infinitos y operaciones aritméticas, puesto que $\pm\infty$ no son números reales, sino símbolos que expresan la idea de cantidades que se hacen arbitrariamente grandes en valor absoluto. Es decir, en principio, no son números sino representaciones simbólicas de límites.

Bajo este punto de vista, la fórmula

$$R + \infty = \infty, \qquad R \in \mathbb{R}$$

se debe interpretar de la siguiente manera: si f, g son dos funciones tales que en un mismo punto a (posiblemente infinito)

$$\lim_{x\to a} f(x) = R, \quad \text{y} \quad \lim_{x\to a} g(x) = \infty, \quad \text{entonces} \quad \lim_{x\to a}\big(f(x) + g(x)\big) = +\infty.$$

Interpretadas de esta manera, se obtienen fórmulas donde aparecen operaciones aritméticas ampliadas con los símbolos $\pm\infty$. De hecho, en la teoría de límites se puede dar un sentido riguroso teórico a $\pm\infty$ como puntos de la llamada **recta real ampliada,** denotada por

$$\boxed{\overline{\mathbb{R}} = [-\infty, +\infty]}$$

y con este concepto, acercarse á $\pm\infty$ es como acercarse a cualquier punto $a \in \mathbb{R}$, unificando entonces la teoría de límite en cualquier **punto** $a \in [-\infty, +\infty]$. El problema es que cuando queremos considerar límites con **valores** en $[-\infty, +\infty]$, no hay una aritmética que funcione como la habitual y que incluya a los símbolos $\pm\infty$. Aún así, se pueden manipular los símbolos $\pm\infty$ aritméticamente en los límites, teniendo en cuenta un hecho fundamental:

No se pueden definir todas las operaciones aritméticas con ∞ de manera consistente

Las combinaciones de operaciones y símbolos cuyo resultado no está definido se denominan **indeterminaciones.**

La tabla de sumar con ∞

$+$	$-\infty$	R	∞
$-\infty$	$-\infty$	$-\infty$	•
R	$-\infty$	R	∞
∞	•	∞	∞

Clave	
R	un número real
•	una indeterminación

La tabla de multiplicar con ∞

$\cdot$	$-\infty$	N	0	P	∞
$-\infty$	∞	∞	•	$-\infty$	$-\infty$
N	∞	P	0	N	$-\infty$
0	•	0	0	0	•
P	$-\infty$	N	0	P	∞
∞	$-\infty$	$-\infty$	•	∞	∞

Clave	
N	un número real negativo
P	un número real positivo
•	una indeterminación

División por ∞

$$\frac{R}{\pm\infty} = 0$$

División por $0^{\pm}$

$$\frac{P}{0^+} = \infty \quad \frac{N}{0^+} = -\infty \quad \frac{P}{0^-} = -\infty \quad \frac{N}{0^-} = \infty \quad \frac{0}{0} = \bullet$$

Clave	
N	un número real negativo
0^-	límite nulo pero acercándose por números negativos
0^+	límite nulo pero acercándose por números positivos
P	un número real positivo
•	una indeterminación

Un límite nulo en el denominador sólo tendrá sentido cuando éste no sea nulo para los puntos de un entorno reducido, de ahí que pongamos las condiciones $0^{\pm}$ que, por otra parte, pueden ser perfectamente límites laterales o en infinito. Tener presente el "prototipo" $\lim_{x\to 0^{\pm}} 1/x = \pm\infty$.

El resultado acerca de "mover" los límites mediante la exponencial y el logaritmo se pueden traspasar a los límites de valor infinito, extendiendo la definición mediante $e^{-\infty} = 0$, $e^{\infty} = \infty$, $\log 0^+ = -\infty$, $\log\infty = \infty$. A las equivalencias enunciadas anteriormente se añaden:

Límites de valor infinito compuestos con la exponencial y logaritmo

Para una función $f(x)$ definida en un entorno reducido (posiblemente lateral) de un punto $a \in [-\infty, \infty]$, se tiene

$$\begin{aligned}
\lim_{x\to a} f(x) = \infty &\iff \lim_{x\to a} e^{f(x)} = \infty \\
\lim_{x\to a} f(x) = -\infty &\iff \lim_{x\to a} e^{f(x)} = 0^+ \\
\lim_{x\to a} f(x) = \infty &\iff \lim_{x\to a} \log f(x) = \infty \\
\lim_{x\to a} f(x) = 0^+ &\iff \lim_{x\to a} \log f(x) = -\infty \quad \text{si } f(x) > 0
\end{aligned}$$

Dicho de otra manera, las igualdades

$$\lim_{x\to a} f(x) = \log\Big(\lim_{x\to a} e^{f(x)}\Big)$$

$$\lim_{x\to a} f(x) = e^{\lim_{x\to a} \log f(x)} \quad \text{si } f(x) > 0$$

siguen siendo válidas como relaciones entre valores en la recta ampliada.

Demostración. Es el mismo razonamiento con la continuidad de la exponencial y el logaritmo que para los límites finitos, teniendo en cuenta sus límites en infinito. □

La observación anterior lleva a fórmulas que permiten manipular $\pm\infty$ en una potencia.

Potencias con ∞

- $\ell^{+\infty} = +\infty$ si $\ell \in \mathbb{R}, \ell > 1$
- $\ell^{-\infty} = 0$ si $\ell \in \mathbb{R}, \ell > 1$
- $\ell^{+\infty} = 0$ si $\ell \in \mathbb{R}, 0 < \ell < 1$
- $\ell^{-\infty} = +\infty$ si $\ell \in \mathbb{R}, 0 < \ell < 1$
- $(+\infty)^{\ell} = +\infty$ si $\ell \in \mathbb{R}, \ell > 0$
- $(+\infty)^{\ell} = 0$ si $\ell \in \mathbb{R}, \ell < 0$
- $(+\infty)^{+\infty} = +\infty$
- $(+\infty)^{-\infty} = 0$

Por ejemplo, $\ell^{\infty} = \infty$ para $\ell > 1$ es equivalente, tomando logaritmos, a $\infty \cdot \log \ell = \infty$, lo cual es cierto ya que $\log \ell > 0$.

Puede observarse que la lista de indeterminaciones es más corta que las de operaciones que sí están bien definidas. Por tanto, una vez asimiladas estas últimas como "de sentido común," es más sencillo tener en cuenta cuáles son las indeterminaciones.

Las indeterminaciones son las combinaciones de límites con operaciones aritméticas para las cuales no se puede decidir, sin aplicar otras técnicas, cuál puede ser el valor del límite, si es que existe.

Tanto 0 como ∞ son causantes de indeterminación, por ejemplo, es bien sabido que no está definida la división de un real por 0. Quitamos los signos en la siguiente tabla para simplificar las expresiones.

Tabla de indeterminaciones

Aditivas	Multiplicativas	Exponenciales
$\infty - \infty$	$0 \cdot \infty$	0^0
	$\frac{0}{0}$	∞^0
	$\frac{\infty}{\infty}$	1^∞

El lector observador puede constatar que tomando logaritmos (con signos adecuados), las indeterminaciones multiplicativas corresponden todas a la aditiva, y las exponenciales a la multiplicativa $0 \cdot \infty$.

• **Ejemplo:** Las expresiones

$$\frac{x^2}{x}, \quad \frac{x}{x^2}, \quad \frac{x}{x}$$

son indeterminaciones de tipo $0/0$ en $a = 0$, cada una con límite distinto: $0, \infty, 1$.

• **Ejemplo:** Que 1^∞ es indeterminación significa que

No existe ℓ tal que si $\lim_{x\to a} f(x) = 1, \lim_{x\to a} g(x) = \infty$, entonces $\lim_{x\to a} f(x)^{g(x)} = \ell$

suponiendo, claro está, que $f(x)^{g(x)}$ está definida en un entorno reducido de a. Lo que ocurre es que **no podemos determinar solamente con la información** $\lim_{x\to a} f(x) = 1, \lim_{x\to a} g(x) = \infty$ si existe o no el límite $\lim_{x\to a} f(x)^{g(x)}$ y cuál es su valor. Un ejemplo bien conocido de esto es el límite

$$e^r = \lim_{x\to\infty} \left(1 + \frac{r}{x}\right)^x$$

donde $f(x) = 1 + \frac{r}{x}$ cumple $\lim_{x\to\infty} f(x) = 1$, y $g(x) = x$ cumple $\lim_{x\to\infty} g(x) = \infty$, pero solo con esta información no se puede deducir el valor de $\lim_{x\to\infty} f(x)^{g(x)}$, que depende de la información adicional proporcionada por el valor de r.

2.7. Crecimiento de funciones elementales

Límites de funciones racionales en infinito

Consideremos una función racional, es decir, un cociente

$$r(x) = \frac{p(x)}{q(x)}$$

con $p(x)$ y $q(x)$ polinomios en x. Abreviamos el grado por "gr". Entonces

$$\lim_{x\to\infty} r(x) = \begin{cases} 0 & \textit{si } \mathrm{gr}(p) < \mathrm{gr}(q) \\ \infty & \textit{si } \mathrm{gr}(p) > \mathrm{gr}(q) \\ \dfrac{a}{b} & \textit{si } \mathrm{gr}(p) = \mathrm{gr}(q) \end{cases}$$

*donde a y b son los coeficientes de los términos de **mayor** grado en $p(x)$ y $q(x)$, respectivamente.*

Para calcular el límite de una función racional en $-\infty$, basta con hacer el cambio de variable $x \to -x$ y calcularlo en ∞.

• **Ejemplo:** Por ejemplo,

$$\lim_{x\to\infty} \frac{6x^3 + 2x + 1}{4x^7 + 3x^6 + 2} = 0, \quad \lim_{x\to\infty} \frac{2x^8 + 2x + 1}{2x^7 + 3x^6 + 2} = \infty, \quad \lim_{x\to\infty} \frac{6x^7 + 2x + 1}{2x^7 + 3x^6 + 2} = \frac{6}{2} = 3.$$

Por otra parte,

$$\lim_{x\to-\infty} \frac{6x^7 + 2x + 1}{2x^7 + 3x^6 + 2} = \lim_{x\to\infty} \frac{6(-x)^7 - 2x + 1}{2(-x)^7 + 3(-x)^6 + 2} = \lim_{x\to\infty} \frac{-6x^7 - 2x + 1}{-2x^7 + 3x^6 + 2} = 3.$$

El fenómeno descrito para funciones racionales extiende a combinaciones lineales de potencias, sean positivas o negativas, aunque éstas no sean números naturales.

Límites de combinaciones de potencias en ∞

Sean $p_1 < p_2 < \cdots < p_n$ y $q_1 < q_2 < \cdots < q_m$, y sean $a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_m \in \mathbb{R}$, con $a_n, b_m \neq 0$. Entonces

$$\lim_{x\to\infty} \frac{a_1 x^{p_1} + a_2 x^{p_2} + \cdots + a_n x^{p_n}}{b_1 x^{q_1} + b_2 x^{q_2} + \cdots + b_m x^{q_m}} = \begin{cases} 0 & \textit{si } p_n < q_m \\ \infty & \textit{si } p_n > q_m \\ \dfrac{a_n}{b_m} & \textit{si } p_n = q_m \end{cases}$$

• Ejemplo:

$$\lim_{x\to\infty} \frac{2\sqrt[4]{x} - 3\sqrt[3]{x} + 4\sqrt{x}}{5\sqrt[7]{x} + 6\sqrt[6]{x}} = \infty,$$

$$\lim_{x\to\infty} \frac{7\sqrt[4]{x} - 8\sqrt[3]{x} + 9\sqrt{x}}{10\sqrt[7]{x} - 11\sqrt[6]{x} + 12\,x} = 0,$$

$$\lim_{x\to\infty} \frac{13\sqrt[4]{x} - 14\sqrt[3]{x} + 15\sqrt{x}}{16\sqrt[7]{x} - 17\sqrt[6]{x} - 18\sqrt{x}} = -\frac{5}{6}.$$

Jerarquía del crecimiento en ∞ de las potencias, logaritmo y exponencial

Para cualquier potencia real p y constante $k > 0$,

$$\lim_{x\to\infty} \frac{x^p}{e^{kx}} = 0, \qquad \lim_{x\to\infty} \frac{e^{kx}}{x^p} = \infty, \qquad \lim_{x\to\infty} \frac{(\log x)^p}{x^k} = 0, \qquad \lim_{x\to\infty} \frac{x^k}{(\log x)^p} = \infty.$$

• Ejemplo: El resultado anterior significa que cualquier exponencial creciente, por pequeña que sea k, tiende a infinito mucho más rápidamente que cualquier potencia, por alta que ésta sea. Igualmente, cualquier potencia positiva, por baja que sea, tiende a infinito mucho más rápidamente que cualquier potencia del logaritmo, por alta que ésta sea. Por ejemplo,

$$\lim_{x\to\infty} \frac{x^{1000000}}{e^{0,0000001x}} = 0,$$

$$\lim_{x\to\infty} \frac{(\log x)^{1000000}}{x^{0,0000001}} = 0.$$

Eso sí, para valores "pequeños" de x, puede que se tarde un rato en observar este comportamiento...

2.7.1. Continuidad de las funciones elementales [1]

Continuidad de la exponencial

La exponencial es continua en $\mathbb{R}$.

Demostración. Sea $a > 0$. Veamos que e^x es continua en a. Trasladamos el límite a 0 y observamos que

$$\lim_{x\to a} e^x = e^a \iff \lim_{h\to 0} e^{a+h} = e^a$$

[1] Se puede omitir en una lectura rápida, dando por hecho el resultado, que es simplemente que las funciones elementales son continuas en sus dominios.

$$\begin{aligned}&\iff \lim_{h\to 0} e^a e^h = e^a\\ &\iff \lim_{h\to 0} e^h = 1\end{aligned}$$

ya que $e^a > 0$ para todo $a \in \mathbb{R}$.

Esto significa que basta con demostrar la continuidad en $a = 0$. Para ello consideramos conocida la fórmula

$$e^x = \lim_{n\to\infty}\left(1 + \frac{x}{n}\right)^n$$

y la fórmula para desarrollar un binomio. Primero suponemos que $x \geq 0$. Entonces

$$\begin{aligned}\left(1+\frac{x}{n}\right)^n &= 1 + n\cdot\frac{x}{n} + \frac{n(n-1)}{2}\cdot\frac{x^2}{n^2} + \cdots + \frac{n(n-1)\ldots(n-k+1)}{k!}\cdot\frac{x^k}{n^k} + \cdots + \frac{x^n}{n^n}\\ &= 1 + x + \frac{n(n-1)}{n^2}\cdot\frac{x^2}{2} + \cdots + \frac{n(n-1)\ldots(n-k+1)}{n^k}\cdot\frac{x^k}{k!} + \cdots + \frac{n!}{n^n}\cdot\frac{x^n}{n!}\end{aligned}$$

y como

$$\frac{n(n-1)\ldots(n-k+1)}{n^k} = \frac{n(n-1)\ldots(n-k+1)}{n\cdot n\ldots n} = \underbrace{\frac{n}{n}\cdot\frac{n-1}{n}\ldots\frac{n-k+1}{n}}_{k} < 1,$$

para $x > 0$ se tiene

$$1 + x < \left(1+\frac{x}{n}\right)^n < 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^k}{k!} + \cdots + \frac{x^n}{n!}$$

y, estimando menos exactamente, como cada denominador es menor o igual que 1,

$$1 + x < \left(1+\frac{x}{n}\right)^n < 1 + x + x^2 + \cdots + x^n.$$

Haciendo $n \to \infty$ queda, usando la fórmula para la **suma de una serie geométrica,**

$$1 + x \leq e^x \leq \frac{1}{1-x}, \qquad (0 < x < 1).$$

Haciendo que $x \to 0^+$ se obtiene $\lim_{x\to 0^+} e^x = 1$. Para el otro límite lateral, cambiando de variable por $x \to -x$, queda

$$\lim_{x\to 0^-} e^x = \lim_{x\to 0^+} e^{-x} = \lim_{x\to 0^+}\frac{1}{e^x} = \frac{1}{\lim_{x\to 0^+} e^x} = 1$$

por tanto $\lim_{x\to 0} e^x = 1$. □

Destacaremos por su utilidad posterior las desigualdades demostradas.

Estimaciones relacionadas con la exponencial

Para $x > 0$ se tiene

$$\begin{aligned}1 + x < \left(1+\frac{x}{n}\right)^n &< 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^k}{k!} + \cdots + \frac{x^n}{n!}\\ &< 1 + x + x^2 + \cdots + x^n.\end{aligned}$$

Continuidad del logaritmo

El logaritmo (en cualquier base) es continuo en $(0, \infty)$.

Demostración. En vez de hacer una demostración particularizada, se puede aplicar el resultado sobre la continuidad de la inversa de una función biyectiva continua, en este caso, la exponencial $x \to e^x : (-\infty, \infty) \to (0, \infty)$, con inversa el logaritmo neperiano $\log : (0, \infty) \to (-\infty, \infty)$. Como el logaritmo en base b es simplemente $\log_b(x) = \log(x)/\log(b)$, la continuidad de cualquier logaritmo es inmediata. □

Continuidad del seno y el coseno

Las funciones $\sin, \cos : \mathbb{R} \to [-1, 1]$ *son continuas.*

Demostración. Una manera elegante de deducir la continuidad de seno y coseno a partir de la continuidad de la exponencial consiste en usar la **Fórmula de Euler**

$$e^{ix} = \cos(x) + i\sin(x),$$

que tiene como consecuencias las expresiones

$$\cos(x) = \frac{e^{ix} + e^{-ix}}{2}, \quad \sin(x) = \frac{e^{ix} - e^{-ix}}{2i}.$$

Esta demostración requiere usar números complejos y justificar las fórmulas. □

También se podría apelar a la definición geométrica del seno y el coseno. Si C es la circunferencia unidad, de centro el origen y radio 1, entonces un punto $p \in C$ tiene coordenadas $p = (\cos\theta, \sin\theta)$ donde θ representa el ángulo polar medido en sentido antihorario desde el eje x positivo, que corresponde a $\theta = 0$. Es razonable pensar que un punto $p \in C$ varía de manera continua con θ, de lo cual obtendríamos la continuidad del seno y el coseno, pero esto no es una demostración.

Las demostraciones de continuidad de las funciones elementales dependen mucho de cómo se definen las funciones. Un método que sirve para todas es definir cada función mediante una serie de potencias. Esto requiere desarrollar de antemano la teoría de tales series.

2.8. La notación asintótica [2]

A menudo al analizar una función nos interesa dar una idea de cómo de grande se puede hacer, comparándola con otras funciones más sencillas cuyo crecimiento usamos como punto de referencia.

[2]Esta sección se puede omitir en una lectura rápida

En estas circunstancias, no interesa expresar todo el detalle inherente en un enunciado de límite, sino solamente la esencia de la comparación de una función con otra. Para ello se ha desarrollado una notación especial.

Orden de magnitud

Sean f, g funciones definidas en un entorno reducido del punto a, que puede ser también $\pm\infty$. Escribimos

$$f(x) = O(g(x)) \quad (x \to a) \quad \text{o respectivamente, } (x \to a^{\pm})$$

que se lee como "f es **"o" grande de** g cuando x tiende á a," ó "f está **acotada asintóticamente** por g," si existe una constante $M \geq 0$ tal que

$$|f(x)| \leq M|g(x)|$$

en un entorno reducido del punto a (respectivamente, $a^{\pm}$).

Recordamos aquí que

- "entorno (reducido o no) de ∞" se refiere a un intervalo de la forma (M, ∞).
- "entorno (reducido o no) de $-\infty$" se refiere a un intervalo de la forma $(-\infty, M)$.
- "entorno reducido de a^+" es un intervalo de la forma $(a, a + \delta)$.
- "entorno reducido de a^-" es un intervalo de la forma $(a - \delta, a)$.

Esta notación, llamada también "O-notación," expresa rigurosamente la idea de que el tamaño de f se puede acotar por el de g, o, en otras palabras, que f no es más grande que g, cerca del punto a.

Por ejemplo,

$$x^7 + 6x^5 + x^3 - 11x^2 + 17 = O(x^7) \quad (x \to \infty)$$

y, en general, para un polinomio

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0 = O(x^n) \quad (x \to \infty).$$

Para una función racional en general,

$$\frac{a_0 + a_1 x + \cdots + a_n x^n}{b_0 + b_1 x + \cdots + b_m x^m} = O(x^{n-m}) \quad (x \to \infty)$$

(aquí se suponen $a_n, b_m \neq 0$). Como puede observarse, estos enunciados son algo menos precisos que los enunciados correspondientes sobre los límites de estas funciones en ∞. Sin embargo, expresan la idea que la fórmula relativamente complicada de la izquierda se acota por la fórmula sencilla a la derecha, dando así una idea de la magnitud de la función "complicada."

El caso particular

$$f(x) = O(1) \quad (x \to a)$$

significa que f está acotada en un entorno (reducido) de a.

A menudo la notación O grande se emplea para indicar que se reemplaza una función por una aproximación más sencilla. Por ejemplo, más adelante veremos que

$$\sin x = x - \frac{x^3}{6} + O(x^5) \quad (x \to 0).$$

El término $O(x^5)$ se interpreta en este contexto como una **cota del error** cometido al reemplazar la función por su aproximación. Por ejemplo, si tenemos $x = 0{,}1$ entonces $x^5 = 0{,}00001$ y así nos hacemos una idea de la precisión que nos da la aproximación.

Hay que recordar que hay una **constante implícita** M en tales cotas ($|f| \leq M|g|$), así que sólo es una idea aproximada, pues la constante podría ser muy grande. De todas formas, a medida que $x \to a$ se hará más precisa la estimación.

Con funciones muy complicadas o ecuaciones derivadas empíricamente, puede que este tipo de estimaciones sea toda la información que tengamos.

Notación infinitesimal

Sean f, g funciones definidas en un entorno reducido del punto a, que puede ser también $\pm\infty$. Escribimos

$$f(x) = o(g(x)) \quad (x \to a) \quad \text{o respectivamente, } (x \to a^{\pm})$$

que se lee como "f es **"o" pequeña de** g cuando x tiende á a," ó "f es **infinitesimal respecto á** g," si, dada **cualquier** cota $\varepsilon > 0$, se tiene

$$|f(x)| \leq \varepsilon |g(x)|$$

en algún entorno reducido del punto a (respectivamente, $a^{\pm}$). Si $g(x) \neq 0$, esto equivale a que

$$\lim_{x \to a} \frac{f(x)}{g(x)} = 0, \quad \text{respectivamente,} \quad \lim_{x \to a^{\pm}} \frac{f(x)}{g(x)} = 0.$$

Esto expresa la idea que f es mucho más pequeña o "descartable" comparada con g:

$$x^2 = o(x) \quad (x \to 0), \qquad x = o(x^2) \quad (x \to \infty).$$

Antes hemos visto, al comparar el crecimiento de las funciones elementales en ∞, que

$$x^p = o(e^{kx}), \quad (\log x)^p = o(x^k) \quad (x \to \infty), \quad \forall\, p \in \mathbb{R},\ k > 0.$$

El caso particular

$$f(x) = o(1) \quad (x \to a)$$

significa que $\lim_{x \to a} f(x) = 0$.

Notación de equivalencia asintótica

Sean f, g funciones definidas en un entorno reducido del punto a, que puede ser también $\pm\infty$, con $f(x), g(x) \neq 0$ en ese entorno. Escribimos

$$f(x) \sim g(x) \quad (x \to a) \quad \text{o respectivamente, } (x \to a^{\pm})$$

que se lee como "f es **asintótica á** g" o "asintóticamente igual/equivalente á g", si

$$\lim_{x \to a} \frac{f(x)}{g(x)} = 1, \quad \text{respectivamente,} \quad \lim_{x \to a^{\pm}} \frac{f(x)}{g(x)} = 1.$$

Se puede relajar la condición de que f y g sean no nulas en la definición de equivalencia asintótica si se toma como definición ampliada

$$f \sim g \iff f - g = o(g) \quad (x \to a, a^{\pm}).$$

Con esta definición más general, $\sim$ sigue siendo una relación de equivalencia, que es la misma que la anterior en el caso de ser f, g no nulas en un entorno reducido.

Por ejemplo,

$$\sin x \sim x \quad (x \to 0)$$

que es simplemente otra manera de expresar el resultado

$$\lim_{x \to 0} \frac{\sin x}{x} = 1.$$

También se tiene

$$\log(1+x) \sim x \quad (x \to 0)$$

pues

$$\lim_{x \to 0} \frac{\log(1+x)}{x} = \lim_{x \to 0} \log(1+x)^{1/x} = \log \lim_{x \to 0} (1+x)^{1/x} = \log \lim_{t \to \pm\infty} \left(1 + \frac{1}{t}\right)^t = \log e = 1.$$

Otro ejemplo es

$$\frac{1 + x - 2x^2 + 3x^3}{5 - 8x - 13x^2 + 21x^3} \sim \frac{3}{21} = \frac{1}{7} \quad (x \to \infty)$$

y en general, para funciones racionales,

$$\frac{a_0 + a_1 x + \cdots + a_n x^n}{b_0 + b_1 x + \cdots + b_n x^n} \sim \frac{a_n}{b_n} \quad (x \to \infty)$$

suponiendo $a_n, b_n \neq 0$.

Uso en solitario de expresiones asintóticas

Es frecuente ver los símbolos $O(g)$ u $o(g)$ como parte de una ecuación de aproximación en un punto a. La interpretación de $O(g)$ es "una función f cualquiera que está acotada asintóticamente por g cuando $x \to a$." Por ejemplo, $O(x)$ representa una función f no especificada que cumple $|f(x)| \leq Mx$ en un entorno reducido de a. De manera análoga, un término $o(1)$ representa una función que tiende á 0 cuando $x \to a$.

Al igual que se hace con los símbolos $\pm\infty$ y los límites, se pueden dar algunas reglas de manipulación de la O-notación, que representan versiones concisas de enunciados sobre acotaciones o límites y, también como ocurre con las indeterminaciones, no hay que confundir estas reglas con identidades aritméticas entre números, pues no lo son. Es una notación, que puede ser más o menos lógica y razonable, y que sirve para resumir ciertos fenómenos que se encuentran frecuentemente al comparar funciones.

Manipulación algebraica de la O-notación

Considerando las siguientes expresiones cuando $x \to a$ para un punto fijado a, se tiene

- $kO(g) = O(g)$ *si* k *es una constante,* $k \neq 0$.
- $O(g_1) \cdot O(g_2) = O(g_1 g_2)$
- $O(g_1) + O(g_2) = O(\max(|g_1|, |g_2|))$, *en particular,* $O(g) + O(g) = O(g)$
- $ko(g) = o(g)$ *si* k *es una constante,* $k \neq 0$.
- $o(g_1) \cdot o(g_2) = o(g_1 \cdot g_2)$
- $o(g_1) + o(g_2) = o(\max(|g_1|, |g_2|))$, *en particular,* $o(g) + o(g) = o(g)$
- $o(g) = O(g)$
- $o(g) \cdot O(h) = o(g \cdot h)$
- $O(o(f)) = o(O(f)) = o(f)$

Advertencia sobre la interpretación de la O-notación

Como puede observarse en la tabla anterior, no se trata de verdaderas identidades aritméticas entre números, sino de abreviaciones para relaciones entre funciones. Por ejemplo

$$O(g) + O(g) = O(g)$$

no significa desde luego que $O(g) = 0$, sino que "la suma de dos funciones acotadas asintóticamente por g es otra función acotada asintóticamente por g." Tampoco la relación

$$o(g) = O(g)$$

significa que sean la misma cosa ambos símbolos, pues no es una relación simétrica. Se lee de izquierda a derecha y significa que "una función que es $o(g)$ también es $O(g)$." De hecho, a veces se usa la notación de pertenencia a un conjunto,

$$\text{``}f \in O(g)''$$

para expresar mejor la idea que conlleva, en vez de la de igualdad, $f = O(g)$, pero la tradición y el uso emplean mayoritariamente la igualdad, por tanto es muy importante tener claras cuáles son las interpretaciones correctas.

• **Ejemplo:** Entendida correctamente, la O-notación revela las razones generales detrás de resultados particulares como $\lim_{x\to 0^+} x\sin(1/x) = 0$, que se puede considerar como un caso particular de la relación $o(1)\cdot O(1) = o(1)$, cuya interpretación es "una función que tiende á 0 multiplicada por otra acotada, da como resultado una función que tiende a 0."

Propiedades de la equivalencia asintótica

Supongamos que las funciones estudiadas son **positivas.** Entonces la equivalencia asintótica es una relación de equivalencia. Es decir, para $x \to a$ con a fijado,

- $f \sim f$ (reflexiva)
- $f \sim g \iff g \sim f$ (simétrica)
- $f \sim g,\ g \sim h \implies f \sim h$ (transitiva)

Además,

- $f \sim g \implies kf \sim kg$ si $k \in \mathbb{R}, k \neq 0$
- $f \sim g \implies \dfrac{1}{f} \sim \dfrac{1}{g}$
- $f_1 \sim g_1,\ f_2 \sim g_2 \implies f_1 + f_2 \sim g_1 + g_2$
- $f_1 \sim g_1,\ f_2 \sim g_2 \implies f_1 f_2 \sim g_1 g_2$
- $f_1 \sim g_1,\ f_2 \sim g_2 \implies \dfrac{f_1}{f_2} \sim \dfrac{g_1}{g_2}$

. Resaltamos la prueba para sumas, que depende de la desigualdad

$$\text{mín}\left(\frac{a}{A},\frac{b}{B}\right) \le \frac{a+b}{A+B} \le \text{máx}\left(\frac{a}{A},\frac{b}{B}\right), \qquad a,b,A,B>0$$

La fracción del medio se llama el **mediante** entre a/A y b/B[3]. El resultado para sumas es consecuencia de esta desigualdad y del principio del sandwich para límites. Las demás son consecuencias directas de las propiedades básicas de los límites. Por ejemplo, para la transitividad, basta con poner $f/h = (f/g)\cdot(g/h)$. □

No es verdad en general que $f_1 \sim g_1,\ f_2 \sim g_2 \implies f_1 - f_2 \sim g_1 - g_2$. Para empezar, haría falta $g_1 - g_2 \neq 0$. Aún así, por ejemplo, para $x \to \infty$,

$$x+1 \sim x+2, \quad x \sim x, \quad \text{pero} \quad 1 \not\sim 2.$$

Por este motivo hay que poner especial cuidado con las equivalencias asintóticas en expresiones donde aparezcan restas.

Cambio de variable con la notación asintótica

Si $y=\varphi(x)$ lleva entornos reducidos (suficientemente pequeños) de a á entornos reducidos de b, o sea, para puntos finitos,

$$\forall\,\varepsilon>0\ \exists\,\delta>0 : 0<|x-a|<\delta \implies 0<|\varphi(x)-b|<\varepsilon$$

entonces

- *$g(y)\in O(f(y))$ cuando $y\to b$ implica $g(\varphi(x))\in O(f(\varphi(x)))$ cuando $x\to a$.*
- *$g(y)\in o(f(y))$ cuando $y\to b$ implica $g(\varphi(x))\in o(f(\varphi(x)))$ cuando $x\to a$.*
- *$g(y)\sim f(y)$ cuando $y\to b$ implica $g(\varphi(x))\sim f(\varphi(x))$ cuando $x\to a$.*

Demostración. Simbólicamente, la condición de llevar entornos reducidos a entornos reducidos es $x\approx a^* \implies y=\varphi(x)\approx b^*$. Si $g(y)\in O(f(y))$ cuando $y\to b$, entonces para $y\approx b^*$ se tiene $|g(y)|\le C|f(y)|$ para alguna constante $C>0$. Por tanto $x\approx a^* \implies y=\varphi(x)\approx b^* \implies |g(\varphi(x))|\le C|f(\varphi(x))|$, o sea $g(\varphi(x))\in O(f(\varphi(x)))$ cuando $x\to a$. El razonamiento para las otras condiciones es igual. □

En particular, si a,b son puntos finitos y φ es continua e invertible en un entorno de a, con $\varphi(a)=b$, cumplirá la condición de llevar entornos reducidos (suficientemente pequeños) de a á entornos reducidos de b.

[3] http://es.wikipedia.org/wiki/Mediante_(matemáticas)

3. Aspectos teóricos de la continuidad [4]

3.1. Conservación del signo

Conservación del signo de una función continua

Sea $f(x)$ continua en a con $f(a) \neq 0$. Entonces hay un entorno de a en el cual el signo de $f(x)$ es el mismo que el signo de $f(a)$. Es decir,

$$\operatorname{sgn} f(x) = \operatorname{sgn} f(a) \quad (x \approx a)$$

Demostración. Es consecuencia de la propiedad correspondiente para límites pues como $f(a) = \lim_{x\to a} f(x)$, si $f(a) \neq 0$ entonces para $x \approx a^*$ (entorno reducido) el signo de $f(x)$ es el mismo que el de $f(a)$, y como evidentemente esto también es cierto para $x = a$, podemos poner $x \approx a$ en vez de $x \approx a^*$. □

Conservación de desigualdades continuas

Si f y g son continuas en a y $f(a) < g(a)$ entonces $f(x) < g(x)$ para $x \approx a$. En particular, si m, M son constantes y $m < f(a) < M$ entonces $m < f(x) < M$ para $x \approx a$.

Demostración. La diferencia $h(x) = f(x) - g(x)$ es continua en a y $h(a) < 0$, por tanto $h(x) < 0$ para $x \approx a$, es decir, $f(x) < g(x)$. □

Se puede observar que estos resultados expresan un fenómeno del tipo "si se da una condición en un punto, se da en un entorno de ese punto." En estos casos, la condición es una desigualdad, y lo expresamos diciendo que se conserva o también que se propaga de un punto a un entorno. Esta es una característica típica de las funciones continuas.

Conservación del signo en un límite continuo.

Si $f(x)$ es continua en a y $f(x) \geq 0$ a un lado de a, es decir, en algún intervalo $(a, a+\delta)$ ó $(a-\delta, a)$, entonces $f(a) = \lim_{x\to a} f(x) \geq 0$. Si $f(x) \leq 0$ a un lado de a, entonces $f(a) \leq 0$.

Demostración. También es consecuencia del resultado correspondiente para límites. □

[4] Esta sección es teórica. Las demostraciones se pueden omitir en una lectura rápida

3.2. El Teorema de Bolzano

Teorema de Bolzano

Sea $f(x)$ continua en el intervalo compacto $[a, b]$ y con ***signos opuestos*** *en a y en b, es decir, $f(a)f(b) < 0$. Entonces existe algún punto $c \in (a, b)$ tal que $f(c) = 0$.*

• **Ejemplo:** Sea $f(x) = x^3 - x - 1$. Se tiene $f(1) = -1$ y $f(2) = 5$. Como los signos son opuestos, el Teorema de Bolzano implica que en algún punto c entre 1 y 2 se tiene $f(c) = 0$.

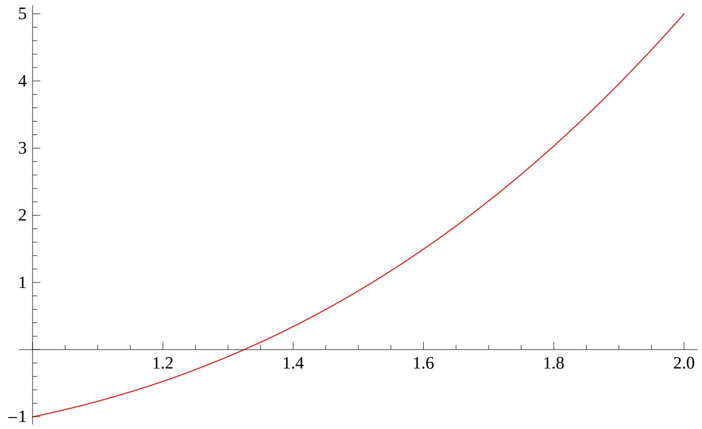

Figura 1 – La función $x^3 - x - 1$ en el intervalo $[1, 2]$.

Un punto c donde la función f se anula,

$$f(c) = 0$$

se llama un **cero** de f o una **raíz** o **solución** de la ecuación $f = 0$. Así pues el Teorema de Bolzano es un teorema de **existencia de soluciones de ecuaciones.**

Nota: Si debilitamos la hipótesis en el Teorema de Bolzano á $f(a)f(b) \leq 0$, entonces se concluye que existe algún $c \in [a, b]$ con $f(c) = 0$: c podría ser uno de los extremos. Efectivamente, o bien $f(a)f(b) < 0$, que es la hipótesis original, o bien $f(a)f(b) = 0$, en cuyo caso $f(a) = 0$ ó $f(b) = 0$ y sirve tomar $c = a$ ó $c = b$ respectivamente.

El Teorema de Bolzano puede dejar de ser válido si la función es discontinua. Por ejemplo,

$$f(x) = \begin{cases} 1 & \text{si } x \geq 0 \\ -1 & \text{si } x < 0 \end{cases}$$

tiene signos opuestos en $a = -1$ y $a = 1$, pero no hay ningún punto c con $f(c) = 0$. Esto expresa la idea que una función continua "no pega saltos."

Demostración del Teorema de Bolzano utilizando el método de bisección. Una demostración que además es algorítmica consiste en subdividir el intervalo $[a,b]$ sucesivamente en mitades de igual longitud y estudiar el patrón de signos en cada mitad.

Supondremos, sin pérdida de generalidad, que el patrón de signos inicial, en a, b, es $-, +$.

- Se pone $a_0 = a$, $b_0 = b$.
- Se toma el punto medio $c_0 = \frac{a_0+b_0}{2}$, que subdivide $[a,b] = \left[a, \frac{a+b}{2}\right] \cup \left[\frac{a+b}{2}, b\right]$.
- Si $f(c_0) = 0$, hemos acabado.
- Si $f(c_0) \neq 0$, entonces el patrón de signos en a_0, c_0, b_0 es uno de estos dos:
  - $-, +, +$: entonces ponemos $I_1 = [a_1, b_1] = [a_0, c_0]$ (mitad izquierda).
  - $-, -, +$: entonces ponemos $I_1 = [a_1, b_1] = [c_0, b_0]$ (mitad derecha).

Es decir, elegimos la mitad en cuyos extremos se reproduce el patrón original de signos. La longitud de I_1 es $(b-a)/2$.

Aplicando el mismo razonamiento al intervalo I_1, o bien hallamos que $f(c_1) = 0$ en su punto medio c_1, ó se obtiene un subintervalo I_2 de longitud $(b-a)/4$ que reproduce en sus extremos el patrón de signos original.

Repitiendo este proceso, o bien se acaba tras un número finito de pasos, siendo c igual al punto medio de uno de los intervalos, o bien continúa sin fin, generando una sucesión infinita de intervalos $I_n = [a_n, b_n]$ cuya longitud es $(b-a)/2^n$ y con $f(a_n) < 0$, $f(b_n) > 0$ para todo n.

Por la propiedad de los intervalos encajados de los números reales $\mathbb{R}$, en el caso infinito la sucesión de intervalos se reduce a un punto $c \in [a,b]$. Es decir, los extremos de los intervalos I_n tienen un límite común $\lim_n a_n = \lim_n b_n = c$.

La continuidad de f implica que

- $f(c) = f(\lim_n a_n) = \lim_n f(a_n) \leq 0$
- $f(c) = f(\lim_n b_n) = \lim_n f(b_n) \geq 0$

Dado que $f(c) \geq 0$ y $f(c) \leq 0$ simultáneamente, debe ser $f(c) = 0$. □

Esta demostración, básica en el Análisis, da lugar a los **algoritmos de bisección** para hallar soluciones de una ecuación $f(x) = 0$ con f continua; por ejemplo, de ecuaciones polinómicas de cualquier grado. La comparación $f = 0?$ es solamente aproximada, determinada por la precisión elegida, y el algoritmo siempre acaba al alcanzar esta precisión.

Debe tenerse en cuenta que el punto de partida, dos puntos el los cuales f tiene signos distintos, no siempre se va a dar, con lo cual se requiere una búsqueda previa para encontrar tales puntos, o abandonar tras un número finito de pasos si no se hallan.

Demostración del Teorema de Bolzano mediante ínfimos y supremos. Si usamos la propiedad que todo conjunto no vacío de reales tiene una máxima cota inferior (ínfimo) y mínima cota superior (supremo), se puede razonar de la siguiente manera. Suponemos, sin pérdida de generalidad, que $f(a) > 0$ y $f(b) < 0$. Sea

$$\alpha = \text{ínf}\{x \in [a,b] : f(x) < 0\}.$$

Veamos que α es el primer cruce de la gráfica de f con el eje x. Intuitivamente, es el primer (menor) punto donde $f(x)$ cambia de signo (pero el cambio puede ser oscilante a la derecha).

Como α es el ínfimo, es un límite de puntos x donde $f(x) < 0$; la continuidad de f implica que $f(\alpha) \leq 0$. Esto implica que $\alpha > a$ ya que $f(a) > 0$.

Por ser α cota inferior, si $f(x) < 0$, entonces $x \geq \alpha$. Por tanto $f(x) \geq 0$ para $a \leq x < \alpha$. Es decir, $f > 0$ a la izquierda de α. Haciendo $x \to \alpha^-$, la continuidad de f implica $f(\alpha) \geq 0$.

Como $f(\alpha) \leq 0$ y $f(\alpha) \geq 0$, debe ser $f(\alpha) = 0$.

De manera similar, si

$$\beta = \sup\{x \in [a,b] : f(x) > 0\}$$

también $f(\beta) = 0$. Es el último cruce de la gráfica de f con el eje x, el último (mayor) punto donde $f(x)$ cambia de signo. □

Esta segunda demostración es de existencia; no dice cómo se puede hallar un punto c con $f(c) = 0$; parece decir "este tipo de punto existe porque existe uno que es el mas pequeño de ellos," dando un aire circular a la demostración (aunque no lo es).

Suele considerarse preferible una demostración constructiva, como la primera que indica un proceso (la bisección) mediante el cual uno puede construir (por aproximación en este caso) el punto buscado.

Constancia del signo

Sea f continua en un intervalo I (de cualquier tipo) y tal que $f(x) \neq 0$ para todo $x \in I$. Entonces $f(x)$ no cambia de signo en I; es decir, o bien

- $f(x) > 0$ *para todo* $x \in I$, *ó*
- $f(x) < 0$ *para todo* $x \in I$.

Demostración. Si $f(x)$ cambiara de signo, habría dos puntos $a, b \in I$ tales que los signos de $f(a)$ y $f(b)$ son opuestos. Aplicando el Teorema de Bolzano á f sobre el subintervalo $[a,b] \subseteq I$, se concluye que existiría $c \in (a,b)$ con $f(c) = 0$. Esta contradicción prueba el resultado. □

El enunciado anterior es el contrapositivo del Teorema de Bolzano, junto con la observación sencilla de que también es válido para cualquier tipo de intervalo.

3.3. El Teorema del Valor Intermedio

Teorema del Valor Intermedio

Sea f continua en el intervalo $[a, b]$. Dado cualquier número λ entre los valores $f(a)$ y $f(b)$, existe $c \in [a, b]$ tal que $f(c) = \lambda$.

Demostración. Es de hecho equivalente al Teorema de Bolzano, que es el caso particular $\lambda = 0$. Sin pérdida de generalidad, $f(a) \leq f(b)$, de modo que $f(a) \leq \lambda \leq f(b)$. Consideramos $\varphi(x) = f(x) - \lambda$, que es también continua. Entonces $\varphi(a) = f(a) - \lambda \leq 0$ y $\varphi(b) = f(b) - \lambda \geq 0$. Por el Teorema de Bolzano, existe $c \in [a, b]$ tal que $\varphi(c) = 0$, o sea, $f(c) = \lambda$. □

3.4. El Teorema del Punto Fijo

Teorema del Punto Fijo

Sea f una función continua que aplica el intervalo $[a, b]$ en sí mismo: $f : [a, b] \to [a, b]$. Existe un punto $c \in [a, b]$ tal que $f(c) = c$.

Demostración. Es una consecuencia del Teorema de Bolzano. Como f aplica $[a, b]$ en sí mismo, se tiene $a \leq f(x) \leq b$ para todo x con $a \leq x \leq b$. En particular, $a \leq f(a), f(b) \leq b$.

Consideramos $\varphi(x) = f(x) - x$, que es continua, con $\varphi(a) = f(a) - a \geq 0$ y $\varphi(b) = f(b) - b \leq 0$. Por tanto, existe $c \in [a, b]$ con $\varphi(c) = 0$, es decir, $f(c) = c$. □

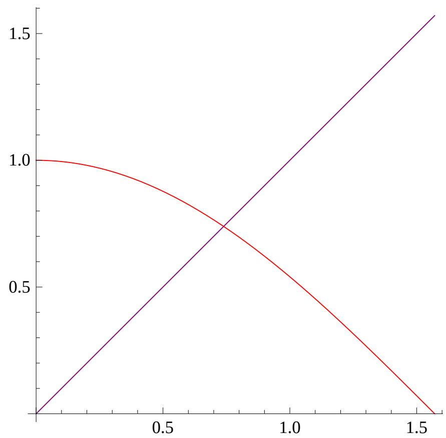

Figura 2 – Un punto fijo de f es un punto de corte de la gráfica de f con la recta $y = x$.

3.5. Acotación

Decimos que una función f definida en $D \subseteq \mathbb{R}$ está **acotada** en D si existe $M > 0$ tal que $|f(x)| \leq M$ para todo $x \in D$. Se dice también que f es una **función acotada** en D.

Por ejemplo, $|\sin x|, |\cos x| \leq 1$ para todo $x \in \mathbb{R}$, luego el seno y el coseno son funciones acotadas en $\mathbb{R}$. En cambio, $f(x) = 1/x$ no está acotada en $(0, 1)$.

f está **acotada superiormente** en D si existe M tal que $f(x) \leq M$ para todo $x \in D$, y **acotada inferiormente** si existe m tal que $m \leq f(x)$ para todo $x \in D$.

Por ejemplo, $1/x$ está acotada inferiormente por 0 en $(0, \infty)$, porque $1/x > 0$, pero no está acotada superiormente, porque $\lim_{x \to 0^+} 1/x = \infty$.

f está acotada si y sólo si está acotada superiormente y acotada inferiormente.

Si f está acotada en un dominio D, está acotada en cualquier subdominio $S \subseteq D$.

Las funciones continuas están localmente acotadas

Sea f continua en a. Entonces f está acotada en un entorno de a.

Demostración. Elegimos cualquier valor $\varepsilon > 0$. Existe $\delta > 0$ tal que $|x - a| < \delta \implies |f(x) - f(a)| < \varepsilon$. Esto significa que $f(c) - \varepsilon < f(x) < f(c) + \varepsilon$ para todo $x \in (a - \delta, a + \delta)$, por tanto f está acotada en el entorno $(a - \delta, a + \delta)$. □

Las funciones continuas en un intervalo compacto están acotadas.

Sea f continua en un intervalo compacto $[a, b]$. Entonces f está acotada en $[a, b]$.

Demostración. Razonamos por contradicción. Supongamos que f es continua en $[a, b]$ pero **no** está acotada. Entonces no está acotada en alguna de las dos mitades en las que $\frac{a+b}{2}$ divide á $[a, b]$. Procediendo recursivamente, se genera una sucesión infinita de intervalos $I_0 = [a, b] \supseteq I_1 = [a_1, b_1] \supseteq \cdots \supseteq I_n = [a_n, b_n]$ tales que f no está acotada en I_n y la longitud de I_n es $(b - a)/2^n$. Por tanto existe un punto $c \in [a, b]$ tal que $\lim_n a_n = \lim_n b_n = c$.

Como f es continua en c, está acotada cerca de c, o sea, en algún intervalo $J = (c - \delta, c + \delta)$. Pero como $a_n, b_n \to c$, se tendrá $c - \delta < a_n < b_n < c + \delta$ si n es suficientemente grande, o sea $I_n \subseteq J$. Por tanto f también está acotada en I_n, lo cual contradice lo anterior.

De esta contradicción resulta que f está acotada en $[a, b]$. □

3.6. Extremos: el Teorema de Weierstrass

Sea $f(x)$ una función definida en $D \subseteq \mathbb{R}$. Un punto $c \in D$ es un

- **máximo global de f en** D si $f(x) \leq f(c)$ para todo $x \in D$.
- **mínimo global de f en** D si $f(x) \geq f(c)$ para todo $x \in D$.

Se dice que $f(x)$ **tiene máximo,** o **alcanza** o **asume** su máximo sobre D, si existe al menos un máximo global de f sobre D. De manera similar se dice que $f(x)$ tiene mínimo sobre D, si existe al menos un mínimo global de f sobre D.

Los máximos y mínimos de una función se denominan sus **extremos,** y los valores correspondientes sus **valores extremos.** El problema de encontrar los máximos y mínimos de una función dada se conoce como **optimización.**

No debe confundirse **máximo** con **valor máximo** ni **mínimo** con **valor mínimo.** Máximo y mínimos son miembros del **dominio** de la función, y puede haber más de uno, incluso infinitos, mientras que los valores máximo y mínimo, si existen, son miembros del **recorrido** de la función, y son **únicos.**

• **Ejemplo:** $f(x) = \sin x$ tiene máximos en los puntos $x = \frac{\pi}{2} + 2\pi n$, donde $n \in \mathbb{Z}$, y mínimos en $x = -\frac{\pi}{2} + 2\pi n$, donde $n \in \mathbb{Z}$. Su valor máximo es 1 y su valor mínimo es -1.

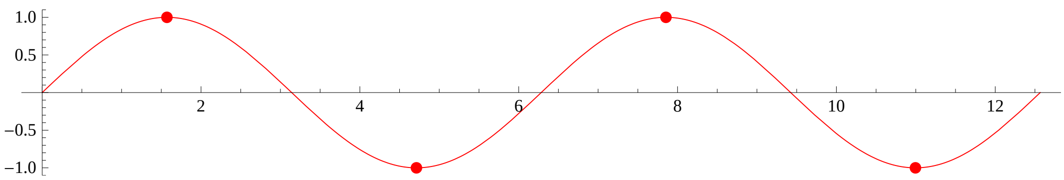

Figura 3 – Parte de la gráfica de $\sin x$ mostrando máximos y mínimos.

Una función continua en general no tiene por qué tener máximos y mínimos globales. Por ejemplo, $f(x) = x$ en $D = \mathbb{R}$ no los tiene. Sin embargo en $D = [0,1]$ sí los tiene. Vemos así que también depende del dominio de la función.

Teorema de Weierstrass

Sea f continua en un intervalo compacto $[a,b]$. Entonces f tiene al menos un máximo y al menos un mínimo sobre $[a,b]$. Es decir, f alcanza sus valores extremos sobre $[a,b]$.

Demostración del Teorema de Weierstrass por bisección. Si f es continua en $[a,b]$, entonces está acotada. Consideremos la mayor cota inferior y la menor cota superior de los valores

$f(x)$ con $x \in [a, b]$, es decir,

$$m = \text{ínf}\{f(x) : x \in [a, b]\}, \qquad M = \sup\{f(x) : x \in [a, b]\}.$$

Como f está acotada, estos son números reales, no infinitos: $-\infty < m \leq M < \infty$.

Podemos proceder una vez más por bisección para encontrar un $c \in [a, b]$ tal que $f(c) = M$, es decir, tal que c sea un máximo de f sobre $[a, b]$. Dividiendo $I_0 = [a, b]$ en mitades, en al menos una de ellas, digamos I_1, el supremo de f sobre I_1 debe ser el mismo que sobre I_0, es decir, M.

Por el proceso de bisección, se genera una sucesión infinita de intervalos encajados $I_n = [a_n, b_n]$ tales que $\sup_{x \in I_n} f(x) = M$. Sea $c = \text{lím}_n a_n = \text{lím}_n b_n$. Entonces debe ser $f(c) = M$, pues ciertamente $f(c) \leq M$ ya que M es el supremo, pero si fuera $f(c) < M$, entonces de hecho $f(c) < M - \varepsilon$ para algún $\varepsilon > 0$, y por tanto en algún intervalo $c - \delta < x < c + \delta$ sería $f(x) < M - \varepsilon$. Pero entonces, si n es suficientemente grande, como $c - \delta < a_n < b_n < c + \delta$, se tiene $\sup_{x \in I_n} f(x) \leq M - \varepsilon$. Esto contradice que $\sup_{x \in I_n} f(x) = M$. Por tanto $f(c) = M$.

La existencia de un mínimo se demuestra de manera similar. □

El Teorema de Weierstrass es fundamental porque proporciona dos condiciones sencillas **(continuidad de la función y compacidad del intervalo)** para que el problema de optimización tenga solución, pero es un teorema de existencia, **no constructivo.**

Un **método para hallar los extremos** es precisamente una de las aplicaciones fundamentales del Cálculo Diferencial.

Quitando cualquiera de las dos condiciones, continuidad o compacidad, el resultado puede dejar de ser cierto.

Por ejemplo, $f(x) = \frac{1}{x}$ sobre $(0, \infty)$ es una función continua en un intervalo no compacto, y no tiene extremos globales sobre este intervalo.

En cambio la función f definida por

$$f(x) = \begin{cases} \tan(x) & x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \\ 0 & x = \pm\frac{\pi}{2} \end{cases}$$

es una función discontinua en el intervalo compacto $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, sin extremos globales.

Una función continua en un intervalo abierto no tiene por qué tener extremos globales, pero hay condiciones sobre los extremos del intervalo que sí implican la existencia de un extremo global, por ejemplo, cuando la función tiende a infinito al acercarse a ellos (ver la figura 4).

Existencia de extremos globales para funciones que tienden a infinito

Sea f continua en un intervalo abierto (a, b) (posiblemente a, b sean infinitos).

- *Si* $\lim_{x\to a^+} f(x) = \lim_{x\to b^-} f(x) = +\infty$, *$f$ tiene un mínimo global en (a, b).*
- *Si* $\lim_{x\to a^+} f(x) = \lim_{x\to b^-} f(x) = -\infty$, *$f$ tiene un máximo global en (a, b).*

Demostración. Basta con demostrar el primero y aplicarlo a $-f$ para obtener el segundo. Supongamos entonces que $\lim_{x\to a^+} f(x) = \lim_{x\to b^-} f(x) = +\infty$.

Sea $c \in (a, b)$ cualquier punto. La infinitud de los límites laterales significa que existe $\delta > 0$ tal que $f(c) < f(x)$ si $(a, a + \delta) \cup (b - \delta, b)$. Suponemos δ suficientemente pequeño para que $a + \delta < b - \delta$, es decir, $\delta < \frac{b-a}{2}$.

Como $f(c) < f(x)$ para $x \in (a, a + \delta) \cup (b - \delta, b)$, debe ser $c \in [a + \delta, b - \delta]$. Por el Teorema de Weierstrass, existe $p \in [a + \delta, b - \delta]$ tal que $f(p) \leq f(x)$ para todo $x \in [a + \delta, b - \delta]$. En particular $f(p) \leq f(c)$. Como $f(p) \leq f(c) < f(x)$ para $x \in (a, a + \delta) \cup (b - \delta, b)$ se tiene $f(p) \leq f(x)$ para todo $x \in (a, b)$. □

• **Ejemplo:** La función $f(x) = x^2$ en $\mathbb{R}$ cumple $\lim_{x\to\pm\infty} f(x) = x^2$, y efectivamente, $c = 0$ es un mínimo global, ya que $f(c) = 0 \leq x^2 = f(x)$ para todo $x \in \mathbb{R}$.

• **Ejemplo:** Del mismo modo, cualquier polinomio $p(x)$ de grado par tiene un mínimo global, pues igualmente $\lim_{x\to\pm\infty} p(x) = \infty$ (dado que conocemos su crecimiento en infinito.)

Demostrar que los polinomios de grado impar no tienen extremos globales sobre $\mathbb{R}$.

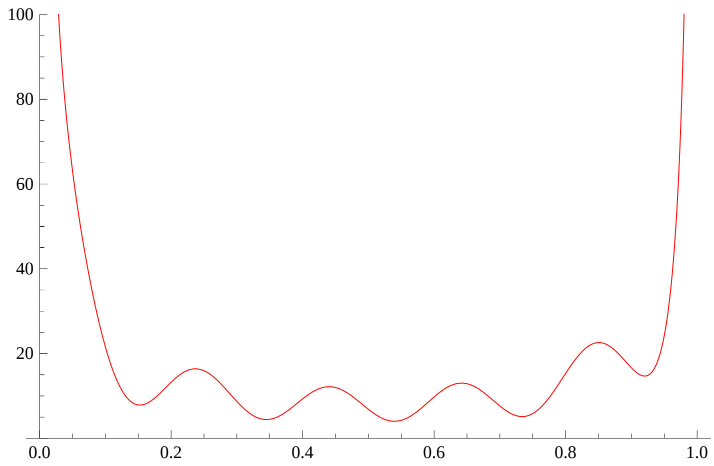

Figura 4 – Una función continua que tiende a ∞ en los extremos de un intervalo tiene un mínimo global sobre él.

3.7. La imagen de una función continua

Juntando todos los resultados anteriores se llega a la siguiente conclusión:

La imagen continua de un intervalo

Sea $f(x)$ continua en el intervalo compacto $[a, b]$. Entonces existen $\alpha, \beta \in [a, b]$ tales que

$$m \overset{\text{def}}{=} f(\alpha) \leq f(x) \leq f(\beta) \overset{\text{def}}{=} M \quad \forall\, x \in [a, b]$$

y la imagen directa del intervalo $[a, b]$ bajo f es el intervalo $[m, M]$.

Demostración. El Teorema de Weierstrass implica que f tiene un mínimo α y un máximo β sobre $[a, b]$. Sean m, M los valores mínimo y máximo. La condición de ser extremos dice que la imagen está contenida en $[m, M]$: $f([a, b]) \subseteq [m, M]$. Por el Teorema del Valor Intermedio, dado cualquier $\lambda \in [m, M]$, existe $c \in [a, b]$ tal que $f(c) = \lambda$, es decir, todo punto $[m, M]$ está en la imagen: $[m, M] \subseteq f([a, b])$. Por tanto, $f([a, b]) = [m, M]$. $\square$

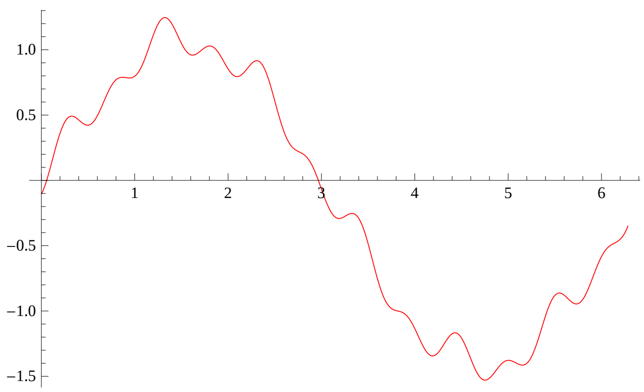

Figura 5 – La imagen continua de un intervalo es el intervalo entre los valores extremos.

3.8. Caracterización abstracta de la continuidad

Caracterización abstracta de la continuidad en un punto

Una función f definida en un intervalo $D \subseteq \mathbb{R}$ es continua en el punto $a \in D$ si y sólo si dado cualquier intervalo abierto $J \ni f(a)$, existe un intervalo abierto $I \ni a$ tal que

$$x \in I \cap D \implies f(x) \in J$$

Demostración. Todo se reduce a la observación que dado un punto a contenido en un intervalo abierto I, existe $\delta > 0$ con $(a-\delta, a+\delta) \subseteq I$. O sea, que se puede "encoger" el intervalo a un intervalo centrado en el punto. □

3.9. Continuidad, invertibilidad y monotonía[5]

Definición de función monótona

Una función f definida en un intervalo $I \subseteq \mathbb{R}$ es

- ***creciente*** *en I si $a, b \in I$, $a \leq b \implies f(a) \leq f(b)$*
- ***estrictamente creciente*** *en I si $a, b \in I$, $a < b \implies f(a) < f(b)$*
- ***decreciente*** *en I si $a, b \in I$, $a \leq b \implies f(a) \geq f(b)$*
- ***estrictamente decreciente*** *en I si $a, b \in I$, $a < b \implies f(a) > f(b)$*

Decimos que f es ***monótona*** *en I si es creciente o decreciente. Decimos que es* ***estrictamente monótona*** *si es estrictamente creciente o estrictamente decreciente.*

Hacemos algunas observaciones sencillas.

- f es estrictamente creciente si y sólo si $a < b \iff f(a) < f(b)$
- f es estrictamente decreciente si y sólo si $a < b \iff f(a) > f(b)$

Es decir, la implicación $\implies$ en la definición de función estrictamente monótona se puede reemplazar por la equivalencia $\iff$.

Demostración. Supongamos, sin pérdida de generalidad, que f es estrictamente creciente. La implicación contraria $f(a) < f(b) \implies a < b$ se deduce de considerar los tres casos posibles: $a = b$, que lleva a $f(a) = f(b)$, $a > b$, que lleva a $f(a) > f(b)$, dejando sólo la posibilidad $a < b$. □

Operaciones sobre funciones monótonas

Las sumas y los ***múltiplos positivos*** *de funciones (estrictamente) crecientes/(estrictamente) decrecientes también son del mismo tipo.*

Demostración. Son inmediatas a partir de las definiciones. □

[5]Se pueden omitir las demostraciones en una lectura rápida, pero es útil considerar lo que dicen los enunciados para un entendimiento más completo del tema.

Funciones monótonas invertibles

Si $f : I \to J$ es invertible, entonces es monótona si y sólo si lo es estrictamente, y eso ocurre si y sólo si su inversa $f^{-1} : J \to I$ es del mismo tipo.

Demostración. Una función invertible, de ser monótona lo debe ser estrictamente, ya que la inyectividad dice que $a \neq b \implies f(a) \neq f(b)$. Esto excluye la posibilidad de igualdad de los valores a partir de $a < b$.

Supongamos, sin pérdida de generalidad, que f es biyectiva y estrictamente creciente. Al ser biyectiva, dos valores $c, d \in J$ son imágenes $c = f(a)$, $d = f(b)$ de dos valores $a, b \in I$. Entonces $c < d \iff f(a) < f(b) \iff a < b \iff f^{-1}(c) < f^{-1}(d)$, o sea, f^{-1} también es estrictamente creciente.

Aplicando el mismo razonamiento a f^{-1}, se deduce el recíproco: si f^{-1} es estrictamente creciente entonces $f = (f^{-1})^{-1}$ es estrictamente creciente. □

El siguiente resultado relaciona la continuidad con la monotonía.

Funciones continuas biyectivas

Sean $I, J \subseteq \mathbb{R}$ dos intervalos de cualquier tipo y supongamos que $f : I \to J$ es biyectiva. Entonces f es continua en I si y sólo si es estrictamente monótona.

Demostración. 1ª parte: si f es biyectiva y continua, entonces es estrictamente monótona

Elegimos dos puntos cualesquiera $a, b \in I$ con $a < b$. Como f es biyectiva, $f(a) \neq f(b)$. Veamos entonces que

- Si $f(a) < f(b)$, entonces f es estrictamente creciente en I.
- Si $f(a) > f(b)$, entonces f es estrictamente decreciente en I.

Demostraremos solo el primer caso pues el razonamiento para el segundo es igual pero cambiando el sentido de las desigualdades.

Consideremos el conjunto “triangular” (posiblemente de lados infinitos)

$$T = \{(x, y) : x \in I,\ y \in I, x < y\},$$

que es un recinto convexo en el semiplano $y > x$. Definimos la función de dos variables

$$\Delta(x, y) = f(y) - f(x), \qquad (x, y) \in T.$$

Que f sea estrictamente creciente sobre I equivale a que $\Delta(x, y) > 0$ para todo $(x, y) \in T$.

Como f es biyectiva, debe ser $\Delta(x, y) \neq 0$ para todo $(x, y) \in T$. Por tanto, lo que necesitamos es un resultado de constancia del signo para Δ.

La hipótesis es que tenemos un punto $(a,b) \in T$ con $\Delta(a,b) > 0$. Sea ahora $(x,y) \in T$ cualquier punto. El segmento que une (a,b) con (x,y) está contenido también en T (por convexidad) y se puede describir por las ecuaciones

$$X(t) = (1-t)a + tx, \quad Y(t) = (1-t)b + ty, \qquad 0 \leq t \leq 1.$$

La función de una variable

$$\delta(t) = \Delta(X(t), Y(t)) = f((1-t)b + t\beta) - f((1-t)a + t\alpha)$$

es continua al ser combinación aritmética y composición de funciones continuas. Como Δ no se anula, δ tampoco lo hace, y al ser continua, tiene un signo constante. Como $\delta(0) = \Delta(a,b) > 0$, debe ser también $\delta(1) = \Delta(x,y) > 0$, como queríamos demostrar.

2ª parte: si f es biyectiva y estrictamente monótona, entonces es continua.

Sin pérdida de generalidad, suponemos que es estrictamente creciente. Entonces la inversa $f^{-1} : J \to I$ también es estrictamente creciente. Por tanto, dados $c, d \in J$,

$$y \in (c,d) \iff c < y < d \iff f^{-1}(c) < f^{-1}(y) < f^{-1}(d) \iff f^{-1}(y) \in \big(f^{-1}(c), f^{-1}(d)\big)$$

Suponemos que $y_0 \in J$ no es un extremo, si no hay que modificar el siguiente razonamiento usando intervalos laterales. La monotonía estricta implica que $x_0 = f^{-1}(y_0)$ tampoco es un extremo de I. Entonces, dado $\varepsilon > 0$, existen puntos $a, b \in I$ con

$$x_0 \in (a,b) \subseteq (x_0 - \varepsilon, x_0 + \varepsilon) = \big(f^{-1}(y_0) - \varepsilon, f^{-1}(y_0) + \varepsilon\big).$$

Poniendo $c = f(a)$, $d = f(b)$ se tiene $c, d \in J$ con $y_0 \in (c,d)$. Para $\delta > 0$ suficientemente pequeño (basta con $\delta \leq \min(y_0 - c, d - y_0)$) , se tiene $(y_0 - \delta, y_0 + \delta) \subseteq (c,d)$ y por tanto

$$\begin{aligned} |y - y_0| < \delta &\implies y \in (c,d) \\ &\implies f^{-1}(y) \in (a,b) = \big(f^{-1}(y_0) - \varepsilon, f^{-1}(y_0) + \varepsilon\big) \\ &\implies |f^{-1}(y) - f^{-1}(y_0)| < \varepsilon, \end{aligned}$$

lo cual pone de manifiesto que f^{-1} es continua en y_0. □

La inversa de una función biyectiva continua es continua

Sean $I, J \subseteq \mathbb{R}$ dos intervalos de cualquier tipo y $f : I \to J$ biyectiva y continua. Entonces $f^{-1} : J \to I$ es continua.

Demostración. Hemos visto que si f es biyectiva y continua, entonces es estrictamente monótona. Entonces f^{-1} es estrictamente monótona del mismo tipo, y como también es biyectiva, el mismo resultado implica que es continua. □

• **Ejemplo:** La exponencial $e^x : (-\infty, \infty) \to (0, \infty)$ es estrictamente creciente, por tanto su inversa, el logaritmo (neperiano) $\log(x) : (0, \infty) \to (-\infty, \infty)$ también lo es. La continuidad de una de ellas implica la de la otra.

• **Ejemplo:** La función $x \mapsto x^2 : [0, \infty) \to [0, \infty)$ es estrictamente creciente, por tanto su inversa, la raíz cuadrada (positiva) $\sqrt{x} : [0, \infty) \to [0, \infty)$ también lo es. La continuidad de una de ellas (la de x^2 es más fácil) implica la de la otra.

4. Cálculo Diferencial

4.1. Definiciones

La **tasa de variación media (tvm)** de una función $f(x)$ en un intervalo $[a,b]$ es

$$\frac{f(b)-f(a)}{b-a}$$

• **Ejemplo:** La siguiente tabla indica el tiempo empleado por Usain Bolt al establecer un récord mundial en los 100m, midiendo cada 20m:[6]

Tiempo (s)	0	2.89	4.64	6.31	7.92	9.58
Distancia (m)	0	20	40	60	80	100

Las tvm del tiempo respecto a la distancia son

- $[0,20]:\frac{2{,}89-0}{20}=0{,}07225$
- $[20,40]:\frac{4{,}64-2{,}89}{20}=0{,}0875$
- $[40,60]:\frac{6{,}31-4{,}64}{20}=0{,}0835$
- $[60,80]:\frac{7{,}92-6{,}31}{20}=0{,}0805$
- $[80,100]:\frac{9{,}58-7{,}92}{20}=0{,}083$

Se observa que el número de segundos por metro fue menor en la salida, subió en el segundo intervalo, decreció en el tercero y cuarto, hasta volver a aumentar en el último.

• **Ejemplo:** Se ha medido la velocidad de un automóvil que está frenando, dando la tabla:[7]

Velocidad (km/h)	120	78	44	18	0
Tiempo (s)	0	2	4	6	8

Las tvm en cada intervalo de tiempo son

- $[0,2]:\dfrac{78-120}{2-0}=\dfrac{-42}{2}=-26$
- $[2,4]:\dfrac{44-78}{4-2}=\dfrac{-34}{2}=-17$
- $[4,6]:\dfrac{18-44}{6-4}=\dfrac{-26}{2}=-13$
- $[6,8]:\dfrac{0-18}{8-6}=\dfrac{-18}{2}=-9$

Estos datos miden la potencia de frenada. Se observa que la tvm es negativa, indicando que, como tiene que ser al frenar, la velocidad va decreciendo. Además, la magnitud de la tvm va disminuyendo a lo largo de la frenada, indicando que la potencia de frenada va disminuyendo con el tiempo.

Consideremos el problema de medir la distancia x recorrida por un objeto, en función del tiempo t. Tenemos un cronómetro que cada Δt segundos hace una medición de la distancia.

[6] http://www.sportsscientists.com/2009/08/analysis-of-bolts-958-wr.html
[7] Ejemplos tomado de Etayo, Colera, Ruiz, Matemáticas 2

Esto va delimitando intervalos de tiempo $[t, t + \Delta t]$ de igual longitud. Si denotamos por

$$\Delta x = x(t + \Delta t) - x(t)$$

a la diferencia entre la distancia en tiempo t y en tiempo $t + \Delta t$, entonces la tvm en el intervalo de tiempo $[t, t + \Delta t]$ es

$$\frac{\Delta x}{\Delta t} = \frac{x(t + \Delta t) - x(t)}{\Delta t}$$

Imaginemos ahora que el cronómetro fuera arbitrariamente preciso, permitiendo ajustar Δt tan pequeño como se quiera. Entonces, en un instante t de tiempo, esta tvm estaría más cerca de medir lo que nuestra intuición dice sería la **variación instantánea** en ese instante. Esta idea lleva directamente a la definición de derivada.

Definición de la derivada

*La **derivada de** $f(x)$ **en el punto** a es el valor del límite (si existe)*

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

*En este caso se dice que f es **derivable** o **diferenciable** en a. Equivalentemente, cambiando de variable por $x = a + h$, la derivada es*

$$f'(a) = \lim_{h \to 0} \frac{f(a + h) - f(a)}{h}$$

*Si $f(x)$ está definida en el intervalo abierto $I = (a, b)$ y existe $f'(x)$ para todo $x \in (a, b)$, decimos que f es **derivable** en (a, b) y llamamos a f' la (función) **derivada de** f **en** I.*

La notación diferencial de Leibnitz

Si $y = f(x)$, se escribe también

$$f'(x) = \frac{dy}{dx} = \frac{df}{dx}$$

Esta notación está sugerida por la utilizada para los cocientes de diferencias $\Delta y / \Delta x$.

La notación diferencial de Newton

Si x es una función de t (habitualmente t mide el tiempo), se puede escribir

$$\dot{x} = x'(t)$$

4.1.1. Derivadas laterales [8]

Definición de las derivadas laterales

Al igual que ocurre con los conceptos de límite y de continuidad, hay también una versión lateral de la definición de derivada:

- *La **derivada lateral derecha** es el valor del límite (si existe)*

$$f'_+(a) = \lim_{x\to a^+} \frac{f(x)-f(a)}{x-a}$$

- *La **derivada lateral izquierda** es el valor del límite (si existe)*

$$f'_-(a) = \lim_{x\to a^-} \frac{f(x)-f(a)}{x-a}$$

Derivadas laterales y derivada

La derivada de una función definida en un entorno de un punto a existe si y sólo si existen las derivadas laterales en a y coinciden:

$$f'(a) = f'_+(a) = f'_-(a).$$

• **Ejemplo:** La función valor absoluto $f(x) = |x|$ es continua en $a = 0$ y tiene

$$f'_+(0) = \lim_{x\to 0^+} \frac{|x|}{x} = \lim_{x\to 0^+} \frac{x}{x} = 1$$

pero

$$f'_-(0) = \lim_{x\to 0^-} \frac{|x|}{x} = \lim_{x\to 0^-} \frac{-x}{x} = -1,$$

por tanto, existen sus derivadas laterales pero, al no coincidir, no existe la derivada $f'(0)$.

No nos vamos a concentrar en este concepto. Es útil para derivar en un extremo del intervalo, cuando sólo podemos acercarnos a ese extremo en una dirección sin salirnos del intervalo.

[8] Esta parte se puede omitir en una lectura rápida

Observamos que hemos escrito

$$f'_+(a) \text{ y no } f'(a^+).$$

El segundo concepto ya ha sido definido y corresponde a un límite lateral, pero es el de la propia función derivada:

$$f'(a^+) = \lim_{x \to a^+} f'(x)$$

y **no** es el mismo concepto que el de derivada lateral, pues éste último requiere dos cosas

- que exista la derivada $f'(x)$ para todo x en un entorno lateral $x \approx a^+$ y
- que exista el límite lateral de la función derivada $f'(x)$ cuando $x \to a^+$

mientras que el concepto de derivada lateral $f'_+(a)$ requiere

- que la función f esté definida para $x \approx a$, en particular para $x = a$,
- que exista el límite lateral del cociente $\displaystyle\lim_{x \to a^+} \frac{f(x) - f(a)}{x - a}$.

• **Ejemplo:** Consideremos la función signo $f(x) = \operatorname{sgn} x$, que es discontinua en 0. Como se puede comprobar,

$$f'(x) = 0 \quad (x \neq 0)$$

por tanto

$$f'(0^+) = f'(0^-) = 0$$

pero

$$f'_+(0) = \lim_{x \to 0^+} \frac{\operatorname{sgn} x}{x} = \lim_{x \to 0^+} \frac{1}{x} = \infty,$$

y

$$f'_-(0) = \lim_{x \to 0^-} \frac{\operatorname{sgn} x}{x} = \lim_{x \to 0^-} -\frac{1}{x} = -\infty$$

por tanto no existe ninguna derivada lateral y por tanto tampoco la derivada.

4.1.2. Interpretación geométrica de la derivada

Si $f(x)$ es lineal, es decir, de la forma $f(x) = mx$, entonces la tvm de f en $[a, x]$ y su derivada son

$$\frac{mx - ma}{x - a} = m, \quad f'(a) = \lim_{x \to a} m = m.$$

La gráfica de f es una recta de pendiente m, luego la derivada es la pendiente.

Para una función general $y = f(x)$, escribiendo Δx para un incremento en x y

$$\Delta y = f(x + \Delta x) - f(x), \quad \therefore \quad f(x + \Delta x) = y + \Delta y$$

La tvm es

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

y representa la pendiente de la cuerda que une (x, y) con $(x + \Delta x, y + \Delta y)$. Haciendo que $\Delta x \to 0$, manteniendo fijo (x, y), esta cuerda va uniendo puntos cada vez más cercanos de la gráfica de f. En el límite, si existe, se llega a la **recta tangente** a la gráfica de f en x.

Ecuación de la recta tangente

La derivada $f'(a)$ representa la pendiente de la recta tangente a la gráfica de f en el punto $(a, f(a))$. La ecuación de esta recta tangente es

$$y = f(a) + f'(a)(x - a)$$

Esto es consecuencia inmediata de la fórmula general $y = m(x - a) + b$ para una recta en el plano que pasa por el punto (a, b) con pendiente m.

4.1.3. La derivada como aproximación

Caracterización de la derivada como aproximación de Taylor lineal

f es derivable en a si y sólo si existe un número λ y una función $\varepsilon(x)$ con $\lim_{x\to a} \varepsilon(x) = 0$ tales que

$$f(x) = f(a) + \lambda \cdot (x - a) + \varepsilon(x)(x - a),$$

siendo entonces $\lambda = f'(a)$. Es decir, se tiene

$$f(x) = f(a) + f'(a) \cdot (x - a) + \varepsilon(x)(x - a),$$

o, en términos del incremento $h = x - a$ que tiende á 0,

$$f(a + h) = f(a) + f'(a)h + \varepsilon(h) \cdot h, \qquad \lim_{h\to 0} \varepsilon(h) = 0.$$

Demostración. Si f es derivable en a, ponemos

$$\varepsilon(x) = \frac{f(x) - f(a)}{x - a} - f'(a).$$

Por definición de derivada, $\lim_{x\to a} \varepsilon(x) = 0$. Reordenando los términos, queda

$$f(x) = f(a) + f'(a) \cdot (x - a) + \varepsilon(x) \cdot (x - a).$$

Recíprocamente, dada una fórmula de aproximación,

$$f(x) = f(a) + \lambda \cdot (x - a) + \varepsilon(x) \cdot (x - a) \implies \frac{f(x) - f(a)}{x - a} = \lambda + \varepsilon(x)$$

y haciendo $x \to a$ queda que f es derivable en a con $f'(a) = \lambda$. □

Observar que la parte afín (término constante y término lineal) de la fórmula anterior, $f(a) + f'(a)(x-a)$, es la ecuación de la recta tangente.

Las funciones derivables son continuas

Si f es derivable en a, entonces es continua en a.

Demostración. Haciendo el límite en la aproximación queda $\lim_{x\to a} f(x) = f(a)$. □

Si f es derivable en un intervalo abierto, la propia función f tiene que ser continua, pero su derivada f' no tiene por qué serlo. Más adelante veremos un ejemplo de derivada discontinua.

El recíproco, si una función continua tiene que ser derivable, es falsa. De hecho existen funciones continuas que no son derivables en **ningún** punto. El primer ejemplo de una tal función se debe a Weierstrass. Es $f(x) = \sum_{n=0}^{\infty} a^n \cos(\pi b^n x)$ con $0 < a < 1$, $b \in \mathbb{N}$ impar con $ab > 1 + \frac{3\pi}{2}$.

El resultado correspondiente para derivadas laterales es

- Si existe $f'_+(a)$ entonces f es continua por la derecha en a.
- Si existe $f'_-(a)$ entonces f es continua por la izquierda en a.

Demostración. Tomar límites laterales en la fórmula de aproximación. □

4.2. Algunas derivadas básicas

La derivada de una función constante es nula.

Demostración. Si $f(x) = k$ (constante) para todo x, entonces

$$\frac{f(x+h) - f(x)}{h} = \frac{k-k}{h} = \frac{0}{h} = 0$$

para todo $x \in \mathbb{R}$ y $h \neq 0$, con lo cual haciendo $h \to 0$ queda $f'(x) = 0$. □

La derivada de una función afín $f(x) = mx + b$ es la pendiente, m.

Demostración. Calculamos

$$\frac{f(x+h) - f(x)}{h} = \frac{(m(x+h)+b) - (mx+b)}{h} = \frac{(\cancel{mx} + mh + \cancel{b} - \cancel{mx} - \cancel{b}}{h} = \frac{mh}{h} = m,$$

con lo cual haciendo $h \to 0$ queda $f'(x) = m$. □

Sea $n \in \mathbb{N}$ y $f(x) = x^n$. Entonces

$$f'(x) = nx^{n-1}$$

Demostración. Usando el teorema del binomio de Newton

$$\begin{aligned}\frac{(x+h)^n - x^n}{h} &= \frac{1}{h}\left(x^n + nx^{n-1}h + \binom{n}{2}x^{n-2}h^2 \dots \binom{n}{k}x^{n-k}h^k + \dots + h^n - x^n\right)\\ &= nx^{n-1} + \binom{n}{2}x^{n-2}h + \dots \binom{n}{k}x^{n-k}h^{k-1} + \dots + h^{n-1}\end{aligned}$$

y, como esto es un polinomio en h, su límite se obtiene haciendo $h = 0$, quedando nx^{n-1}. $\square$

La derivada del logaritmo (natural o neperiano) en un $x > 0$ es

$$\log'(x) = \frac{1}{x}$$

Demostración. Usando las propiedades del logaritmo,

$$\begin{aligned}\frac{\log(x+h) - \log x}{h} &= \frac{1}{h}\log\left(\frac{x+h}{x}\right) = \frac{1}{h}\log\left(1 + \frac{h}{x}\right)\\ &= \frac{1}{x}\cdot\frac{x}{h}\cdot\log\left(1+\frac{h}{x}\right) = \frac{1}{x}\cdot\log\left(1+\frac{h}{x}\right)^{x/h}.\end{aligned}$$

Como $x > 0$ es fijo, haciendo el cambio de variable $t = x/h$, si $h \to 0^+$ entonces $t \to \infty$, y tenemos por tanto

$$\log\left(1+\frac{h}{x}\right)^{x/h} = \log\left(1+\frac{1}{t}\right)^t \to \log e = 1 \quad (h \to 0).$$

Por tanto

$$\log'(x) = \lim_{h\to 0}\frac{\log(x+h) - \log x}{h} = \lim_{h\to 0}\frac{1}{x}\cdot\log\left(1+\frac{h}{x}\right)^{x/h} = \frac{1}{x}.$$

$\square$

La derivada de la función exponencial en cualquier punto $x \in \mathbb{R}$ es ella misma

$$\exp'(x) = \exp(x)$$

Demostración. Sea $a \in \mathbb{R}$. Entonces

$$\lim_{h\to 0}\frac{e^{a+h} - e^a}{h} = \lim_{h\to 0}\frac{e^a e^h - e^a}{h} = e^a \lim_{h\to 0}\frac{e^h - 1}{h}$$

lo cual muestra que basta con demostrar que $\exp'(0) = 1$. Esto puede verse a partir de esta desigualdad, válida para $0 < x < 1$:

$$1 + x < \left(1 + \frac{x}{n}\right)^n < 1 + x + x^2 + \cdots + x^n.$$

Usando la fórmula para la suma de una serie geométrica, queda, separando los primeros dos términos,

$$1 + x \le e^x \le 1 + x + \frac{x^2}{1-x} \quad \therefore \quad 1 \le \frac{e^x - 1}{x} \le 1 + \frac{x}{1-x} \qquad (0 < x < 1),$$

y entonces, haciendo $x \to 0^+$, obtenemos $\lim_{x\to 0^+} \frac{e^x-1}{x} = 1$. Para el otro límite lateral, cambiamos de variable

$$\lim_{x\to 0^-} \frac{e^x - 1}{x} = \lim_{x\to 0^+} \frac{e^{-x} - 1}{(-x)} = \lim_{x\to 0^+} \frac{1 - e^x}{(-x)e^x} = \lim_{x\to 0^+} \frac{e^x - 1}{x} \lim_{x\to 0^+} \frac{1}{e^x} = 1,$$

donde en el último paso hemos usado la continuidad de la exponencial. □

Las derivadas de las funciones trigonométricas seno y coseno son

$$\sin' x = \cos x, \qquad \cos' x = -\sin x.$$

Demostración. Suponemos la continuidad de $\sin x$ y $\cos x$. Entonces

$$\begin{aligned}\lim_{h\to 0} \frac{\sin(x+h) - \sin(x)}{h} &= \lim_{h\to 0} \frac{\sin x \cos h + \sin h \cos x - \sin x}{h} \\ &= \lim_{h\to 0} \left(\sin x \cdot \frac{\cos h - 1}{h} + \cos x \cdot \frac{\sin h}{h}\right)\end{aligned}$$

El resultado depende de saber

$$\lim_{h\to 0} \frac{\sin h}{h} = 1, \qquad \lim_{h\to 0} \frac{\cos h - 1}{h^2} = \frac{1}{2}, \quad \therefore \quad \lim_{h\to 0} \frac{\cos h - 1}{h} = 0.$$

Estos límites de pueden demostrar geométricamente o usando las fórmulas para seno y coseno como series de potencias. Los pasos para $\cos x$ son similares. □

La fórmula $\cos' x = -\sin x$ se puede deducir de $\sin' x = \cos x$ teniendo en cuenta que $\cos x = \sin\left(\frac{\pi}{2} - x\right)$, pero antes hace falta, al igual que hemos hecho con los límites y la continuidad, estudiar la relación entre la derivada y las distintas operaciones sobre funciones. Esto proporcionará un método más eficiente para calcular derivadas de funciones construidas a partir de otras mediante esas operaciones.

4.3. Las reglas de derivación

Reglas aritméticas de derivación

Sean f, g dos funciones derivables en a. Entonces kf $(k \in \mathbb{R})$, $f+g$, $f-g$, fg son derivables en a, y f/g lo es si suponemos $g(a) \neq 0$, teniendo entonces

- ***Múltiplos*** $(kf)'(a) = kf'(a)$
- ***Sumas*** $(f+g)'(a) = f'(a) + g'(a)$
- ***Restas*** $(f-g)'(a) = f'(a) - g'(a)$
- ***Regla del Producto*** $(fg)'(a) = f'(a)g(a) + f(a)g'(a)$
- ***Regla del Cociente*** $\left(\dfrac{f}{g}\right)'(a) = \dfrac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}$

En particular, las funciones derivables en un intervalo abierto forman un espacio vectorial y la derivada es una operación lineal:

$$(\alpha f + \beta g)' = \alpha f' + \beta g', \quad \alpha, \beta \in \mathbb{R}.$$

Cuando se considera la derivada como una operación sobre funciones, a veces se usa la notación $Df = f'$. Por ejemplo, la linealidad se escribe $D(\alpha f + \beta g) = \alpha Df + \beta Dg$.

Demostración. La estrategia común para demostrar las reglas es expresar la tasa de variación media del resultado en términos de las tasas de cada parte y hacer el límite, usando las reglas aritméticas para límites. Por ejemplo, para sumas es inmediato que

$$\frac{(f+g)(x) - (f+g)(a)}{x-a} = \frac{f(x) - f(a)}{x-a} + \frac{g(x) - g(a)}{x-a}$$

lo cual demuestra entonces que $(f+g)'(a) = f'(a) + g'(a)$.

Para productos (y múltiplos), escribimos

$$\begin{aligned}\frac{(f \cdot g)(x) - (f \cdot g)(a)}{x-a} &= \frac{f(x)g(x) - f(a)g(a)}{x-a} \\ &= \frac{f(x)g(x) - f(a)g(x) + f(a)g(x) - f(a)g(a)}{x-a} \\ &= \frac{f(x) - f(a)}{x-a} \cdot g(x) + f(a) \cdot \frac{g(x) - g(a)}{x-a}\end{aligned}$$

y usamos la continuidad de las funciones derivables.

Para cocientes, basta con demostrar que $(1/g)' = -g'/g^2$ y aplicar la Regla del Producto a $(f/g) = f \cdot (1/g)$. En cuanto a $1/g$, si $g(a) \neq 0$, la continuidad de g en a implica que $g(x) \neq 0$ en un entorno de a y podemos escribir

$$\frac{\dfrac{1}{g(x)} - \dfrac{1}{g(a)}}{x-a} = \frac{g(a) - g(x)}{(x-a)g(x)g(a)} = -\frac{g(x) - g(a)}{x-a} \cdot \frac{1}{g(x)g(a)}$$

que al hacer el límite nos da $(1/g)'(a) = -g'(a)/g(a)^2$.

$\square$

La Regla de la Cadena

Supongamos que f, g son dos funciones que se pueden componer. Si f es derivable en a y g es derivable en $f(a)$, entonces la composición $g \circ f$ es derivable en a, con

$$(g \circ f)'(a) = g'(f(a)) \cdot f'(a)$$

Demostración que uno quisiera que fuera correcta pero no lo es. Sea $b = f(a)$. Hacemos la sustitución $k = f(a+h) - f(a)$, de modo que $b + k = f(a+h)$. Entonces

$$\begin{aligned}\frac{g(f(a+h)) - g(f(a))}{h} = \frac{g(b+k) - g(b)}{h} &= \frac{g(b+k) - g(b)}{k} \cdot \frac{k}{h} \\ &= \frac{g(b+k) - g(b)}{k} \cdot \frac{f(a+h) - f(a)}{h}.\end{aligned}$$

Como f es derivable en a, es continua, luego $h \to 0 \implies k \to 0$ y cuando hacemos el límite queda $(g \circ f)'(a) = g'(b)f'(a) = g'(f(a))f'(a)$.

¿Por qué no es correcta? porque aun siendo $h \neq 0$, nada nos garantiza que $k = f(a+h) - f(a)$ no sea nula, en cuyo caso el cociente por k no estaría definido. $\square$

Demostración que sí es correcta pero es mas larga. Vamos a usar la interpretación de la derivada como fórmula de aproximación con incrementos. Sea $b = f(a)$. Tenemos, para $h, k \approx 0$,

$$\begin{aligned} f(a+h) &= f(a) + f'(a)h + \alpha(h)h, & \lim_{h \to 0} \alpha(h) = 0, \\ g(b+k) &= g(b) + g'(b)k + \beta(k)k, & \lim_{k \to 0} \beta(k) = 0. \end{aligned}$$

Definimos $\alpha(0) = \beta(0) = 0$ para tener de este modo funciones continuas en 0 y extender la validez de las fórmulas de aproximación al propio punto a.

Sustituimos $k = f(a+h) - f(a) = f(a+h) - b$, de modo que $b + k = f(a+h)$ y

$$(g \circ f)(a+h) = (g \circ f)(a) + g'(b)k + \beta(k)k$$

pero como además $k = f'(a)h + \alpha(h)h$, queda

$$\begin{aligned}(g \circ f)(a+h) &= (g \circ f)(a) + g'(b)f'(a)h + g'(b)\alpha(h)h + \beta(k)k \\ &= (g \circ f)(a) + g'(f(a))f'(a)h + \gamma(h)h\end{aligned}$$

donde

$$\gamma(h)h = g'(b)\alpha(h)h + \beta(k)k \quad \therefore \quad \gamma(h) = g'(b)\alpha(h) + \beta(k)\frac{k}{h}$$

y para acabar la demostración, tenemos que ver que $\lim_{h\to 0} \gamma(h) = 0$.

El primer sumando $g'(b)\alpha(h)$ sí tiende a cero pues lo hace $\alpha(h)$. En cuanto al segundo, lo escribimos para $h \neq 0$ como

$$\beta(k)\frac{k}{h} = \beta(k)\frac{f(a+h)-f(a)}{h} = \beta(k)\big(f'(a)+\alpha(h)\big)$$

y acabamos observando que $\lim_{h\to 0} k = \lim_{h\to 0}(f(a+h)-f(a)) = 0$ por continuidad de f en a, luego al combinar los límites queda $\lim_{h\to 0} \beta(k) = 0$ y con ello hemos demostrado que el segundo sumando también tiende a cero. □

La regla de la cadena en notación diferencial

Si y es función de x y z es función de y, entonces

$$\frac{dz}{dx} = \frac{dz}{dy}\cdot\frac{dy}{dx}$$

Efectivamente, si denotamos las composiciones como cambios de variable, escribiendo por ejemplo $y = f(x)$, $z = g(y)$, entonces $z = h(x)$ con $h = g \circ f$ y en la notación diferencial,

$$f'(x) = \frac{dy}{dx}, \quad g'(y) = \frac{dz}{dy}, \quad h'(x) = \frac{dz}{dx},$$

y la regla de la cadena $h'(x) = g'(f(x)) \cdot f'(x) = g'(y)f'(x)$ adquiere la forma indicada de una “cancelación” simbólica.

La derivada de una inversa

Sea $f : (a,b) \to (c,d)$ una función invertible, con inversa $g = f^{-1} : (c,d) \to (a,b)$. Si f y f^{-1} son ambas derivables en sus dominios, entonces

$$(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}$$

para todo $y \in (c,d)$.

Demostración. Poniendo $y = f(x)$, $x = f^{-1}(y)$. Por la Regla de la Cadena,

$$g(f(x)) = x \implies g'(f(x))\cdot f'(x) = 1 \implies g'(y) = \frac{1}{f'(x)} = \frac{1}{f'(f^{-1}(y))}.$$

□

La regla de la inversa en notación diferencial

$$\frac{dx}{dy} = \frac{1}{\frac{dy}{dx}}$$

Efectivamente, escribiendo $y = f(x)$, $x = g(y)$, donde $g = f^{-1}$, tenemos

$$f'(x) = \frac{dy}{dx}, \quad g'(y) = \frac{dx}{dy}, \quad \therefore \quad g'(y) = \frac{1}{f'(x)} \iff \frac{dx}{dy} = \frac{1}{\frac{dy}{dx}}.$$

Una función biyectiva y derivable puede tener inversa que no sea derivable en todo punto. Por ejemplo, $f(x) = x^3 : \mathbb{R} \to \mathbb{R}$ es biyectiva y derivable pero su inversa $f^{-1}(x) = \sqrt[3]{x}$ no es derivable en $x = 0$.

Derivabilidad de la función inversa

Sean $I, J \subseteq \mathbb{R}$ intervalos abiertos y $f : I \to J$ biyectiva y derivable en $a \in I$ con $f'(a) \neq 0$. Entonces la inversa $f^{-1} : J \to I$ es derivable en $b = f(a) \in J$, siendo $(f^{-1})'(b) = 1/f'(a)$.

Demostración. Sea $b \in J$ con $b = f(a)$, $a \in I$. Dado $k \neq 0$, consideramos el cociente

$$\frac{f^{-1}(b+k) - f^{-1}(b)}{k} = \frac{h}{f(a+h) - f(a)}$$

donde $h = f^{-1}(b+k) - f^{-1}(b)$. Este cambio de variable es invertible, con $k = f(a+h) - f(a)$. Esto significa que $h \neq 0 \iff k \neq 0$, luego el denominador a la derecha no es nulo.

Como f es biyectiva y derivable, también es biyectiva y continua, por tanto f^{-1} también es biyectiva y continua. Esto significa que además $h \to 0 \iff k \to 0$, luego haciendo el límite se concluye que f^{-1} es derivable en b y además $(f^{-1})'(b) = 1/f'(a)$. □

4.3.1. Aplicaciones al cálculo de derivadas

• **Ejemplo:** Poniendo $y = \log x$, $x = \exp y$ obtenemos otra vez la derivada de exp que habíamos visto antes en la sección 4.2:

$$\frac{dy}{dx} = \frac{1}{x} \quad \therefore \quad \frac{dx}{dy} = \frac{1}{\frac{1}{x}} = x = \exp(y), \quad \therefore \quad \frac{d}{dy} \exp y = \exp y.$$

Derivadas de las funciones trigonométricas

- $\dfrac{d}{dx}\sin x = \cos x$
- $\dfrac{d}{dx}\cos x = -\sin x$
- $\dfrac{d}{dx}\tan x = \dfrac{1}{\cos^2 x} = \sec^2 x = 1 + \tan^2 x$
- $\dfrac{d}{dx}\cot x = -\dfrac{1}{\sin^2 x} = -\csc^2 x = -(1 + \cot^2 x)$
- $\dfrac{d}{dx}\sec x = \sec x \tan x$
- $\dfrac{d}{dx}\csc x = -\csc x \cot x$

• Sabiendo que $\sin'(x) = \cos x$, derivando la relación $\cos x = \sin\left(\frac{\pi}{2} - x\right)$ obtenemos

$$\cos'(x) = \sin'\left(\frac{\pi}{2} - x\right) \cdot (-x)' = \cos\left(\frac{\pi}{2} - x\right) \cdot (-1) = -\sin x.$$

Usando la regla del cociente,

$$\tan'(x) = \frac{\sin' x \cos x - \sin x \cos' x}{\cos^2 x} = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x} = \sec^2 x$$

y también es igual a $1 + \tan^2 x$.

Demostrar las demás fórmulas en el cuadro anterior, partiendo de las dos fórmulas para las derivadas del seno y el coseno.

Derivadas de las funciones trigonométricas inversas

- $\dfrac{d}{dx}\arcsin x = \dfrac{1}{\sqrt{1 - x^2}} \quad (|x| < 1)$
- $\dfrac{d}{dx}\arccos x = -\dfrac{1}{\sqrt{1 - x^2}} \quad (|x| < 1)$
- $\dfrac{d}{dx}\arctan x = \dfrac{1}{1 + x^2}$
- $\dfrac{d}{dx}\cot^{-1} x = -\dfrac{1}{1 + x^2}$
- $\dfrac{d}{dx}\sec^{-1} x = \dfrac{1}{|x|\sqrt{x^2 - 1}} \quad (|x| > 1)$
- $\dfrac{d}{dx}\csc^{-1} x = -\dfrac{1}{|x|\sqrt{x^2 - 1}} \quad (|x| > 1)$

Demostración. Por ejemplo,

$$
\begin{aligned}
y = \tan x,\ \ x = \arctan y \implies& \frac{dy}{dx} = 1 + \tan^2 x \\
\implies& \frac{dx}{dy} = \frac{1}{1+\tan^2 x} = \frac{1}{1+y^2}
\end{aligned}
$$

y

$$
\begin{aligned}
y = \sin x,\ \ x = \arcsin y \implies& \frac{dy}{dx} = \cos x = \sqrt{1-\sin^2 x} = \sqrt{1-y^2} \\
\implies& \frac{dx}{dy} = \frac{1}{\sqrt{1-y^2}}.
\end{aligned}
$$

Los otros son similares. □

Como siempre, es importante tener en cuenta el dominio y recorrido de las funciones. En el último ejemplo se trata de invertir

$$\sin : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \to [-1,1], \quad \therefore \quad \arcsin : [-1,1] \to \left[-\frac{\pi}{2}, \frac{\pi}{2}\right].$$

Cuando ponemos $\cos x = \sqrt{1-\sin^2 x}$ a partir de $\sin^2 x + \cos^2 x = 1$, estamos teniendo en cuenta que es la raíz cuadrada positiva la que se toma, pues $\cos x \geq 0$ en $[-\pi/2, \pi/2]$.

Recordar que las funciones trigonométricas hiperbólicas se definen a partir de

$$\sinh x = \frac{e^x - e^{-x}}{2}, \qquad \cosh x = \frac{e^x + e^{-x}}{2}.$$

Derivadas de las funciones trigonométricas hiperbólicas

- $\dfrac{d}{dx} \sinh x = \cosh x$
- $\dfrac{d}{dx} \cosh x = \sinh x$
- $\dfrac{d}{dx} \tanh x = \dfrac{1}{\cosh^2 x} = \operatorname{sech}^2 x = 1 - \tanh^2 x$
- $\dfrac{d}{dx} \coth x = -\dfrac{1}{\sinh^2 x} = -\operatorname{csch}^2 x = -(1 - \coth^2 x)$
- $\dfrac{d}{dx} \operatorname{sech} x = -\operatorname{sech} x \tanh x$
- $\dfrac{d}{dx} \operatorname{csch} x = -\operatorname{csch} x \coth x$

Derivadas de las funciones trigonométricas hiperbólicas inversas

- $\dfrac{d}{dx}\sinh^{-1}x = \dfrac{1}{\sqrt{1+x^2}}$
- $\dfrac{d}{dx}\cosh^{-1}x = -\dfrac{1}{\sqrt{x^2-1}} \quad (|x|>1)$
- $\dfrac{d}{dx}\tanh^{-1}x = \dfrac{1}{1-x^2} \quad (|x|<1)$
- $\dfrac{d}{dx}\coth^{-1}x = -\dfrac{1}{1-x^2} \quad (|x|>1)$
- $\dfrac{d}{dx}\operatorname{sech}^{-1}x = -\dfrac{1}{x\sqrt{1-x^2}} \quad (0<|x|<1)$
- $\dfrac{d}{dx}\operatorname{csch}^{-1}x = -\dfrac{1}{|x|\sqrt{1+x^2}} \quad (x\neq 0)$

4.3.2. Derivación logarítmica

Dado que el logaritmo y la exponencial son derivables en sus dominios de definición, si una función $f(x)$ es derivable, entonces también $e^{f(x)}$ es derivable, y si además $f(x)>0$, entonces $\log f(x)$ es derivable. Por tanto una función f cualquiera es derivable si y sólo si lo es e^f, y una función positiva cualquiera f es derivable si y sólo si lo es $\log f$.

Estas sencillas observaciones sirven para simplificar el cálculo de derivadas de potencias.

Derivada de una potencia (exponente fijo)

Sea $p\in\mathbb{R}$. Entonces para $x>0$ se tiene

$$\frac{d}{dx}(x^p) = px^{p-1}$$

Demostración. Tomando logaritmos y derivando:

$$\log(x^p) = p\log x \implies \frac{(x^p)'}{x^p} = \frac{p}{x} \implies (x^p)' = \frac{p}{x}\cdot x^p = px^{p-1}.$$

□

Derivada de una potencia (base fija)

Para cualquier base $b > 0$ se tiene

$$\frac{d}{dx} b^x = (\log b) b^x$$

Demostración. Tomando logaritmos y derivando,

$$\log b^x = x \log b \implies \frac{(b^x)'}{b^x} = \log b \implies (b^x)' = (\log b) b^x.$$

□

• **Ejemplo:** De manera similar, para derivar algo como $f(x) = x^x$, es más fácil si consideramos que

$$\log f(x) = x \log x \implies \frac{f'(x)}{f(x)} = \log x + x \cdot \frac{1}{x} \implies f'(x) = x^x(\log x + 1)$$

Estas ideas se pueden aplicar en general para derivar f^g.

Derivada de potencias funcionales

Si $f > 0$ y g son derivables, entonces

$$(f^g)' = f^g \left(g' \log f + g \frac{f'}{f} \right)$$

Demostración.

$$\log f^g = g \log f \implies \frac{(f^g)'}{f^g} = g' \log f + g \frac{f'}{f} \implies (f^g)' = f^g \left(g' \log f + g \frac{f'}{f} \right)$$

□

No merece mucho la pena memorizar fórmulas de este tipo, sino más bien conocer el método de derivación logarítmica para obtenerlas.

Utilizar la fórmula anterior para deducir la regla del cociente para $f^{-1} = 1/f$ y las fórmulas para las derivadas de b^x con $b > 0$ y x^p con $x > 0$.

Se utiliza la misma notación f^{-1} para dos cosas distintas. Una es la función inversa y otra el recíproco $1/f$. En estos casos hay que tener claro el contexto.

Como ya hemos utilizado en los cálculos anteriores, la regla de la cadena implica que si $f(x) > 0$, entonces $\log f(x)$ es derivable, con derivada dada por

$$\frac{d}{dx} \log f(x) = \frac{f'(x)}{f(x)}.$$

Esto sirve de base a una definición general.

La **derivada logarítmica** de una función positiva $f(x)$ es

$$\operatorname{dlog} f(x) \stackrel{\text{def}}{=} \frac{f'(x)}{f(x)}$$

Propiedades de la derivada logarítmica

Si f, g son funciones positivas, entonces

- $\operatorname{dlog}(fg) = \operatorname{dlog} f + \operatorname{dlog} g$
- $\operatorname{dlog}(f/g) = \operatorname{dlog} f - \operatorname{dlog} g$
- $\operatorname{dlog}(1/f) = -\operatorname{dlog} f$
- $\operatorname{dlog} f^p = p \operatorname{dlog} f$ *para* $p \in \mathbb{R}$

• En general, si $f = f_1 f_2 \cdots f_n$, tomando la derivada logarítmica se obtiene

$$\operatorname{dlog} f = \sum_{k=1}^{n} \operatorname{dlog} f_k, \quad \text{y por tanto} \quad f' = f \sum_{k=1}^{n} \operatorname{dlog} f_k$$

por inducción en el número de factores.

Al igual que el logaritmo convierte productos en sumas y potencias en productos, la derivada logarítmica simplifica las reglas de derivación.

4.4. La regla de L'Hôpital

La regla de L'Hôpital para las indeterminaciones $0/0$ y ∞/∞

Sean f, g funciones definidas en un entorno reducido de a y derivables en él, tales que $g(x) \neq 0$, $g'(x) \neq 0$ para $x \approx a^*$, y se tiene

- $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0$, *ó*
- $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = \pm\infty$.

Es decir, el límite del cociente es una indeterminación de los tipos

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{0}{0} \quad \text{ó} \quad \frac{\pm\infty}{\pm\infty}.$$

Entonces

$$\lim_{x\to a} \frac{f'(x)}{g'(x)} = \ell \implies \lim_{x\to a} \frac{f(x)}{g(x)} = \ell$$

Nota 1: *a y ℓ pueden ser iguales a $\pm\infty$.*

Nota 2: *la regla también es aplicable a límites laterales.*

Nota 3: *De hecho, la hipótesis $g(x) \neq 0$ siempre es redundante.*

Demostración. Demostraremos la versión para la indeterminación $0/0$ en un punto finito a. La hipótesis de derivabilidad en un entorno reducido de a implica la continuidad en ese entorno reducido, y la condición $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0$ significa que podemos definir $f(a) = g(a) = 0$ y de este modo hacerlas continuas también en a.

Entonces, dado $x \approx a^*$, las funciones f, g son continuas en el intervalo compacto $[a, x]$ y derivables en el abierto (a, x). El Teorema del Valor Medio de Cauchy implica que existe $c = c_x$ (depende de x) tal que

$$(f(x) - f(a))g'(c_x) = (g(x) - g(a))f'(c_x)$$

lo cual, como $f(a) = g(a) = 0$, equivale a

$$f(x)g'(c_x) = g(x)f'(c_x).$$

Entonces podemos hacer el cociente para obtener

$$\frac{f(x)}{g(x)} = \frac{f'(c_x)}{g'(c_x)}.$$

Como $c_x \in (a, x)$, cuando $x \to a$, también $c_x \to a$, siendo $c_x \neq a$, y por tanto

$$\lim_{x\to a} \frac{f'(x)}{g'(x)} = \ell \implies \lim_{x\to a} \frac{f'(c_x)}{g'(c_x)} = \ell \implies \lim_{x\to a} \frac{f(x)}{g(x)} = \ell,$$

como queríamos demostrar. $\square$

Como comentamos en el enunciado del Teorema del Valor Medio de Cauchy, la hipótesis que g' no se anula cerca de a implica que $g(x) \neq 0$, de ahí que sea redundante esta hipótesis en el caso $0/0$. Esto es una consecuencia del Teorema del Valor Medio ordinario: si fuera $g(x) = 0$, sería

$$0 = g(x) = g(x) - g(a) = (x - a)g'(\alpha)$$

para algún $\alpha \in (a, x)$ y por tanto $g'(\alpha) = 0$.

En el caso ∞/∞ es todavía más obvio: como $\lim_{x\to a} |g(x)| = \infty$, claramente $g(x) \neq 0$ para $x \approx a^*$.

• **Ejemplo:** Se tiene

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

A veces hace falta aplicar la regla más de una vez:

$$\lim_{x\to 0} \frac{1 - \cos x}{x^2} = \lim_{x\to 0} \frac{\sin x}{2x} = \lim_{x\to 0} \frac{\cos x}{2} = \frac{1}{2}.$$

La regla de L'Hôpital se puede seguir aplicando mientras el límite resultante de derivar numerador y denominador siga cumpliendo las hipótesis. En particular, tiene que seguir siendo indeterminación del tipo $0/0$ ó ∞/∞.

• **Ejemplo:** Mediante la manipulación aritmética de ∞ es posible cambiar un tipo de indeterminación distinto en una del tipo $0/0$ ó ∞/∞ al cual se puede intentar aplicar la regla de L'Hôpital. Por ejemplo,

$$\underbrace{\lim_{x\to 0^+} x \log x}_{(\text{tipo } 0\cdot\infty)} = \underbrace{\lim_{x\to 0^+} \frac{\log x}{1/x}}_{(\text{tipo } \pm\infty/\infty)} = \underbrace{\lim_{x\to 0^+} \frac{1/x}{-1/x^2}}_{\text{regla de L'Hôpital}} = \lim_{x\to 0^+} (-x) = 0$$

y por tanto

$$\underbrace{\lim_{x\to 0^+} x^x}_{\text{tipo } 0^0} = \lim_{x\to 0^+} e^{x\log x} = e^{\lim_{x\to 0^+} x \log x} = e^0 = 1.$$

• **Ejemplo:** La regla de L'Hôpital hace posible resolver indeterminaciones de forma sencilla:

$$\lim_{x\to 0} \frac{\sin(\tan x)}{\sin x} = \lim_{x\to 0} \frac{(\sec^2 x)\cos(\tan x)}{\cos x} = \frac{(\sec^2 0)\cos(\tan 0)}{\cos 0} = \frac{(\sec^2 0)\cos 0}{\cos 0} = 1.$$

• **Ejemplo:** En las ocasiones cuando hace falta derivar más de una vez, como

$$\lim_{x\to 0} \frac{a^{x^2} - 1}{\sin^2 x} = \lim_{x\to 0} \frac{e^{x^2\log a} - 1}{\sin^2 x} = \lim_{x\to 0} \frac{2x(\log a)\, a^{x^2}}{2\sin x\cos x}$$

$$= \lim_{x\to 0} \frac{2(\log a)\, a^{x^2} + (2x(\log a))^2\, a^{x^2}}{2\cos^2 x - 2\sin^2 x} = \log a$$

se suele poder simplificar separando términos:

$$\lim_{x\to 0} \frac{\cancel{2}\, x\,(\log a)\, a^{x^2}}{\cancel{2} \sin x \cos x} = (\log a) \lim_{x\to 0} \frac{x}{\sin x} = (\log a) \lim_{x\to 0} \frac{1}{\cos x} = \log a.$$

• **Ejemplo:** A veces, el número de pasos depende de un parámetro. El límite

$$\lim_{x\to\infty} \frac{x^p}{e^x}$$

sólo es indeterminación ∞/∞ cuando $p > 0$. Si $p = 0$, es $1/\infty = 0$ y si $p < 0$, es $0/\infty = 0$. Para $p > 0$, la regla de L'Hôpital lleva a considerar

$$\lim_{x\to\infty} \frac{p\, x^{p-1}}{e^x}$$

que será 0 si $p \le 1$ y del tipo ∞/∞ si $p > 1$, y aplicar la regla de L'Hôpital otra vez lleva a

$$\lim_{x\to\infty} \frac{p(p-1)\, x^{p-2}}{e^x}$$

que será 0 si $p \le 2$, o seguirá siendo indeterminación de tipo ∞/∞ si $p > 2$. Como hay un número natural n tal que

$$n - 1 < p \le n,$$

aplicar la regla de L'Hôpital nos dará una indeterminación ∞/∞ las primeras $n-1$ veces, pero a la n-ésima tendremos el límite

$$\lim_{x\to\infty} \frac{x^{p-n}}{e^x} = 0.$$

Por tanto, hemos demostrado que

$$\lim_{x\to\infty} \frac{x^p}{e^x} = 0$$

para todo $p \in \mathbb{R}$.

• **Ejemplo:** Cambiando de variable en una indeterminación conocida nos lleva a otra. Por ejemplo, en el límite anterior, poniendo $x = \log t$, queda

$$\lim_{t\to\infty} \frac{\log^p t}{t} = 0$$

para todo $p \in \mathbb{R}$, y poniendo ahora $t = x^q$ con $q > 0$,

$$0 = \lim_{t\to\infty} \frac{\log^p t^q}{t^q} = \lim_{t\to\infty} \frac{q^p \log^p t}{t^q} = \lim_{t\to\infty} \frac{\log^p t}{t^q} \qquad \forall\, p, q \in \mathbb{R} : q > 0.$$

• **Ejemplo:** No olvidar ir simplificando a la vez que se aplica la regla de L'Hôpital y así ahorrar pasos:

$$\lim_{x\to 0} \frac{\log \cos ax}{\log \cos bx} = \lim_{x\to 0} \frac{\frac{a \sin ax}{\cos ax}}{\frac{b \sin bx}{\cos bx}} = \lim_{x\to 0} \frac{a}{b} \cdot \frac{\cos bx}{\cos ax} \cdot \frac{\sin ax}{\sin bx} = \frac{a}{b} \cdot 1 \cdot \lim_{x\to 0} \frac{\sin ax}{\sin bx}$$

$$= \frac{a}{b} \cdot 1 \cdot \lim_{x\to 0} \frac{a \cos ax}{b \cos bx} = \frac{a}{b} \cdot 1 \cdot \frac{a}{b} = \frac{a^2}{b^2}.$$

Hay casos en los que, por más que se derive, la expresión sigue siendo una indeterminación del mismo tipo. Entonces hay que apelar a otros métodos. Por ejemplo,

$$\lim_{x\to 0^+} \frac{e^{-1/x}}{x}$$

es una indeterminación de tipo $0/0$. Derivando queda

$$\lim_{x\to 0^+} \frac{x^{-2}e^{-1/x}}{1} = \lim_{x\to 0^+} \frac{e^{-1/x}}{x^2}$$

que sigue siendo indeterminación del tipo $0/0$. Derivando otra vez, queda

$$\lim_{x\to 0^+} \frac{x^{-2}e^{-1/x}}{2x} = \lim_{x\to 0^+} \frac{e^{-1/x}}{2x^3}$$

y, después de derivar n veces, por inducción, vemos que queda

$$\lim_{x\to 0^+} \frac{e^{-1/x}}{n!\, x^{n+1}}$$

con lo cual así nunca vamos a poder determinar el límite. Sin embargo, cambiando de variable con $t = 1/x$ en el límite original, éste se transforma en la indeterminación ∞/∞

$$\lim_{t\to\infty} \frac{e^{-t}}{t^{-1}} = \lim_{t\to\infty} \frac{t}{e^t}$$

y este último límite sí se calcula con una sola aplicación de la regla de L'Hôpital:

$$\lim_{t\to\infty} \frac{t}{e^t} = \lim_{t\to\infty} \frac{1}{e^t} = 0.$$

La regla de L'Hôpital funciona tan bien que a veces uno se olvida que sólo se aplica a las indeterminaciones del tipo $0/0, \infty/\infty$. Por ejemplo,

$$\lim_{x\to 0} \frac{6x+1}{3x+1} \neq \lim_{x\to 0} \frac{6}{3} = 2$$

pero este límite no era indeterminación, y el valor correcto es 1. La regla de L'Hôpital ¡no es un sustituto para el razonamiento!

4.5. Optimización I: puntos críticos

En la sección 3.6 hemos tratado el tema de la optimización de una función, es decir, de hallar puntos donde alcanza un valor que es el menor o mayor posible, pero de un modo teórico. El resultado principal era el Teorema de Weierstrass, pero éste es un resultado de existencia: dice que existen puntos extremos, pero no enseña **cómo encontrarlos.**

El Teorema de Weierstrass se refiere a extremos globales y sólo es aplicable a intervalos compactos. Dada la variedad de fenómenos que se encuentran en la práctica, conviene generalizar esta situación, considerando tanto otros tipos de extremos como de intervalos.

Clasificación de extremos

Sea f una función definida en un intervalo I. Un punto $c \in I$ es

- ***máximo local*** de f *si* $f(x) \leq f(c)$ *para* $x \in I, x \approx c$
- ***máximo local*** **estricto** de f *si* $f(x) < f(c)$ *para* $x \in I, x \approx c, x \neq c$
- ***máximo*** **global** de f *sobre* I *si* $f(x) \leq f(c)$ *para* **todo** $x \in I$.
- ***máximo*** **global** **estricto** de f *sobre* I *si* $f(x) < f(c)$ *para* **todo** $x \in I, x \neq c$.
- ***mínimo local*** de f *si* $f(x) \geq f(c)$ *para* $x \in I, x \approx c$
- ***mínimo local*** **estricto** de f *si* $f(x) > f(c)$ *para* $x \in I, x \approx c, x \neq c$
- ***mínimo*** **global** de f *sobre* I *si* $f(x) \geq f(c)$ *para* **todo** $x \in I$.
- ***mínimo*** **global** **estricto** de f *sobre* I *si* $f(x) > f(c)$ *para* **todo** $x \in I, x \neq c$.

En general, si c es un máximo o mínimo (local/global/estricto), decimos que es un ***extremo*** de f *(respectivamente local/global/estricto).*

No debe confundirse un extremo c, que es un punto del **dominio** de la función f, con el **valor** correspondiente $f(c)$, que es un punto en su **imagen.**

Un extremo global es en particular también un extremo local, pero no al revés, en general.

Si el punto c en las definiciones anteriores es un extremo del intervalo I, la condición $x \in I, x \approx c$ es **lateral.** Por ejemplo, si $I = [a, b]$ y $c = a$, quiere decir que x está en un subintervalo de I de la forma $[a, a + \delta)$ y si $c = b$, en uno de la forma $(b - \delta, b]$.

• **Ejemplo:** La función $f(x) = \sin x$ tiene máximos globales sobre $\mathbb{R}$ en los puntos $\pi/2 + 2\pi n, n \in \mathbb{Z}$, siendo 1 el valor máximo, y mínimos globales en $-\pi/2 + 2\pi n, n \in \mathbb{Z}$, siendo -1 el valor mínimo. Ninguno de ellos es estricto.

• **Ejemplo:** $f(x) = x^2$ tiene un mínimo global estricto en $x = 0$ con valor mínimo 0.

Razonar: una función puede tener a lo sumo un extremo global estricto.

Extremos en términos de signos

Un punto c es un extremo local de f si y sólo si la diferencia $f(x) - f(c)$ *no cambia de signo para $x \approx c$.*

Demostración. Es de comprobación inmediata. □

Optimización

El problema de optimización para una función f definida en un intervalo I consiste en hallar sus extremos locales, si existen, y clasificarlos de acuerdo con las definiciones anteriores, siendo de especial interés determinar si tiene extremos globales.

El signo de la derivada en un punto y la relación de valores en un entorno

Sea f una función definida en un entorno del punto c y derivable en c.

$$f'(c) > 0 \implies \begin{cases} f(x) < f(c) & \text{para } x \approx c^- \\ f(x) > f(c) & \text{para } x \approx c^+ \end{cases}$$

$$f'(c) < 0 \implies \begin{cases} f(x) > f(c) & \text{para } x \approx c^- \\ f(x) < f(c) & \text{para } x \approx c^+ \end{cases}$$

En particular, si $f'(c) \neq 0$, entonces c no es un extremo local de f.

Demostración. Haremos la demostración cuando $f'(c) > 0$. El otro caso es similar. Por conservación del signo en los límites,

$$f'(c) = \lim_{x \to c} \frac{f(x) - f(c)}{x - c} > 0 \implies \frac{f(x) - f(c)}{x - c} > 0 \text{ para } x \approx c^*.$$

Para que este cociente sea positivo, según a qué lado estamos de c debe darse

$$\begin{aligned} x \approx c^- &\implies x - c < 0 \implies f(x) - f(c) < 0 \implies f(x) < f(c) \\ x \approx c^+ &\implies x - c > 0 \implies f(x) - f(c) > 0 \implies f(x) > f(c), \end{aligned}$$

lo cual demuestra el enunciado.

Si $f'(c) \neq 0$, vemos que según lo anterior, la diferencia $f(x) - f(c)$ cambia de signo según a qué lado de c está x, por tanto c no puede ser un extremo local de f. $\square$

alternativa usando la función signo. La conservación del signo en los límites implica que

$$\begin{aligned} f'(c) \neq 0 &\implies \operatorname{sgn} \frac{f(x) - f(c)}{x - c} = \operatorname{sgn} f'(c) && x \approx c^* \\ &\implies \operatorname{sgn}(f(x) - f(c)) = \operatorname{sgn} f'(c) \cdot \operatorname{sgn}(x - c) && x \approx c^* \end{aligned}$$

El lado derecho claramente alterna en signo según sea $x \approx c^{\pm}$, luego $f(x) - f(c)$ también cambia de signo, de acuerdo con el patrón señalado, y c no puede ser un extremo. $\square$

Sea f una función en el intervalo I. Un punto $c \in I$ es un **punto crítico** de f si se dan una de las siguientes condiciones:

- c es un extremo de I,
- f no es derivable en c,
- f es derivable en c con $f'(c) = 0$.

Al valor correspondiente $f(c)$ se le llama **valor crítico** de f.

Teorema del Punto Crítico

Sea f una función definida en un intervalo I y $c \in I$ un punto ***interior*** *que es extremo local de f. Si f es derivable en c, entonces $f'(c) = 0$. Por tanto los extremos de f se hallan entre sus puntos críticos en I.*

Demostración. El resultado anterior dice que si f es derivable en c y $f'(c) \neq 0$, entonces c no es un extremo. El contrapositivo de esto dice que si c es un extremo en el cual f sea derivable, entonces $f'(c) = 0$.

Como cabe la posibilidad de que un extremo de f se halle justo en un extremo del intervalo o en un punto donde f no sea derivable, no pudiendo entonces aplicar el resultado anterior, tenemos que añadir éstos a la lista de puntos especiales dónde posiblemente se halle algún extremo de f. $\square$

Intuitivamente, el teorema dice que para una función **derivable,** los extremos son puntos de **estabilidad,** donde la función "no varía" (derivada nula) o tiene "velocidad instantánea nula", o "está en equilibrio". Geométricamente, son **picos y valles** de la gráfica.

El Teorema del Punto Crítico dice que los extremos locales de una función se hallan **entre** sus puntos críticos. **No todo punto crítico tiene por qué ser un extremo de la función.** Su importancia es que proporciona un punto de partida **dónde buscar** los extremos.

No debe confundirse extremo de una función con extremo de un intervalo. Son dos conceptos distintos. Los extremos del intervalo pueden ser o no extremos de la función y viceversa. El contexto sirve para distinguirlos pero se debe ser preciso.

El Teorema del Punto Crítico y el Teorema de Weierstrass (3.6) proporcionan un **método completo** para determinar los valores mínimo y máximo de una función f definida en un intervalo compacto $[a, b]$, cuya existencia enuncia este último teorema.

Optimización de una función en un intervalo compacto

Sea f una función definida en el intervalo compacto $[a, b]$, tal que

- *f es continua en $[a, b]$,*
- *f es derivable en (a, b) salvo posiblemente en un número finito de puntos.*

Para hallar los valores extremos de f,

- *Hacemos la lista $\mathcal{C}$ de los puntos críticos, que recordamos son*
    - *los extremos a, b del intervalo*
    - *los puntos de (a, b) donde f no es derivable*
    - *los puntos $c \in (a, b)$ donde f es derivable con $f'(c) = 0$*
- *Para cada punto crítico $c \in \mathcal{C}$, calculamos el valor crítico $f(c)$.*

Entonces

- *El valor mínimo de f sobre $[a, b]$ es el mínimo de los valores críticos.*
- *El valor máximo de f sobre $[a, b]$ es el máximo de los valores críticos.*

• **Ejemplo:** Para hallar los valores extremos de $f(x) = x^2 - x + 6$ en $[-1, 1]$, calculamos $f'(x) = 2x - 1$. El único punto con $f'(x) = 0$ es $1/2$. Hacemos las listas de puntos y valores críticos:

punto crítico c	-1	$1/2$	1
valor crítico $f(c)$	8	$5\frac{3}{4}$	6

Concluimos que -1 es el único máximo global, siendo el valor máximo 8, y $1/2$ es el único mínimo global, siendo el valor mínimo $5\frac{3}{4} = 23/4$.

• **Ejemplo:** Para hallar los valores extremos de $f(x) = x^3 - 3x + 1$ en $[-2, 2]$, calculamos $f'(x) = 3x^2 - 3$, vemos que los puntos con $f'(x) = 0$ son ± 1, y hacemos las listas de puntos y valores críticos:

punto crítico c	-2	-1	1	2
valor crítico $f(c)$	-1	3	-1	3

Concluimos que $-2, 1$ son los mínimos globales, siendo el valor mínimo -1, y $-1, 2$ son los máximos globales, siendo el valor máximo 3.

Aunque f sea derivable, si la estudiamos en un intervalo no compacto, puede que no tenga extremos de ningún tipo (por ejemplo, $f(x) = x$ en $\mathbb{R}$). Típicamente habrá **extremos locales pero no globales** que debemos hallar. El Teorema del Punto Crítico dice que éstos se hallan **entre** los puntos críticos. Necesitamos pues un procedimiento general para, **dado un punto crítico, analizar si es o no un extremo local.** Hace falta profundizar en el significado de la derivada.

Procedimiento general para optimizar una función

Sea f una función en un intervalo I. Para hallar los extremos de f en I:

(1) Hallar los puntos críticos de f en I

(2) Aplicar una serie de ***criterios*** *que analizan un punto crítico para intentar determinar si es o no extremo, y de qué tipo.*

El resultado inicial de esta sección sobre el signo de la derivada en un punto es el punto de partida para estudiar no sólo si un punto es o no extremo de una función, que sigue siendo uno de nuestros objetivos principales, sino que, al dejar variar el punto a lo largo de un intervalo y estudiar cómo van variando los signos de la derivada, este estudio proporciona bastante más información acerca del comportamiento de la función en ese intervalo aparte de la existencia de extremos.

4.6. Derivada y crecimiento

Definición de función monótona

Una función f definida en un intervalo $I \subseteq \mathbb{R}$ es

- ***creciente*** *en I si $x, y \in I$, $x < y \implies f(x) \leq f(y)$*
- ***estrictamente creciente*** *en I si $x, y \in I$, $x < y \implies f(x) < f(y)$*
- ***decreciente*** *en I si $x, y \in I$, $x < y \implies f(x) \geq f(y)$*
- ***estrictamente decreciente*** *en I si $x, y \in I$, $x < y \implies f(x) > f(y)$*

Decimos que f es ***monótona*** *en I si es creciente o decreciente. Decimos que es* ***estrictamente monótona*** *si es estrictamente creciente o estrictamente decreciente.*

Una función constante es a la vez creciente y decreciente, pero no estrictamente.

• Ejemplo:

- x es estrictamente creciente en $\mathbb{R}$
- x^2 es estrictamente creciente en $[0, \infty)$
- $\sin x$ es estrictamente creciente en $[0, \pi/2]$

4.6.1. El signo de la tasa de variación

Sea f una función definida en algún intervalo I. Dados $a, b \in I$ con $a < b$, consideremos su tasa de variación media en $[a, b]$:

$$v(a, b) = \frac{f(b) - f(a)}{b - a}.$$

Es inmediato que

- f es estrictamente creciente si y sólo si $v(a, b) > 0$ para todos los pares $a, b \in I$ con $a < b$,
- f es estrictamente decreciente si y sólo si $v(a, b) < 0$ para estos pares.
- f es constante si y sólo si $v(a, b) = 0$ para estos pares.

Parece razonable pues que, si f es derivable, el signo de la derivada determine si f crece o decrece. Más precisamente, este vínculo estará establecido si **las tasas de variación media son derivadas.** Esto es así y constituye un resultado central del Cálculo Diferencial.

4.6.2. El Teorema del Valor Medio

Teorema del Valor Medio del Cálculo Diferencial

Sea f continua en el intervalo compacto $[a, b]$ y derivable en el interior (a, b). Entonces existe $c \in (a, b)$ tal que

$$\frac{f(b) - f(a)}{b - a} = f'(c).$$

Es decir, existe un punto intermedio en el cual la derivada es igual a la tvm.

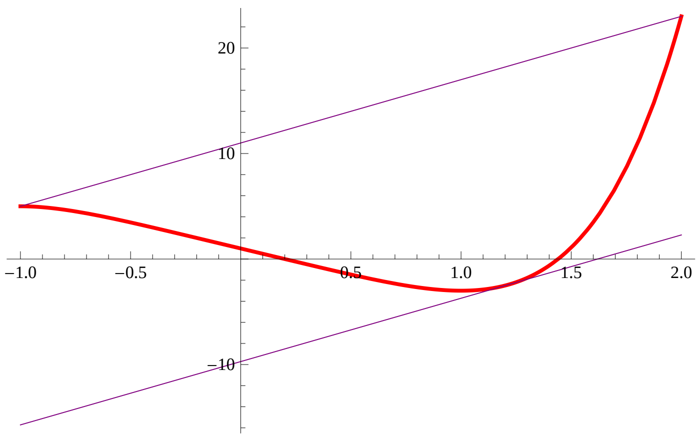

Figura 6 – Ilustración del Teorema del Valor Medio para $x^5 - 5x + 1$ en $[-1, 2]$.

Teorema de Rolle

Sea f es continua en el intervalo compacto $[a, b]$ y derivable en (a, b). Si $f(a) = f(b)$, entonces existe $c \in (a, b)$ tal que $f'(c) = 0$.

Demostración del Teorema de Rolle. Sea f continua en $[a, b]$ y derivable en (a, b) con $f(a) = f(b) \stackrel{\text{def}}{=} k$. Por el Teorema de Weierstrass, existen $\alpha, \beta \in [a, b]$ tales que

$$f(\alpha) \le f(x) \le f(\beta) \quad \forall\, x \in [a, b].$$

Se pueden dar dos posibilidades:

(1) α, β son ambos extremos del intervalo, es decir, iguales á a ó b.

(2) Al menos uno de α, β está en el interior (a, b).

En el primer caso, la hipótesis $f(a) = f(b) = k$ implica que $f(\alpha) = f(\beta) = k$, y por tanto

$$k \le f(x) \le k \quad \forall\, x \in [a, b], \quad \therefore \quad f(x) = k \quad \forall\, x \in [a, b].$$

Como f es constante, $f' = 0$ en todo (a, b), y cualquier $c \in (a, b)$ vale.

En el segundo caso, sea c cualquiera de los puntos α, β que esté en (a, b). Como c es un extremo local, y f es derivable en c, debe ser un punto crítico, por tanto $f'(c) = 0$.

En términos geométricos, si la gráfica de f representa una cuerda, sujeta a la misma altura $(f(a) = f(b))$ en a, b, o bien está tensada formando un segmento (función constante) o si no, sobresale por arriba o por abajo en algún punto interior, y entonces el máximo o mínimo global (donde más sobresale) es un punto interior de derivada nula. □

Demostración del Teorema del Valor Medio a partir del Teorema de Rolle. Sea f continua en el intervalo compacto $[a, b]$ y derivable en el interior (a, b). La ecuación de la recta que pasa por $\big(a, f(a)\big)$ y $\big(b, f(b)\big)$ es

$$\frac{y - f(a)}{x - a} = m, \qquad m = \frac{f(b) - f(a)}{b - a} \quad \therefore \quad y = \ell(x) = f(a) + m(x - a)$$

que coincide con f en a y b: $\ell(a) = f(a)$, $\ell(b) = f(b)$. Entonces,

$$\varphi(x) = f(x) - \ell(x)$$

satisface las hipótesis del Teorema de Rolle: es continua en $[a, b]$, derivable en (a, b), y $\varphi(a) = \varphi(b) = 0$. El Teorema de Rolle implica que existe $c \in (a, b)$ tal que $\varphi'(c) = 0$. Pero

$$\varphi'(c) = f'(c) - m = 0 \iff f'(c) = m = \frac{f(b) - f(a)}{b - a},$$

que es precisamente la conclusión deseada. □

El Teorema de Rolle es el caso particular $f(a) = f(b)$ del Teorema del Valor Medio, cuando la tasa de variación es $v(a, b) = 0$.

Relación entre el signo de la derivada y la monotonía en un intervalo

Sea f continua en $[a, b]$ y derivable en (a, b).

+ *Si $f'(x) > 0$ para todo $x \in (a, b)$, entonces f es estrictamente creciente en $[a, b]$.*

+ *Si $f'(x) \geq 0$ para todo $x \in (a, b)$, entonces f es creciente en $[a, b]$.*

− *Si $f'(x) < 0$ para todo $x \in (a, b)$, entonces f es estrictamente decreciente en $[a, b]$.*

− *Si $f'(x) \leq 0$ para todo $x \in (a, b)$, entonces f es decreciente en $[a, b]$.*

0 *Si $f'(x) = 0$ para todo $x \in (a, b)$, entonces f es constante en $[a, b]$.*

Demostración. Supongamos por ejemplo que $f' > 0$. Los demás casos de desigualdades son similares. Sean x, y con $a \leq x < y \leq b$. El Teorema del Valor Medio es aplicable a la restricción de f al intervalo $[x, y]$, por tanto existe $c \in (x, y)$ tal que

$$\frac{f(y) - f(x)}{y - x} = f'(c) > 0, \quad \therefore \quad f(y) - f(x) > 0, \quad \therefore \quad f(x) < f(y),$$

luego f es estrictamente creciente.

El caso donde se supone $f' = 0$, lleva a la conclusión $f(x) = f(y)$ para todo x, y, es decir, f es constante. □

Aunque el Teorema del Valor Medio es suficiente para demostrar la relación clave entre signo de la derivada y crecimiento, hay otra versión más general que a veces es necesaria.

4.6.3. El Teorema del Valor Medio de Cauchy

Teorema del Valor Medio de Cauchy

Sean f, g continuas en el intervalo compacto $[a, b]$ y derivables en el abierto (a, b). Entonces existe algún $c \in (a, b)$ tal que

$$(f(b) - f(a))g'(c) = (g(b) - g(a))f'(c).$$

En particular, si $g'(x) \neq 0$ para todo $x \in (a, b)$, entonces $g(b) \neq g(a)$ y

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)}$$

para algún $c \in (a, b)$.

Demostración. Si $g(b) = g(a)$, entonces por el Teorema del Valor Medio ordinario, existe $c \in (a, b)$ tal que $g'(c) = 0$. Dicho punto hace que ambos lados de la ecuación

$$(f(b) - f(a))g'(c) = (g(b) - g(a))f'(c)$$

sean nulos, con lo cual es válida, pero trivial.

Por tanto podemos suponer que $g(b) \neq g(a)$. Entonces procedemos de manera parecida como en la demostración del Teorema del Valor Medio ordinario, considerando

$$\varphi(x) = f(x) - mg(x), \qquad m = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Esta función es continua en $[a, b]$ y derivable en (a, b), y cumple la hipótesis del Teorema de Rolle, ya que

$$\varphi(a) = \varphi(b) \iff f(a) - mg(a) = f(b) - mg(b) \iff f(a) - f(b) = m(g(a) - g(b))$$

por tanto existe algún $c \in (a, b)$ tal que $\varphi'(c) = 0$, lo cual equivale a

$$\begin{aligned} f'(c) = mg'(c) \iff f'(c) &= \left(\frac{f(b) - f(a)}{g(b) - g(a)}\right)g'(c) \\ \iff (g(b) - g(a))f'(c) &= (f(b) - f(a))g'(c) \end{aligned}$$

que es la fórmula indicada.

Como ya hemos comentado, si $g(b) = g(a)$ entonces por el Teorema del Valor Medio ordinario, existe $c \in (a, b)$ tal que $g'(c) = 0$. Por tanto, si $g'(x) \neq 0$ para todo $x \in (a, b)$, entonces $g(b) \neq g(b)$ y obtenemos la versión con los cocientes. ☐

4.7. Optimización II: patrones de signos de la derivada

Extremos y patrones de signos

Sea f una función continua en $(c - \delta, c + \delta)$, y derivable salvo posiblemente en c.

$+, -$ *: Si $f' > 0$ en $(c - \delta, c)$ y $f' < 0$ en $(c, c + \delta)$, c es un máximo local estricto de f.*

$-, +$ *: Si $f' < 0$ en $(c - \delta, c)$ y $f' > 0$ en $(c, c + \delta)$, c es un mínimo local estricto de f.*

Demostración. Por ejemplo, en el primer caso, f es estrictamente creciente en $(c - \delta, c]$ y estrictamente decreciente en $[c, c + \delta)$. Esto implica que $f(x) < f(c)$ para $x \in (c - \delta, c)$ y $f(c) > f(x)$ para $x \in (c, c + \delta)$, por tanto c es un máximo local estricto. ☐

En cuanto a los otros dos restantes patrones de signos posibles, éstos llevan a una definición adicional, en el caso de un punto crítico.

Punto crítico de inflexión

Sea f derivable en $(c-\delta, c+\delta)$ con $f'(c)=0$. Decimos que c es un punto de inflexión si:

$+,+$ *: $f'(x) > 0$ para $x \in (c-\delta, c+\delta), x \neq c$, ó*

$-,-$ *: $f'(x) < 0$ para $x \in (c-\delta, c+\delta), x \neq c$,*

*concretando que es **ascendente** en el primer caso y **descendente** en el segundo.*

Más adelante daremos una definición general de punto de inflexión que es aplicable también a puntos que no son críticos.

Constancia del signo en un intervalo abierto sin puntos críticos

Si f es derivable en (a,b), con derivada f' continua, y (a,b) no contiene ningún punto crítico de f, entonces f' no cambia de signo entre a y b.

Demostración. Es una consecuencia inmediata del Teorema de Bolzano: si f' cambiara de signo entre a y b, en algún punto de (a,b) habría un punto c con $f'(c)=0$, es decir, un punto crítico entre a y b, contradiciendo la hipótesis. $\square$

El resultado anterior no debe llevar a pensar que todas las funciones derivables tienen intervalos de signo constante de la derivada. Los puntos críticos pueden ser mucho más complicados de lo que la regla de signos sugiere.

• Ejemplo:

La mera existencia de la derivada f' no implica que ésta sea continua. Un ejemplo habitual es la función

$$f(x) = \begin{cases} x^2 \sin \dfrac{1}{x} & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$$

Es un ejemplo de función derivable con derivada discontinua con infinitos cambios de signo a cada lado de un punto crítico. No se puede aplicar el criterio del patrón de signos.

• En efecto, se tiene

$$f'(x) = \begin{cases} 2x \sin \dfrac{1}{x} + x^2 \left(\cos \dfrac{1}{x}\right) \cdot \left(-\dfrac{1}{x^2}\right) = 2x \sin \dfrac{1}{x} - \cos \dfrac{1}{x} & \text{si } x \neq 0 \\ \lim\limits_{x \to 0} \dfrac{f(x)-f(0)}{x-0} = \lim\limits_{x \to 0} \dfrac{x^2 \sin \frac{1}{x}}{x} = \lim\limits_{x \to 0} x \sin \dfrac{1}{x} = 0 & \text{si } x = 0. \end{cases}$$

Por tanto f es derivable en todo $\mathbb{R}$; $c = 0$ es un punto crítico, pero $f'(x)$ no es continua en 0, pues no existe $\lim_{x\to 0} f'(x)$, ya que el primer término cumple $\lim_{x\to 0} 2x \sin \frac{1}{x} = 0$ pero el segundo, $\cos \frac{1}{x}$, no tiene límite, sino que oscila entre -1 y $+1$ a medida que $x \to 0$, lo cual hace que $f'(x)$ cambie infinitas veces de signo al acercarse a 0, en cualquier dirección.

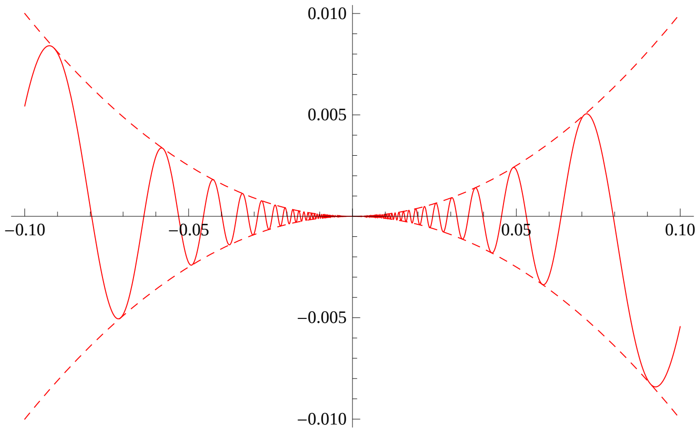

Figura 7 – Gráfica de $x^2 \sin \dfrac{1}{x}$.

De hecho, se puede comprobar que f tiene infinitos puntos críticos (las crestas y valles de la oscilación atenuada) a cada lado en cualquier entorno de 0. No hay manera de usar el orden de $\mathbb{R}$ para ordenar los puntos críticos empezando con $c = 0$. ¿Cuál sería el “siguiente”?

Por el mismo motivo, no debe malinterpretarse el resultado sobre el signo de la derivada en un punto confundiéndolo con el resultado sobre el signo en un intervalo. **No** es verdad, por ejemplo, que si f es derivable y en un cierto punto c se tiene $f'(c) > 0$, entonces f es creciente en un intervalo alrededor de c. **Es cierto si f' es continua** por conservación del signo, pero si no, puede ser falso.

Intervalos de signo constante de la derivada primera

Sea f una función derivable en un intervalo abierto (a, b) tal que

- *La **derivada** f' es una función **continua***
- *f sólo tiene un **número finito de puntos críticos** $a < c_1 < c_2 < \cdots < c_n < b$.*

Entonces

- *El signo de la derivada es constante en cada subintervalo determinado por los puntos críticos:*

$$I_0 = (a, c_1), \quad I_k = (c_k, c_{k+1}),\ 1 \le k \le n-1, \quad I_n = (c_n, b).$$

- *f es estrictamente monótona en cada uno de los subintervalos I_k anteriores.*
- *El criterio del patrón de signos de f' siempre es capaz de determinar el tipo de cada punto crítico c_k.*

Demostración. La constancia del signo en cada I_k es consecuencia inmediata del resultado anterior sobre la constancia del signo en un intervalo sin puntos críticos. La monotonía de f en cada I_k es consecuencia de la relación entre la monotonía y el signo de la derivada. Finalmente, como el signo está bien determinado a cada lado de cualquier punto crítico, el criterio del patrón de signos siempre determinará el tipo del punto crítico (máximo, mínimo o punto de inflexión). □

• **Ejemplo:** Los polinomios cumplen las condiciones anteriores, en cualquier intervalo. Las funciones racionales las cumplen en los intervalos que no contengan ceros del denominador. El ejemplo que vimos de función con derivada discontinua **no** las cumple.

Determinación sencilla del signo de la derivada

• *Si f es derivable con derivada continua y un número finito de puntos críticos, para determinar cuál es el signo de f' en uno de los subintervalos determinados por los puntos críticos, basta con escoger un punto cualquiera x del intervalo y ver qué signo tiene $f'(x)$. En caso de analizar una función en un intervalo infinito, como (c, ∞), también sirve analizar el signo del límite $\lim_{x\to\infty} f(x)$, si éste límite existe.*

• *Si $f'(x)$ factoriza como un producto de un cierto número de términos, el signo de $f'(x)$ será el producto de los signos de cada factor. A veces, el proceso de determinar los puntos críticos requiere una tal factorización y de paso nos sirve también para determinar los signos de $f'(x)$.*

• **Ejemplo:** $f(x) = x^3$ tiene un punto de inflexión en 0, con patrón $+, +$, ya que $f'(x) = 3x^2 > 0$ si $x \neq 0$. En cambio $f(x) = -x^3$ tiene un punto de inflexión descendente en 0.

• **Ejemplo:** De manera similar, $f(x) = x^n$ tiene un punto de inflexión ascendente en 0 para todo $n \geq 3$ impar, pues entonces $f'(x) = nx^{n-1} > 0$ si $x \neq 0$.

• **Ejemplo:** Sea $f(x) = x^3 - 3x + 1$. Entonces

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

Los puntos críticos son $-1, 1$. Se tiene

intervalo	$(-\infty, -1)$	$(-1, 1)$	$(1, \infty)$
signo de f'	$+$	$-$	$+$

por tanto -1 es un máximo local estricto y 1 es un mínimo local estricto. Como f es derivable en todo $\mathbb{R}$, no hay más puntos críticos.

• **Ejemplo:** Sea $f(x) = x^4 + 4x^3 + 1$. Entonces

$$f'(x) = 4x^3 + 12x^2 = 4x^2(x+3).$$

Los puntos críticos son $c = -3, 0$. Se tiene

intervalo	$(-\infty, -3)$	$(-3, 0)$	$(0, \infty)$
signo de f'	$-$	$+$	$+$

por tanto -3 es un mínimo local estricto y 0 es un punto de inflexión ascendente.

4.8. Derivadas de Orden Superior

Sea f derivable en un intervalo abierto I. Su derivada es una nueva función, f', en I. A su vez, f' puede ser derivable. Su derivada se denota por

$$(f')' = f''$$

Si f'' es derivable, escribimos

$$(f'')' = f'''$$

para su derivada. Si este proceso se puede repetir n veces, decimos que f es **derivable n veces,** y definimos por inducción

$$\begin{aligned} f^{(0)} &= f \\ f^{(n)} &= \left(f^{(n-1)}\right)' \quad (n \geq 1). \end{aligned}$$

Decimos que $f^{(n)}$ es la **derivada de f de orden n**. Si f se puede derivar cualquier número n de veces, decimos que es **infinitamente derivable.**

• **Ejemplo:** Como $\exp'(x) = \exp(x)$, se tiene $\exp^{(n)}(x) = \exp(x)$ para todo $n \in \mathbb{N}$.

• **Ejemplo:** Para la función seno, tenemos

$$\begin{aligned} f(x) &= \sin x, & f'(x) &= \cos x, & f''(x) &= -\sin x, & f'''(x) &= -\cos x, \\ f^{(4)}(x) &= f(x) = \sin x, & f^{(5)} &= f^{(1)}(x), & f^{(6)} &= f^{(2)}(x), & f^{(7)} &= f^{(3)}(x), \end{aligned}$$

y en general $f^{(4n+r)}(x) = f^{(r)}(x)$ para $n \in \mathbb{N}, r = 0, 1, 2, 3$.

4.9. Optimización III: criterio de la derivada segunda

Criterio de la derivada segunda

Sea f derivable en un entorno de c, con $f'(c) = 0$, y dos veces derivable en c.

- *Si $f''(c) > 0$, c es un mínimo local estricto de f*
- *Si $f''(c) < 0$, c es un máximo local estricto de f*
- *Si $f''(c) = 0$, c no se obtiene información acerca de si c es extremo.*

Demostración. Supongamos que $f''(c) > 0$. Como $f'(c) = 0$, se tiene

$$f''(c) = \lim_{x \to c} \frac{f'(x) - f'(c)}{x - c} = \lim_{x \to c} \frac{f'(x)}{x - c} > 0.$$

Por conservación del signo en los límites, debe ser $\dfrac{f'(x)}{x - c} > 0$ para $x \approx c^*$. Entonces

$$x \approx c^-, x \neq c \implies x - c < 0 \implies f'(x) < 0$$

$$x \approx c^+, x \neq c \implies x - c > 0 \implies f'(x) > 0$$

luego f' exhibe el patrón de signos $-, +$ alrededor de c, y por tanto c es un mínimo local estricto. El razonamiento para $f''(c) < 0$ es similar.

Dicho en palabras, como el denominador $x - c$ cambia de signo a uno y otro lado de c pero el cociente es positivo, el numerador debe tener el mismo signo que el denominador, llevando al patrón señalado.

Comparar esta demostración con la del resultado sobre el signo de f' en un punto que lleva al Teorema del Punto Crítico. □

La demostración anterior no revela la razón más profunda de por qué se tienen resultados de este tipo. Esa razón hay que buscarla en la teoría de aproximación, que veremos en la sección 4.12. La idea consiste en comparar una función con un polinomio cuadrárico (geométricamente, una parábola).

Si $f''(c) = 0$ en un punto crítico c, éste puede ser de cualquier tipo:

- $f(x) = x^4$ en $c = 0$ tiene $f''(c) = 0$ y c es mínimo local estricto.
- $f(x) = -x^4$ en $c = 0$ tiene $f''(c) = 0$ y c es máximo local estricto.
- $f(x) = x^3$ en $c = 0$ tiene $f''(c) = 0$ y c es punto de inflexión ascendente.
- $f(x) = -x^3$ en $c = 0$ tiene $f''(c) = 0$ y c es punto de inflexión descendente.

• **Ejemplo:** Sea $f(x) = x^3 - 3x + 1$. Entonces

$$f'(x) = 3x^2 - 3 = 3(x + 1)(x - 1), \qquad f''(x) = 6x.$$

Los puntos críticos son $-1, 1$. Se tiene

punto crítico c	-1	1
$f''(c)$	-6	6
signo de $f''(c)$	$-$	$+$

Por tanto -1 es un máximo local estricto y 1 es un mínimo local estricto. $f(x)$ no puede tener extremos globales, pues $\lim_{x \to -\infty} f(x) = -\infty$ y $\lim_{x \to \infty} f(x) = \infty$.

• **Ejemplo:** Sea $f(x) = x^4 + 4x^3 + 1$. Entonces

$$f'(x) = 4x^3 + 12x^2 = 4x^2(x+3), \qquad f''(x) = 12x^2 + 24x = 12x(x+2).$$

Los puntos críticos son $c = -3, 0$. Se tiene

punto crítico c	-3	0
$f''(c)$	36	0
signo de $f''(c)$	$+$	0

por tanto -3 es un mínimo local estricto. Como $f''(0) = 0$, no podemos concluir nada con este criterio. Sin embargo, como vimos antes, analizando los signos de f' alrededor de 0 se concluye que es un punto de inflexión. Como $\lim_{x\to\pm\infty} f(x) = \infty$, se deduce que $f(x)$ tiene un mínimo global estricto, que debe estar en -3.

4.10. Convexidad

Los resultados anteriores sobre puntos críticos, crecimiento y decrecimiento, e incluso la derivada segunda, ayudan a **dibujar las gráficas de las funciones.**

Definición de función convexa

Decimos que la función $f(x)$ es ***convexa*** *en un intervalo I si, para todo par de puntos $a, b \in I$, la gráfica de $f(x)$ se halla* ***por debajo*** *de la cuerda que une los puntos $(a, f(a))$ y $(b, f(b))$. Esto equivale a*

$$f\big((1-t)a + tb\big) \leq (1-t)f(a) + tf(b) \quad \forall t \in [0,1].$$

Si la desigualdad es estricta, es decir, si para $a \neq b$ se tiene

$$f\big((1-t)a + tb\big) < (1-t)f(a) + tf(b) \quad \forall t \in (0,1),$$

decimos que es ***estrictamente convexa.***

Decimos que f es (estrictamente) **cóncava** si $-f$ es (estrictamente) convexa.

En Geometría, un conjunto A es convexo si para cada par de puntos $p, q \in A$, el segmento que une p con q está contenido en A, es decir, si al viajar en línea recta desde p hasta q, no se sale fuera del conjunto A. La relación entre este concepto y la convexidad de funciones es la siguiente:

$f : I \to \mathbb{R}$ es una función convexa si y sólo si el conjunto de puntos que se hallan **por encima** de su gráfica,

$$A = \{(x, y) : y \geq f(x)\},$$

es un conjunto convexo en el sentido geométrico.

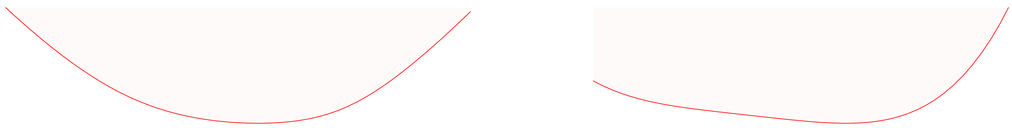

Figura 8 – Funciones convexas.

A la vista de dibujos como el de la Figura 8, la terminología puede parecer contradictoria con el uso coloquial de los términos, pues en el lenguaje común se llamaría cóncavo a un recipiente con la forma dada, siendo concavidad una hendidura. Realmente es cuestión de perspectiva. Por ello también se dice de una función convexa que es **cóncava hacia arriba** y de una función cóncava que es **cóncava hacia abajo.** En Matemáticas, la definición que hemos dado de función convexa es la aceptada universalmente.

Criterio de convexidad de la segunda derivada

- *Si f es dos veces derivable en un intervalo abierto I entonces f es convexa en I si y sólo si $f'' \geq 0$.*
- *Si $f'' > 0$ en I entonces f es estrictamente convexa en I, pero el recíproco no es cierto en general.*

• **Ejemplo:** La función $f(x) = x^2$ es estrictamente convexa en $\mathbb{R}$ ya que $f''(x) = 2 > 0$.

• **Ejemplo:** La función $f(x) = -x^2$ es estrictamente cóncava en $\mathbb{R}$ ya que $f''(x) = -2 < 0$.

• **Ejemplo:** La función $f(x) = x^4$ es estrictamente convexa en $\mathbb{R}$ aunque $f''(x) = 12x^2 > 0$ si $x \neq 0$ pero $f''(0) = 0$. Lo mismo ocurre con $f(x) = x^{2n}$ para $n \geq 2$ natural.

• **Ejemplo:** Sea $f(x) = x^4 + 4x^3 + 1$. Entonces

$$f'(x) = 4x^3 + 12x^2 = 4x^2(x+3), \qquad f''(x) = 12x^2 + 24x = 12x(x+2).$$

Por tanto los signos de f'' vienen dados por

intervalo	$(-\infty, -2)$	$(-2, 0)$	$(0, \infty)$
signo de f''	$+$	$-$	$+$
convexidad	convexa	cóncava	convexa

Entonces $f(x)$ es convexa en $x < -2$ y $x > 0$, y cóncava en $-2 < x < 0$.

Definición geométrica de un punto de inflexión

Un punto c es un ***punto de inflexión*** *de una función f si en c, f cambia de convexa a cóncava o viceversa. Geométricamente, si cambia de "abrirse hacia arriba" á "abrirse hacia abajo" o viceversa. El punto es* ***ascendente*** *si la primera derivada es positiva (f es creciente) a ambos lados de c y* ***descendente*** *(f es decreciente) si es negativa.*

Para saber que todo punto de inflexión es ascendente o descendente, hace falta demostrar que en un punto de inflexión, la primera derivada de hecho tiene un signo fijo a ambos lados. Esto se puede hacer basándose en el siguiente resultado.

Criterios para puntos de inflexión

Supongamos que f es dos veces derivable en un entorno de c.

- *Si f'' cambia de signo en c, es decir, $f''(c) = 0$ y f'' tiene signos distintos a cada lado de c, entonces c es un punto de inflexión de f.*
- *Si c es un punto de inflexión de f, entonces $f''(c) = 0$, pero esta condición no es en general suficiente. Si f es derivable n veces y se tiene*

$$f''(c) = \cdots = f^{(n-1)}(c) = 0, \text{ pero } f^{(n)}(c) \neq 0,$$

entonces c es un punto de inflexión si n es impar, ascendente si $f'(c) > 0$ y descendente si $f'(c) < 0$.

• **Ejemplo:** Por ejemplo, $f(x) = x^3$ tiene un punto de inflexión en 0 ya que $f''(0) = 0$ pero $f'''(0) = 6 \neq 0$, y 3 es impar.

• **Ejemplo:** Sin embargo, $f(x) = x^4$ también cumple $f''(0) = 0$, pero esta vez $f'''(0) = 0$ y $f^{(4)}(0) = 24 \neq 0$, y 0 no es un punto de inflexión, es un mínimo.

4.11. Aplicaciones geométricas: dibujos de gráficas

Los diversos resultados anteriores sirven también para hacerse una idea cualitativa de la variación de una función, permitiendo dibujar su gráfica a grandes trazos, sin necesidad de evaluarla en muchos puntos.

Vamos a revisitar algunos de los ejemplos que hemos visto anteriormente, aplicando toda la información obtenida en las secciones anteriores para analizar la variación de las funciones. Las funciones que vamos a ver cumplen las condiciones de constancia del signo de la derivada primera. De hecho, aplicando el mismo criterio a f' en vez de f, vemos que si f'' es continua y tiene un número finito de ceros, también tiene signo constante entre ellos.

• **Ejemplo:** Sea $f(x) = x^4 + 4x^3 + 1$. Entonces

- $f'(x) = 4x^3 + 12x^2 = 4x^2(x + 3)$
- $f''(x) = 12x^2 + 24x = 12x(x + 2)$

Resolviendo $f'(x) = 0$ nos da la lista de puntos críticos: $-3, 0$. Estos puntos separan $\mathbb{R}$ en tres intervalos, y los signos de f' vienen dados por

intervalo	$(-\infty, -3)$	$(-3, 0)$	$(0, \infty)$
punto/valor	$f'(-\infty) = -\infty$	$f'(-1) = 8$	$f'(\infty) = \infty$
signo de f'	$-$	$+$	$+$
crecimiento	decreciente	creciente	creciente

Hemos usado el método de elegir un punto arbitrario en cada intervalo para evaluar f' y así determinar su signo, que es constante en el intervalo. También podíamos considerar cada factor por separado, prestando especial atención a los cuadrados pues nunca son negativos:

intervalo	$(-\infty, -3)$	$(-3, 0)$	$(0, \infty)$
signo de x^2	$+$	$+$	$+$
signo de $x+3$	$-$	$+$	$+$
signo de f'	$-$	$+$	$+$

El análisis de los patrones de signos de f' en los puntos críticos nos lleva a la clasificación

punto crítico	-3	0
patrón	$-+$	$++$
tipo	mínimo	inflexión ascendente

El análisis de los puntos críticos mediante el criterio de la derivada segunda lleva a

punto crítico c	-3	0
$f''(c)$	36	0
tipo	mínimo	indeterminado

De hecho, -3 es el mínimo global estricto.

• En cuanto a la derivada segunda, sus ceros en $-2, 0$ separan $\mathbb{R}$ en otros tres intervalos, y sus signos vienen dados por

intervalo	$(-\infty, -2)$	$(-2, 0)$	$(0, \infty)$
punto/valor	$f''(-\infty) = \infty$	$f''(-1) = -12$	$f''(\infty) = \infty$
signo de f''	$+$	$-$	$+$
convexidad	convexa	cóncava	convexa

Los candidatos a puntos de inflexión son los ceros de f'', es decir, -2 y 0. Usando los criterios para puntos de inflexión, dado que tanto en -2 como en 0 la derivada segunda f'' cambia de signo, ambos son puntos de inflexión. Mirando los signos de la primera derivada, se concluye que ambos son puntos de inflexión ascendentes.

• Finalmente, esta información se ve reflejada en la gráfica de f:

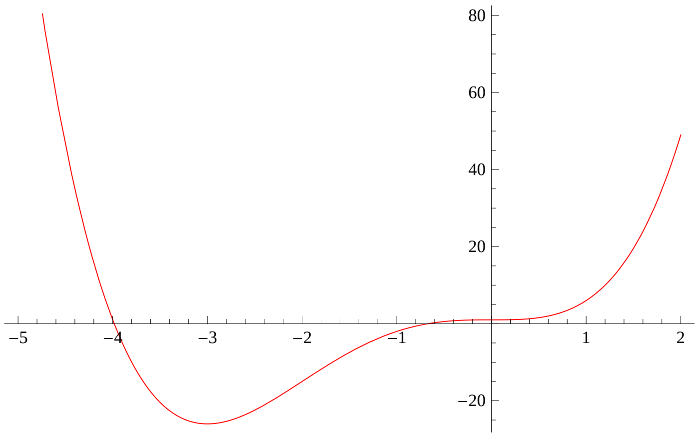

Figura 9 – gráfica de $x^4 + 4x^3 + 1$

• **Ejemplo:** Sea $f(x) = x^3 - 3x + 1$. Entonces

- $f'(x) = 3x^2 - 3 = 3(x+1)(x-1)$
- $f''(x) = 6x$

Resolviendo $f'(x) = 0$ nos da la lista de puntos críticos: $-1, 1$. Estos puntos separan $\mathbb{R}$ en tres intervalos,

intervalo	$(-\infty, -1)$	$(-1, 1)$	$(1, \infty)$
punto/valor	$f'(-\infty) = \infty$	$f'(0) = -3$	$f'(\infty) = \infty$
signo de f'	$-$	$-$	$+$
crecimiento	creciente	decreciente	creciente

Para determinar los signos también podíamos considerar cada factor por separado:

intervalo	$(-\infty, -1)$	$(-1, 1)$	$(1, \infty)$
signo de $x + 1$	$-$	$+$	$+$
signo de $x - 1$	$-$	$-$	$+$
signo de f'	$+$	$-$	$+$

El análisis de los patrones de signos de f' en los puntos críticos nos lleva a la clasificación

punto crítico	-1	1
patrón	$+-$	$-+$
tipo	máximo	mínimo

El análisis de los puntos críticos mediante el criterio de la derivada segunda lleva a

punto crítico c	-1	1
$f''(c)$	-6	6
tipo	máximo	mínimo

En cuanto a la derivada segunda, su cero en 0 separa $\mathbb{R}$ en dos intervalos, y sus signos vienen dados por

intervalo	$(-\infty, 0)$	$(0, \infty)$
punto/valor	$f''(-\infty) = -\infty$	$f''(\infty) = \infty$
signo de f''	$-$	$+$
convexidad	cóncava	convexa

Entonces $f(x)$ es cóncava para $x < 0$ y convexa para $x > 0$. Los candidatos a puntos de inflexión son los ceros de f'', es decir, 0. Usando los criterios para puntos de inflexión, dado que tanto en 0 la derivada segunda f'' cambia de signo, es un punto de inflexión. Mirando el signo de la primera derivada, se concluye que es un punto de inflexión descendente.

Finalmente, esta información se ve reflejada en la gráfica de f:

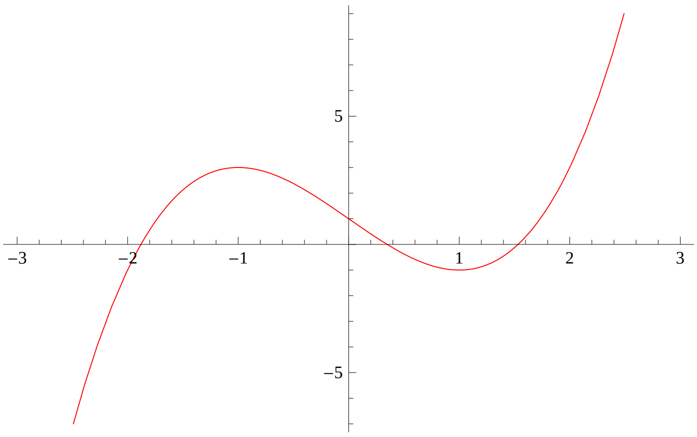

Figura 10 – gráfica de $x^3 - 3x + 1$

4.11.1. Asíntotas

En Geometría, una recta asintótica a una curva es una recta a la cual se aproxima la curva pero sin llegar nunca a alcanzarla. Cuando la curva es la gráfica de una función de una variable, las asíntotas corresponden a límites infinitos de la función en puntos finitos y a límites finitos en puntos infinitos.

Decimos que f tiene una **asíntota vertical** en a si $\lim_{x\to a^\pm} f(x) = \pm\infty$. Geométricamente, esto se refleja en que la gráfica de f se acerca a la recta vertical $x = a$ a medida que x se acerca á a.

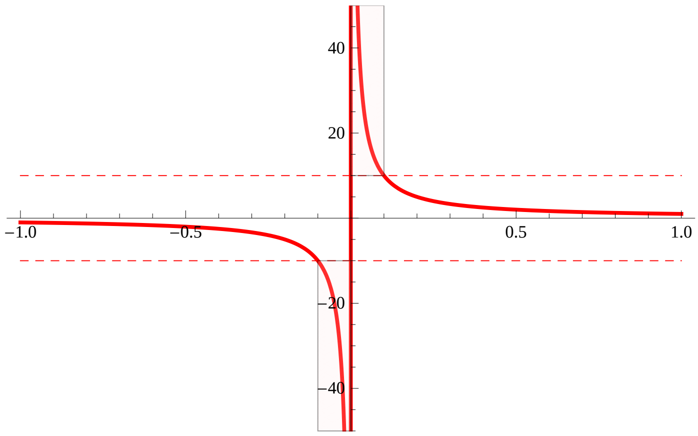

Figura 11 – Una asíntota vertical

Es la expresión geométrica de la definición del límite: si por ejemplo $\lim_{x\to a^+} f(x) = \infty$, dada una cota arbitrariamente grande M, existe $\delta > 0$ tal que $a < x < a+\delta \implies f(x) > M$. Esto significa que la parte de la gráfica de f sobre $(a, a+\delta)$ se halla dentro de la banda vertical de anchura δ dada por $\{(x,y) : a < x < a+\delta, y > M\}$ (ver la figura 11).

• **Ejemplo:** $f(x) = \dfrac{1}{x}$ tiene una asíntota vertical en 0.

Decimos que f tiene una **asíntota horizontal** si $\lim_{x\to\pm\infty} f(x) = k^{\pm}$ (se acerca á k por arriba o por abajo) para algún $k \in \mathbb{R}$. Geométricamente, esto se refleja en que la gráfica de f se acerca a la recta horizontal $y = k$ a medida que x se aleja a infinito.

Si por ejemplo $\lim_{x\to\infty} f(x) = k^+$ entonces, dado $\varepsilon > 0$, existe una cota $M > 0$ tal que $x > M \implies k < f(x) < k+\varepsilon$. Es decir, la gráfica de f sobre (M, ∞) se halla dentro de la banda horizontal de anchura ε dada por $\{(x,y) : x > M, k < y < k+\varepsilon\}$.

• **Ejemplo:** $f(x) = \dfrac{1}{x}$ tiene $y = 0$ como asíntota horizontal.

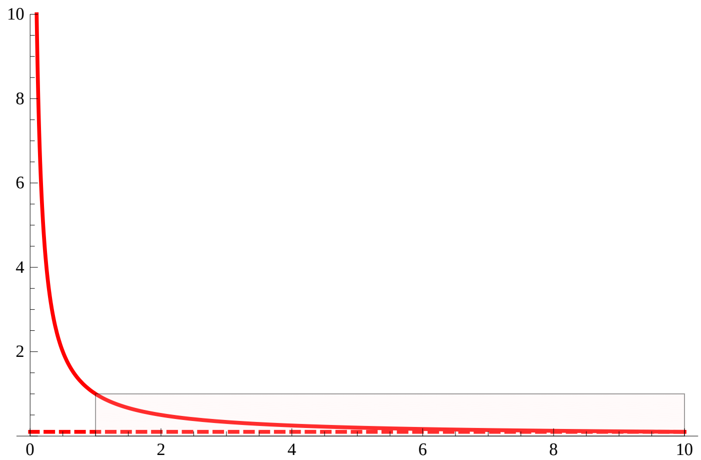

Figura 12 – Una asíntota horizontal

Técnicamente, con esta definición, una función con $\lim_{x\to\pm\infty} f(x) = k$ pero que oscila por encima y por debajo de k infinitas veces no se considera que tiene una asíntota horizontal, pues en este caso la gráfica cruza la recta infinitas veces.

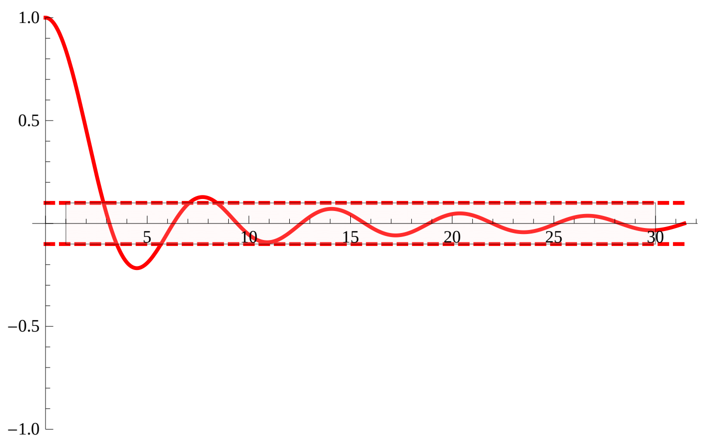

Figura 13 – Aproximación oscilatoria a una recta horizontal.

Decimos que $f(x)$ tiene una **asíntota oblicua** si existen $m \neq 0, k \in \mathbb{R}$ tales que

$$\lim_{x\to\pm\infty} (f(x) - mx) = k^{\pm}$$

Geométricamente, esto se refleja en que la gráfica de f se acerca a la recta $y = mx + k$ por arriba o por debajo a medida que x se aleja a infinito.

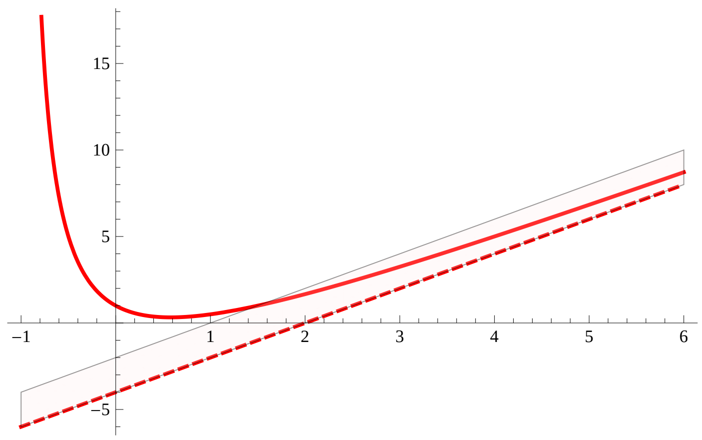

Figura 14 – Asíntota oblicua.

Esto significa geométricamente que, dado $\varepsilon > 0$, la gráfica de f sobre un intervalo (M, ∞) con M suficientemente grande, se halla dentro de la banda con base en la recta $y = mx + k$ y altura ε dada por $\{(x, y) : x > M, mx + k < y < mx + k + \varepsilon\}$.

Para determinar si una función tiene una asintota oblicua, por ejemplo cuando $x \to \infty$, observamos que

$$\lim_{x\to\infty} (f(x) - mx - k) = 0 \implies \lim_{x\to\infty} \frac{f(x)}{x} = m.$$

Esto nos determina la pendiente m, si existe, y luego

$$k = \lim_{x\to\infty} (f(x) - mx)$$

• **Ejemplo:** Sea $f(x) = \dfrac{2x^2+1}{x+1}$. Se tiene

$$\lim_{x\to\pm\infty} \frac{f(x)}{x} = \lim_{x\to\pm\infty} \frac{2x^2+1}{x^2+x} = 2,$$

y luego

$$\lim_{x\to\pm\infty} (f(x) - 2x) = \lim_{x\to\pm\infty} \left(\frac{2x^2+1}{x+1} - 2x\right) = \lim_{x\to\pm\infty} \left(\frac{2x^2+1-2x^2-2x}{x+1}\right)$$

$$= \lim_{x\to\pm\infty} \left(\frac{1-2x}{x+1}\right) = -2$$

por tanto $y = 2x - 2$ es una asíntota oblicua de f tanto en ∞ como en $-\infty$. Hagamos un análisis completo de esta función para realizar un dibujo de su gráfica.

• También tiene $x = -1$ como asíntota vertical, ya que $\lim_{x\to-1^{\pm}} = \pm\infty$.

• Se tiene

$$f'(x) = \frac{4x(x+1) - (2x^2+1)}{(x+1)^2} = \frac{2x^2+4x-1}{(x+1)^2}$$

por tanto sus puntos críticos son

$$x = \frac{-4 \pm \sqrt{24}}{4} = -1 \pm \frac{\sqrt{6}}{2}.$$

Denotemos $\alpha = -1 - \frac{\sqrt{6}}{2} \approx -2{,}2247$ y $\beta = -1 + \frac{\sqrt{6}}{2} \approx 0{,}2247$ de modo que

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

Hay que excluir el punto $x = -1$ del dominio de f ya que se hace infinita en él. Obtenemos la siguiente distribución de signos de la primera derivada:

intervalo	$(-\infty, \alpha)$	$(\alpha, -1)$	$(-1, \beta)$	(β, ∞)
punto/valor	$f'(-\infty) = \infty$	$f'(-2) = -\frac{1}{9}$	$f'(0) = -1$	$f'(\infty) = \infty$
signo de f'	$+$	$-$	$-$	$+$
crecimiento	creciente	decreciente	decreciente	creciente

De esto se deduce que α es un máximo y β un mínimo. Esto también se podía ver analizando la derivada segunda, que simplifica bastante:

$$
\begin{aligned}
f''(x) &= \frac{(4x+4)(x+1)^2 - (2x^2+4x-1)(2(x+1))}{(x+1)^4} \\
&= \frac{4(x+1)^3 - (-2+6x+12x^2+4x^3)}{(x+1)^4} = \frac{4(x+1)^3 - 2(1+x)(-1+4x+2x^2)}{(x+1)^4} \\
&= \frac{4x^2+8x+4+2-8x-4x^2}{(x+1)^3} = \frac{6}{(x+1)^3}.
\end{aligned}
$$

Entonces $f''(\alpha) < 0$ y $f''(\beta) > 0$, lo cual confirma que α es máximo y β mínimo.

punto crítico c	α	β
$f''(c)$	$\frac{6}{(-\sqrt{6}/2)^3}$	$\frac{6}{(\sqrt{6}/2)^3}$
tipo	máximo	mínimo

En cuanto a convexidad, $f''(x) \neq 0$ así que no hay puntos de inflexión, y debemos quitar el punto -1 para estudiar los signos:

intervalo	$(-\infty, -1)$	$(-1, \infty)$
punto/valor	$f''(-\infty) = -\infty$	$f''(\infty) = \infty$
signo de f''	$-$	$+$
convexidad	cóncava	convexa

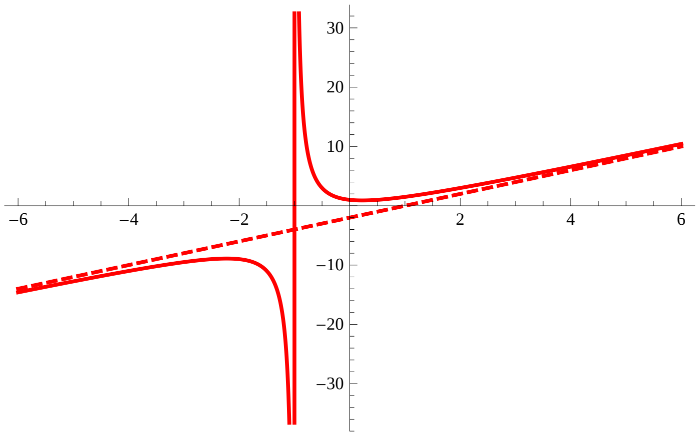

Figura 15 – Gráfica de $\dfrac{2x^2+1}{x+1}$ mostrando su asíntota oblicua $y = 2x - 2$.

El objetivo de estos análisis sobre las funciones es adquirir la capacidad de "ver" una ecuación y poder hacerse una idea cualitativa de cómo es su gráfica, es decir, de cómo varía la función. Es cierto que uno puede dibujar gráficas con cierta precisión siguiendo el método de dibujar suficientes puntos de su gráfica, pero ante la presencia de fenómenos como oscilaciones comprimidas a un punto o asíntotas, no es tan fácil, ya que para hacer una gráfica correcta se precisarían cada vez más y más puntos, y sin los conocimientos teóricos, no se puede saber qué puntos de la gráfica van a causar este tipo de problemas. Por todo esto, sigue siendo extremadamente útil entender los principios expuestos para entender la variación de las funciones.

Si multiplicamos la función del ejemplo anterior por $\sin x$, ¿cómo cambia su comportamiento?

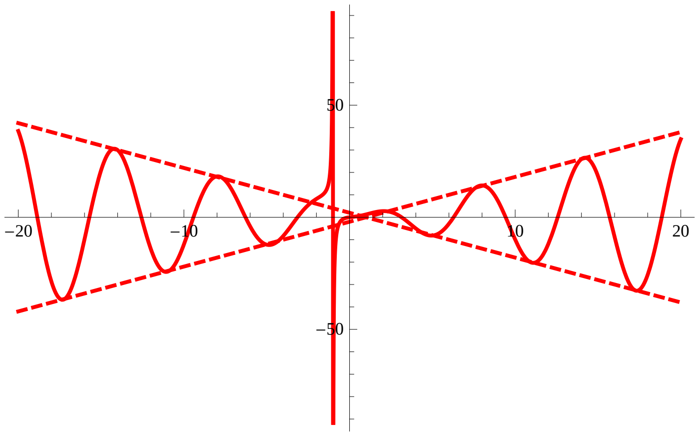

Figura 16 – Gráfica de $\dfrac{2x^2+1}{x+1}\sin x$.

4.12. El desarrollo de Taylor

Sea f una función definida en el intervalo $(a-\delta, a+\delta)$ y derivable n veces. El **polinomio de Taylor** de orden n de f en a es

$$p_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k$$

• Se puede añadir el caso $n=0$, entendiendo entonces que el polinomio de Taylor de orden 0 es simplemente la función constante con valor $f(a)$.

• El polinomio de Taylor de orden 1 es la ecuación de la recta tangente:

$$p_1(x) = f(a) + f'(a)(x-a)$$

Caracterización del polinomio de Taylor

El polinomio de Taylor de $f(x)$ en $x=a$ es el ***único*** *polinomio $p(x)$ de grado $\leq n$ cuyas derivadas en a hasta orden n coinciden con las de f:*

$$f^{(k)}(a) = p^{(k)}(a), \quad 0 \leq k \leq n.$$

Demostración. Se basa en que, dado un polinomio de la forma

$$p(x) = a_0 + a_1(x-a) + \cdots + a_n(x-a)^n,$$

siempre se cumple que

$$a_k = \frac{p^{(k)}(a)}{k!}, \quad 0 \le k \le n.$$

Aplicado al polinomio de Taylor, nos dice que $f^{(k)}(a) = p^{(k)}(a)$ para $0 \le k \le n$. □

Teorema de Taylor

Si f está definida en $(a-\delta, a+\delta)$, es $n-1$ veces derivable en este intervalo, y derivable n veces en a, con polinomio de Taylor $p_n(x)$, entonces existe una función $\varepsilon(x)$ tal que

$$f(x) = p_n(x) + (x-a)^n \varepsilon(x), \qquad \lim_{x\to a} \varepsilon(x) = 0.$$

Esta fórmula es la ***aproximación de Taylor.*** *Abreviada con la O-notación (2.8):*

$$f(x) = p_n(x) + o((x-a)^n) \quad (x \to a).$$

Si la función es derivable $n+1$ veces en $(a-\delta, a+\delta)$ y $f^{(n+1)}$ es continua, entonces

$$f(x) = p_n(x) + O((x-a)^{n+1}) \quad (x \to a)$$

Demostración. Para $n = 1$, la hipótesis es que f sea continua en el intervalo y derivable en a. Poniendo

$$\varepsilon(x) = \frac{f(x) - f(a)}{x-a} - f'(a),$$

se obtiene la función requerida. En general, para f como en el enunciado del teorema, queremos demostrar que

$$\lim_{x\to a} \frac{f(x) - p_n(x)}{(x-a)^n} = 0.$$

Este límite es una indeterminación $0/0$, y como $f^{(k)}(a) = p_n^{(k)}(a),\ 0 \le k \le n$, lo sigue siendo durante $n-1$ aplicaciones de la regla de L'Hôpital, llegando finalmente al límite equivalente

$$\lim_{x\to a} \frac{f^{(n-1)}(x) - p_n^{(n-1)}(x)}{n!(x-a)} = 0.$$

Observando que

$$p_n(x) = \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{f^{(n-1)}(a)}{(n-1)!}(x-a)^{n-1} + \cdots$$

se tiene

$$p_n^{(n-1)}(x) = f^{(n)}(a)(x-a) + f^{(n-1)}(a)$$

y por tanto basta con demostrar

$$\lim_{x\to a} \frac{f^{(n-1)}(x) - f^{(n-1)}(a) - f^{(n)}(a)(x-a)}{(x-a)} = 0,$$

pero esto último es consecuencia de la existencia de $f^{(n)}(a)$:

$$\lim_{x\to a} \frac{f^{(n-1)}(x) - f^{(n-1)}(a)}{(x-a)} = f^{(n)}(a).$$

Dejaremos sin demostrar la segunda parte correspondiente a las funciones que son $n+1$ veces derivables con derivada $(n+1)$-ésima continua. □

Si $p_n(x)$ es el polinomio de Taylor de orden n de una función f en a, la diferencia

$$R_n(x) = f(x) - p_n(x)$$

se denomina el **resto** de orden n en la aproximación. Representa una medida del **error** que se comete al usar $p_n(x)$ como valor aproximado de $f(x)$.

Hemos visto que en las condiciones mínimas, se tiene

$$R_n(x) = o((x-a)^n) \quad (x \to a)$$

y de hecho se puede mejorar a

$$R_n(x) = O((x-a)^{n+1}) \quad (x \to a)$$

si se supone f derivable $n+1$ veces con derivada $(n+1)$-ésima continua. En las aplicaciones a la aproximación numérica de funciones, son útiles expresiones más concretas para el resto.

Forma de Lagrange para el resto en la aproximación de Taylor

Sea f una función $n+1$ veces derivable en un intervalo I y $a, x \in I$. Sea $p_n(x)$ el polinomio de Taylor de f en a de orden n. Entonces existe un punto $c \in (a, x)$ tal que

$$f(x) = p_n(x) + \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}$$

En particular, si $|f^{(n+1)}(x)| \leq M$ para todo $x \in I$, se tiene

$$|f(x) - p_n(x)| \leq \frac{M}{(n+1)!}|x-a|^{n+1} \quad \forall\, x \in I.$$

Tal cota existe por ejemplo si se supone $f^{(n+1)}$ continua, como consecuencia del Teorema de Weierstrass.

• **Ejemplo:** Sea $f(x) = e^x$ y consideremos $a = 0$. Como $f^{(k)} = f$ para todo $k \geq 0$, el polinomio de Taylor de f de orden n en 0 es

$$p_n(x) = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}.$$

Consideremos un intervalo $I = [-b, b]$. Entonces $|f^{(n+1)}(x)| = e^x \leq e^b$ para todo $x \in I$. Por tanto

$$|e^x - p_n(x)| \leq \frac{e^b}{(n+1)!}|x|^{n+1} \quad \forall\, x \in [-b, b].$$

Fijados b y x, se puede demostrar que esta cota tiende a 0 cuando $n \to \infty$, por lo cual

$$e^x = \lim_{n\to\infty} p_n(x) \quad \forall\, x \in \mathbb{R}.$$

Es decir, el valor de e^x en cualquier x se puede calcular como límite de los valores en x de sus aproximaciones de Taylor en $a = 0$.

• **Ejemplo:** Como las derivadas de $\sin x$ y $\cos x$ son sucesivamente $\pm\sin x$ y $\pm\cos x$, para $f(x) = \sin x$ ó $\cos x$ se tiene $|f^{(n+1)}(x)| \leq 1$ para todo n y por tanto, en un intervalo $[-x, x]$ centrado en 0, se tiene

$$|\sin x - s_n(x)| = |S_n(x)| \leq \frac{|x|^{n+1}}{(n+1)!}$$

donde $s_n(x)$ es el polinomio de Taylor de $\sin x$ de orden n en 0 y $S_n(x)$ es el resto, y de manera completamente análoga

$$|\cos x - c_n(x)| = |C_n(x)| \leq \frac{|x|^{n+1}}{(n+1)!}$$

donde $c_n(x)$ es el polinomio de Taylor de $\cos x$ de orden n en 0 y $C_n(x)$ es el resto. En particular, $\sin x = \lim_n s_n(x)$ y $\cos x = \lim_n c_n(x)$ para todo $x \in \mathbb{R}$.

En los ejemplos anteriores, la función es infinitamente derivable y por tanto tiene aproximaciones de Taylor de cualquier orden.

Si una función $f(x)$ es infinitamente derivable en $(a-\delta, a+\delta)$, su **serie de Taylor en** a es la serie infinita

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

Cuando la serie de Taylor **converge** a la función, decimos que la función es **analítica.**

Resulta que las funciones elementales son analíticas.

Series de Taylor en 0 de algunas funciones elementales

- $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$
- $\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$
- $-\log(1-x) = \sum_{n=1}^{\infty} \frac{x^n}{n}$
- $(1+x)^p = \sum_{n=0}^{\infty} \binom{p}{n} x^n$ (∗)
- $\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$

(∗) El coeficiente binomial se define como

$$\binom{p}{n} = \frac{p(p-1)(p-2)\cdots(p-n+1)}{n!}, \quad p \in \mathbb{R}, \quad n \in \mathbb{N}.$$

La función de Gauss

Hay un ejemplo famoso de función infinitamente derivable pero no analítica debido a ***Gauss.*** *Se trata de*

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

Se puede demostrar que $g(x)$ es infinitamente derivable en $\mathbb{R}$ con $g^{(k)}(0) = 0$ para todo $k \geq 0$. Por tanto su serie de Taylor en $a = 0$ es la serie nula, que converge a la función nula. Sin embargo, g no es la función nula a la derecha del origen.

La serie de Taylor nos da el polinomio de Taylor de orden n con sólo truncar la serie en ese grado.

• **Ejemplo:** Por ejemplo, las primeras aproximaciones del seno y coseno cerca de 0 son

$$\sin x = x + O(x^3), \quad \sin x = x - \frac{x^3}{6} + O(x^5),$$
$$\cos x = 1 + O(x^2), \quad \cos x = 1 - \frac{x^2}{2} + O(x^4).$$

En particular, de estas aproximaciones se pueden deducir los límites

$$\lim_{x\to 0} \frac{\sin x}{x} = 1, \qquad \lim_{x\to 0} \frac{1-\cos x}{x^2} = \frac{1}{2}.$$

• **Ejemplo:** Para la exponencial, es útil tener en cuenta que, cerca de 0

$$e^x = 1 + x + O(x^2), \quad \therefore \quad \lim_{x\to 0} \frac{e^x - 1}{x} = 1.$$

• **Ejemplo:** Para el logaritmo, cerca de 0

$$\log(1+x) = x + O(x^2)$$

Al igual que ocurre con los límites y las reglas de derivación, existen fórmulas para obtener los polinomios de Taylor de sumas, restas, productos, cocientes y composiciones de funciones. Estas fórmulas generalizan las reglas de derivación.

Cálculo algebraico de los polinomios de Taylor

Dadas dos funciones, $f(x), g(x)$, con polinomios de Taylor de orden n en un punto a iguales a $p(x)$ y $q(x)$ respectivamente,

- *El polinomio de Taylor de $kf(x)$ con $k \in \mathbb{R}$, es $k\,p(x)$.*
- *El polinomio de Taylor de $f(x)+g(x)$ es $p(x)+q(x)$.*
- *El polinomio de Taylor de $f(x)\cdot g(x)$ consta de los términos hasta grado n en $p(x)\cdot q(x)$.*
- *Si $g(a)\neq 0$, el polinomio de Taylor de $f(x)/g(x)$ consta de los términos hasta grado n en el cociente $p(x)/q(x)$ (que en general da una serie infinita).*

En cuanto a la composición, si $f(x)$ tiene polinomio de Taylor $p(x)$ en a y $g(y)$ tiene polinomio de Taylor $q(y)$ en $b=f(a)$, entonces el polinomio de Taylor de $(g\circ f)(x)$ consta de los términos hasta $(x-a)^n$ (grado n) en la composición $q(p(x))$.

• **Ejemplo:** Los polinomios de Taylor de orden 3 de $\cos x$ en 0 y de $\log y$ en 1 son

$$c_3(x) = 1-\frac{x^2}{2}, \qquad \ell_3(y) = (y-1)-\frac{(y-1)^2}{2}+\frac{(y-1)^3}{3}.$$

Por tanto, el polinomio de Taylor de orden 3 de $f(x)=\log(\cos x)$ en 0 se obtiene de

$$\ell_3(c_3(x)) = (c_3(x)-1)-\frac{(c_3(x)-1)^2}{2}+\frac{(c_3(x)-1)^3}{3} = -\frac{x^2}{2}-\frac{1}{2}\Big(-\frac{x^2}{2}\Big)^2+\frac{1}{3}\Big(-\frac{x^2}{2}\Big)^3$$
$$= -\frac{x^2}{2} - \frac{x^4}{8} - \frac{x^6}{24},$$

descartando los términos de grado 3 o más, dejando $-x^2/2$. Esto se puede comprobar calculando directamente

$$\begin{aligned} &f(x)=\log\cos x, && f'(x)=\frac{-\sin x}{\cos x}=-\tan x, && f''(x)=-1-\tan^2 x,\\ &f(0)=0, && f'(0)=0, && f''(0)=-1, \end{aligned}$$

y volvemos a obtener $-x^2/2$ para el polinomio de Taylor.

• **Ejemplo:** Los polinomios de Taylor de orden 3 de $\sin x$ y $\cos x$ en 0 son

$$s_3(x) = x-\frac{x^3}{6}, \qquad c_3(x) = 1-\frac{x^2}{2}.$$

Por tanto el polinomio de Taylor de orden 3 de $2\sin x\cos x$ se obtiene multiplicando

$$2s_3(x)c_3(x) = 2\Big(x-\frac{x^3}{6}\Big)\Big(1-\frac{x^2}{2}\Big) = \Big(x-\frac{x^3}{6}\Big)(2-x^2) = 2x-x^3-\frac{x^3}{3}+\frac{x^5}{6}$$
$$= 2x-\frac{4}{3}x^3+\frac{1}{6}x^5$$

y quitando los términos de grado 4 o más, dejando $2x-\frac{4}{3}x^3$. Como $2\sin x\cos x=\sin 2x$, esto debe corresponder a $s_3(2x)$, y efectivamente $s_3(2x)=2x-\frac{(2x)^3}{6}=2x-\frac{4}{3}x^3$.

4.12.1. Aplicación a la optimización

El siguiente criterio se puede demostrar a partir de la aproximación de Taylor.

Criterio de la derivada n-ésima

Sea f derivable n veces en $(a-\delta, a+\delta)$, con $n \geq 2$, y a un punto crítico de f con

$$f'(a) = f''(a) = \cdots = f^{(n-1)}(a) = 0, \quad f^{(n)}(a) \neq 0,$$

entonces

- *Si n es* ***par,*** *entonces*
  - *Si $f^{(n)}(a) > 0$, a es un mínimo local estricto.*
  - *Si $f^{(n)}(a) < 0$, a es un máximo local estricto.*
- *Si n es* ***impar,*** *a es un punto de inflexión*
  - *ascendente si $f^{(n)}(a) > 0$,*
  - *descendente si $f^{(n)}(a) < 0$.*

Demostración. Usando el Teorema de Taylor las hipótesis implican que

$$f(x) = f(a) + \frac{f^{(n)}(a)}{n!}(x-a)^n + o((x-a)^{n+1}),$$

es decir,

$$f(x) - f(a) = (x-a)^n \left(\frac{f^{(n)}(a)}{n!} + \varepsilon(x) \right)$$

donde $\varepsilon(x)$ es una función definida en un entorno de a tal que $\lim_{x \to a} \varepsilon(x) = 0$. Como $f^{(n)}(a) \neq 0$, el signo del segundo factor es el mismo que el de esta derivada si $x \approx a$, mientras que el signo de $(x-a)^n$ para $x \approx a^*$ es positivo si n es par. Por tanto

$$\operatorname{sgn}(f(x) - f(a)) = \operatorname{sgn}(f^{(n)}(a)) \text{ si } n \text{ es par y } x \approx a^*.$$

de lo cual se deduce que para $x \approx a^*$, si $f^{(n)}(a) > 0$, entonces $f(x) > f(a)$, o sea, a es un mínimo, y si $f^{(n)}(a) < 0$ entonces $f(x) < f(a)$, o sea, a es un máximo.

Si n es impar entonces $n-1$ es par. Aplicando el razonamiento anterior a la aproximación de Taylor de orden $n-1$ de f', se tiene

$$\operatorname{sgn} f'(x) = \operatorname{sgn}(f'(x) - f'(a)) = \operatorname{sgn}(f^{(n)}(a)) \text{ si } n \text{ es impar y } x \approx a^*,$$

de lo cual se deduce que para $x \approx a^*$, si $f^{(n)}(a) > 0$, entonces el patrón de signos de f' en a es $+, +$, o sea, a es un punto de inflexión ascendente, y si $f^{(n)}(a) < 0$, el patrón de signos es $-, -$, o sea, a es un punto de inflexión descendente. $\square$

• **Ejemplo:** $f(x) = x^4$ en $a = 0$ tiene $f'(0) = f''(0) = f'''(0) = 0$ pero $f^{(4)}(0) = 24 > 0$, luego 0 es un mínimo local estricto.

• **Ejemplo:** $f(x) = x^5$ en $a = 0$ tiene $f'(0) = f''(0) = f^{(3)}(0) = f^{(4)}(0) = 0$ pero $f^{(5)}(0) = 120 \neq 0$, luego 0 es un punto de inflexión ascendente.

4.12.2. Aplicación al cálculo de límites

Con el desarrollo de Taylor se resuelven indeterminaciones que pueden ser demasiado incluso para la regla de L'Hôpital.

• **Ejemplo:** Para calcular

$$\lim_{x\to 0} \frac{(a^x - 1)\log(1 - x^2)}{((1 - x^2)^m - 1)\arcsin x} = \frac{\log a}{m}$$

podemos aproximar, cuando $x \to 0$

$$\begin{aligned} a^x - 1 &= e^{x \log a} - 1 = x \log a + O(x^2) = x\,(\log a + O(x)), \\ \log(1 - x^2) &= -x^2 + O(x^4) = x^2(-1 + O(x^2)) \\ (1 - x^2)^m - 1 &= -mx^2 + O(x^4) = x^2(-m + O(x^2)) \\ \arcsin x &= x + O(x^3) = x(1 + O(x^2)) \end{aligned}$$

y por tanto

$$\frac{(a^x - 1)\log(1 - x^2)}{((1 - x^2)^m - 1)\arcsin x} = \frac{x^3(\log a + O(x))(-1 + O(x))}{x^3(-m + O(x^2))(1 + O(x^2))} = \frac{-\log a + O(x)}{-m + O(x)} = \frac{\log a}{m} + O(x)$$

• **Ejemplo:** Para calcular

$$\lim_{x\to 0} \frac{1 - \cos x}{x^2 - a \sin x^2} = \frac{1}{2(1 - a)} \qquad (a \neq 1)$$

podemos aproximar, cuando $x \to 0$

$$\frac{1 - \cos x}{x^2 - a \sin x^2} = \frac{\frac{x^2}{2} + O(x^3)}{x^2 - a(x^2 + O(x^6))} = \frac{\frac{1}{2} + O(x)}{1 - a + O(x^4)} = \frac{1}{2(1 - a)} + O(x)$$

• **Ejemplo:** Para calcular

$$\lim_{x\to 0} \frac{a^{x^2} - 1}{\sin^2 x} = \log a \qquad (a > 0)$$

podemos aproximar, cuando $x \to 0$

$$\frac{a^{x^2} - 1}{\sin^2 x} = \frac{e^{x^2 \log a} - 1}{\sin^2 x} = \frac{x^2 \log a + O(x^3)}{x^2 + O(x^3)} = \frac{\log a + O(x)}{1 + O(x)} = \log a + O(x)$$

• **Ejemplo:** Para calcular

$$\lim_{x\to 0}\left(\frac{x-2}{x^2+x-2}\right)^{1+\cot^2 x} = \sqrt{e},$$

que es una indeterminación del tipo 1^∞, llamando $f(x)$ a la función, consideramos

$$\begin{aligned}
\log f(x) &= (1+\cot^2 x)\log\left(\frac{x-2}{x^2+x-2}\right) = \csc^2 x \log\left(\frac{x^2+x-2-x^2}{x^2+x-2}\right)\\
&= \frac{\log\left(1-\dfrac{x^2}{x^2+x-2}\right)}{\sin^2 x} = \frac{\log\left(1-\dfrac{x^2}{-2+O(x)}\right)}{\sin^2 x} = \frac{\log\left(1-x^2\left(-\frac{1}{2}+O(x)\right)\right)}{\sin^2 x}\\
&= \frac{\log\left(1+\dfrac{x^2}{2}+O(x^3)\right)}{\sin^2 x} = \frac{\dfrac{x^2}{2}+O(x^3)}{x^2+O(x^4)} = \frac{\dfrac{1}{2}+O(x)}{1+O(x^2)} = \frac{1}{2}+O(x)
\end{aligned}$$

y por tanto $\lim_{x\to 0} f(x) = \lim_{x\to 0}\exp\log f(x) = \exp(\lim_{x\to 0}\log f(x)) = e^{1/2}$.

4.13. Derivación implícita

Consideremos la ecuación $x^2+y^2=1$. Como subconjunto de puntos (x,y) del plano, representa la circunferencia de radio 1 con centro en el origen $(0,0)$. De hecho, podemos despejar y en función de x:

$$x^2+y^2=1 \implies y^2 = 1-x^2 \implies y = \pm\sqrt{1-x^2}.$$

Estas expresiones son válidas para $|x|\le 1$ pues estamos trabajando con números reales y la raíz cuadrada real sólo está para números no negativos. Si dibujamos las gráficas de las dos funciones resultantes, vemos que corresponden a las partes de arriba y de abajo de la circunferencia.

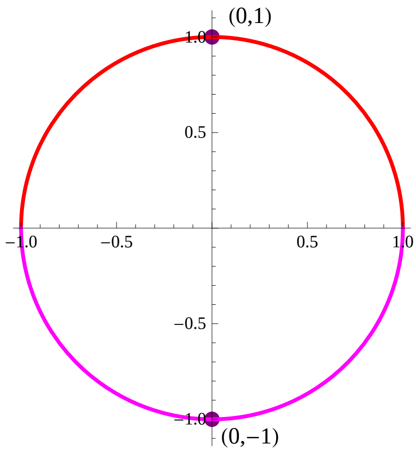

Figura 17 – La circunferencia x^2+y^2 dividida en dos semicírculos.

Si ponemos $x = 0$ obtenemos $y = \pm 1$. La elección de signo determina cuál de las dos partes consideramos. Con signo positivo es la de arriba, que pasa por el punto $(0, 1)$ y con signo negativo es la de abajo, que pasa por $(0, -1)$.

Como hemos podido despejar y en función de x, pudiendo entonces denotar $y = y(x)$, podemos también calcular fácilmente las derivadas

$$y(x) = \sqrt{1-x^2} \implies y'(x) = -\frac{x}{\sqrt{1-x^2}}$$
$$y(x) = -\sqrt{1-x^2} \implies y'(x) = \frac{x}{\sqrt{1-x^2}}.$$

Volvamos atrás a la ecuación original

$$x^2 + y^2 = 1.$$

Sabiendo ya que y es función de x, podríamos derivar esta ecuación respecto á x, considerándola simplemente como una igualdad entre funciones de x. Podemos poner provisionalmente $y = f(x)$. Entonces la ecuación es

$$x^2 + f(x)^2 = 1$$

y derivando se obtiene

$$2x + 2f(x)f'(x) = 0, \quad \text{abreviado por} \quad 2x + 2yy' = 0.$$

Suponiendo que $y \neq 0$, podemos agrupar términos para obtener

$$\boxed{y' = -\frac{x}{y}.}$$

Observar que esta fórmula expresa la derivada en términos de la variable x y de la propia función y. Si lo pensamos, vemos que no necesitábamos despejar y en función de x para llegar a esta fórmula. Sólo necesitábamos saber que se podía despejar! Por supuesto, una vez despejada, es inmediato verificar que efectivamente

$$y = \sqrt{1-x^2} \implies y' = -\frac{x}{y} = -\frac{x}{\sqrt{1-x^2}}$$

y que la misma fórmula es válida para la otra función que resulta al depsejar

$$y = -\sqrt{1-x^2} \implies y' = -\frac{x}{y} = \frac{x}{\sqrt{1-x^2}}.$$

En este sentido, la fórmula $y' = -x/y$ es mejor porque vale cualquiera que sea la función y que ha resultado al despejar.

Podemos intentar hacer lo mismo con otras ecuaciones de este tipo, por ejemplo con las secciones cónicas tales como elipses o hipérbolas,

$$\frac{(x-1)^2}{4} + \frac{(y+2)^2}{9} = 1, \quad x^2 - y^2 = 1,$$

etc. o con otras más complicadas, como

$$x^2 + y^2 + \sin 4x + \sin 4y = 2$$

en las cuales ya no va a ser tan fácil o siquiera posible despejar y en términos de funciones sencillas (ver la figura 18).

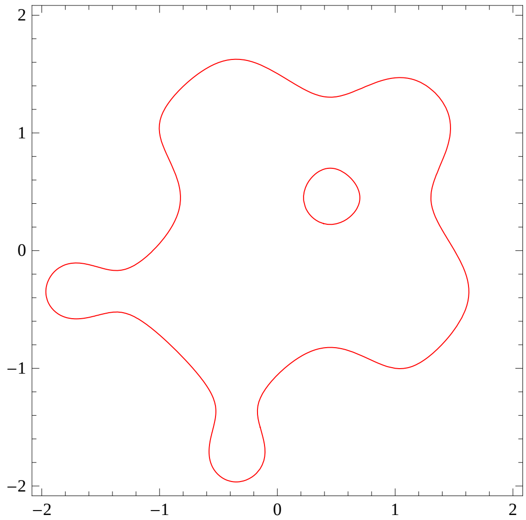

Figura 18 – La curva implícita $x^2+y^2+\sin 4x+\sin 4y = 2$.

Las consideraciones anteriores llevan a las siguientes definiciones.

Decimos que una función y está **dada implícitamente como función de** x por una ecuación si existe alguna función f de dos variables y una constante k tal que

$$f(x, y) = k.$$

Decimos que esta es la **ecuación implícita** que determina y.

No es posible entrar aquí en la teoría de estas ecuaciones, por ejemplo para justificar que efectivamente, dadas ciertas circunstancias, una ecuación implícita permite de hecho despejar una variable en función de la otra. Simplemente queremos resaltar que en estos casos, es posible hallar las derivadas de y en términos de otras funciones de x, y, derivando la ecuación implícita. Este método se conoce como derivación implícita.

El resultado de derivar una ecuación

$$f(x, y) = k$$

que define á y como función de x, resulta en una ecuación del tipo

$$y' = g(x, y).$$

La justificación teórica de este resultado conlleva hablar de derivadas de funciones de más de una variable, que es un tema que no vamos a tocar aquí. No es difícil entender por qué resulta una tal ecuación: si f se compone de distintos términos, al ir derivándolos aplicando la regla de la cadena, saldrá y' fuera de algunas de las expresiones, luego se reúnen todos los términos con y' y quedará una ecuación como la descrita.

• **Ejemplo:** Supongamos que $2x^2y - 4y^3 = 4$ define á y como función de x. Se puede poner provisionalmente $y = f(x)$ y la ecuación es

$$2x^2 f(x) - 4f(x)^3 = 4.$$

Derivando implícitamente respecto á x, se obtiene

$$4xf(x) + 2x^2 f'(x) - 12f(x)^2 f'(x) = 0,$$

aunque lo habitual es escribir todo de forma más sencilla en términos de y', y, x:

$$2x^2y - 4y^3 = 4 \quad \therefore \quad 4xy + 2x^2y' - 12y^2y' = 0, \quad \therefore \quad y' = \frac{2xy}{6y^2 - x^2}.$$

Si conocemos una solución concreta $x = x_0, y = y_0$ de la ecuación implícita:

$$f(x_0, y_0) = k,$$

podemos obtener el valor concreto de la derivada sustituyendo en la ecuación

$$y' = g(x, y)$$

que resulta al derivar implícitamente.

Por ejemplo, $x = 2, y = 1$ satisface la ecuación $2x^2y - 4y^3 = 4$. Sustituyendo $x = 2, y = 1$ en la fórmula

$$y' = \frac{2xy}{6y^2 - x^2}$$

se obtiene $y' = 2$.

Derivando implícitamente más veces obtenemos las derivadas implícitas de orden superior. De este modo, conociendo soluciones concretas $x = x_0, y = y_0$ de la ecuación original, es posible analizar puntos críticos, etc. de la función implícita aún no pudiendo despejar explícitamente.

Por ejemplo, derivando la ecuación obtenida para y' y simplificando,

$$y'' = \frac{6y(12y^4 - 4x^2y^2 - x^4)}{(6y^2 - x^2)^3}.$$

Sustituyendo $x = 2, y = 1, y' = 2$, se obtiene $y'' = -15$.

Con la misma ecuación, vemos que $x = 0, y = -1$ es otra solución, y en este caso $y' = 0$, luego la función implícita $y(x)$ tiene un punto crítico en $x = 0$. Además, $y'' = -72/6^3 < 0$, luego es un máximo local. Todo esto lo hemos obtenido sin tener ni idea de qué curva representa la ecuación. Es la mostrada en la figura 19.

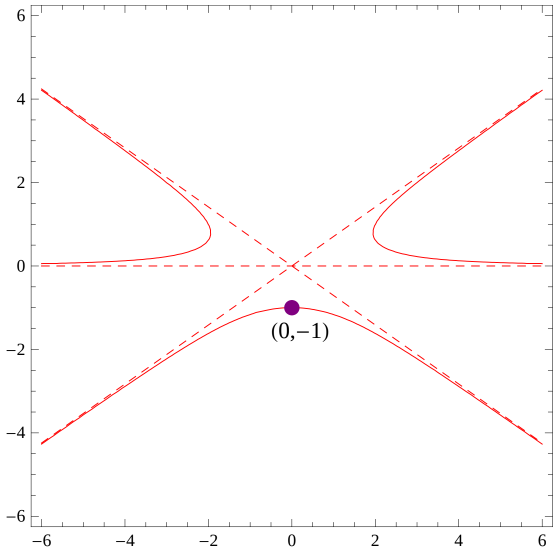

Figura 19 – La curva $2x^2y - 4y^3 = 4$.

Como puede verse, tiene tres partes y además hay rectas asintóticas, que de hecho se obtienen igualando a 0 en vez de á 4,

$$2x^2y - 4y^3 = 0 \iff y(x^2 - 2y^2) = 0 \iff y = 0, x = \pm\sqrt{2}y,$$

pero las razones de esto superan el contenido de este tema.

5. Cálculo Integral

5.1. Introducción

La integral surgió, al igual que la derivada, de la Geometría clásica, esta vez, del problema de calcular el área limitada por una curva. Arquímedes descubrió el **método exhaustivo** para calcular un área curva a partir de **aproximaciones poligonales.** Esto sigue siendo el punto de partida de la definición actual de la integral.

Sea f una función definida en un intervalo compacto $I = [a, b]$. Consideremos su gráfica, que es el conjunto de puntos

$$\{(x, f(x)) : a \leq x < b\},$$

siendo en general una curva en el plano.

La **figura limitada por la gráfica de** f es el subconjunto del plano

$$\{(x, y) : x \in [a, b],\ \ y \in [0, f(x)]\ \},$$

limitado por la gráfica de f junto con el eje horizontal $y = 0$ y las rectas verticales $x = a$, $x = b$, donde se considera por definición que $[0, c] \overset{\text{def}}{=} [c, 0]$ si $c < 0$ y $[0, c]$ es el único punto $\{0\}$ si $c = 0$.

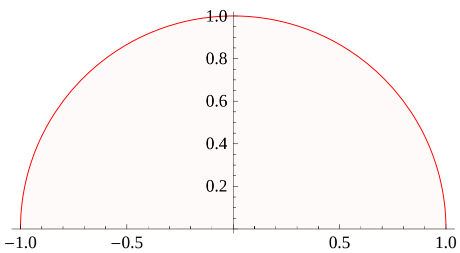

Figura 20 – El semidisco limitado por $y = \sqrt{1 - x^2}$.

Si $f(x) \geq 0$, la figura está debajo de la gráfica, es decir, entre el eje horizontal y la gráfica (figura 20) pero en general, puede haber partes tanto por encima como por debajo del eje horizontal (figura 21).

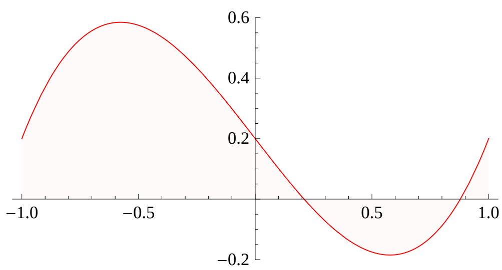

Figura 21 – Figura plana limitada por la gráfica de una función.

Una figura poligonal sencilla y particularmente adecuada para aproximar figuras curvas es el rectángulo. La idea de Arquímedes, trasladada a esta situación, consiste en trazar rectángulos con bases en el eje horizontal y alturas que aproximan a la función f por debajo y por arriba, es decir, por dentro y por fuera de la figura limitada por su gráfica.

Por ejemplo, para calcular el área limitada por la parábola $f(x) = x^2$ entre 0 y1, podemos empezar trazando 6 rectángulos como en la figura 22.

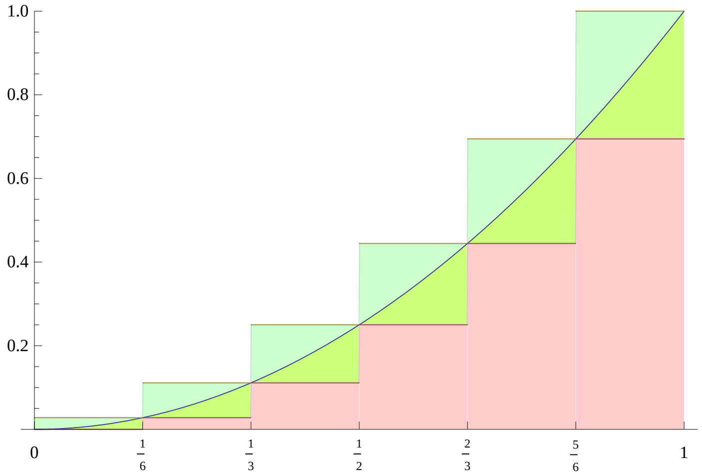

Figura 22 – Aproximación al área de la parábola $y = x^2$ entre $x = 0, x = 1$.

La suma de las áreas de los rectángulos sombreados en rojo aproxima el área con un defecto y la suma de las áreas de los rectángulos que incluyen las partes en verde la aproxima por exceso. La idea es que, si aumentamos el número de rectángulos, la diferencia entre el exceso y el defecto debería tender a 0 y así ambas aproximaciones se acercarán a un valor límite que será el área exacta limitada por la parábola (figura 23).

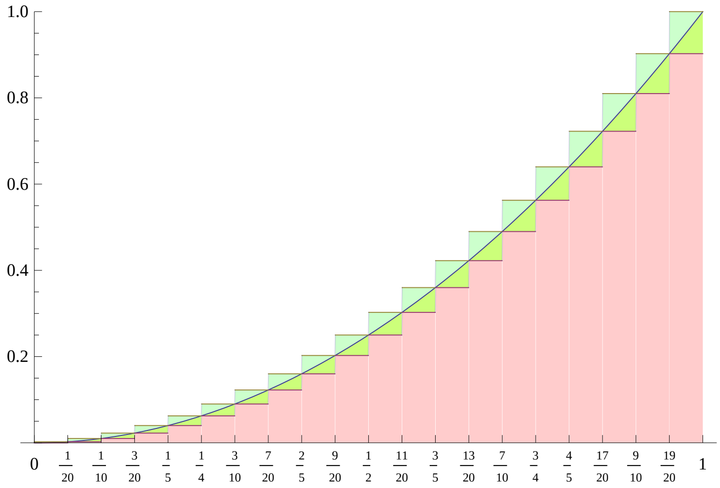

Figura 23 – Aproximación mejorada al área de la parábola $y = x^2$ entre $x = 0, x = 1$.

5.2. Definiciones básicas

Definición de partición de un intervalo

Sea $I = [a, b]$ un intervalo compacto. Una ***partición*** *de I es una* ***colección ordenada*** *finita de puntos de I, que incluye a los extremos:*

$$\mathfrak{p} : a = x_0 < x_1 < x_2 < \cdots < x_n = b.$$

Una partición con $n + 1$ puntos divide al intervalo I en n ***subintervalos***

$$I_k = [x_{k-1}, x_k], \quad 1 \leq k \leq n,$$

llamados ***(sub)intervalos de la partición.***

Los intervalos de la partición **no se solapan** y juntos **cubren el intervalo completo.** Recíprocamente, dada una colección finita de subintervalos de $I = [a, b]$ que no se solapan y lo cubren, sus extremos definen una partición. Por tanto la partición se puede considerar equivalentemente de dos maneras:

- como la propia **colección de puntos** $\{x_k\}_{k=0}^n$, ó
- como **la colección de subintervalos** I_k que delimitan los puntos.

Cuando adoptamos el primer punto de vista, escribimos $x_k \in \mathfrak{p}$ para referirnos a un punto de la partición. Cuando adoptamos el segundo punto de vista, escribimos también

$$J \in \mathfrak{p}$$

para referirnos a un subintervalo $J = I_k$.

Siguiendo las ideas expresadas en la sección anterior, los subintervalos determinados por una partición serán las bases de rectángulos que aproximarán el área limitada por la gráfica de una función. La aproximación por defecto es la obtenida moviendo el subintervalo perpendicular al eje hasta que toca la gráfica, y la aproximación por exceso es la obtenida al seguir moviendo el subintervalo de esta manera hasta que deja de tocar la gráfica. En términos precisos, de lo que se trata es de estudiar el **ínfimo** y el **supremo** de la función en cada subintervalo:

$$\begin{aligned} m_k &= \operatorname{ínf}\{f(x) : x \in [x_{k-1}, x_k]\} = \operatorname{ínf} f(I)_k, \\ M_k &= \sup\{f(x) : x \in [x_{k-1}, x_k]\} = \sup f(I)_k, \end{aligned}$$

que determinan las **alturas** de los rectángulos que aproximan.

Usamos “altura” en un sentido más amplio que el geométrico, pues nuestras alturas son **negativas** si los valores correspondientes de la función lo son. En esta situación se habla de **altura dirigida** y de **área dirigida.**

Las áreas que aproximan corresponden a la suma de las áreas de los rectángulos determinados por estos datos. Como el área de un rectángulo es igual a la longitud de su base por su altura, se obtienen la suma por defecto

$$m_1(x_1 - x_0) + m_2(x_2 - x_1) + \cdots + m_n(x_n - x_{n-1}) = \sum_{k=1}^{n} m_k(x_k - x_{k-1})$$

y la suma por exceso

$$M_1(x_1 - x_0) + M_2(x_2 - x_1) + \cdots + M_n(x_n - x_{n-1}) = \sum_{k=1}^{n} M_k(x_k - x_{k-1}).$$

Cuando se emplea este tipo de suma también es habitual usar la notación

$$\Delta x_k \stackrel{\text{def}}{=} x_k - x_{k-1}$$

para las longitudes.

Sumas de Riemann inferiores y superiores

Sea f una función **acotada** *sobre el intervalo compacto $I = [a, b]$ y*

$$\mathfrak{p} : a = x_0 < x_1 < x_2 < \cdots < x_n = b$$

una partición de I.

- *La* ***suma de Riemann superior de f respecto á*** $\mathfrak{p}$ *es*

$$\overline{S}(f, \mathfrak{p}) \stackrel{\text{def}}{=} \sum_{k=1}^{n} M_k(x_k - x_{k-1}), \quad M_k = \sup\{f(x) : x \in I_k \stackrel{\text{def}}{=} [x_{k-1}, x_k]\}$$

- *La* ***suma de Riemann inferior de f respecto á*** $\mathfrak{p}$ *es*

$$\underline{S}(f, \mathfrak{p}) \stackrel{\text{def}}{=} \sum_{k=1}^{n} m_k(x_k - x_{k-1}), \quad m_k = \inf\{f(x) : x \in I_k \stackrel{\text{def}}{=} [x_{k-1}, x_k]\}.$$

Las sumas de Riemann inferior y superior son una herramienta más teórica que práctica, pues calcularlas requiere conocer los ínfimos y supremos de la función en cada subintervalo de la partición, es decir, resolver $2n$ problemas de optimización para una partición con $n + 1$ puntos.

Otra opción para determinar alturas de rectángulos aproximantes es simplemente elegir un punto cualquiera de cada subintervalo y usar el valor de f en ese punto como altura:

$$c_k \in I_k = [x_{k-1}, x_k] \to f(c_k)$$

Muestra de una partición

Sea $I = [a, b]$ *un intervalo compacto y*

$$\mathfrak{p} : a = x_0 < x_1 < x_2 < \cdots < x_n = b$$

una partición de I. *Llamamos* ***muestra*** *de* $\mathfrak{p}$ *o simplemente* $\mathfrak{p}$***-muestra*** *a una colección de puntos* $\mathcal{M} = \{c_k : 1 \le k \le n\}$ *tal que* $c_k \in I_k = [x_{k-1}, x_k]$ *para cada* k, $1 \le k \le n$.

Suma de Riemann respecto a una muestra

Sea $I = [a, b]$ *un intervalo compacto y*

$$\mathfrak{p} : a = x_0 < x_1 < x_2 < \cdots < x_n = b$$

una partición de I *y* $\mathcal{M} = \{c_k\}_{k=1}^{n}$ *una* $\mathfrak{p}$***-muestra***. *Si* f *es una función acotada sobre* I, *la* ***suma de Riemann de*** f ***respecto á*** $\mathfrak{p}$ ***y*** $\mathcal{M}$ *es*

$$S(f, \mathfrak{p}, \mathcal{M}) \overset{\text{def}}{=} \sum_{k=1}^{n} f(c_k)(x_k - x_{k-1}) = \sum_{k=1}^{n} f(c_k)\Delta x_k.$$

Con estas definiciones tenemos las nociones rigurosas de sumas rectangulares que aproximan al área. Ahora necesitamos dar una definición de límite de las sumas que exprese la existencia del "área acotada por la gráfica de f."

Funciones integrables

Sea f *una función acotada sobre el intervalo compacto* $I = [a, b]$. *Decimos que* f *es* ***integrable*** *en* I *si, dado* $\varepsilon > 0$, *existe alguna partición* $\mathfrak{p}$ *de* I *tal que*

$$\Delta(f, \mathfrak{p}) \overset{\text{def}}{=} \overline{S}(f, \mathfrak{p}) - \underline{S}(f, \mathfrak{p}) < \varepsilon$$

Geométricamente, esto expresa que la diferencia entre el área por exceso y el área por defecto se puede hacer arbitrariamente pequeña. La definición que hemos dado de integrabilidad es la habitual, pero sigue sin ser demasiado práctica. Para empezar, ¡en la definición ni siquiera se habla del área exacta!

Para convertir la definición de función integrable en algo más útil hace falta estudiar más a fondo el comportamiento de las sumas de Riemann. Al final se llega a los siguientes resultados.

Integral de una función integrable

Sea f una función integrable sobre el intervalo compacto $I = [a, b]$. Entonces existe un único número A tal que

$$\underline{S}(f, \mathfrak{p}) \leq A \leq \overline{S}(f, \mathfrak{p})$$

para cualquier partición $\mathfrak{p}$ de I. Decimos que A es la ***integral definida*** *de f sobre $I = [a, b]$, y escribimos*

$$A = \int_a^b f \quad \text{o también} \quad \int_a^b f(x)\, dx$$

La notación de Leibniz $\int_a^b f(x)\, dx$ está ideada para sugerir el proceso de suma: el símbolo de integral es una "S" estilizada y dx representa el cambio en x correspondiente a la longitud de un intervalo de una partición y denotado por $\Delta x_k = x_k - x_{k-1}$.

Al menos ahora ha aparecido ya el número que representa la noción de área limitada por la gráfica de f. Sin embargo, no tenemos un procedimiento claro para calcular este número.

Anchura de una partición

Sea $\mathfrak{p} : a = x_0 < x_1 < x_2 < \cdots < x_n = b$ una partición del intervalo $[a, b]$. La ***anchura*** *de $\mathfrak{p}$, denotada por $\|\mathfrak{p}\|$, es el máximo de las longitudes de los subintervalos de $\mathfrak{p}$:*

$$\|\mathfrak{p}\| = \text{máx}\{x_k - x_{k-1} : 1 \leq k \leq n\}.$$

La integral como límite de sumas

Sea f una función integrable sobre el intervalo compacto $I = [a, b]$. Dado $\varepsilon > 0$, existe $\delta > 0$ tal que si $\mathfrak{p}$ es una partición de $[a, b]$, entonces

$$\|\mathfrak{p}\| < \delta \implies \left| S(f, \mathfrak{p}, \mathcal{M}) - \int_a^b f \right| < \varepsilon$$

para cualquier $\mathfrak{p}$-muestra $\mathcal{M}$.

Es decir, la integral se puede calcular tomando una partición lo suficientemente **fina** (con anchura suficientemente pequeña) y calculando cualquier suma de Riemann con esa partición, pudiendo elegir una muestra arbitrariamente.

Por tanto, si f es integrable en $[a, b]$, su integral se puede calcular como el límite de sumas

$$\int_a^b f = \lim_{n \to \infty} S(f, \mathfrak{p}_n, \mathcal{M}_n)$$

donde $\mathfrak{p}_n$ denota una sucesión de particiones cuya anchura tiende a 0 y $\mathcal{M}_n$ es cualquier muestra tomada de $\mathfrak{p}_n$.

En cuanto a elección de muestra, es habitual tomar los extremos del intervalo, tanto los extremos de la izquierda

$$\mathcal{I} = \{x_{k-1} : 1 \leq k \leq n\}$$

como los de la derecha

$$\mathcal{D} = \{x_k : 1 \leq k \leq n\}.$$

Este último resultado al menos proporciona un método práctico de aproximar la integral numéricamente, pues es fácil sistemáticamente construir particiones con una anchura pequeña. Para ello se suele recurrir a la subdivisión en partes iguales del intervalo.

Particiones aritméticas

Sea $I = [a, b]$ un intervalo compacto y $n \in \mathbb{N}$. La ***partición aritmética*** *(o también partición* ***regular****) de orden n de $[a, b]$ es aquella cuyos puntos vienen dados por la progresión aritmética de diferencia*

$$\frac{b-a}{n}$$

es decir, la n-ésima parte de la longitud total de $[a, b]$. Se denota por $\mathfrak{a}_n$ y sus puntos vienen dados por la fórmula para el término general de una progresión aritmética:

$$x_k = a + \frac{b-a}{n} k, \quad 1 \leq k \leq n.$$

En particular, $x_0 = a$, $x_n = b$, y $\Delta x_k = x_k - x_{k-1} = \dfrac{b-a}{n}$ para todo $k : 1 \leq k \leq n$.

La partición aritmética $\mathfrak{a}_n$ de $[a, b]$ de orden n tiene anchura

$$\|\mathfrak{a}_n\| = \frac{b-a}{n},$$

por tanto $\lim_{n\to\infty} \|\mathfrak{a}_n\| = 0$.

Por tanto hemos llegado al primer resultado práctico para intentar calcular integrales:

La integral como límite de sumas regulares

Sea f una función integrable en el intervalo compacto $[a, b]$. Entonces

$$\int_a^b f = \lim_{n\to\infty} \sum_{k=1}^{n} f\left(a + \frac{b-a}{n} k\right) \cdot \frac{b-a}{n}$$

En general,

$$\int_a^b f = \lim_{n\to\infty} S(f, \mathfrak{a}_n, \mathcal{M}_n) = \lim_{n\to\infty} \sum_{k=1}^{n} f(c_k) \cdot \frac{b-a}{n}$$

donde $\mathcal{M}_n = \{c_k\}_{k=1}^{n}$ es cualquier muestra de la partición aritmética $\mathfrak{a}_n$.

• **Ejemplo:** Para una función constante $f(x) = k$,

$$\boxed{\int_a^b k = k(b-a)}$$

pues cualquier suma de Riemann tiene este valor:

$$\sum_{j=1}^{n} f(c_j)(x_j - x_{j-1}) = \sum_{j=1}^{n} k(x_j - x_{j-1}) = k\sum_{j=1}^{n}(x_j - x_{j-1}) = k(b-a).$$

En particular, $\int_a^b 0 = 0$. Como la figura limitada por una función constante es un rectángulo de altura k y base $b-a$, este primer resultado confirma que, al menos en este caso, la integral es igual al área limitada por la gráfica.

Hemos usado que la suma de las longitudes de los subintervalos de una partición es igual a la longitud total del intervalo. Este hecho es "obvio" geométricamente porque los subintervalos no se solapan y cubren al intervalo total, pero algebraicamente depende de que es una **suma telescópica:**

$$\begin{aligned}\sum_{j=1}^{n}(x_j - x_{j-1}) &= (x_n - x_{n-1}) + (x_{n-1} - x_{n-2}) + \cdots + (x_2 - x_1) + (x_1 - x_0)\\ &= (x_n - \cancel{x_{n-1}}) + (\cancel{x_{n-1}} - \cancel{x_{n-2}}) + \cdots + (\cancel{x_2} - \cancel{x_1}) + (\cancel{x_1} - x_0)\\ &= x_n - x_0\\ &= b - a.\end{aligned}$$

• **Ejemplo:** Sea $f(x) = x$ en el intervalo $[0, b]$. La figura limitada por f es un triángulo de base y altura iguales a b, por tanto la integral debería ser igual al área de este triángulo, $b^2/2$. Poniendo $a = 0$, estamos considerando las sumas

$$\sum_{k=1}^{n}\left(\frac{bk}{n}\right)\cdot\frac{b}{n} = \frac{b^2}{n^2}\sum_{k=1}^{n} k.$$

Necesitamos el valor de la suma

$$\sum_{k=1}^{n} k = 1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2} \sim \frac{n^2}{2},$$

(re)descubierto por Gauss a la edad de cinco años. Entonces

$$\int_0^b x\,dx = \lim_{n\to\infty}\frac{b^2}{n^2}\sum_{k=1}^{n} k = \lim_{n\to\infty}\frac{b^2}{n^2}\cdot\frac{n(n+1)}{2} = \frac{b^2}{2}.$$

• **Ejemplo:** Vamos a intentar calcular, como hizo Arquímedes, el área limitada por la parábola $y = x^2$ entre $x = 0$ y un punto $x = b$ con $b > 0$ cualquiera, utilizando esta fórmula. Poniendo $a = 0$, tenemos que considerar las sumas

$$\sum_{k=1}^{n}\left(\frac{bk}{n}\right)^2\cdot\frac{b}{n} = \frac{b^3}{n^3}\sum_{k=1}^{n} k^2.$$

Necesitamos el valor de la suma

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

o al menos un método de estimar su orden de magnitud asintótico. Habiendo hecho esto, queda finalmente

$$\int_0^b x^2\,dx = \lim_{n\to\infty} \frac{b^3}{n^3} \sum_{k=1}^{n} k^2. = \lim_{n\to\infty} \frac{b^3}{n^3} \cdot \frac{n(n+1)(2n+1)}{6} = \lim_{n\to\infty} \frac{b^3}{n^3} \cdot \frac{n^3}{3} = \frac{b^3}{3}.$$

Calcular las sumas

$$1^p + 2^p + 3^p + \cdots + n^p$$

es un problema bastante conocido en la historia de las Matemáticas. Véase por ejemplo http://es.wikipedia.org/wiki/Formula_de_Faulhaber

El problema de calcular el área limitada por $y = x^p$ para cualquier potencia natural p usando este tipo de fórmulas lleva, tras algo de trabajo, a la conclusión

$$\boxed{\int_0^b x^p\,dx = \frac{b^{p+1}}{p+1}}$$

y de hecho esta fórmula es válida para cualquier potencia real $p \geq 0$.

Existen otros tipos de particiones relativamente sencillas aparte de la aritmética, por ejemplo, una partición que tiene sus puntos en progresión geométrica. Con más esfuerzo se puede conseguir integrar otras funciones, como la exponencial, el seno o el coseno, que dan respectivamente

$$\int_a^b e^x\,dx = e^b - e^a$$

y

$$\int_a^b \cos x\,dx = \sin b - \sin a, \qquad \int_a^b \sin x\,dx = \cos a - \cos b.$$

5.3. Integrabilidad

Antes de proseguir, debemos plantearnos una cuestión básica: ¿Cuáles son las funciones integrables? Al menos, ¿Qué tipo de funciones son integrables?

Las funciones monótonas son integrables

Sea f acotada y monótona (creciente o decreciente) en el intervalo compacto $[a, b]$. Entonces f es integrable en $[a, b]$.

Demostración. Supongamos que f es creciente. La demostración es parecida si es decreciente. Consideremos la partición aritmética $\mathfrak{a}_n$ del intervalo $[a, b]$:

$$\mathfrak{a}_n : x_k = a + \frac{b-a}{n} k$$

Como f es creciente, sus valores mínimo y máximo sobre el subintervalo $I_k = [x_{k-1}, x_k]$ son sus valores en los extremos del subintervalo; respectivamente

$$m_k = \min f(I_k) = f(x_{k-1}), \qquad M_k = \max f(I_k) = f(x_k).$$

Por tanto las sumas inferior y superior de Riemann de f respecto á $\mathfrak{a}_n$ son

$$\underline{S}(f, \mathfrak{a}_n) = \sum_{k=1}^{n} f(x_{k-1}) \Delta x_k, \quad \overline{S}(f, \mathfrak{a}_n) = \sum_{k=1}^{n} f(x_k) \Delta x_k$$

y su diferencia es

$$\begin{aligned} \Delta(f, \mathfrak{a}_n) = \overline{S}(f, \mathfrak{a}_n) - \underline{S}(f, \mathfrak{a}_n) &= \sum_{k=1}^{n} (f(x_k) - f(x_{k-1}))\, \Delta x_k \\ &= \sum_{k=1}^{n} (f(x_k) - f(x_{k-1}))\, \frac{b-a}{n} \\ &= \frac{b-a}{n} \sum_{k=1}^{n} (f(x_k) - f(x_{k-1})) = \frac{(b-a)(f(b) - f(a))}{n} \end{aligned}$$

pues la última suma es telescópica. Claramente $\lim_{n\to\infty} \Delta(f, \mathfrak{a}_n) = 0$, por tanto f es integrable. □

Las funciones continuas son integrables

Sea f continua en el intervalo compacto $[a, b]$. Entonces f es integrable en $[a, b]$.

Demostración. Se omite. □

El siguiente resultado dice que las funciones integrables se pueden “estropear” en un número finito de puntos sin afectar ni a su integrabilidad ni al valor de su integral.

Modificación de una función integrable en un número finito de puntos

Sea f integrable en $[a, b]$ y $c_1, c_2, \dots, c_n \in [a, b]$. La función $\tilde{f}$ obtenida modificando arbitrariamente el valor de f es los puntos c_k es integrable en $[a, b]$ y su integral es igual a la de f:

$$\int_a^b f = \int_a^b \tilde{f}.$$

Demostración. Se omite. □

• **Ejemplo:** Sea $f(x)$ una función que vale 0 salvo en un número finito de puntos. Entonces f es integrable en cualquier intervalo y su integral vale 0.

Se dice que una función f en el intervalo $[a,b]$ es **continua a trozos** si

- Es una función acotada.
- Tiene un número finito de discontinuidades en $[a,b]$. Es decir, existe una partición

$$\mathfrak{p} : a_0 = a < a_1 < a_2 < \cdots < a_n = b$$

de $[a,b]$ tal que f es continua en cada subintervalo abierto (a_{k-1}, a_k).

Las funciones continuas a trozos son integrables

Demostración. Se omite. □

Dado el resultado anterior, no importa como esté definida $f(x)$ en el número finito de puntos donde no es continua. De hecho, podemos cambiar su definición si queremos, sin afectar al valor de la integral.

• **Ejemplo:** La función $\sin(1/x)$ es continua a trozos en cualquier intervalo pues está acotada y solo tiene una discontinuidad en $x=0$. Por tanto es integrable. Esto puede parecer paradójico ya que su gráfica cruza infinitas veces el eje horizontal al acercarse á 0 y es difícil ver cuál puede ser el área que limita. Sin embargo se tiene, por ejemplo

$$\int_0^\pi \sin\frac{1}{x}\,dx \approx 1{,}57594$$

que curiosamente no está lejos de $\pi/2 \approx 1{,}5708$.

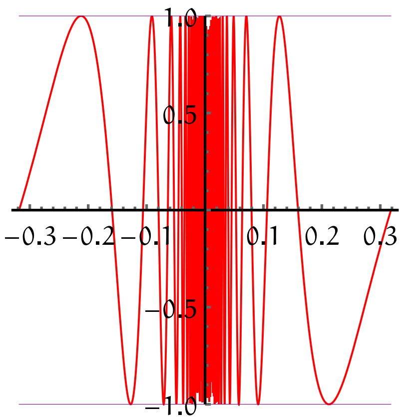

Figura 24 – Gráfica de $\sin\dfrac{1}{x}$ en $[0,\pi]$.

• **Ejemplo:** Un tipo de función continua a trozos más fácil de visualizar en cuanto á area viene dado por la **función suelo,** definida por

$$\lfloor x \rfloor = n, \text{ donde } n \text{ es el único entero con } n \le x < n+1$$

Por ejemplo, $\lfloor \pi \rfloor = 3$, $\lfloor e \rfloor = 2$, es decir, se **redondea hacia abajo** al entero más próximo.

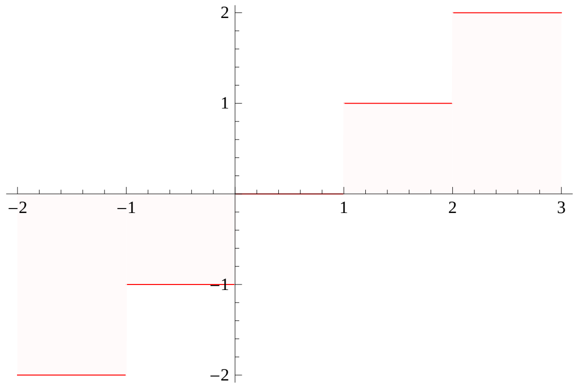

Figura 25 – Gráfica de la función suelo $\lfloor x \rfloor$.

Se dice que una función f en el intervalo $[a, b]$ es **monótona a trozos** si

- Es una función acotada.
- Existe alguna partición

$$\mathfrak{p} : a_0 = a < a_1 < a_2 < \cdots < a_n = b$$

de $[a, b]$ tal que f es monótona en cada subintervalo abierto (a_{k-1}, a_k).

Las funciones monótonas a trozos son integrables

A todo esto, uno puede plantear la pregunta: ¿Cuál sería un ejemplo de función que **no** es integrable? La respuesta la da un ejemplo famoso.

La **función de Dirichlet** es la función en $\mathbb{R}$ definida por

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

es decir, vale 1 si x es un número **racional** y 0 si es **irracional.**

La función de Dirichlet no es integrable en ningún intervalo.

Demostración. Como hay infinitos números racionales y también infinitos números irracionales en cualquier intervalo, dada cualquier partición $\mathfrak{p}$ de un intervalo $[a, b]$, se tendrá

$$m_k = \text{ínf } f(I_k) = 0, \quad M_k = \sup f(I_k) = 1.$$

Por tanto las sumas de Riemann inferior y superior serán respectivamente

$$\underline{S}(f, \mathfrak{p}) = 0, \quad \overline{S}(f, \mathfrak{p}) = b - a$$

y por tanto cualquier diferencia de sumas $\Delta(f, \mathfrak{p}) = b - a$. Evidentemente esta cantidad no se puede hacer arbitrariamente pequeña ya que es una constante positiva (la longitud del intervalo). Por tanto f no es integrable. $\square$

Intuitivamente, la función de Dirichlet está oscilando instantáneamente entre 0 y 1 en cualquier punto. En particular, tampoco es continua en ningún punto. No se le puede asignar un área en el sentido clásico de la palabra. Sin embargo, en la teoría de integración desarrollada a comienzos del siglo XX y que supera a lo que estamos exponiendo aquí, sí se admite la función de Dirichlet como función integrable.

5.4. Propiedades básicas de la integral

Calcular integrales mediante sumas de Riemann lleva enseguida a la realización que calcular sumas explícitamente es bastante complicado, y existen relativamente pocas fórmulas. Como ocurre al estudiar límites, funciones continuas, y derivadas, es deseable tener algún procedimiento mediante el cual se pueden calcular integrales de combinaciones de funciones conociendo las integrales de las funciones originales. Sin embargo, para integrales no hay reglas para productos, cocientes o composiciones tan directas como para derivadas.

Propiedades básicas de la integral

Sean f, g funciones integrables en $[a, b]$. Entonces se tiene:

- ***Linealidad:*** *$f + g$ es integrable en $[a, b]$, con*

$$\int_a^b (f+g) = \int_a^b f + \int_a^b g$$

- *kf es integrable en $[a, b]$ para cualquier constante k, con*

$$\int_a^b kf = k \int_a^b f$$

- *$f \cdot g$ es integrable sobre $[a, b]$.*
- *f/g es integrable sobre $[a, b]$ si existe $m > 0$ tal que $|g(x)| \geq m$ para todo $x \in [a, b]$.*
- ***Monotonía:*** *Si $f \leq g$ entonces $\int_a^b f \leq \int_a^b g$.*
- ***Desigualdad triangular:*** *$|f|$ es integrable en $[a, b]$, con*

$$\left| \int_a^b f \right| \leq \int_a^b |f|$$

En particular, las funciones integrables sobre un intervalo forman un espacio vectorial real y la integral es un funcional lineal sobre este espacio.

No es cierto que si $|f|$ es integrable entonces f también es integrable.

En particular destacamos las siguientes consecuencias:

Cotas para una integral

Sea f una función integrable en $[a,b]$. Si m, M son constantes tales que $m \leq f(x) \leq M$ para todo $x \in [a,b]$ entonces

$$m(b-a) \leq \int_a^b f \leq M(b-a)$$

Si $|f(x)| \leq M$, entonces

$$\left|\int_a^b f\right| \leq M(b-a)$$

Hay otra serie de propiedades relacionadas con cambiar el intervalo de integración manteniendo la misma función.

Aditividad de la integral

Sea f una función acotada en $[a,b]$ y $c \in (a,b)$. Entonces f es integrable en $[a,b]$ si y sólo si lo es en cada intervalo $[a,c]$ y $[c,b]$, y en ese caso se tiene

$$\int_a^b f = \int_a^c f + \int_c^b f.$$

Por inducción, si

$$\mathfrak{p} : a_0 = a < a_1 < a_2 < \cdots < a_n = b$$

es cualquier partición de $[a,b]$, se tiene que f es integrable en $[a,b]$ si y sólo si lo es en cada subintervalo $[a_{k-1}, a_k]$, y en ese caso se tiene

$$\int_a^b f = \int_{a_0}^{a_1} f + \int_{a_1}^{a_2} f + \cdots + \int_{a_{n-1}}^{a_n} f$$

Esta propiedad corresponde a la intuición geométrica que el área de una figura formada por varias figuras que no se solapan es la suma de las áreas de sus partes.

En la notación de integral estamos siempre suponiendo que integramos desde un punto a hasta otro b mayor que él. Por comodidad conviene considerar también los casos $a \geq b$.

Extensión de la notación integral

Si $a = b$, definimos $\int_a^a f = 0$ y si $a < b$, definimos $\int_b^a f = -\int_a^b f$.

Con esta extensión de la notación, la aditividad es cierta cualquiera que sea el orden de los puntos:

$$\int_{a_0}^{a_n} f = \int_{a_0}^{a_1} f + \int_{a_1}^{a_2} f + \cdots + \int_{a_{n-1}}^{a_n} f$$

si $\{a_0, a_1, a_2, \ldots, a_n\}$ es cualquier conjunto de puntos y f es integrable en el intervalo $[\text{mín}\, a_k, \text{máx}\, a_k]$. Equivalentemente,

$$\int_{a_0}^{a_1} f + \int_{a_1}^{a_2} f + \cdots + \int_{a_{n-1}}^{a_n} f + \int_{a_n}^{a_0} f = 0.$$

Podemos resumir la situación anterior considerando los puntos como una cadena circular, que empieza y acaba en el mismo punto:

$$a_0 \to a_1 \to a_2 \cdots \to a_n \to a_0.$$

De este modo, integrando sucesivamente en cada eslabón y sumando siempre da 0.

La linealidad de la integral $\int_a^b$ es válida cualquiera que sea la posición relativa de a y b, pero la monotonía y la desigualdad triangular requieren que $a \leq b$.

• **Ejemplo:** Sabiendo que $\displaystyle\int_0^b x^p\, dx = \frac{b^{p+1}}{p+1}$ para cualquier $b > 0$ y potencia $p \geq 0$, entonces para $0 \leq a \leq b$ tenemos

$$\int_0^b x^p\, dx = \int_0^a x^p\, dx + \int_a^b x^p\, dx$$

y entonces

$$\int_a^b x^p\, dx = \int_0^b x^p\, dx - \int_0^a x^p\, dx$$

y por tanto

$$\boxed{\int_a^b x^p\, dx = \frac{b^{p+1} - a^{p+1}}{p+1}}$$

Con esta fórmula y la linealidad, podemos calcular la integral de cualquier polinomio, al menos en intervalos de números positivos:

$$\int_a^b \sum_{p=0}^{n} k_p x^p\, dx = \sum_{p=0}^{n} k_p \frac{b^{p+1} - a^{p+1}}{p+1} = \sum_{p=0}^{n} \frac{k_p}{p+1}\, b^{p+1} - \sum_{p=0}^{n} \frac{k_p}{p+1}\, a^{p+1}$$

La ausencia de reglas claras para productos, cocientes o composiciones crea dificultades a la hora de poder calcular integrales sin tener que recurrir a las sumas de Riemann. Para poder avanzar más es necesario encontrar un método general para calcular integrales que sea capaz de suplir las carencias señaladas.

5.5. El Teorema Fundamental del Cálculo

Basándonos en los ejemplos de las integrales concretas que hemos mencionado hasta ahora, podemos llegar a observar un fenómeno curioso: parece ser que al calcular una integral, obtenemos siempre una fórmula del tipo

$$\int_a^b f(x)\,dx = F(b) - F(a)$$

donde $F(x)$ es otra función. Además, si nos fijamos bien, de hecho en todos los ejemplos se tiene una relación sencilla entre F y f:

$$F'(x) = f(x).$$

Es decir, al integrar una función f aparece una función F cuya derivada es la función original f que estamos integrando. Esta observación es el hecho central del Cálculo Integral. Procedemos a estudiarlo.

Teorema Fundamental del Cálculo

*Sea f una función integrable en un intervalo compacto $[a, b]$. Para $x \in [a, b]$, definimos la función **integral parcial** de f por*

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

Entonces

- *F es una función continua en $[a, b]$.*
- *Si f es continua, entonces F es derivable en (a, b), con*

$$F'(x) = f(x) \quad \forall\, x \in (a, b).$$

- ***La Regla de Barrow:** para cualquier función Φ que sea continua en $[a, b]$ y derivable en (a, b) con $\Phi' = f$, se tiene*

$$\int_a^b f = \Phi(b) - \Phi(a).$$

Esta diferencia se suele simbolizar por $\Big[\Phi(x)\Big]_{x=a}^{x=b}$ ó simplemente $\Big[\Phi(x)\Big]_a^b$.

Definición de primitiva

*Dada una función f en un intervalo $[a, b]$, se llama **primitiva** ó **antiderivada** ó **integral indefinida** de f a cualquier función F tal que F es continua en $[a, b]$ y derivable en (a, b) con $F' = f$.*

Por tanto, el Teorema Fundamental dice que:

- Toda función continua f tiene una primitiva F (la integral parcial).
- Cualquier primitiva F sirve para calcular la integral mediante la Regla de Barrow.

Relación entre las primitivas de una función

Sean F, G ambas primitivas de f en $[a, b]$. Entonces $G - F$ es una constante.

Demostración. La diferencia $H = G - F$ también es continua en $[a, b]$ y derivable en (a, b), con $H' = G' - F' = f - f = 0$, por tanto H es una constante. □

Se dice que todas las primitivas de una función son iguales salvo una constante. Si ignoramos esta distinción, podemos pensar que la primitiva es esencialmente única.

Notación para integrales indefinidas

Si f es continua, escribimos

$$F(x) = \int f(x)\,dx + C$$

para indicar una integral indefinida de f, es decir, una función con $F' = f$. A menudo se omite mención de la constante arbitraria C.

No debe olvidarse prestar atención al dominio de la función a la hora de escribir una fórmula para la antiderivada.

• **Ejemplo:** Es fácil ver que

$$\int x^p\,dx = \frac{x^{p+1}}{p+1} + C$$

si p es cualquier potencia con $p \neq -1$. Según el valor de p, tendremos distintos dominios. Por ejemplo, si $p \in \mathbb{N}$, la fórmula es válida en cualquier intervalo y nos lleva a

$$\int_a^b x^n\,dx = \frac{b^{n+1} - a^{n+1}}{n+1}, \quad a, b \in \mathbb{R}, \quad n \in \mathbb{N},$$

como ya teníamos antes para $0 \leq a < b$.

Intepretación del Teorema: integrar y derivar son operaciones inversas.

Sea f continua en un intervalo abierto I y sea $a \in I$ cualquier punto fijado.

(1) Primero se integra f, es decir, se hace su integral parcial desde a:

$$f(x) \to F(x) = \int_a^x f.$$

(2) Al derivar la función obtenida se termina con la f original:

$$F(x) \to F'(x) = f(x)$$

En orden inverso, sea F continua en $[a, b]$ y derivable en (a, b).

(1) Primero derivamos F, obteniendo su función derivada, que llamamos f:

$$F(x) \to F'(x) = f(x).$$

(2) Al integrar f, haciendo su integral parcial desde a, como F es por definición una primitiva de f, la Regla de Barrow dice que

$$\int_a^x f = F(x) - F(a)$$

y se termina con la función original F salvo la constante $F(a)$.

Si ignoramos el tema de las constantes, tenemos una relación exacta de inversas.

Para demostrar el Teorema Fundamental es necesario desarrollar más la teoría.

5.5.1. La integral como una media

Sea f integrable sobre $[a, b]$. Vimos que podemos calcular $\int_a^b f$ como límite de sumas regulares, eligiendo cualquier muestra $\mathcal{M}_n = \{c_k\}_{k=1}^n$:

$$\begin{aligned}\int_a^b f = \lim_{n\to\infty} \sum_{k=1}^n f(c_k)\, \frac{b-a}{n} &= (b-a) \cdot \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^n f(c_k) \\ &= (b-a) \cdot \lim_{n\to\infty} \mu_n\end{aligned}$$

donde

$$\mu_n \stackrel{\text{def}}{=} \frac{1}{n} \sum_{k=1}^n f(c_k) = \frac{f(c_1) + f(c_2) + \cdots + f(c_n)}{n}$$

es la **media aritmética** de los valores $f(c_k)$ en los puntos c_k que hemos elegido al azar, uno en cada subintervalo de la partición aritmética $\mathfrak{a}_n$ de $[a, b]$. Por tanto, el límite anterior se puede reescribir como

$$\boxed{\lim_n \mu_n = \frac{1}{b-a} \int_a^b f}$$

Este resultado motiva la siguiente definición.

Valor medio de una función

*Sea f una función integrable en $[a, b]$. El **valor medio** de f sobre $[a, b]$ es*

$$\mu(f) = \frac{1}{b-a}\int_a^b f.$$

También escribiremos $\mu_{[a,b]}(f)$ ó $\mu(f, [a, b])$ si queremos hacer referencia al intervalo.

La definición del valor medio no cambia si $a > b$ en vez de $a < b$.

El valor medio no está definido si $a = b$.

• **Ejemplo:** El valor medio de x sobre $[0, b]$ es

$$\frac{1}{b}\int_0^b x\,dx = \frac{1}{b}\cdot\frac{b^2}{2} = \frac{b}{2}.$$

• **Ejemplo:** El valor medio de x^2 sobre $[0, b]$ es

$$\frac{1}{b}\int_0^b x^2\,dx = \frac{1}{b}\cdot\frac{b^3}{3} = \frac{b^2}{3}.$$

Las propiedades básicas de la integral se traducen en propiedades del valor medio:

Propiedades del valor medio integral

El valor medio de una función integrable cumple:

- ***Linealidad:*** $\mu(f + g) = \mu(f) + \mu(g)$ *y* $\mu(kf) = k\mu(f)$ *si k es una constante.*
- ***Monotonía:*** *si* $f \leq g$ *entonces* $\mu(f) \leq \mu(g)$.
- ***Desigualdad triangular:*** $|\mu(f)| \leq \mu(|f|)$.

Además, se tiene

- ***La media de una constante es ella misma:*** $\mu(k) = k$.
- *Si m, M son constantes con $m \leq f(x) \leq M$, entonces $m \leq \mu(f) \leq M$.*

El valor medio de una función no es siempre un valor de ésta, a pesar de lo que pueda indicar la terminología. Por ejemplo, el valor medio de la función suelo en $[0, 2]$ es

$$\frac{1}{2-0}\int_0^2 \lfloor x \rfloor\,dx = \frac{1}{2}\cdot 1 = \frac{1}{2}$$

pero $1/2$ no es un valor de esta función, que sólo toma valores enteros.

Teorema del Valor Medio del Cálculo Integral

Sea f continua en $[a, b]$. Entonces existe un punto $c \in (a, b)$ tal que

$$\int_a^b f(x)\,dx = f(c)\,(b-a).$$

Es decir, el valor medio de f es de hecho el valor de f en algún punto intermedio.

Demostración. Por el Teorema de Weierstrass, f asume su valor mínimo m y su valor máximo M sobre $[a, b]$. Además, la imagen de $[a, b]$ bajo f es $[m, M]$. En particular, por monotonía se tiene

$$m \leq f(x) \leq M \implies m(b-a) \leq \int_a^b f \leq M(b-a) \implies m \leq \mu(f) = \frac{1}{b-a}\int_a^b f \leq M$$

por tanto el valor medio $\mu(f)$ se halla entre m y M y por tanto, como f es continua, por el Teorema del Valor Intermedio, existe algún $c \in [a, b]$ tal que $f(c) = \mu(f)$.

Hace falta pensar un poco más para ver que de hecho se puede conseguir $c \in (a, b)$. Esta parte de la demostración la omitimos. □

5.5.2. Demostración del Teorema Fundamental

Vamos a demostrar el Teorema Fundamental usando el Teorema del Valor Medio.

Primera parte: si f es integrable, la integral parcial F es continua.

Sea f integrable en $[a, b]$ y consideremos la intergal parcial

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

Por aditividad de la integral,

$$F(x) - F(c) = \int_a^x f - \int_a^c f = \int_c^x f \quad \forall\, c, x \in [a, b].$$

Como f está acotada, existe una constante $K > 0$ tal que $|f(x)| \leq K$ para todo $x \in [a, b]$. Por tanto, suponiendo $c \leq x$,

$$|F(x) - F(c)| = \left|\int_c^x f\right| \leq \int_c^x |f| \leq \int_c^x K = K(x-c).$$

y de manera similar, si $x \leq c$, se tiene $|F(x) - F(c)| \leq K(c-x)$. Por tanto

$$|F(x) - F(c)| \leq K|x-c| \quad \forall\, c, x \in [a, b]$$

El Principio del Sandwich implica que $\lim_{x\to c} F(x) = F(c)$, o sea, F es continua en $[a, b]$.

Segunda parte: si f es continua, la integral parcial F es una primitiva de f.

Supongamos ahora que f es continua en $[a, b]$. Sea $x \in (a, b)$ un punto fijado. Dado un incremento $h \neq 0$ tal que $x + h \in (a, b)$, tenemos por aditividad

$$\frac{F(x+h) - F(x)}{h} = \frac{1}{h} \int_x^{x+h} f.$$

La clave de todo el teorema reside en reconocer que esto es el valor medio de f sobre $[x, x+h]$. Por el Teorema del Valor Medio, existe algún punto $c_h \in [x, x+h]$ tal que

$$\frac{F(x+h) - F(x)}{h} = \mu_{[x,x+h]}(f) = f(c_h).$$

Está claro que cuando $h \to 0$, el intervalo $[x, x+h]$ se comprime al punto x, por tanto, aunque no sepamos qué punto es c_h, como $x \leq c_h \leq x + h$, sabemos por el Principio del Sandwich que $\lim_{h\to 0} c_h = x$. Como f es continua, se tiene

$$F'(x) = \lim_{h\to 0} \frac{F(x+h) - F(x)}{h} = \lim_{h\to 0} f(c_h) = \lim_{c\to x} f(c) = f(x).$$

Esto demuestra que F es una primitiva de f en $[a, b]$.

Tercera parte: Regla de Barrow.

Por definición,

$$F(b) = \int_a^b f, \quad F(a) = \int_a^a f = 0,$$

por tanto

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

Ahora, si Φ es cualquier primitiva, es decir, una función continua en $[a, b]$ y derivable en (a, b) con $\Phi'(x) = f(x)$ para todo $x \in (a, b)$, hemos visto que $\Phi - F$ es una constante C. Pero entonces

$$\begin{aligned}\Phi(b) - \Phi(a) &= (F(b) + C) - (F(a) + C) \\ &= (F(b) + \cancel{C}) - (F(a) + \cancel{C}) = F(b) - F(a) = \int_a^b f,\end{aligned}$$

y queda demostrada la Regla de Barrow.

Esto finaliza la demostración del Teorema Fundamental del Cálculo.

El Teorema Fundamental facilita el cálculo de integrales, reduciendo el proceso a calcular antiderivadas, es decir, realizar el proceso inverso de la derivación. En particular, suplirá la ausencia de fórmulas directas para productos y composiciones con técnicas basadas en reconocer antiderivadas.

Primero, debemos tener una base desde la cual partir, es decir, tablas de antiderivadas de las funciones elementales.

5.6. Algunas antiderivadas comunes

- $\displaystyle\int x^p\,dx = \frac{x^{p+1}}{p+1} + C \qquad (p \neq -1)$
- $\displaystyle\int \frac{dx}{x} = \log x + C$
- $\displaystyle\int e^x\,dx = e^x + C$
- $\displaystyle\int b^x\,dx = \frac{b^x}{\log b} + C \qquad (b > 0, b \neq 1)$

Antiderivadas de las funciones trigonométricas

- $\displaystyle\int \sin x\,dx = -\cos x + C$
- $\displaystyle\int \cos x\,dx = \sin x + C$
- $\displaystyle\int \tan x\,dx = -\log|\cos x| + C$
- $\displaystyle\int \cot x\,dx = \log|\sin x| + C$
- $\displaystyle\int \csc x\,dx = -\log|\csc x + \cot x| + C = \log\left|\tan\frac{x}{2}\right| + C$
- $\displaystyle\int \sec x\,dx = \log|\sec x + \tan x| + C = \log\left|\tan\left(\frac{x}{2} + \frac{\pi}{4}\right)\right| + C$

Antiderivadas de las funciones trigonométricas inversas

- $\displaystyle\int \arcsin x\,dx = x\arcsin x + \sqrt{1-x^2} + C \qquad (|x| < 1)$
- $\displaystyle\int \operatorname{arc\,cos} x\,dx = x\operatorname{arc\,cos} x - \sqrt{1-x^2} + C \qquad (|x| < 1)$
- $\displaystyle\int \arctan x\,dx = x\arctan x - \frac{1}{2}\log(1+x^2) + C$
- $\displaystyle\int \operatorname{arc\,cot} x\,dx = x\operatorname{arc\,cot} x + \frac{1}{2}\log(1+x^2) + C$
- $\displaystyle\int \operatorname{arc\,csc} x\,dx = x\operatorname{arc\,csc} x + \log(x + \sqrt{x^2-1}) + C \qquad (x > 1)$
- $\displaystyle\int \operatorname{arcsec} x\,dx = x\operatorname{arcsec} x - \log(x + \sqrt{x^2-1}) + C \qquad (x > 1)$

Antiderivadas de las funciones trigonométricas hiperbólicas

- $\displaystyle\int \sinh x\,dx = \cosh x + C$
- $\displaystyle\int \cosh x\,dx = \sinh x + C$
- $\displaystyle\int \tanh x\,dx = \log\cosh x + C$
- $\displaystyle\int \coth x\,dx = \log|\sinh x| + C$
- $\displaystyle\int \operatorname{csch} x\,dx = \log\tanh\frac{x}{2} + C = -\log|\operatorname{csch} x + \coth x| + C$
- $\displaystyle\int \operatorname{sech} x\,dx = \arcsin\tanh x + C = 2\arctan\tanh\frac{x}{2} + C$

Antiderivadas de las funciones trigonométricas hiperbólicas inversas

- $\displaystyle\int \operatorname{arcsinh} x\,dx = x\operatorname{arcsinh} x - \sqrt{1+x^2} + C$
- $\displaystyle\int \operatorname{arc}\cosh x\,dx = x\operatorname{arc}\cosh x - \sqrt{x^2-1} + C \qquad (x>1)$
- $\displaystyle\int \operatorname{arctanh} x\,dx = x\operatorname{arctanh} x + \frac{1}{2}\log(1-x^2) + C \qquad (|x|<1)$
- $\displaystyle\int \operatorname{arc}\coth x\,dx = x\operatorname{arc}\coth x + \frac{1}{2}\log(x^2-1) + C \qquad (|x|>1)$
- $\displaystyle\int \operatorname{arc}\operatorname{csch} x\,dx = x\operatorname{arc}\operatorname{csch} x + \operatorname{arcsinh} x + C \qquad (x>0)$
- $\displaystyle\int \operatorname{arcsech} x\,dx = x\operatorname{arcsech} x + \arcsin x + C \qquad (0<x<1)$

5.7. Técnicas de Integración: Sustitución

5.7.1. Sustitución directa

El método de sustitución directa

Se aplica cuando la función f que hay que integrar es de la forma

$$\boxed{f(x) = g(\varphi(x)) \cdot \varphi'(x)}$$

Es decir, f ha de tener dos factores:

- *Un factor es la* ***composición*** *de alguna función g con otra, φ.*
- *El otro factor es la* ***derivada*** *φ' de φ.*

Para calcular la integral de f:

(1) Calculamos una antiderivada G de g: $G' = g$.

(2) Por la Regla de la Cadena, $F(x) = G(\varphi(x))$ es una antiderivada de f:

$$(G \circ \varphi)'(x) = G'(\varphi(x)) \cdot \varphi'(x) = g(\varphi(x)) \cdot \varphi'(x) = f(x).$$

(3) Por tanto $\displaystyle\int f(x)\,dx = G(\varphi(x)) + C$.

• **Uso práctico del método:** la notación diferencial $\int f(x)\,dx$ de Leibniz en los integrandos es más que un modo de designar el proceso límite que lleva a la integral. Sugiere manipulaciones que sirven para calcularlas. Por ejemplo, si queremos calcular la integral de una función del tipo considerado, hacemos una **sustitución** o **cambio de variable**, poniendo

$$\boxed{y = \varphi(x), \quad dy = \varphi'(x)\,dx}$$

y la integral se calcula integrando g en vez de f:

$$\int f(x)\,dx = \int g(\varphi(x)) \cdot \varphi'(x)\,dx = \int g(y)\,dy = G(y) + C$$
$$= G(\varphi(x)) + C = F(x) + C,$$

• **Ejemplo:** $\displaystyle \underbrace{\int 2xe^{x^2}\,dx}_{y = x^2} = \int e^y\,dy = e^y + C = e^{x^2} + C.$

• **Ejemplo:** $\displaystyle \int \tan x\,dx = \underbrace{\int \frac{\sin x}{\cos x}\,dx}_{y = \cos x} = \int -\frac{dy}{y} = -\log y + C = -\log\cos x + C.$

• **Ejemplo:** $\int e^{\sin^2 x} \sin 2x\, dx$ se calcula sustituyendo

$$y = \sin^2 x,\;\; dy = 2\sin x \cos x\, dx = \sin 2x\, dx,$$

quedando entonces $\int e^{\sin^2 x} \sin 2x\, dx = \int e^y\, dy = e^y + C = e^{\sin^2 x} + C.$

También sirve si se puede transformar el integrando en uno que es de la forma requerida, por ejemplo, sumando y restando o multiplicando y dividiendo algún término.

• **Ejemplo:** $\int x\sqrt{1-x^2}\, dx$ se puede calcular sustituyendo $y = 1 - x^2$. Entonces $dy = -2x\, dx$, y al multiplicar y dividir por -2 se obtiene

$$\int x\sqrt{1-x^2}\, dx = -\frac{1}{2}\int -2x\sqrt{1-x^2}\, dx = -\frac{1}{2}\int \sqrt{y}\, dy = -\frac{1}{2}\frac{y^{\frac{3}{2}}}{\frac{3}{2}} = -\frac{1}{3}y^{3/2} = -\frac{1}{3}(1-x^2)^{3/2}.$$

En la práctica, manipulamos las diferenciales algebraicamente, simplemente escribiendo

$$dy = -2x\, dx \implies x\, dx = -\frac{1}{2}dy.$$

La noción de **diferencial**, dx, no ha sido rigurosamente definida, pero gran parte del éxito del Cálculo radica en que se puede manejar **"por sentido común"** y llegar a resultados correctos, como en la sustitución directa. Sin embargo, hay que estar atentos pues hay situaciones en las cuales el sentido común no va a funcionar. En general, es mejor entender la manipulación de las diferenciales como una **receta** o una **ayuda mnemotécnica** para recordar los procedimientos.

• **Ejemplo:** El método también sirve para obtener fórmulas abstractas. Si k es una constante y F es una antiderivada de f, entonces

$$\boxed{\int f(kx)\, dx = \frac{1}{k}F(kx) + C} \quad \text{y recíprocamente} \quad \boxed{\int f\Big(\frac{x}{k}\Big)\, dx = kF\Big(\frac{x}{k}\Big) + C}$$

pues sustituyendo $y = kx$ se tiene $dy = k\, dx$ y

$$\int f(kx)\, dx = \int f(y)\,\frac{dy}{k} = \frac{1}{k}\int f(y)\, dy = \frac{1}{k}F(y) + C = \frac{1}{k}F(kx) + C.$$

• **Ejemplo:** $\int \sin 3x\, dx = -\frac{1}{3}\cos 3x + C.$

5.7.2. Sustitución inversa

El método de sustitución inversa

Dependiendo de qué integrando sea más fácil de calcular, a veces es útil ir al revés, ***convirtiendo*** *la integral $\int f(x)\,dx$ en $\int f(\varphi(u))\,\varphi'(u)$ mediante la sustitución $x = \varphi(u)$.*

Para calcular la integral de f:

- *Calculamos una antiderivada H de $h(u) = f(\varphi(u)) \cdot \varphi'(u)$.*
- *Queda $\int f(x)\,dx = \int h(u)\,du = H(u) + C$.*
- *Para poder expresar la integral original en términos de su variable original x, necesitamos que φ sea* ***invertible*** *para poder hacer la* ***sustitución inversa***

$$u = \varphi^{-1}(x).$$

Entonces finalmente $F(x) = H(\varphi^{-1}(x))$ es una antiderivada de $f(x)$:

$$\int f(x)\,dx = \int h(u)\,du = H(u) + C = H(\varphi^{-1}(x)) + C = F(x) + C.$$

Puede parecer que una sustitución inversa nos está complicando la integral, pero eso es solo debido a la notación. En la práctica, se desarrolla un repertorio de sustituciones para distintos tipos de factores en un integrando, y el método facilita bastante el cálculo de integrales.

• **Ejemplo:** El factor $\sqrt{1-x^2}$ es un buen candidato para la **sustitución trigonométrica** $x = \sin\theta$, ya que $1 - \sin^2\theta = \cos^2\theta$, ó $x = \cos\theta$. Por ejemplo,

$$\begin{aligned}
\int \sqrt{1-x^2}\,dx &= \int \sqrt{1-\sin^2\theta} \cdot \cos\theta\,d\theta = \int \sqrt{\cos^2\theta} \cdot \cos\theta\,d\theta = \int \cos\theta \cdot \cos\theta\,d\theta \\
&= \int \cos^2\theta\,d\theta \\
&= \int \frac{1+\cos 2\theta}{2}\,d\theta = \frac{\theta}{2} + \frac{\sin 2\theta}{4} + C = \frac{\theta}{2} + \frac{\sin\theta\cos\theta}{2} + C
\end{aligned}$$

donde hemos usado las fórmulas de duplicación del ángulo para $\sin, \cos$. La sustitución inversa es $\theta = \arcsin x$, con lo cual queda

$$\begin{aligned}
\int \sqrt{1-x^2}\,dx = \frac{\theta}{2} + \frac{\sin\theta\cos\theta}{2} + C &= \frac{\theta}{2} + \frac{\sin\theta\sqrt{1-\sin^2\theta}}{2} + C \\
&= \frac{\arcsin x}{2} + \frac{x\sqrt{1-x^2}}{2} + C.
\end{aligned}$$

Hay un detalle que no debemos dejar pasar: al poner $\sqrt{\cos^2 x} = \cos x$, estamos suponiendo que $\cos x \geq 0$. Esto no es cierto en cualquier intervalo. Hay que tener cuidado con los dominios de las funciones. En este caso, todo funciona como lo hemos puesto si $x \in [-1, 1]$ y $\theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. En cambio, si estuviésemos considerando un intervalo donde $\cos x \leq 0$, tendríamos que haber puesto $\sqrt{\cos^2 x} = -\cos x$.

No olvidar que el método de sustitución inversa requiere que la función $y = \varphi(x)$ que realiza el cambio de variable sea invertible.

• **Ejemplo:** Un factor $\sqrt{1+x^2}\,dx$ se puede convertir en trigonométrico hiperbólico con la sustitución $x = \sinh u$, ya que $1 + \sinh^2 u = \cosh^2 u$. Además, $\cosh u \geq 0$ para todo u así que no tenemos que preocuparnos del signo en la raíz cuadrada. La sustitución inversa es $u = \operatorname{arcsinh} x$, que está definida en todo $\mathbb{R}$. Por ejemplo,

$$\int \sqrt{1+x^2}\,dx = \int \sqrt{\cosh^2 u}\,\cosh u\,du = \int \cosh^2 u\,du = \int \frac{1+\cosh 2u}{2}\,du$$

$$= \frac{u}{2} + \frac{\sinh 2u}{4} + C = \frac{u}{2} + \frac{\sinh u \cosh u}{2} + C = \frac{u}{2} + \frac{\sinh u\sqrt{1+\sinh^2 u}}{2} + C$$

$$= \frac{\operatorname{arcsinh} x}{2} + \frac{x\sqrt{1+x^2}}{2} + C.$$

• **Ejemplo:** Podemos cambiar el 1 en las dos integrales anteriores por un término k^2 suponiendo $k > 0$:

$$f(x) = \sqrt{1-x^2}, \qquad \int f(x)\,dx = F(x) + C, \qquad F(x) \stackrel{\text{def}}{=} \frac{\arcsin x}{2} + \frac{x\sqrt{1-x^2}}{2}$$

por tanto

$$\int \sqrt{k^2 - x^2}\,dx = \int k\sqrt{1 - \left(\frac{x}{k}\right)^2}\,dx = k\int f\left(\frac{x}{k}\right)dx = k \cdot kF\left(\frac{x}{k}\right) + C$$

$$= k^2 \cdot \left(\frac{\arcsin \frac{x}{k}}{2} + \frac{\frac{x}{k}\sqrt{1 - \frac{x^2}{k^2}}}{2}\right) + C = \frac{k^2}{2}\arcsin\frac{x}{k} + \frac{x\sqrt{k^2-x^2}}{2} + C,$$

y de manera similar

$$\int \sqrt{k^2+x^2}\,dx = \frac{k^2}{2}\operatorname{arcsinh}\frac{x}{k} + \frac{x\sqrt{k^2+x^2}}{2} + C.$$

• **Ejemplo: (no olvidar las constantes de antiderivación).** Dos sustituciones similares, sobre todo trigonométricas, pueden llevar a resultados en apariencia distintos:

$$\int \frac{dx}{\sqrt{1-x^2}} = \int \frac{d(\sin t)}{\sqrt{1-\sin^2 t}} = \int \frac{\cos t\,dt}{\sqrt{\cos^2 t}} = \int dt = t = \arcsin x$$

$$\int \frac{dx}{\sqrt{1-x^2}} = \int \frac{d(\cos t)}{\sqrt{1-\cos^2 t}} = \int \frac{-\sin t\,dt}{\sqrt{\sin^2 t}} = \int -dt = -t = -\arccos x$$

pero son iguales salvo constante; aquí $\arcsin x + \arccos x = \frac{\pi}{2}$ (ángulos complementarios).

5.7.3. La fórmula de cambio de variable para integrales definidas

Teorema del Cambio de Variable

Sean dados

- *Dos intervalos abiertos I, J.*
- *Una función $\varphi : I \to J$ derivable y con derivada continua en I.*
- *Una función f continua en J.*

Entonces, para todos los puntos $a, b \in I$, se tiene

$$\int_{\varphi(a)}^{\varphi(b)} f = \int_a^b (f \circ \varphi) \cdot \varphi'$$

Demostración. Sea F una primitiva de f en J, es decir, con $F' = f$. Esta existe por el Teorema Fundamental. La composición $F \circ \varphi$ está definida ya que φ aplica I en J. Por tanto

$$\begin{aligned}
\int_{\varphi(a)}^{\varphi(b)} f &= F(\varphi(b)) - F(\varphi(a)) && \text{(Regla de Barrow, aplicada a } F \text{ en } [\varphi(a), \varphi(b)]) \\
&= \int_a^b (F \circ \varphi)' && \text{(Regla de Barrow, pero aplicada a } (F \circ \varphi) \text{ en } [a, b]) \\
&= \int_a^b (F' \circ \varphi)\, \varphi' && \text{(Regla de la cadena)} \\
&= \int_a^b (f \circ \varphi)\, \varphi' && \text{(porque } F \text{ es primitiva de } f)
\end{aligned}$$

□

Método de empleo del cambio de variable en una integral definida

- Como sustitución **directa** se entiende

  $$\int_a^b f(\varphi(x)) \cdot \varphi'(x)\, dx = \int_{\varphi(a)}^{\varphi(b)} f(y)\, dy$$

  que se ha **quitado** $\varphi(x)$ del integrando, sustituyéndola por y; a cambio hay que **poner** φ en los **límites de integración.**

- Como sustitución **inversa** se entiende

  $$\int_{\varphi(a)}^{\varphi(b)} f(y)\, dy = \int_a^b f(\varphi(x)) \cdot \varphi'(x)\, dx$$

  que se ha **puesto** $\varphi(x)$ en el integrando, haciendo $y = \varphi(x)$; a cambio hay que **quitar** φ de los **límites de integración.**

De hecho no se requiere que φ sea invertible. Si lo es, podemos escribir en sentido inverso

$$\int_c^d f(y)\,dy = \int_{\varphi^{-1}(c)}^{\varphi^{-1}(d)} f(\varphi(x))\cdot\varphi'(x)\,dx$$

para dos puntos c, d en el dominio de f.

• **Ejemplo:** Para calcular

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

sustituyendo en sentido directo

$$y = \varphi(x) = 1 - x^2,\ \ dy = -2x\,dx,$$

y teniendo en cuenta que $\varphi(0) = 1$ y $\varphi(1) = 0$, por tanto

$$\begin{aligned}\int_0^1 x\sqrt{1-x^2}\,dx &= -\frac{1}{2}\int_1^0 \sqrt{y}\,dy \\ &= \frac{1}{2}\int_0^1 \sqrt{y}\,dy \\ &= \frac{1}{2}\cdot\left[\frac{2}{3}y^{3/2}\right]_0^1 = \frac{1}{3}.\end{aligned}$$

• **Ejemplo:** Calculamos

$$\int_0^{2\pi} \log(2-\log(2+\sin x))\,\cos x\,dx$$

sustituyendo $y = \varphi(x) = \sin x,\ \ dy = \cos x\,dx$; como $\sin 0 = \sin 2\pi = 0$,

$$\int_0^{2\pi} \log(2-\log(2+\sin x))\,\cos x\,dx = \int_0^0 \log(2-\log(2+y))\,dy = 0$$

algo que quizá no sea obvio mirando el integrando, pero que se aclara algo al ver su gráfica (figura 26).

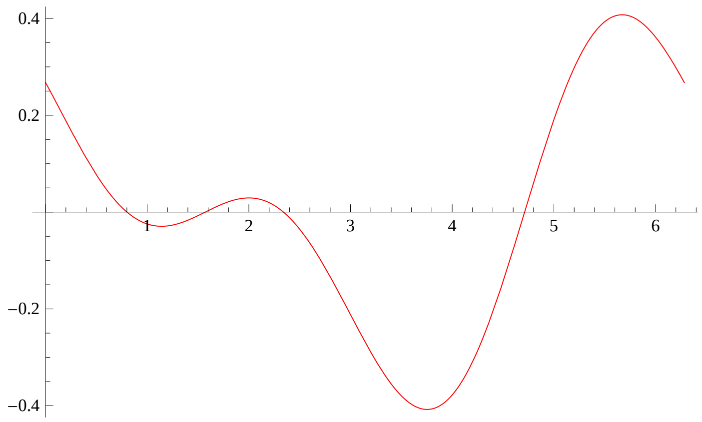

Figura 26 – gráfica de $\log(2-\log(2+\sin x))\,\cos x$

• **Ejemplo:** Consideremos la integral definida $\int_{-1}^{1} \sqrt{1-x^2}\,dx$. Para que se cumplan las hipótesis del teorema, *extendemos* el integrando:

$$f(x) = \begin{cases} \sqrt{1-x^2} & x \in [-1,1] \\ 0 & x \notin [-1,1], \end{cases}$$

que es continua en todo $\mathbb{R}$, ya que $\lim_{|x|\to 1} \sqrt{1-x^2} = 0$. Sustituimos como hicimos antes con la integral indefinida, $x = \varphi(\theta) = \sin\theta$. Formalmente, φ aplica $I = \mathbb{R}$ en $[-1,1] \subseteq \mathbb{R} = J$, es derivable con derivada continua, y f está definida en todo $\mathbb{R}$ y es continua.

Nos damos cuenta que hay infinitas elecciones de a, b con $\varphi(a) = -1,\ \varphi(b) = 1$,

$$1 = \sin\Big(\frac{\pi}{2} + 2\pi n\Big) \qquad -1 = \sin\Big(-\frac{\pi}{2} + 2\pi n\Big) \quad (n \in \mathbb{Z}).$$

Elegimos, por ejemplo $a = -\pi/2,\ b = \pi/2$. Usando la primitiva hallada anteriormente, obtenemos

$$\int_{-1}^{1} \sqrt{1-x^2}\,dx = \left[\frac{\theta}{2} + \frac{\sin 2\theta}{4}\right]_{-\pi/2}^{\pi/2} = \frac{\pi}{2}.$$

Ahora, vamos a tomar $a = -\pi/2,\ b = 5\pi/2$. Ahora

$$\int_{-1}^{1} \sqrt{1-x^2}\,dx = \left[\frac{\theta}{2} + \frac{\sin 2\theta}{4}\right]_{-\pi/2}^{5\pi/2} = \frac{3\pi}{2}.$$

¿Cuál es el fallo? Según el Teorema de Cambio de Variable, el valor de la integral no puede depender de la elección de a, b. Pues está en lo que habíamos advertido: al cambiar de variable, $\sqrt{\cos^2 x} = \cos x$ es válido sólo cuando $\cos x \geq 0$, lo cual no es cierto en $[-\pi/2, 5\pi/2]$. En general, tenemos que poner

$$\sqrt{\cos^2 x} = |\cos x|$$

y precisamente, en $[-\pi/2, 5\pi/2]$ se cancelan $|\cos x| \cos x = -\cos^2 x$ sobre $[\pi/2, 3\pi/2]$ y $|\cos x| \cos x = \cos^2 x$ sobre $[3\pi/2, 5\pi/2]$, quedando solamente la integral sobre $[-\pi/2, \pi/2]$, que es $\pi/2$.

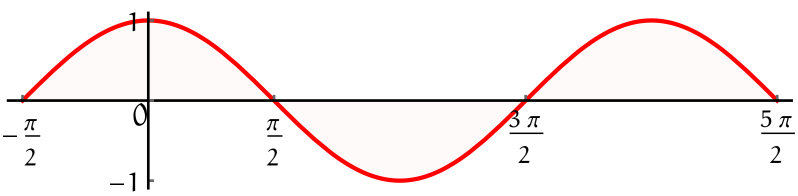

Figura 27 – Gráfica de $\cos x$ en $[\pi/2, 5\pi/2]$.

5.8. Técnicas de Integración: Integración por partes

En las propiedades de la integral se nota la ausencia de una regla para calcular la integral de un producto de funciones, $\int f(x)g(x)\,dx$. Esta carencia se suple con el Teorema Fundamental y el que puede ser el método que más éxito tiene para calcular integrales, la llamada integración por partes.

El método de integración por partes

Aplicable: *cuando el integrando es producto de dos factores y uno de ellos se reconoce como la derivada de una función:*

$$\int f(x)\,dx = \int u(x)\, v'(x)\,dx$$

Método: *se transforma la integral* ***pasando la derivada al otro factor:***

$$\int u(x)\, v'(x)\,dx = u(x)v(x) - \int u'(x)v(x)\,dx$$

Validez: *por ejemplo, cuando todos los factores que intervienen son continuos, es decir,* u, v *son derivables* ***con derivadas continuas.***

La justificación es el Teorema Fundamental y la Regla del Producto para derivadas:

$$(uv)' = u'v + uv' \implies uv' = (uv)' - u'v$$

por tanto si $G(x)$ es una primitiva de $u'(x)v(x)$ entonces $F(x) = u(x)v(x) - G(x)$ es una primitiva de $u(x)v'(x)$.

Notación para la integración por partes Es habitual designar a los factores sin referencia a la variable y en notación de diferenciales:

$$\int u\,dv = uv - \int v\,du$$

donde se está simbolizando

$$\begin{aligned} u &= u(x) \qquad & du &= u'(x)\,dx \\ v &= v(x) \qquad & dv &= v'(x)\,dx. \end{aligned}$$

La habilidad con el método consiste en saber reconocer qué factor conviene que sea "u" y cuál "dv." Puede haber más de una elección, algunas de las cuales no llevan a nada y otras nos dan rápidamente la integral que buscamos. En general:

- Buscamos que dv sea fácil de integrar, al menos, con integral conocida.
- Buscamos que u sea fácil de derivar, o sea, que du no sea demasiado complicada.

• **Ejemplo:**

$$\int \underbrace{x}_{u}\, \underbrace{e^x\,dx}_{dv} \quad \text{con} \quad \begin{matrix} u = x & \qquad dv = e^x\,dx \\ du = dx & \qquad v = e^x \end{matrix}$$

por tanto

$$\int x\, e^x\, dx = uv - \int v\, du = x\, e^x - \int e^x\, dx$$
$$= x\, e^x - e^x$$
$$= (x-1)\, e^x$$

• **Ejemplo:** A veces la propia dx vale como dv:

$$\int \underbrace{\log x}_{u}\, \underbrace{dx}_{dv} \quad \text{con} \quad \begin{array}{cc} u = \log x & dv = dx \\ du = \dfrac{dx}{x} & v = x \end{array}$$

por tanto

$$\int \log x\, dx = uv - \int v\, du = x\, \log x - \int x\, x^{-1}\, dx$$
$$= x\, \log x - x$$
$$= x\, (\log x - 1)$$

• **Nota:** por comodidad, estamos omitiendo la constante arbitraria C en las antiderivadas.

• **Ejemplo:** Con frecuencia, una elección que sirve en una integral sirve para otras integrales parecidas. Para $n = 0, 1, 2, 3, \ldots$, consideremos

$$I_n \stackrel{\text{def}}{=} \int \underbrace{x^n}_{dv}\, \underbrace{\log x\, dx}_{u} \quad \text{con} \quad \begin{array}{cc} u = \log x & dv = x^n\, dx \\ du = \dfrac{dx}{x} & v = \dfrac{x^{n+1}}{n+1} \end{array}$$

Entonces,

$$I_n = \int x^n\, \log x\, dx = \frac{x^{n+1}}{n+1} \log x - \int \frac{x^{n+1}}{n+1} \frac{dx}{x} = \frac{x^{n+1}}{n+1} \log x - \frac{1}{n+1} \int x^n\, dx$$
$$= \frac{x^{n+1}}{n+1} \log x - \frac{x^{n+1}}{(n+1)^2} = \frac{x^{n+1}}{n+1} \left(\log x - \frac{1}{n+1} \right)$$

Las integrales donde además de la variable intervienen ciertas constantes, como un entero n ó un número real t, se llaman **integrales paramétricas.** Con un sólo parámetro, tienen la forma

$$I(t) = \int_a^b f(t, x)\, dx$$

donde t es el parámetro, x la variable, y la función f depende de ambos. La integración por partes es especialmente adecuada para este tipo de integrales.

• **Ejemplo:** A veces hace falta integrar por partes más de una vez, sobre todo con las integrales donde aparecen funciones trigonométricas, que suelen resolverse por pares:

$$A(t) := \int e^x \sin tx\, dx = \int \underbrace{\sin tx}_{u} \underbrace{e^x\, dx}_{dv,\, v=e^x} = e^x \sin tx - t \int e^x \cos tx\, dx$$

$$B(t) \overset{\text{def}}{=} \int e^x \cos tx\, dx = \int \underbrace{\cos tx}_{u} \underbrace{e^x\, dx}_{dv,\, v=e^x} = e^x \cos tx + t \int e^x \sin tx\, dx$$

por tanto, sustituyendo mutuamente en cada expresión,

$$\int e^x \sin tx\, dx = e^x \sin tx - te^x \cos tx - t^2 \int e^x \sin tx\, dx$$

$$\Longrightarrow A(t) = e^x \sin tx - te^x \cos tx - t^2 A(t)$$

$$\Longrightarrow (1+t^2)A(t) = e^x \sin tx - te^x \cos tx$$

$$\Longrightarrow A(t) = e^x \frac{\sin tx - t\cos tx}{t^2+1}$$

y de manera similar

$$\int e^x \cos tx\, dx = e^x \cos tx + te^x \sin tx - t^2 \int e^x \cos tx\, dx$$

$$\Longrightarrow \int e^x \cos tx\, dx = e^x \frac{\cos tx + t\sin tx}{t^2+1}.$$

• **Ejemplo:** A veces al integrar por partes una integral parámetrica, se obtiene una fórmula de recurrencia en vez de uns fórmula completa. Para $n \in \mathbb{Z}, n \geq 0$, consideremos

$$I_n = \int \sin^n x\, dx.$$

Se tiene

$$I_0 = \int dx = x, \qquad I_1 = \int \sin x\, dx = -\cos x$$

y para $n \geq 2$, si integramos por partes,

$$I_n = \int \underbrace{\sin^{n-1} x}_{u} \underbrace{\sin x\, dx}_{dv}$$

con

$$u = \sin^{n-1} x \qquad\qquad dv = \sin x\, dx$$
$$du = (n-1)\sin^{n-2} x\, \cos x \qquad v = -\cos x$$

por tanto

$$\begin{aligned} I_n &= -\sin^{n-1} x \cos x + (n-1) \int \sin^{n-2} x\, \cos^2 x\, dx \\ &= -\sin^{n-1} x \cos x + (n-1) \int \sin^{n-2} x\, (1-\sin^2 x)\, dx \\ &= -\sin^{n-1} x \cos x + (n-1) I_{n-2} - (n-1) I_n \end{aligned}$$

$$\Longrightarrow nI_n = -\sin^{n-1} x \cos x + (n-1) I_{n-2}$$

quedando la recurrencia

$$\boxed{I_0 = x,\ I_1 = -\cos x,\ I_n = -\frac{1}{n}\sin^{n-1} x\cos x + \frac{n-1}{n} I_{n-2} \quad (n \geq 2)}$$

que, si bien no nos da una fórmula completa, nos permite calcular los I_n sucesivamente I_n, para n par, partiendo de $I_0 \to I_2 \to I_4 \to \cdots$ y para n impar, partiendo de $I_1 \to I_3 \to I_5 \to \cdots$.

• **Ejemplo:** Para $n \in \mathbb{Z}, n \geq 0$, consideremos

$$I_n \stackrel{\text{def}}{=} \int \underbrace{x^n}_{u}\,\underbrace{e^x\,dx}_{dv} \quad \text{con} \quad \begin{array}{ll} u = x^n & dv = e^x\,dx \\ du = nx^{n-1}\,dx & v = e^x \end{array}$$

Entonces,

$$I_n = \int x^n\, e^x\, dx = x^n\, e^x - n\int x^{n-1}e^x\,dx = x^n\,e^x - nI_{n-1}.$$

Esto permite calcular recursivamente

$$I_0 = \int e^x\,dx = e^x$$

$$I_1 = xe^x - 1\cdot I_0 = xe^x - e^x = (x-1)e^x$$

$$I_2 = x^2e^x - 2I_1 = x^2e^x - 2(x-1)e^x = (x^2-2x+2)e^x.$$

En este caso un poco de reflexión nos hace observar que hay un patrón

$$I_n = p_n(x)\,e^x$$

donde $p_n(x)$ es un polinomio de grado n. Además, la recurrencia para I_n implica que

$$I_n = p_n(x)\,e^x = x^n\,e^x - np_{n-1}(x)\,e^x = (x^n - np_{n-1}(x))\,e^x,$$

luego p_n se puede calcular recursivamente por

$$\boxed{p_0(x) = 1,\ p_n(x) = x^n - np_{n-1}(x) \quad (n \geq 1)}$$

De hecho, en este caso se tiene una fórmula completa:

$$p_n(x) = (-1)^n\,n!\sum_{k=0}^{n}\frac{(-x)^k}{k!}$$

que se puede demostrar por inducción.

• **Ejemplo:** La integración por partes también tiene un uso teórico. Sea f una función derivable y con derivada continua en el intervalo $[a,b]$. Entonces

$$\lim_{n\to\infty}\int_a^b f(x)\left\{\begin{matrix}\cos nx\\ \sin nx\end{matrix}\right\}dx = 0$$

pues, integrando por partes,

$$\int_a^b \underbrace{f(x)}_{u} \underbrace{\begin{Bmatrix} \cos nx \\ \sin nx \end{Bmatrix} dx}_{dv} = \frac{1}{n} f(x) \begin{Bmatrix} \sin nx \\ -\cos nx \end{Bmatrix} - \frac{1}{n} \int_a^b f'(x) \begin{Bmatrix} \sin nx \\ -\cos nx \end{Bmatrix} dx$$

$|\cos nx|, |\sin nx| \leq 1$, y tanto f como f' están acotadas sobre $[a, b]$ al ser continuas. Sea

$$M = \text{máx}\{|f(x)| : x \in [a, b]\}, \quad M' = \text{máx}\{|f'(x)| : x \in [a, b]\}.$$

Entonces, aplicando la desigualdad triangular,

$$\begin{aligned} \left| \int_a^b f(x) \begin{Bmatrix} \cos nx \\ \sin nx \end{Bmatrix} dx \right| &\leq \frac{1}{n} \cdot M \cdot 1 + \left| \frac{1}{n} \int_a^b f'(x) \begin{Bmatrix} \sin nx \\ -\cos nx \end{Bmatrix} dx \right| \\ &\leq \frac{1}{n} \cdot M + \frac{1}{n} \int_a^b \left| f'(x) \begin{Bmatrix} \sin nx \\ -\cos nx \end{Bmatrix} \right| dx \\ &= \frac{1}{n} \cdot M + \frac{b-a}{n} \cdot M' \to 0 \quad (n \to \infty) \end{aligned}$$

En la práctica, al calcular una integral puede hacer falta usar más de una técnica y más de una vez. Además, no siempre hay un único camino directo hacia la solución, puede haber muchas maneras de llegar al mismo resultado igualmente válidas.

• **Ejemplo:** Vamos a calcular

$$\int \sin(\log y)\, dy$$

(suponiendo $y > 0$) primero sustituyendo

$$y = e^x, \quad dy = e^x\, dx,$$

y después integrando por partes:

$$\begin{aligned} \int \sin(\log y)\, dy &= \int \underbrace{\sin(x)}_{u} \underbrace{e^x\, dx}_{dv} \qquad (u = \sin x,\ v = e^x) \\ &= e^x \sin x - \int e^x \cos x\, dx \\ &= e^x \sin x - \int \underbrace{\cos x}_{u} \underbrace{e^x\, dx}_{dv} \qquad (u = \cos x,\ v = e^x) \\ &= e^x \sin x - \left(e^x \cos x + \int e^x \sin x\, dx \right) \\ &= e^x \sin x - e^x \cos x - \int e^x \sin x\, dx \end{aligned}$$

por tanto

$$\int e^x \sin x\, dx = \frac{1}{2} e^x (\sin x - \cos x)$$

y haciendo la sustitución inversa $x = \log y$ queda finalmente

$$\int \sin(\log y)\, dy = \frac{y}{2}\,(\sin \log y - \cos \log y)$$

5.9. Integral de la Función Inversa

Integral definida de la inversa

Sean I, J intervalos abiertos, $f : I \to J$ y $f^{-1} : J \to I$ funciones mutuamente inversas, derivables con derivadas continuas, por tanto estrictamente monótonas. Entonces

$$\int_a^b f + \int_{f(a)}^{f(b)} f^{-1} = b\, f(b) - a\, f(a)$$

para cualquier par de puntos $a, b \in I$.

Demostración. En la integral

$$\int_{f(a)}^{f(b)} f^{-1}(y)\, dy$$

se puede sustituir la propia f:

$$y = f(x), \quad dy = f'(x)\, dx,$$

resultando una expresión que se puede integrar por partes:

$$\begin{aligned}
\int_{f(a)}^{f(b)} f^{-1}(y)\, dy &= \int_a^b f^{-1}(f(x))\, f'(x)\, dx \\
&= \int_a^b \underbrace{x}_{u}\, \underbrace{f'(x)\, dx}_{dv} \qquad u = x,\ v = f(x) \\
&= [x\, f(x)]_a^b - \int_a^b f(x)\, dx \\
&= b\, f(b) - a\, f(a) - \int_a^b f(x)\, dx
\end{aligned}$$

□

Integral indefinida de la inversa

Sean f, f^{-1} funciones mutuamente inversas, derivables con derivadas continuas. Entonces

$$\int f(x)\, dx + \int f^{-1}(y)\, dy = x\, y + C$$

donde $y = f(x)$.

Demostración. La misma, sin hacer referencia a límites de integración. □

• **Ejemplo:** Consideramos $y = e^x$, $x = \log y$. Entonces

$$\int e^x \, dx + \int \log y \, dy = xy + C,$$

por tanto

$$\int \log y \, dy = xy - e^x + C = y \log y - y + C$$

en $(0, +\infty)$.

• **Ejemplo:** Consideremos $y = \sin x$, $x = \arcsin y$ entre los intervalos $(-\pi/2, \pi/2)$ y $(-1, 1)$. Se tiene

$$\int \sin x \, dx + \int \arcsin y \, dy = xy + C,$$

por tanto

$$\int \arcsin y \, dy = xy + \cos x + C = y \arcsin y + \sqrt{1 - y^2} + C$$

• **Ejemplo:** Consideremos $y = \tan x$, $x = \arctan y$ entre los intervalos $(-\pi/2, \pi/2)$ y $(-\infty, \infty)$. Se tiene

$$\int \tan x \, dx + \int \arctan y \, dy = xy + C,$$

por tanto

$$\begin{aligned} \int \arctan y \, dy = xy + \log|\cos x| + C &= y \arctan y + \log\left(\frac{1}{1+y^2}\right)^{1/2} + C \\ &= y \arctan y - \tfrac{1}{2} \log(1 + y^2) + C \end{aligned}$$

donde hemos usado que

$$1 + y^2 = 1 + \tan^2 x = \frac{1}{\cos^2 x} \quad \therefore \quad |\cos x| = \sqrt{\frac{1}{1+y^2}}.$$

5.10. Integrales de funciones racionales

Consideraremos algunos tipos de integrales que aparecen con cierta frecuencia, empezando con las funciones racionales, es decir, cocientes de polinomios, tales como

$$\int \frac{dx}{x^2+1}, \quad \int \frac{3x^2+4x+3}{x^3+2x^2+3x+4} \, dx, \quad \int \frac{x^6-2x^2+1}{2x^3-6x^2+x+10} \, dx$$

La técnica principal que se emplea para encontrar antiderivadas para funciones racionales se basa en una serie de resultados algebraicos que exponemos a continuación.

5.10.1. Factorización de polinomios

Polinomio irreducible

Sea $p(x)$ un polinomio con coeficientes reales, no nulo. Decimos que $p(x)$ es ***irreducible*** *si* ***no*** *se puede escribir como un producto*

$$p(x) = p_1(x)p_2(x)$$

donde ninguno de los dos factores $p_1(x), p_2(x)$ es constante.

Clasificación de los polinomios irreducibles reales

Los únicos polinomios irreducibles $p(x)$ con coeficientes reales son los

(1) ***lineales:*** *$\ell(x) = ax + b$, siendo $a \neq 0$, o los*

(2) ***cuadráticos de discriminante negativo:***

$$q(x) = ax^2 + bx + c, \quad a \neq 0, \quad \Delta \stackrel{\text{def}}{=} b^2 - 4ac < 0.$$

Equivalentemente, los cuadráticos ***sin raíces reales.***

Los polinomios constantes (no nulos) no se consideran irreducibles, pues la noción de irreducibilidad sólo tiene que ver con descomponer en la variable x. En Álgebra, los polinomios constantes se llaman **unidades** porque son precisamente los polinomios $p(x)$ que tienen recíproco, o sea, tales que hay otro polinomio $q(x)$ tal que $p(x)q(x) = 1$.

• **Ejemplo:** $x + 1, 2x + 2, 3x + 3$ son todos irreducibles lineales.

• **Ejemplo:** $x^2 + 1$ es irreducible cuadrático: no tiene raíces reales. Su discriminante es -4.

• **Ejemplo:** $x^2 + x + 1$ es irreducible cuadrático. Su discriminante es -3.

• **Ejemplo:** $2x^2+6x+4$ no es irreducible pues $2x^2+6x+4 = 2(x^2+3x+2) = 2(x+1)(x+2)$. Alternativamente, su discriminante es $6^2 - 4 \cdot 2 \cdot 4 = 36 - 32 = 4 > 0$.

• **Ejemplo:** $x^3 + 1$ no es irreducible pues $x^3 + 1 = (x + 1)(x^2 - x + 1)$.

• **Ejemplo:** $x^n - 1$ no es irreducible si $n \geq 2$ pues $x^n - 1 = (x-1)(x^{n-1} + x^{n-2} + \cdots + x + 1)$.

Un polinomio es **mónico** si el coeficiente de su monomio de mayor grado es 1.

• **Ejemplo:** $x^6 + 3x^3 + 2x^2 + 6$ es mónico. $2x^3 - x + 1$ y $-x^3 + x + 1$ no lo son.

Factorización única de polinomios

Todo polinomio $p(x)$ no nulo con coeficientes reales

$$p(x) = c_n x^n + c_{n-1}x^{n-1} + \cdots + c_1 x + c_0 \quad (c_n \neq 0)$$

es producto de un número finito de polinomios mónicos irreducibles. Reuniendo los irreducibles repetidos se obtiene una ***factorización***

$$p(x) = c\, f_1(x)^{m_1} f_2(x)^{m_2} \cdots f_r(x)^{m_r}$$

donde c es una constante no nula, $r, m_1, \ldots m_r \in \mathbb{N}$ y los $f_j(x)$ son mónicos irreducibles distintos. La factorización es única salvo por el orden de los factores. Es decir, la colección de datos $c, r, f_j(x), m_j$ está determinada de manera única por $p(x)$, y lo único que podemos alterar es el orden en que aparecen los $f_j(x)$.

- La constante c se denomina el **coeficiente principal** de $p(x)$. Es igual a c_n.
- m_j se denomina la **multiplicidad** u el **orden** del factor irreducible $f_j(x)$.

• **Ejemplo:** La factorización de $p(x) = 2x^2 - 6x + 4 = 2(x^2 - 3x + 2)$ es $2(x-1)(x-2)$.

• **Ejemplo:** La factorización de $p(x) = x^4 - 1$ es $(x-1)(x+1)(x^2+1)$.

• **Ejemplo:** La factorización de $p(x) = x^3 + 1$ es $(x+1)(x^2 - x + 1)$.

• **Ejemplo:** La factorización de $p(x) = x^6 - 1$ es $(x-1)(x+1)(x^2+x+1)(x^2-x+1)$.

El papel del coeficiente principal es secundario, en el sentido que también se suele llamar factorización a cualquier descomposición del tipo

$$p(x) = f_1(x)^{m_1} f_2(x)^{m_2} \cdots f_r(x)^{m_r}$$

donde los $f_j(x)$ son irreducibles distintos pero no necesariamente mónicos. Esto introduce ambigüedad en la unicidad de la factorización al poder "mover" constantes. Por ejemplo,

$$2x^2 - 6x + 4 = 2(x-1)(x-2) = (2x-2)(x-2) = (x-1)(2x-4) = (3x-3)\left(\tfrac{2}{3}x - \tfrac{4}{3}\right)$$

son todas factorizaciones con esta definición ampliada.

La factorización de un polinomio real arbitrario es un problema complicado puesto que en general no existe ninguna fórmula algebraica que nos diga cuáles son las raíces de un polinomio en función de sus coeficientes. En la práctica se emplean métodos numéricos para aproximar las raíces, como el método de bisección basado en el Teorema de Bolzano.

La factorización única es válida para cocientes de polinomios, es decir, para funciones racionales. Lo único que cambia es que los exponentes m_j pueden ser negativos. Por convención, la multiplicidad de un factor se define como el valor absoluto $|m_j|$.

Factorización única de funciones racionales

Toda función racional $r(x)$ no nula

$$r(x) = \frac{a_0 + a_1 x + \cdots + a_n x^n}{b_0 + b_1 x + \cdots + b_m x^m}, \quad (a_n, b_m \neq 0)$$

con coeficientes reales factoriza de manera única salvo por el orden de los factores como

$$r(x) = c\, f_1(x)^{m_1} f_2(x)^{m_2} \cdots f_r(x)^{m_r}$$

donde $c = a_n/b_m$ es el coeficiente principal, $r, m_1, \ldots m_r \in \mathbb{Z}$ (son negativos si están en el denominador) y los $f_j(x)$ son irreducibles mónicos distintos.

• **Ejemplo:** $\dfrac{2x^2-2}{3x^2-15x+18} = \dfrac{2}{3}(x-1)(x+1)(x-2)^{-1}(x-3)^{-1}.$

El resultado principal que sirve para el cálculo de integrales es el siguiente.

5.10.2. La descomposición de una función racional en fracciones parciales

Descomposición en Fracciones Parciales

Sea $r(x) = \dfrac{N(x)}{D(x)}$ una función racional, con numerador $N(x)$ y denominador $D(x)$, es decir

$$r(x) = \frac{N(x)}{D(x)}$$

siendo N, D polinomios en x. Si el grado del numerador es estrictamente menor que el grado del denominador, entonces $r(x)$ es una suma de términos (las "fracciones parciales") de dos tipos, a la cual contribuyen solamente los factores del denominador $D(x)$:

(1) Un factor irreducible lineal $x - \lambda$ de $D(x)$ de multiplicidad m contribuye términos

$$\frac{a_1}{(x-\lambda)} + \frac{a_2}{(x-\lambda)^2} + \cdots \frac{a_m}{(x-\lambda)^m}$$

para algunas constantes $a_1, a_2, \ldots, a_m \in \mathbb{R}$.

(2) Un factor irreducible cuadrático $ax^2 + bx + c$ de $D(x)$ de multiplicidad m contribuye

$$\frac{a_1 + b_1 x}{ax^2+bx+c} + \frac{a_2 + b_2 x}{(ax^2+bx+c)^2} + \cdots + \frac{a_m + b_m x}{(ax^2+bx+c)^m}$$

para algunas constantes $a_1, b_1, a_2, b_2, \ldots, a_m, b_m \in \mathbb{R}$.

A los términos que aparecen en una descomposición en fracciones parciales,

$$\frac{a_n}{(x-\lambda)^n}, \quad \frac{a_n + b_n x}{(ax^2+bx+c)^n}$$

se les denomina **fracciones simples.** A las fracciones simples donde $n = 1$,

$$\frac{a_1}{x-\lambda}, \quad \frac{a_1 + b_1 x}{ax^2+bx+c}$$

las llamaremos **fracciones simples básicas** o **tipos básicos,** lineales o cuadráticos.

En particular, la descomposición en fracciones parciales implica que, para calcular la antiderivada de una función racional $r(x)$, basta con saber cómo calcular la antiderivada de una fracción simple.

Antes de proseguir con esta idea debemos dar un método para poder calcular la descomposición en fracciones parciales. El proceso es puramente algebraico y depende de factorizar el denominador, con lo cual en la práctica se necesitan métodos numéricos, pero aquí en general vamos a suponer que tenemos dada de antemano dicha factorización o considerar solamente ejemplos donde sea fácil de obtener.

5.10.3. Procedimiento algebraico para hallar las fracciones parciales

• Suponemos dada una función racional, cociente de dos polinomios $N(x), D(x)$

$$\boxed{r(x) = \frac{N(x)}{D(x)}}$$

$N(x)$: numerador

$D(x)$: denominador

• **Ejemplo:** Vamos a considerar

$$r(x) = \frac{x^6+3}{2x^6-4x^4-14x^2-8}, \qquad \begin{aligned} N(x) &= x^6+3 \\ D(x) &= 2x^6-4x^4-14x^2-8. \end{aligned}$$

Paso 1: división

Si $\operatorname{gr} N(x) \geq \operatorname{gr} D(x)$, *debemos dividir* N *por* D *con cociente y resto:*

$$N(x) = C(x)D(x) + R(x), \quad \operatorname{gr} R(x) < \operatorname{gr} D(x),$$

siendo el cociente y el resto $C(x), R(x)$ *polinomios en* x. *Se tiene*

$$r(x) = C(x) + \frac{R(x)}{D(x)}$$

En vez de la función original $r(x)$ *pasamos a considerar la función racional* $r^*(x) = \dfrac{R(x)}{D(x)}$, *que cumple la condición sobre los grados* $\operatorname{gr} R < \operatorname{gr} D$.

• **Ejemplo:** En nuestro ejemplo el resultado de la división con resto da

$$\frac{x^6+3}{2(x^6-2x^4-7x^2-4)} = \frac{1}{2} + \frac{2x^4+7x^2+7}{2(x^6-2x^4-7x^2-4)}$$

y pasamos a considerar en vez de la función original $r(x)$, el segundo término en la suma:

$$r(x) :\to \frac{2x^4+7x^2+7}{2(x^6-2x^4-7x^2-4)} \stackrel{\text{def}}{=} r^*(x)$$

Paso 2: factorización del denominador

Factorizamos el denominador $D(x)$

$$D(x) = c\,(x-\lambda_1)^{m_1}\cdots(x-\lambda_L)^{m_L} q_1(x)^{n_1}\cdots q_K(x)^{n_K}$$

en producto de

- *el coeficiente principal c.*
- *L factores de tipo lineal: $(x-\lambda_l)$, con su multiplicidad m_l.*
- *K factores de tipo cuadrático: $q_k(x)$, con su multiplicidad n_k.*

• **Ejemplo:** En nuestro ejemplo, el denominador factoriza como

$$D(x) = \left(2x^6-4x^4-14x^2-8\right) = 2(x-2)(x+2)\left(x^2+1\right)^2.$$

Como ya hemos mencionado, este paso puede ser complicado de realizar en la práctica. Tenemos

- Coeficiente principal: $c = 2$.
- $L = 2$ factores lineales, $x-2$ y $x+2$, ambos de multiplicidad 1.
- $K = 1$ factor cuadrático, x^2+1, de multiplicidad 2.

Paso 3: planteamiento de la descomposición

Escribimos las fracciones simples buscadas en términos de incógnitas:

(1) m_l incógnitas por el factor lineal $x-\lambda_l$ de multiplicidad m_l:

$$\frac{a_{l,1}}{x-\lambda_l} + \frac{a_{l,2}}{(x-\lambda_l)^2} + \cdots + \frac{a_{l,m_l}}{(x-\lambda_l)^{m_l}}, \quad 1 \le l \le L.$$

(2) $2n_k$ incógnitas por el factor cuadrático $q_k(x)^{n_k}$ de multiplicidad n_k:

$$\frac{a_{k,1}+b_{k,1}x}{q_k(x)} + \frac{a_{k,2}+b_{k,2}x}{q_k(x)^2} + \cdots + \frac{a_{k,n_k}+b_{k,n_k}x}{q_k(x)^{n_k}}, \quad 1 \le k \le K.$$

igualando la suma $S(x)$ de todas ellas a la función $r^*(x) = \frac{R(x)}{D(x)}$

Para expresar el caso general necesitamos una notación matricial con dos índices para las incógnitas, uno para indicar el factor y otro para indicar su potencia. En la práctica resulta más cómodo simplemente denotar las incógnitas en orden alfabético.

• **Ejemplo:** En nuestro ejemplo, la descomposición en fracciones parciales tiene la forma

$$S(x) = \frac{a}{x-2} + \frac{b}{x+2} + \frac{c+dx}{x^2+1} + \frac{e+fx}{(x^2+1)^2} = \frac{2x^4+7x^2+7}{2(x^6-2x^4-7x^2-4)}$$

puesto que

- El factor $(x-2)$ de multiplicidad 1 lleva una incógnita a.
- El factor $(x+2)$ de multiplicidad 1 lleva una incógnita b.
- EL factor cuadrático (x^2+1) de multiplicidad 2 lleva $2 \cdot 2 = 4$ incógnitas, c, d, e, f.

Paso 4: planteamiento del sistema lineal equivalente

Multiplicamos cada sumando en la suma con incógnitas $S(x)$ por $D(x)$, obteniendo

$$S(x) = \frac{N^*(x)}{D(x)} = \frac{R(x)}{D(x)} \stackrel{\text{def}}{=} r^*(x),$$

donde el numerador $N^*(x)$ es un polinomio en x con coeficientes que son funciones de las incógnitas. Se deduce que

$$N^*(x) = R(x)$$

y al igualar coeficientes del mismo grado en ambos lados de esta ecuación, resulta un sistema lineal del mismo número de incógnitas que ecuaciones, en concreto, tantas como el grado de $D(x)$. Resolviendo este sistema obtenemos la descomposición en fracciones parciales.

Si $D(x)$ no es mónico, puede ser conveniente separar su coeficiente principal c y tomar $\frac{1}{c}D(x)$, que sí es mónico, como denominador común en vez de $D(x)$. Si hacemos esto tendremos que dividir también los numeradores por c:

$$S(x) = \frac{\frac{1}{c}N^*(x)}{\frac{1}{c}D(x)} = \frac{\frac{1}{c}R(x)}{\frac{1}{c}D(x)}$$

y estaremos igualando $\frac{1}{c}N^*(x) = \frac{1}{c}R(x)$, que es claramente lo mismo que $N^*(x) = R(x)$.

• **Ejemplo:** En nuestro ejemplo, ponemos la suma

$$S(x) = \frac{a}{x-2} + \frac{b}{x+2} + \frac{c+dx}{x^2+1} + \frac{e+fx}{(x^2+1)^2} = \frac{2x^4+7x^2+7}{2(x^6-2x^4-7x^2-4)}$$

bajo el denominador común mónico $\frac{1}{2}D(x)$, multiplicando por él, llevando a

$$S(x) = \frac{N^*(x)}{D(x)}, \qquad \begin{aligned} N^*(x) = (a+b+d)x^5 + (2a-2b+c)x^4 + (2a+2b-3d+f)x^3 \\ +(4a-4b-3c+e)x^2 + (a+b-4d-4f)x \\ +2a-2b-4c-4e, \end{aligned}$$

lo cual al igualar los términos de igual grado en este numerador y en el conocido, también dividido por el coeficiente principal $c=2$, es decir $\frac{1}{2}(2x^4+7x^2+7)$, lleva al sistema de ecuaciones lineales

$$\begin{aligned} 2a-2b-4c-4e &= \tfrac{7}{2} && \text{(grado 0)} \\ a+b-4d-4f &= 0 && \text{(grado 1)} \\ 4a-4b-3c+e &= \tfrac{7}{2} && \text{(grado 2)} \\ 2a+2b-3d+f &= 0 && \text{(grado 3)} \\ 2a-2b+c &= 1 && \text{(grado 4)} \\ a+b+d &= 0 && \text{(grado 5)} \end{aligned}$$

de 6 ecuaciones con 6 incógnitas, ya que $\operatorname{gr} D(x) = 6$, cuyas soluciones son

$$\left\{a \to \frac{67}{200}, b \to -\frac{67}{200}, c \to -\frac{17}{50}, d \to 0, e \to -\frac{1}{5}, f \to 0\right\}.$$

Por tanto la descomposición en fracciones parciales de la función $r^*(x)$ obtenida en el Paso 1 a partir del resto al dividir, es

$$\frac{2x^4+7x^2+7}{2(x^6-2x^4-7x^2-4)} = \frac{67}{200(x-2)} - \frac{67}{200(x+2)} - \frac{17}{50(x^2+1)} - \frac{1}{5(x^2+1)^2}$$

y la descomposición en fracciones parciales de la función original $r(x)$ es

$$\frac{x^6+3}{2(x^6-2x^4-7x^2-4)} = \frac{1}{2} + \frac{67}{200(x-2)} - \frac{67}{200(x+2)} - \frac{17}{50(x^2+1)} - \frac{1}{5(x^2+1)^2}.$$

Como puede observarse, los pasos que hay que dar pueden ser largos, aunque del todo mecánicos, como la resolución del sistema lineal. En ejemplos sencillos (de grado bajo) hay menos cuentas.

- **Ejemplo:** Sea $r(x) = \dfrac{1}{x^2-1}$. Es $N(x) = 1, D(x) = x^2-1$.

- **Paso 1:** $\operatorname{gr} N = 0 < 2 = \operatorname{gr} D$ así que no necesitamos dividir.
- **Paso 2:** El denominador factoriza fácilmente:

$$x^2-1 = (x-1)(x+1).$$

Hay dos factores lineales, ambos de multiplicidad 1, y ninguno cuadrático.

- **Paso 3:** La descomposición en fracciones parciales tiene la forma

$$\frac{a}{x-1}+\frac{b}{x+1}=\frac{1}{x^2-1}.$$

- **Paso 4:** Puesto sobre el denominador común mónico x^2-1 da

$$(a-b)+(a+b)x=1$$

lo cual al igualar coeficientes lleva al sistema de ecuaciones lineales

$$\begin{aligned} a-b&=1\\ a+b&=0 \end{aligned}$$

cuyas soluciones son

$$\left\{a\to\frac{1}{2}, b\to-\frac{1}{2}\right\}.$$

Por tanto la descomposición en fracciones parciales es

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

5.10.4. Integral de una fracción simple de tipo lineal

Esta integral es inmediata.

Integral racional simple de tipo lineal $\displaystyle\int\frac{A}{(x-\lambda)^n}\,dx$

$$\int\frac{dx}{(x-\lambda)^n}=\begin{cases}\log|x-\lambda| & \textit{si } n=1 \textit{ (antiderivada básica logarítmica)}\\ -\dfrac{1}{(n-1)(x-\lambda)^{n-1}} & \textit{si } n>1\end{cases}$$

5.10.5. Integral de una fracción simple básica de tipo cuadrático

Vamos a estudiar la integral

$$\int\frac{A+Bx}{ax^2+bx+c}\,dx \quad (b^2-4ac<0).$$

Por linealidad,

$$\int\frac{A+Bx}{ax^2+bx+c}\,dx=A\int\frac{dx}{ax^2+bx+c}+B\int\frac{x\,dx}{ax^2+bx+c},$$

luego basta con calcular estas dos integrales. De hecho, lo primero que haremos será ver que basta con saber calcular la primera de ellas, correspondiente al caso $A=1, B=0$. Vamos a denotar

$$\boxed{q(x)=ax^2+bx+c}$$

La técnica de reducción consiste en hacer aparecer la derivada $q'(x)$ observando que

$$q'(x) = 2ax + b, \quad \therefore \quad \boxed{x = \frac{q'(x) - b}{2a}}$$

y por tanto se tiene

$$\int \frac{x}{ax^2 + bx + c}\, dx = \frac{1}{2a} \int \frac{q'(x) - b}{q(x)}\, dx = \frac{1}{2a} \underbrace{\int \frac{q'(x)}{q(x)}\, dx}_{y=q(x)} - \frac{b}{2a} \int \frac{dx}{q(x)}$$

$$= \frac{1}{2a} \log q(x) - \frac{b}{2a} \int \frac{dx}{q(x)}.$$

En general, por linealidad queda:

Reducción de $\int \frac{A + Bx}{ax^2 + bx + c}\, dx$ al caso $\int \frac{dx}{ax^2 + bx + c}$:

$$\int \frac{A + Bx}{ax^2 + bx + c}\, dx = \frac{B}{2a} \log |ax^2 + bx + c| + \frac{2Aa - bB}{2a} \int \frac{dx}{ax^2 + bx + c}$$

En la práctica resulta más útil repetir las cuentas que hemos hecho para cada caso particular dado, sabiendo que hay que emplear la técnica de hacer aparecer la derivada, que aprender de memoria fórmulas como la anterior.

• **Ejemplo:** Reducimos "a ojo"

$$\int \frac{x + 1}{x^2 + 1}\, dx = \frac{1}{2} \int \frac{2x}{x^2 + 1}\, dx + \int \frac{dx}{x^2 + 1} = \frac{1}{2} \log |x^2 + 1| + \int \frac{dx}{x^2 + 1}$$

• **Ejemplo:**

$$\int \frac{x + 2}{x^2 + 3x + 2}\, dx = \frac{1}{2} \int \frac{2x + 4}{x^2 + 3x + 2}\, dx = \frac{1}{2} \int \frac{2x + 3}{x^2 + 3x + 2}\, dx + \frac{1}{2} \int \frac{dx}{x^2 + 3x + 2}$$

$$= \frac{1}{2} \log |x^2 + 3x + 2| + \frac{1}{2} \int \frac{dx}{x^2 + 3x + 2}.$$

• Habiendo mostrado cómo se reduce a él, ahora debemos estudiar el caso particular

$$\int \frac{dx}{ax^2 + bx + c}, \quad b^2 - 4ac < 0.$$

Para su resolución se usa la técnica de completar el cuadrado y una sustitución. Sea

$$q(x) \stackrel{\text{def}}{=} ax^2 + bx + c, \quad a \neq 0.$$

Completar el cuadrado consiste en dar estos pasos:

$$q(x) = a\left(x^2 + \frac{b}{a}x + \frac{c}{a}\right) = a\left(x^2 + \frac{b}{a}x + \frac{b^2}{4a^2} + \frac{c}{a} - \frac{b^2}{4a^2}\right) = a\left(\left(x + \frac{b}{2a}\right)^2 + \frac{4ac - b^2}{4a^2}\right)$$

$$= a\left(\left(x + \frac{b}{2a}\right)^2 - \frac{\Delta}{4a^2}\right)$$

que llevan a la fórmula cuadrática para las raíces. Como en nuestro caso es $\Delta < 0$, no hay raíces reales. Teniendo esto en cuenta, transformamos la última expresión un poco más:

$$q(x) = a\left[\left(x+\frac{b}{2a}\right)^2 + \left(\frac{\sqrt{-\Delta}}{2a}\right)^2\right] = a\cdot\left(\frac{\sqrt{-\Delta}}{2a}\right)^2\cdot\left[\left(\frac{2a}{\sqrt{-\Delta}}\right)^2\left(x+\frac{b}{2a}\right)^2+1\right]$$

$$= \frac{-\Delta}{4a}\cdot\left[\left(\frac{2ax+b}{\sqrt{-\Delta}}\right)^2+1\right].$$

Resumiendo, queda

$$\boxed{q(x) = \frac{-\Delta}{4a}(y^2+1)} \quad \text{donde} \quad \boxed{y = \frac{2ax+b}{\sqrt{-\Delta}} = \frac{q'(x)}{\sqrt{-\Delta}}}$$

es la sustitución que nos va a servir para calcular la integral. Para las diferenciales,

$$\boxed{dy = \frac{2a}{\sqrt{-\Delta}}\,dx \qquad dx = \frac{\sqrt{-\Delta}}{2a}\,dy}$$

Al ser suma de cuadrados, queda una arco tangente:

$$\int \frac{dx}{ax^2+bx+c} = \frac{4a}{-\Delta}\cdot\frac{\sqrt{-\Delta}}{2a}\int\frac{dy}{y^2+1} = \frac{2}{\sqrt{-\Delta}}\arctan(y) = \frac{2}{\sqrt{-\Delta}}\arctan\left(\frac{2ax+b}{\sqrt{-\Delta}}\right).$$

Integral racional básica cuadrática $\displaystyle\int \frac{A+Bx}{ax^2+bx+c}\,dx$

Sea $q(x) = ax^2+bx+c$ *con* $a \neq 0$ *y* $\Delta = b^2-4ac < 0$. *Entonces*

$$\int\frac{dx}{q(x)} = \frac{2}{\sqrt{-\Delta}}\arctan\left(\frac{q'(x)}{\sqrt{-\Delta}}\right)$$

y en general

$$\int\frac{A+Bx}{ax^2+bx+c}\,dx = \frac{B}{2a}\log|ax^2+bx+c| + \frac{2aA-bB}{a\sqrt{-\Delta}}\arctan\left(\frac{2ax+b}{\sqrt{-\Delta}}\right).$$

En vez de memorizar estas fórmulas es más recomendable recordar el método empleado para llegar a ellas:

- Reducir al caso $\int\frac{dx}{q(x)}$ haciendo aparecer la derivada $q'(x)$.
- Completar el cuadrado en $q(x)$.
- Sustituir para que aparezcan arctan y $\log|q(x)|$.

• Ejemplo:

$$\int \frac{dx}{4x^2+4x+10} = \int \frac{dx}{4x^2+4x+1+9} = \underbrace{\int \frac{dx}{(2x+1)^2+3^2}}_{u=\frac{2x+1}{3}}$$
$$= \frac{3}{2}\int \frac{du}{(3u)^2+3^2} = \frac{3}{2\cdot 9}\int \frac{du}{u^2+1} = \frac{1}{6}\arctan(u)$$
$$= \frac{1}{6}\arctan\left(\frac{2x+1}{3}\right).$$

5.10.6. Integral de una fracción simple cualquiera de tipo cuadrático

Pasamos a estudiar la integral general

$$\int \frac{A+Bx}{(ax^2+bx+c)^n}\,dx \quad (b^2-4ac<0) \quad (n \geq 1).$$

Acabamos de resolver el caso $n=1$ y ahora veremos que para $n \geq 2$ se puede proceder por inducción, mediante una recurrencia. Intentamos repetir los pasos empleados cuando $n=1$. Consideramos primero el caso particular

$$\boxed{I_n(x) \overset{\text{def}}{=} \int \frac{dx}{(ax^2+bx+c)^n} \qquad (n \geq 2)}$$

Con esta integral bastará, pues con el mismo método que para el caso $n=1$, se tiene

$$\int \frac{x}{(ax^2+bx+c)^n}\,dx = \frac{1}{2a}\int \frac{q'(x)-b}{q(x)^n}\,dx = \frac{1}{2a}\int \frac{q'(x)}{q(x)^n}\,dx - \frac{b}{2a}\int \frac{dx}{q(x)^n}$$
$$= -\frac{1}{2a(n-1)q(x)^{n-1}} - \frac{b}{2a}\int \frac{dx}{q(x)^n}.$$

Reducción de $\int \frac{A+Bx}{(ax^2+bx+c)^n}\,dx$ al caso $\int \frac{dx}{(ax^2+bx+c)^n}$:

Para $n \geq 2$ se tiene

$$\int \frac{A+Bx}{(ax^2+bx+c)^n}\,dx = -\frac{B}{2a(n-1)q(x)^{n-1}} + \frac{2Aa-bB}{2a}\int \frac{dx}{(ax^2+bx+c)^n}$$

Más que la fórmula, es la técnica de hacer aparecer $q'(x)$ lo que se utiliza en la práctica.

A partir de ahora vamos a considerar sólo el caso particular indicado arriba por $I_n(x)$. Al igual que hicimos para $n=1$, primero expresamos $q(x)=ax^2+bx+c$ como suma de cuadrados

$$q(x) = \frac{-\Delta}{4a}\left(y^2+1\right), \quad y = \frac{2ax+b}{\sqrt{-\Delta}}, \quad dx = \frac{\sqrt{-\Delta}}{2a}\,dy,$$

pero ahora vamos a replantear esta suma de cuadrados de otra forma. Como

$$q'(x) = 2ax + b,$$

se tiene

$$q(x) = \frac{-\Delta}{4a}\left(\frac{q'(x)^2}{-\Delta} + 1\right) = \frac{q'(x)^2}{4a} - \frac{\Delta}{4a} \quad \therefore \quad 4aq(x) = q'(x)^2 - \Delta, \quad \Delta = q'(x)^2 - 4a\,q(x)$$

y por tanto[9]

$$\boxed{\Delta = q'(x)^2 - 4aq(x)}$$

Teniendo esto en cuenta, sustituimos $dx = \dfrac{-4aq(x) + q'(x)^2}{\Delta}\,dx$ en la integral:

$$\begin{aligned} I_n(x) = \int \frac{dx}{q(x)^n} &= -\frac{4a}{\Delta}\int \frac{q(x)}{q(x)^n}\,dx + \frac{1}{\Delta}\int \frac{q'(x)^2}{q(x)^n}\,dx \\ &= -\frac{4a}{\Delta} I_{n-1}(x) + \frac{1}{\Delta}\int \frac{q'(x)^2}{q(x)^n}\,dx. \end{aligned}$$

La segunda integral se integra por partes, tomando

$$\begin{aligned} u &= q'(x) & dv &= \frac{q'(x)}{q(x)^n}\,dx \\ du &= q''(x)\,dx = 2a\,dx & v &= -\frac{1}{(n-1)q(x)^{n-1}}. \end{aligned}$$

Queda entonces

$$I_n(x) = -\frac{4a}{\Delta} I_{n-1}(x) + \frac{1}{\Delta}\left(-\frac{q'(x)}{(n-1)q(x)^{n-1}} + \frac{2a}{n-1}\,I_{n-1}(x)\right)$$

lo cual, reuniendo términos, da

$$\begin{aligned} I_n(x) &= \frac{q'(x)}{(n-1)(-\Delta)q(x)^{n-1}} + \left(\frac{2a}{(n-1)\Delta} - \frac{4a}{\Delta}\right) I_{n-1}(x) \\ &= \frac{q'(x)}{(n-1)(-\Delta)q(x)^{n-1}} + \frac{2a}{-\Delta}\left(2 - \frac{1}{n-1}\right) I_{n-1}(x) \\ &= \frac{q'(x)}{(n-1)(-\Delta)q(x)^{n-1}} + \frac{(2n-3)(2a)}{(n-1)(-\Delta)}\,I_{n-1}(x). \end{aligned}$$

[9] Esto es un ejemplo de ecuación diferencial.

Fórmula de recurrencia para $I_n(x) \stackrel{\text{def}}{=} \int \frac{dx}{(ax^2+bx+c)^n}$:

Sea $q(x) = ax^2 + bx + c$ con $\Delta = b^2 - 4ac < 0$. Entonces

- *Para $n = 1$,*
$$I_1(x) = \int \frac{dx}{q(x)} = \frac{2}{\sqrt{-\Delta}} \arctan\left(\frac{q'(x)}{\sqrt{-\Delta}}\right)$$

- *Para $n \geq 2$ se tiene*
$$I_n(x) = \frac{q'(x)}{(n-1)(-\Delta)q(x)^{n-1}} + \frac{(2n-3)(2a)}{(n-1)(-\Delta)} I_{n-1}(x)$$

También aquí es recomendable seguir los pasos que hemos dado para deducir esta recurrencia, más que memorizarla. En todo caso, se basa en la fórmula para el discriminante en términos de $q(x), q'(x)$.

• **Ejemplo:** Como caso particular, si

$$I_n(x) = \int \frac{dx}{(1+x^2)^n},$$

entonces $I_1(x) = \arctan(x)$ y para $n \geq 2$ se tiene

$$I_n(x) = \frac{x}{2(n-1)(1+x^2)^{n-1}} + \frac{2n-3}{2n-2} I_{n-1}(x).$$

Escogiendo $n = 2$ se obtiene

$$I_2(x) = \int \frac{dx}{(1+x^2)^2} = \frac{x}{2(1+x^2)} + \frac{1}{2}\arctan(x).$$

y para $n = 3$

$$I_3(x) = \int \frac{dx}{(1+x^2)^3} = \frac{x\left(3x^2+5\right)}{8\left(x^2+1\right)^2} + \frac{3}{8}\arctan(x).$$

En general por inducción se ve que $I_n(x)$ será un múltiplo racional de $\arctan(x)$ más una función racional de x con denominador $(1+x^2)^{n-1}$. Como el caso de un polinomio cuadrático irreducible general $q(x)$ es cómo este salvo cambios de variable lineales, se puede asegurar que en general

$$I_n(x) = \int \frac{dx}{(ax^2+bx+c)^n}$$

será un múltiplo del arco tangente

$$\arctan \frac{q'(x)}{\sqrt{-\Delta}}$$

más una función racional con denominador $(ax^2 + bx + c)^{n-1}$. La integral de una fracción simple cuadrática en general

$$\int \frac{A + Bx}{(ax^2+bx+c)^n}\, dx$$

tiene la misma forma, excepto en el caso $n = 1$ donde puede contener además un término logarítmico $\log q(x)$.

Esto concluye la discusión del método de integración para funciones racionales. Como puede verse, consta de bastantes pasos, esencialmente algebraicos, por lo cual en general este tipo de integral es adecuado para métodos algorítmicos. Su importancia, aparte de la dada, reside en que muchos tipos de integrales se pueden transformar en racionales.

Resumen del método de integración de una función racional

Dada una función racional $r(x)$, para hallar su integral

(1) Descomponemos $r(x)$ en fracciones parciales.

(2) Las correspondientes a factores lineales (5.10.4) son inmediatos y dan denominadores que son potencias de $(x - \lambda)$ o $\log |x - \lambda|$.

(3) Las correspondientes a factores cuadráticos $q(x)$ se reducen por recurrencia (5.10.6) al caso de multiplicidad $n = 1$ que se calcula haciendo aparecer la derivada y completando el cuadrado (5.10.5). Dan denominadores que son potencias de $q(x)$, factores con la arco tangente, o factores $\log |q(x)|$.

5.11. Integrales de funciones trigonométricas

5.11.1. Polinomios en $\sin x, \cos x$

Se trata de $\int p(\sin x, \cos x)\, dx$ donde $p(x, y)$ es una combinación lineal de términos de la forma $x^m y^n$. Por linealidad, se reduce a calcular las integrales

$$\boxed{I(n, m) = \int \cos^n(x) \sin^m(x)\, dx \quad (n, m \in \mathbb{Z}, n, m \geq 0)}$$

• **Ejemplo:** Las primeras son fáciles:

$$I(0, 0) = \int dx = x, \quad I(1, 0) = \int \cos(x)\, dx = \sin(x), \quad I(0, 1) = \int \sin(x)\, dx = -\cos(x).$$

En general, se procede de manera distinta según la paridad de n, m.

Cálculo de $\int \cos^n(x)\sin^m(x)\,dx$ si n es impar

Si $n = 2k+1$, se separa un coseno y los $2k$ restantes se convierten a senos usando $\cos^2(x) = 1 - \sin^2(x)$:

$$\begin{aligned} I(2k+1,m) = \int \cos^{2k+1}(x)\sin^m(x)\,dx &= \int \cos^{2k}(x)\sin^m(x)\cos(x)\,dx \\ &= \int (1-\sin^2(x))^k \sin^m(x)\cos(x)\,dx \end{aligned}$$

Sustituyendo $s = \sin(x)$, $ds = \cos(x)\,dx$ la integral se transforma en

$$I(2k+1,m) = \int (1-s^2)^k s^m\,ds$$

que es la integral de un polinomio, por tanto inmediata al expandir el binomio.

Cálculo de $\int \cos^n(x)\sin^m(x)\,dx$ si m es impar

Si $m = 2k+1$, se separa un seno y los $2k$ restantes se convierten a cosenos usando $\sin^2(x) = 1 - \cos^2(x)$:

$$\begin{aligned} I(n,2k+1) = \int \cos^n(x)\sin^{2k+1}(x)\,dx &= \int \cos^n(x)\sin^{2k}(x)\sin(x)\,dx \\ &= \int \cos^n(x)(1-\cos^2(x))^k \sin(x)\,dx \end{aligned}$$

Sustituyendo $c = \cos(x)$, $dc = -\sin(x)\,dx$ la integral se transforma en

$$I(2k+1,m) = -\int (1-c^2)^k c^m\,dc$$

que es la integral de un polinomio, por tanto inmediata al expandir el binomio.

• **Ejemplo:** $I(3,0) = \int \cos^3(x)\,dx$: como el exponente 3 del coseno es impar,

$$\begin{aligned} \int \cos^3(x)\,dx = \int \cos^2(x)\cos(x)\,dx &= \int (1-\sin^2(x))\cos(x)\,dx \\ &= \int \left(1-s^2\right)ds = s - \frac{s^3}{3} \\ &= \sin(x) - \frac{\sin^3(x)}{3}. \end{aligned}$$

• **Ejemplo:** $I(0,3) = \int \sin^3(x)\,dx$: como el exponente 3 del seno es impar,

$$\begin{aligned}\int \sin^3(x)\,dx &= \int \sin^2(x)\sin(x)\,dx = \int (1-\cos^2(x))\sin(x)\,dx \\ &= -\int (1-c^2)\,dc = \int \left(c^2-1\right)dc = \frac{c^3}{3} - c \\ &= \frac{\cos^3(x)}{3} - \cos(x).\end{aligned}$$

Cuando n, m son **ambos** impares, para reducir el número de cuentas, conviene escoger el **menor** de n, m para aplicar el método.

• **Ejemplo:** $I(1,3) = \int \cos(x)\sin^3(x)$: elegiendo el coseno, se integra inmediatamente:

$$\int \cos(x)\sin^3(x)\,dx = \int s^3\,ds = \frac{s^4}{4} = \frac{\sin^4(x)}{4},$$

y eligiendo el seno,

$$\begin{aligned}\int \cos(x)\sin^3(x)\,dx &= \int \cos(x)\sin^2(x)\sin(x)\,dx = \int \cos(x)(1-\cos^2(x))\sin(x)\,dx \\ &= -\int c(1-c^2)\,dc = \int \left(c^3-c\right)dc = \frac{c^4}{4} - \frac{c^2}{2} \\ &= \frac{\cos^4(x)}{4} - \frac{\cos^2(x)}{2}\end{aligned}$$

Dado que las primitivas son iguales salvo constante, sabemos que

$$\frac{\cos^4(x)}{4} - \frac{\cos^2(x)}{2} = \frac{\sin^4(x)}{4} + C,$$

aunque esto no sea obvio de antemano. De hecho, poniendo $x = 0$ se obtiene $C = -\frac{1}{4}$.

Deducir esta identidad trigonométrica a partir de la identidad $\cos^2(x) + \sin^2(x) = 1$.

• **Ejemplo:** $I(31,3) = \int \cos^{31}(x)\sin^3(x)\,dx$: elegimos trabajar con el 3:

$$\begin{aligned}\int \cos^{31}(x)\sin^3(x)\,dx &= \int \cos^{31}(x)\sin^2(x)\sin(x)\,dx = \int \cos^{31}(x)(1-\cos^2(x))\sin(x)\,dx \\ &= -\int c^{31}(1-c^2)\,dc = \int \left(c^{33}-c^{31}\right)dc = \frac{c^{34}}{34} - \frac{c^{32}}{32} \\ &= \frac{\cos^{34}(x)}{34} - \frac{\cos^{32}(x)}{32}.\end{aligned}$$

Si hubiéramos elegido el 31, tendríamos que expandir un binomio de 16 términos:

$$\int \cos^{31}(x)\sin^3(x)\,dx = \int \cos^{30}(x)\sin^3(x)\cos(x)\,dx = \int (1-\sin^2(x))^{15}\sin^3(x)\cos(x)\,dx$$

$$= \int (1-s^2)^{15} s^3\, ds = -\tfrac{1}{34}s^{34} + \tfrac{15}{32}s^{32} - \tfrac{7}{2}s^{30} + \tfrac{65}{4}s^{28} - \tfrac{105}{2}s^{26} + \tfrac{1001}{8}s^{24} - \tfrac{455}{2}s^{22} + \tfrac{1287}{4}s^{20}$$
$$- \tfrac{715}{2}s^{18} + \tfrac{5005}{16}s^{16} - \tfrac{429}{2}s^{14} + \tfrac{455}{4}s^{12} - \tfrac{91}{2}s^{10} + \tfrac{105}{8}s^{8} - \tfrac{5}{2}s^{6} + \tfrac{1}{4}s^{4}$$

que nos da una expresión de 16 potencias del seno. Llamemos $C(x), S(x)$ respectivamente a estas antiderivadas obtenidas mediante el coseno y el seno. El Teorema Fundamental asegura que son iguales salvo una constante. La constante se puede averiguar sustituyendo cualquier valor de x, por ejemplo $x=0$. Queda

$$S(0) - C(0) = 0 - \frac{1}{34} + \frac{1}{32} = \frac{1}{544} \quad \therefore \quad S(x) = C(x) + \frac{1}{544},$$

pero comprobar esto directamente a partir de las expresiones de C, S no es fácil.

Queda todavía un caso de la integral, cuando ambos exponentes son pares.

Cálculo de $\int \cos^n(x)\sin^m(x)\,dx$ si m,n son ambos pares

Si $n = 2N$ y $m = 2M$, usamos las fórmulas

$$\cos^2(x) = \frac{1+\cos(2x)}{2}, \qquad \sin^2(x) = \frac{1-\cos(2x)}{2}$$

para reducir la integral

$$I(2N,2M) = \int \cos^{2N}(x)\sin^{2M}(x)\,dx = \int \left(\frac{1+\cos(2x)}{2}\right)^N \left(\frac{1-\cos(2x)}{2}\right)^M dx.$$

Al expandir los binomios queda una combinación lineal de integrales de potencias $\cos^k(2x)$ con $k \le N+M = \frac{1}{2}(m+n)$, luego son potencias más bajas. Seguimos aplicando éste y los otros dos métodos a las integrales resultantes hasta calcularla completamente.

Repitiendo el empleo de estos métodos, las potencias pares se van reduciendo sucesivamente, por tanto tras un número finito de pasos se halla el valor de la integral, aunque se vaya también duplicando sucesivamente la variable.

• **Ejemplo:** $I(2,2) = \int \cos^2(x)\sin^2(x)\,dx$:

$$\int \cos^2(x)\sin^2(x)\,dx = \int \left(\frac{1+\cos(2x)}{2}\right)\left(\frac{1-\cos(2x)}{2}\right) dx = \int \left(\frac{1}{4} - \frac{1}{4}\cos^2(2x)\right) dx$$
$$= \frac{x}{4} - \frac{1}{4}\int \cos^2(2x)\,dx = \frac{x}{4} - \frac{1}{4}\int \left(\frac{1+\cos(4x)}{2}\right) dx$$
$$= \frac{x}{8} - \frac{1}{32}\sin(4x).$$

En esta integral también se puede usar la identidad de duplicación del ángulo para el seno:

$$\sin(2x) = 2\sin(x)\cos(x),$$

por tanto

$$\int \cos^2(x)\sin^2(x)\,dx = \int (\sin(x)\cos(x))^2\,dx = \frac{1}{4}\int \sin^2(2x)\,dx = \frac{1}{4}\int \frac{1-\cos(4x)}{2}\,dx$$
$$= -\frac{x}{8} - \frac{1}{32}\sin(4x).$$

En este caso ambos métodos llevan a expresiones idénticas, no como en el ejemplo anterior.

5.11.2. Polinomios en $\sin(ax), \cos(bx)$

• Se tratan como las anteriores, empleando las fórmulas para separar productos en sumas:

Fórmulas para productos de senos y cosenos

Para todo A, B se tiene

$$\sin A \cos B = \tfrac{1}{2}(\sin(A+B) + \sin(A-B))$$
$$\sin A \sin B = \tfrac{1}{2}(\cos(A-B) - \cos(A+B))$$
$$\cos A \cos B = \tfrac{1}{2}(\cos(A+B) + \cos(A-B))$$

Estas fórmulas son fácilmente deducibles de las fórmulas de suma de ángulos.

• Ejemplo:

$$\int \sin(2x)\cos(3x)\,dx = \frac{1}{2}\int (\sin(5x) + \sin(-x))\,dx = \frac{1}{2}\int (\sin(5x) - \sin(x))\,dx$$
$$= -\frac{\cos(5x)}{10} + \frac{\cos(x)}{2}.$$

• Ejemplo:

$$\int \sin(x)\cos^2(2x)\,dx = \int \sin(x)\left(\frac{1+\cos(4x)}{2}\right)dx = -\frac{\cos(x)}{2} + \frac{1}{2}\int \sin(x)\cos(4x)\,dx$$
$$= -\frac{\cos(x)}{2} + \frac{1}{4}\int (\sin(5x) - \sin(3x)) = -\frac{\cos(x)}{2} - \frac{\cos(5x)}{20} + \frac{\cos(3x)}{12}.$$

En este ejemplo también se puede proceder así:

$$\begin{aligned}\int \sin(x)\cos^2(2x)\,dx &= \int \sin(x)\cos(2x)\cos(2x)\,dx = \frac{1}{2}\int (\sin(3x)-\sin(x))\cos(2x)\,dx \\ &= \frac{1}{2}\int \sin(3x)\cos(2x)\,dx - \frac{1}{2}\int \sin(x)\cos(2x)\,dx \\ &= \frac{1}{4}\int (\sin(5x)+\sin(x))\,dx - \frac{1}{4}\int (\sin(3x)-\sin(x))\,dx \\ &= \frac{1}{2}\int \sin(x)\,dx + \frac{1}{4}\int \sin(5x)\,dx - \frac{1}{4}\int \sin(3x)\,dx \\ &= -\frac{\cos(x)}{2} - \frac{\cos(5x)}{20} + \frac{\cos(3x)}{12}.\end{aligned}$$

Como suele pasar con las funciones trigonométricas, muchos caminos llevan a la solución, aunque a veces dos soluciones no se parezcan mucho pero sin embargo sean iguales, debido a la gran cantidad de identidades trigonométricas.

5.11.3. Funciones racionales de seno y coseno

Consideramos ahora una función racional $r(x, y)$ en dos variables, es decir, el cociente de dos polinomios en x, y:

$$r(x,y) = \frac{p(x,y)}{q(x,y)} = \frac{\sum_{m,n} a_{mn}x^m y^n}{\sum_{m,n} b_{mn}x^m y^n}.$$

Queremos calcular la integral

$$I = \int r(\cos x, \sin x)\,dx.$$

Racional impar en coseno

Esto quiere decir que $r(-x, y) = -r(x, y)$. En este caso se podrá poner en la forma

$$r(x,y) = \frac{x\,p(x^2,y)}{q(x^2,y)}$$

y se puede emplear el mismo método de separar $\cos(x)\,dx$ y sustituir $s = \sin(x)$ que vimos para estas integrales, transformándola en una función racional de s.

Racional impar en seno

Esto quiere decir que $r(x, -y) = -r(x, y)$. En este caso se podrá poner en la forma

$$r(x,y) = \frac{y\,p(x,y^2)}{q(x,y^2)}$$

y se puede emplear el mismo método de separar $\sin(x)\,dx$ y sustituir $c = \cos(x)$ que vimos para estas integrales, transformándola en una función racional de c.

• **Ejemplo:** La integral de la secante

$$\int \sec(x)\,dx = \int \frac{dx}{\cos(x)}$$

corresponde a $r(x, y) = 1/x$. Es impar en coseno. Sustituyendo $s = \sin(x)$, $ds = \cos(x)\,dx$ y "separando" un $\cos(x)\,dx$, se obtiene

$$\begin{aligned}\int \frac{dx}{\cos(x)} &= \int \frac{\cos(x)\,dx}{\cos^2(x)}\\ &= \int \frac{ds}{1-s^2} = \frac{1}{2}\int \frac{ds}{1+s} + \frac{1}{2}\int \frac{ds}{1-s}\\ &= \frac{1}{2}\log(1+s) - \frac{1}{2}\log(1-s) = \frac{1}{2}\log\frac{1+s}{1-s}\\ &= \frac{1}{2}\log\frac{1+\sin(x)}{1-\sin(x)}.\end{aligned}$$

• **Ejemplo:** $\int \frac{\cos^2(x)}{\sin(x)}\,dx$ es impar en el seno. Sustituimos $c = \cos(x)$ y separamos $\sin(x)\,dx$:

$$\begin{aligned}\int \frac{\cos^2(x)}{\sin(x)}\,dx &= \int \frac{\cos^2(x)}{\sin^2(x)}\sin(x)\,dx = \int \frac{-c^2}{1-c^2}\,dc = \int \frac{1-c^2-1}{1-c^2}\,dc\\ &= \int \left(1 - \frac{1}{1-c^2}\right)dc = c - \frac{1}{2}\log\frac{1+c}{1-c} = \cos(x) - \frac{1}{2}\log\frac{1+\cos(x)}{1-\cos(x)}.\end{aligned}$$

La integral racional es esencialmente la misma que en el ejemplo anterior.

La sustitución $t = \tan\frac{x}{2}$

Esta tiene la ventaja de que ***sirve para cualquier función racional*** *aunque no sea impar ni en seno ni en coseno. Se tiene*

$$dx = \frac{2\,dt}{1+t^2}, \quad \cos(x) = \frac{1-t^2}{1+t^2}, \quad \sin(x) = \frac{2t}{1+t^2},$$

por tanto la integral $\int r(\cos x, \sin x)\,dx$ *se transforma en la integral de una función racional en* t*.*

La desventaja es que las funciones racionales en t que resultan suelen ser algo complicadas y por tanto las integrales son largas de calcular.

• **Ejemplo:** Volvemos a calcular $\int \sec(x)\,dx$ pero con la sustitución $t = \tan\frac{x}{2}$:

$$\int \frac{dx}{\cos(x)} = \int \frac{\frac{2\,dt}{1+t^2}}{\frac{1-t^2}{1+t^2}} = \int \frac{2\,dt}{1-t^2} = \log\frac{1+t}{1-t} = \log\frac{1+\tan\frac{x}{2}}{1-\tan\frac{x}{2}}.$$

Ha vuelto a salir la misma integral racional, aunque la expresión final sea con tangentes. Se puede comprobar que las dos primitivas obtenidas para la secante,

$$\frac{1}{2}\log\frac{1+\sin(x)}{1-\sin(x)}, \qquad \log\frac{1+\tan\frac{x}{2}}{1-\tan\frac{x}{2}},$$

de hecho son iguales.

• **Ejemplo:** Calcular $\displaystyle\int \frac{dx}{2+\cos x}$.

$$\begin{aligned}\int \frac{dx}{2+\cos x} &= \int \frac{2\,dt}{(1+t^2)\left(2+\dfrac{1-t^2}{1+t^2}\right)} = \int \frac{2\,dt}{2+2t^2+1-t^2}\\ &= \underbrace{\int \frac{2\,dt}{3+t^2}}_{t=\sqrt{3}u} = \int \frac{2\sqrt{3}\,du}{3+3u^2} = \frac{2}{\sqrt{3}}\arctan u = \frac{2}{\sqrt{3}}\arctan\frac{t}{\sqrt{3}}\\ &= \frac{2}{\sqrt{3}}\arctan\left(\frac{\tan\frac{x}{2}}{\sqrt{3}}\right).\end{aligned}$$

• **Ejemplo:** Calcular $\displaystyle\int \frac{1}{\sin^3(x)\cos^5(x)}\,dx$

Es impar en seno y coseno, pero las tres sustituciones habituales llevan a funciones racionales bastante complicadas:

- Haciendo $c=\cos(x)$ la transforma en

$$\int \frac{\sin(x)\,dx}{\sin^4(x)\cos^5(x)} = \int \frac{-dc}{(1-c^2)^2c^5}$$

- Haciendo $s=\sin(x)$ la transforma en

$$\int \frac{\cos(x)\,dx}{\sin^3(x)\cos^6(x)} = \int \frac{ds}{s^3(1-s^2)^3}$$

- Haciendo $t=\tan\frac{x}{2}$ la transforma en

$$\int \frac{2\,dt}{\dfrac{8t^3}{(1+t^2)^3}\dfrac{(1-t^2)^5}{(1+t^2)^5}(1+t^2)} = \int \frac{(1+t^2)^7\,dt}{4t^3(1-t^2)^5},$$

  que es todavía peor que las anteriores.

Se podría trabajar con la primera, pero hay una sustitución menos habitual que aquí funciona mejor: $t=\tan(x)$. Entonces

$$dx=\frac{dt}{1+t^2}, \quad \cos^2(x)=\frac{1}{1+t^2}, \quad \sin^2(x)=\frac{t^2}{1+t^2}.$$

Para despejar $\sin(x)$ y $\cos(x)$ tendríamos que preocuparnos de estar en un intervalo en el cual tengan un signo definitivo. Aquí no hace falta, pues multiplicando y dividiendo por $\cos^3(x)$ queda

$$\begin{aligned}\int \frac{dx}{\sin^3(x)\cos^5(x)} &= \int \frac{dx}{\tan^3(x)\cos^8(x)} = \int \frac{dt}{t^3 \dfrac{1}{(1+t^2)^4}(1+t^2)} = \int \frac{(1+t^2)^3}{t^3}\,dt \\ &= \int \frac{1+3t^2+3t^4+t^6}{t^3}\,dt = \int \left(t^{-3}+3t^{-1}+3t+t^3\right)dt \\ &= -\frac{1}{2t^2} + 3\log t + \frac{3}{2}t^2 + \frac{t^4}{4} \\ &= -\frac{1}{2\tan^2(x)} + 3\log(\tan(x)) + \frac{3\tan^2(x)}{2} + \frac{\tan^4(x)}{4},\end{aligned}$$

que es mucho más fácil que las otras.

Con integrales de la forma $\displaystyle\int \frac{dx}{\cos^n(x)\sin^m(x)}$ donde m, n son ambos impares, probar

$$t = \tan x, \quad dx = \frac{dt}{1+t^2}, \quad \cos^2(x) = \frac{1}{1+t^2}, \quad \sin^2(x) = \frac{t^2}{1+t^2}.$$

Trigonométricas hiperbólicas

Los métodos vistos para integrales de seno y coseno sirven también para sus versiones hiperbólicas, con las modificaciones necesarias. En concreto, la identidad básica $\sin^2 x + \cos^2 x = 1$ *ahora es*

$$\cosh^2 x - \sinh^2 x = 1$$

y las relaciones de derivación son $\sinh'(x) = \cosh(x)$ *y* $\cosh'(x) = \sinh(x)$*. En cuanto a la sustitución* $t = \tan\frac{x}{2}$*, su versión hiperbólica es*

$$t = \tanh\frac{x}{2}, \quad dx = \frac{2dt}{1-t^2}, \quad \cosh x = \frac{1+t^2}{1-t^2}, \quad \sinh x = \frac{2t}{1-t^2}.$$

• **Ejemplo:** Calcular $\displaystyle\int \frac{dx}{2+\cosh x}$.

Son esencialmente las mismas cuentas que en el ejemplo anterior, cambiando signos. Al final queda una expresión casi idéntica, sólo que con las funciones hiperbólicas:

$$\int \frac{dx}{2+\cosh x} = \frac{2}{\sqrt{3}}\operatorname{arctanh}\left(\frac{\tanh\frac{x}{2}}{\sqrt{3}}\right).$$

La explicación de la similitud entre las funciones trigonométricas y sus variantes hiperbólicas es algebraica. Se debe a la introducción de números complejos, que unifican las funciones trigonométricas y la exponencial mediante la relación resumida en la

$$\boxed{\textbf{FórmuladeEuler} : e^{ix} = \cos(x) + i\sin(x)}$$

siendo $i = \sqrt{-1}$, de la cual se deduce que

$$\boxed{\cos(x) = \frac{e^{ix} + e^{-ix}}{2} \qquad \sin(x) = \frac{e^{ix} - e^{-ix}}{2i}}$$

y finalmente

$$\boxed{\sinh(x) = -i\sin(ix) \qquad \cosh(x) = \cos(ix)}$$

Muchos métodos, incluidos los trigonométricos y las fracciones parciales, son más sencillos con números complejos. Quizá por inercia, el Cálculo se introduce sólo con reales.

5.11.4. Observaciones teóricas sobre la tangente del ángulo medio [10]

• Si $t = \tan\frac{x}{2}$, la demostración de las fórmulas

$$\boxed{dx = \frac{2\,dt}{1+t^2}, \quad \cos(x) = \frac{1-t^2}{1+t^2}, \quad \sin(x) = \frac{2t}{1+t^2}}$$

se puede hacer de la siguiente manera. Aplicando la sustitución inversa queda

$$x = 2\arctan t \quad \therefore \quad dx = \frac{2\,dt}{1+t^2}.$$

Las otras son consecuencia de la relación

$$1 + \tan^2\theta = \frac{1}{\cos^2\theta}$$

y las fórmulas de duplicación del ángulo: con $t = \tan\frac{x}{2}$ queda

$$1 + t^2 = \frac{1}{\cos^2\frac{x}{2}} = \frac{2}{1+\cos x} \quad \therefore \quad \cos x = \frac{2}{1+t^2} - 1 = \frac{1-t^2}{1+t^2}$$

y

$$\sin x = 2\sin\frac{x}{2}\cos\frac{x}{2} = 2\tan\frac{x}{2}\cos^2\frac{x}{2} = \frac{2t}{1+t^2}.$$

Hay una razón más profunda que explica geométricamente cómo y por qué precisamente esta sustitución transforma $\cos(x)$ y $\sin(x)$ en expresiones que son funciones racionales de $t = \tan\frac{x}{2}$. Está relacionada con una construcción geométrica importante, y además con un problema muy conocido y antiguo de las matemáticas.

[10] Esta sección es una digresión respecto al tema principal; se puede omitir en una lectura rápida.

La sustitución $t = \tan \frac{x}{2}$ como proyección estereográfica

La sustitución $t = \tan \frac{x}{2}$ representa la proyección estereográfica de la circunferencia unidad $x^2 + y^2 = 1$ desde el punto $(-1, 0)$ sobre el eje y (Figura 28).

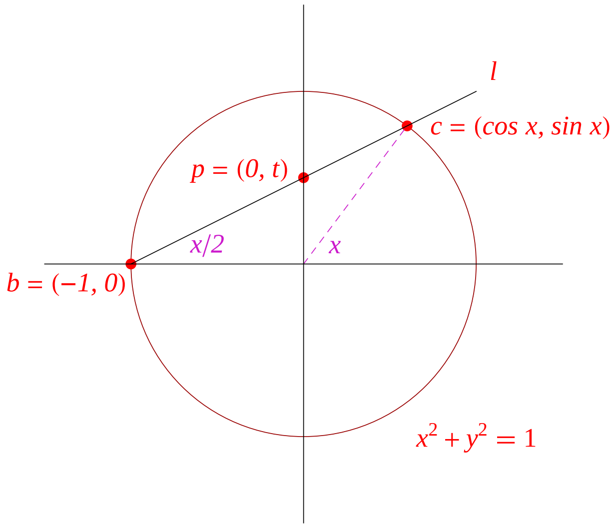

Figura 28 – Proyección estereográfica del círculo unidad sobre el eje y desde $(-1, 0)$.

Demostración. La proyección estereográfica de un punto c de la circunferencia unidad es el punto p del eje y donde la recta ℓ que une c con el punto base $b = (-1, 0)$ corta al eje.

Sea $c = (\cos x, \sin x)$, donde x es el ángulo que hace el radio en c con el eje x. Dado que el triángulo de vértices en b, c y el origen es isósceles (dos de sus lados son radios), vemos que los ángulos en b y c son iguales, digamos a α. Como el ángulo en el origen es $\pi - x$, se tiene $2\alpha + \pi - x = \pi$, por tanto $\alpha = \frac{x}{2}$.

Ahora, poniendo $p = (0, t)$ y considerando el triángulo con vértices en b, p y el origen, vemos que $t = \tan \frac{x}{2}$. Las relaciones

$$\cos(x) = \frac{1 - t^2}{1 + t^2}, \quad \sin(x) = \frac{2t}{1 + t^2}$$

aparecen cuando calculamos las coordenadas de c en términos de t. Pongamos $p = (X, Y)$. La ecuación de la recta ℓ que pasa por $b = (-1, 0)$ y $p = (0, t)$ es

$$\frac{Y - 0}{X - (-1)} = \frac{t - 0}{0 - (-1)} \quad \therefore \quad \frac{Y}{X + 1} = t \quad \therefore \quad Y = t(X + 1).$$

Si (X, Y) se halla tanto en la recta como en la circunferencia, entonces satisface

$$X^2 + Y^2 = 1, \quad Y = t(X + 1), \quad \therefore \quad X^2 + t^2(X + 1)^2 = 1.$$

Esto lleva a una ecuación cuadrática para X:

$$X^2 + t^2(X + 1)^2 = 1 \iff (1 + t^2)\, X^2 + 2t^2\, X + (t^2 - 1) = 0,$$

cuyo discriminante es, algo sorprendentemente,

$$\Delta = 4t^4 - 4(1+t^2)(t^2-1) = 4t^4 - 4(t^4-1) = 4,$$

por tanto sus raíces son

$$X = \frac{-2t^2 \pm 2}{2(1+t^2)} = \frac{-t^2 \pm 1}{1+t^2}.$$

Eligiendo el signo de $-$ queda $X = -1$, luego $Y = 0$, es decir, obtenemos el punto base, que efectivamente es un punto de corte de la recta con la circunferencia, pero no es el que nos interesa. Por tanto, tomamos el signo de $+$, y queda p:

$$\cos x = X = \frac{1-t^2}{1+t^2}, \qquad \sin x = Y = t(X+1) = \frac{2t}{1+t^2},$$

como habíamos dicho. $\square$

La relación entre el punto $c = (\cos x, \sin x)$ de la circunferencia unidad y su proyección estereográfica $p = (0, t)$ sobre el eje y se puede describir como una función invertible

$$c(t) = \left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right), \qquad c^{-1}(\cos x, \sin x) = \left(0, \tan\frac{x}{2}\right)$$

c aplica $\mathbb{R}$ biyectivamente sobre los puntos de la circunferencia unidad menos el punto base de la proyección, que es el $(-1, 0)$. Por ejemplo,

$$c(0) = (1,0), \quad c(1) = (0,1), \quad c(-1) = (0,-1).$$

A medida que el parámetro t recorre $(-\infty, \infty)$, el punto $c(t) = \left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right)$ recorre la circunferencia unidad menos el punto base $(-1,0)$. El punto base sólo se alcanza como límite en infinito:

$$\lim_{t\to\pm\infty} c(t) = (-1, 0).$$

En Geometría se dice que el punto base $(-1,0)$ corresponde al **punto en infinito** de la recta $x = 0$.

Pasemos ahora a otra consideración, de naturaleza aritmética. Si restringimos el valor de t a números racionales poniendo $t = \frac{m}{n}$, con m, n enteros sin factor común y $n > 0$,

$$(X,Y) = c\left(\frac{m}{n}\right) = \left(\frac{1-\dfrac{m^2}{n^2}}{1+\dfrac{m^2}{n^2}}, \frac{2\dfrac{m}{n}}{1+\dfrac{m^2}{n^2}}\right) = \left(\frac{n^2-m^2}{n^2+m^2}, \frac{2mn}{n^2+m^2}\right)$$

con lo cual X, Y también son racionales. En sentido contrario, si $X, Y \in \mathbb{Q}$ entonces también $t = \frac{Y}{X+1} \in \mathbb{Q}$. Es decir,

$$\boxed{t \in \mathbb{Q} \iff X, Y \in \mathbb{Q}}$$

Un punto en una curva con coordenadas racionales se llama simplemente punto racional de la curva. Por tanto, hemos demostrado el siguiente resultado.

Los puntos racionales de la circunferencia unidad menos el punto $(-1, 0)$ están en correspondencia biyectiva con los puntos racionales de la recta $x = 0$, mediante la proyección estereográfica de la circunferencia sobre esta recta desde $(-1, 0)$.

Volviendo a la fórmula para X, Y en términos de $t = \frac{m}{n}$, observamos que:

$$X = \frac{n^2 - m^2}{n^2 + m^2} = \frac{a}{c} \qquad Y = \frac{2mn}{n^2 + m^2} = \frac{b}{c}$$

donde

$$a = n^2 - m^2 \qquad b = 2mn \qquad c = n^2 + m^2$$

son enteros. Sabemos que $c(t) = (X, Y)$ es un punto de la circunferencia unidad, es decir, $X^2 + Y^2 = 1$. Por tanto

$$\boxed{a^2 + b^2 = c^2}$$

Una **terna pitagórica** es una terna (a, b, c) de enteros positivos tales que $a^2 + b^2 = c^2$.

Si $n > m > 0$, o sea, $0 < t = \frac{m}{n} < 1$, entonces $a, b, c > 0$. Los puntos de la circunferencia unidad que proyectan en $(0, t)$ con $0 < t < 1$ son los del primer cuadrante. Esto demuestra por tanto que los puntos racionales en el primer cuadrante de la circunferencia unidad dan lugar a ternas pitagóricas. La asociación se puede denotar por la función

$$\boxed{T(n, m) \stackrel{\text{def}}{=} (n^2 - m^2, 2mn, n^2 + m^2) = (a, b, c)}$$

Estamos diciendo que, si damos valores enteros á n y m, siendo $n > m > 0$, entonces $T(n, m)$ es una terna pitagórica.

- **Ejemplo:** $T(2, 1) = (3, 4, 5)$.

- **Ejemplo:** $T(3, 1) = (8, 6, 10)$. Es la anterior multiplicada por 2 y con a, b intercambiados.

- **Ejemplo:** $T(3, 2) = (5, 12, 13)$.

Cabe plantearse ahora si este método nos genera todas las posibles ternas pitagóricas. Es cierto, con algunos pequeños matices.

Decimos que la terna pitagóricas (a, b, c) es **primitiva** si a, b, c no tienen divisor común.

Las ternas pitagóricas no primitivas se obtienen de las primitivas simplemente multiplicando cada término por un factor común, luego basta con saber cuáles son las ternas primitivas para tenerlas todas. Por ejemplo, $(6, 8, 10)$ no es primitiva pero proviene de $(3, 4, 5)$ que sí es primitiva.

En toda terna pitagórica (a, b, c) primitiva, c siempre es impar, uno de a, b es impar y el otro par. Si a es impar y b par, existen enteros positivos n, m sin factor común tales que

$$a = n^2 - m^2 \qquad b = 2mn \qquad c = n^2 + m^2$$

Es decir, $(a, b, c) = T(n, m)$, y por tanto toda terna pitagórica primitiva se obtiene a partir de un punto racional en el primer cuadrante de la circunferencia unidad mediante la proyección estereográfica sobre el eje y.

Demostración. La dejamos para el lector. □

Este resultado que explica cómo se generan todas las ternas pitagóricas es uno de los más antiguos resultados conocidos de la teoría de números. Data al menos de la Grecia Clásica, y puede ser anterior, ya que muchos historiadores de las Matemáticas argumentan que ya se encuentra en una tableta de arcilla de la era babilónica, conocida como Plimpton 322 y datada en unos ¡4000 años!

5.12. Integrales de irracionalidades algebraicas [11]

5.12.1. Racionales en x y $\sqrt[n]{\frac{ax+b}{cx+d}}$

Vamos a estudiar una última clase de integrales reducibles a integrales de funciones racionales mediante cambios de variable.

Integrales de la forma $\displaystyle\int R\left(x, \sqrt[n]{\frac{ax+b}{cx+d}}\right) dx$ **con** $R(x, y)$ **racional**

Sustituyendo

$$t = \sqrt[n]{\frac{ax+b}{cx+d}}$$

e invirtiendo para expresar x en función de t, la integral se convierte en racional. Se tiene

$$dx = \frac{(ad-bc)\, nt^{n-1}}{(a-c\,t^n)^2}\, dt = (ad-bc)\cdot d\left(\frac{1}{c(a-c\,t^n)}\right)$$

(no confundir el coeficiente d con la d en una diferencial).

Si bien es cierto que el método reduce estas integrales a racionales, algebraicamente puede ser complicado descomponer la función racional que resulta salvo para casos relativamente sencillos.

• **Ejemplo:** Calcular $\displaystyle\int \frac{x+1}{x\sqrt{x+4}}\, dx$.

[11] Esta sección se puede omitir en una lectura rápida

Siguiendo el método señalado, sustituimos

$$t = \sqrt{x+4} \quad \therefore \quad x = t^2 - 4, \; dx = 2t\,dt$$

y entonces la integral se transforma en racional:

$$\int \frac{x+1}{x\sqrt{x+4}}\,dx = \int \frac{t^2-3}{(t^2-4)t} \cdot 2t\,dt = 2\int \frac{t^2-3}{(t^2-4)}\,dt.$$

Procedemos a integrar la función racional que resulta. Como el grado del numerador no es menor que el del denominador, dividimos y separamos el cociente del resto:

$$\frac{t^2-3}{t^2-4} = 1 + \frac{1}{t^2-4}.$$

El denominador factoriza como

$$t^2 - 4 = (t-2)(t+2)$$

por tanto la descomposición en fracciones parciales tiene la forma

$$\frac{a}{t-2} + \frac{b}{t+2} = \frac{1}{t^2-4}$$

que puesto sobre el denominador común da

$$2a - 2b + (a+b)t = 1$$

lo cual equivale al sistema de ecuaciones lineales

$$\begin{aligned} 2a - 2b &= 1 \\ a + b &= 0 \end{aligned}$$

cuyas soluciones son

$$\left\{a \to \frac{1}{4}, b \to -\frac{1}{4}\right\}.$$

Por tanto la descomposición en fracciones parciales es

$$\frac{t^2-3}{t^2-4} = -\frac{1}{4(t+2)} + \frac{1}{4(t-2)} + 1$$

y la integral es

$$\int \frac{t^2-3}{t^2-4}\,dt = t + \frac{1}{4}\log|t-2| - \frac{1}{4}\log|t+2|$$

con lo cual la integral original es

$$\int \frac{x+1}{x\sqrt{x+4}}\,dx = 2\sqrt{x+4} + \frac{1}{2}\log\left|\sqrt{x+4}-2\right| - \frac{1}{2}\log\left|\sqrt{x+4}+2\right|.$$

• **Ejemplo:** Calcular $\displaystyle\int \frac{x+1}{x(x+8)^{1/3}}\,dx$. Sustituimos

$$t = (x+8)^{1/3}, \quad x = t^3 - 8, \quad dx = 3t^2\,dt$$

y entonces
$$\int \frac{x+1}{x(x+8)^{1/3}}\,dx = \int \frac{t^3-7}{(t^3-8)t}\,3t^2\,dt = \int \frac{3t(t^3-7)}{(t^3-8)}\,dt.$$
Ahora tenemos que calcular la integral racional que ha resultado. Como el grado del numerador no es menor que el del denominador, dividimos y separamos el cociente del resto:
$$\frac{3t\left(t^3-7\right)}{t^3-8} = 3t + \frac{3t}{t^3-8}.$$
El denominador factoriza como
$$t^3-8 = (t-2)\left(t^2+2t+4\right)$$
por tanto la descomposición en fracciones parciales tiene la forma
$$\frac{a}{t-2} + \frac{b+ct}{t^2+2t+4} = \frac{3t}{t^3-8}$$
que puesto sobre el denominador común da
$$(a+c)t^2 + (2a+b-2c)t + 4a - 2b = 3t$$
lo cual equivale al sistema de ecuaciones lineales
$$\begin{aligned} 4a-2b&=0\\ 2a+b-2c&=3\\ a+c&=0 \end{aligned}$$
cuyas soluciones son
$$\left\{a \to \frac{1}{2}, b \to 1, c \to -\frac{1}{2}\right\}.$$
Por tanto la descomposición en fracciones parciales es
$$\frac{3t\left(t^3-7\right)}{t^3-8} = 3t + \frac{2-t}{2(t^2+2t+4)} + \frac{1}{2(t-2)}$$
y la integral es (usando la fórmula para fracciones simples básicas cuadráticas)
$$\int \frac{3t\left(t^3-7\right)}{t^3-8}\,dt = \frac{3t^2}{2} - \frac{1}{4}\log\left|t^2+2t+4\right| + \frac{1}{2}\log|2-t| + \frac{1}{2}\sqrt{3}\arctan\left(\frac{t+1}{\sqrt{3}}\right)$$
con lo cual la integral original es
$$\begin{aligned}\int \frac{x+1}{x(x+8)^{1/3}}\,dx = \frac{3}{2}(x+8)^{2/3} + \frac{1}{2}\log\left|2-\sqrt[3]{x+8}\right| - \frac{1}{4}\log\left|(x+8)^{2/3} + 2\sqrt[3]{x+8}+4\right|\\ + \frac{1}{2}\sqrt{3}\arctan\left(\frac{\sqrt[3]{x+8}+1}{\sqrt{3}}\right).\end{aligned}$$

La función definida por una matriz 2×2 (con segunda fila no nula) mediante
$$M = \begin{pmatrix} a & b\\ c & d\end{pmatrix} \mapsto [M](x) = \frac{ax+b}{cx+d}$$
se llama la **transformación de Möbius** asociada a M.

Propiedades de las transformaciones de Möbius

- *La transformación de Möbius asociada a un múltiplo no nulo kM de una matriz M es la misma que la asociada a M:*

$$[kM](x) = [M](x)$$

- *Si M, N son matrices, entonces*

$$([M] \circ [N])(x) = [MN](x),$$

es decir, la composición de las transformaciones asociadas a dos matrices es la transformación asociada al producto de las matrices.

- *La transformación asociada a la matriz identidad I es la función identidad:*

$$[I](x) = x.$$

- *Si M es invertible, la transformación asociada a la matriz inversa M^{-1} es la inversa de la transformación asociada a M:*

$$[M^{-1}](x) = [M]^{-1}(x).$$

• **Ejemplo:** $\int \sqrt{\frac{x+2}{3x+4}}\, dx$. Sustituimos

$$t = \sqrt{[M](x)}, \quad M = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad M^{-1} = -\frac{1}{2}\begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix}$$

y por tanto

$$x = [M^{-1}](t^2) = \left[\begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix}\right](t^2) = \frac{4t^2-2}{-3t^2+1}, \qquad dx = -\frac{4t}{(3t^2-1)^2}\, dt$$

y la integral se transforma en

$$\int \sqrt{\frac{x+2}{3x+4}}\, dx = \int t \cdot \frac{-4t}{(3t^2-1)^2}\, dt = -\int \frac{4t^2}{(3t^2-1)^2}\, dt$$

$$= \frac{2t}{3(3t^2-1)} - \frac{\log\left|3t-\sqrt{3}\right|}{3\sqrt{3}} + \frac{\log\left|3t+\sqrt{3}\right|}{3\sqrt{3}}$$

con lo cual, tras simplificar un poco, queda

$$\int \sqrt{\frac{x+2}{3x+4}}\, dx = \frac{(3x+4)\sqrt{\frac{x+2}{3x+4}}}{3} - \frac{\log\left|3\sqrt{\frac{x+2}{3x+4}}-\sqrt{3}\right|}{3\sqrt{3}} + \frac{\log\left|3\sqrt{\frac{x+2}{3x+4}}+\sqrt{3}\right|}{3\sqrt{3}}.$$

5.12.2. Racionales en x y $\sqrt{ax^2+bx+c}$

Primera sustitución de Euler

• Si $a>0$, las sustituciones

$$t=\sqrt{ax^2+bx+c}\pm\sqrt{a}x$$

la transforman en racional, pues al invertir se considera que

$$\begin{aligned}\sqrt{ax^2+bx+c}=\mp\sqrt{a}x+t &\implies ax^2+bx+c=ax^2\mp 2\sqrt{a}xt+t^2\\ &\implies (b\pm 2\sqrt{a}t)x=t^2-c\\ &\implies x=\frac{t^2-c}{b\pm 2\sqrt{a}t}\end{aligned}$$

por tanto x es función racional de t, luego $\sqrt{ax^2+bx+c}=\pm\sqrt{a}x+t$ también lo es, y al derivar se ve que dx es dt por una función racional de t.

Segunda sustitución de Euler

• Si $c>0$, las sustituciones

$$\sqrt{ax^2+bx+c}=xt\pm\sqrt{c}$$

la transforman en racional, pues al invertir se considera que

$$\begin{aligned}ax^2+bx+c=x^2t^2\pm 2\sqrt{c}xt+c &\implies ax^2+bx=x^2t^2\pm 2\sqrt{c}xt\\ &\implies ax+b=xt^2\pm 2\sqrt{c}t\\ &\implies b\mp 2\sqrt{c}t=(t^2-a)x\\ &\implies x=\frac{b\mp\sqrt{c}t}{t^2-a}\end{aligned}$$

por tanto x es función racional de t, luego $\sqrt{ax^2+bx+c}=xt\pm\sqrt{c}$ también lo es, y al derivar se ve que dx es dt por una función racional de t.

Tercera sustitución de Euler

• Si $b^2-4ac\geq 0$, es decir, si $q(x)=ax^2+bx+c$ tiene dos raíces reales distintas, r y s, entonces podemos factorizar

$$\sqrt{q(x)}=\sqrt{a(x-r)(x-s)}=(x-r)\sqrt{a\frac{x-s}{x-r}}=(x-s)\sqrt{a\frac{x-r}{x-s}}$$

que es del tipo estudiado antes con transformadas de Möbius.

Como siempre se puede completar el cuadrado en un polinomio cuadrático

$$q(x) = ax^2 + bx + c = a\left(\left(x + \frac{b}{2a}\right)^2 - \frac{\Delta}{4a^2}\right), \qquad \Delta \overset{\text{def}}{=} b^2 - 4ac,$$

cualquiera que sea el discriminante Δ, el tipo de integral estudiado se reduce mediante $x \to x - \frac{b}{2a}$ a uno de los casos particulares:

- $q(x) = a^2 - x^2$, donde se puede sustituir $x = a\sin t$ ó $x = a\cos t$,
- $q(x) = a^2 + x^2$, donde se puede sustituir $x = a\tan t$ ó $x = a\sinh t$,
- $q(x) = x^2 - a^2$, donde se puede sustituir $x = a\csc t$ ó $x = \cosh t$,

transformando las integrales correspondientes en integrales racionales trigonométricas o trigonométricas hiperbólicas, que se pueden resolver con los métodos de la sección 5.11.

• **Ejemplo:** Calcular $\displaystyle\int \frac{dx}{\sqrt{x^2+3x-4}}$

Como $q(x) = x^2 + 3x - 4 = (x+4)(x-1)$ tiene dos raíces reales, podemos usar la tercera sustitución de Euler:

$$\sqrt{x^2+3x-4} = (x+4)\sqrt{\frac{x-1}{x+4}} = (x+4)\,t,$$

siendo

$$t = \sqrt{\frac{x-1}{x+4}} = \sqrt{[M](x)}, \quad M = \begin{pmatrix} 1 & -1 \\ 1 & 4 \end{pmatrix},$$

con lo cual

$$M^{-1} = \frac{1}{5}\begin{pmatrix} 4 & 1 \\ -1 & 1 \end{pmatrix}, \qquad \begin{aligned} x &= [M]^{-1}(t^2) = [M^{-1}](t^2) = [5M^{-1}](t^2) = \frac{4t^2+1}{1-t^2} \\ dx &= \frac{10t}{(t^2-1)^2}\,dt, \end{aligned}$$

y entonces

$$\int \frac{dx}{\sqrt{x^2+3x-4}} = \int \frac{1}{\left(\frac{4t^2+1}{1-t^2}+4\right)t}\cdot\frac{10t}{(t^2-1)^2}\,dt = \int \frac{1}{\left(\frac{5}{1-t^2}\right)t}\cdot\frac{10t}{(t^2-1)^2}\,dt$$

$$= \int \frac{2\,dt}{1-t^2} = \log(t+1) - \log(1-t) = \log\frac{1+t}{1-t}.$$

Ahora al poner $t = \sqrt{\frac{x-1}{x+4}}$ se puede simplificar algo la expresión resultante:

$$\log\frac{1+t}{1-t} = \log\frac{(1+t)^2}{1-t^2} = \log\frac{1+2\sqrt{\frac{x-1}{x+4}}+\frac{x-1}{x+4}}{1-\frac{x-1}{x+4}}$$

y, multiplicando numerador y denominador por $x-4$, queda

$$\begin{aligned}\log\frac{x+4+2\sqrt{(x-1)(x+4)}+x-1}{(x+4)-(x-1)} &= \log\left(\frac{2x+3+2\sqrt{x^2+3x-4}}{5}\right)\\ &= \log(2x+3+2\sqrt{x^2+3x-4})-\log 5\end{aligned}$$

con lo cual, como la antiderivada está determinada salvo constante, podemos quitar el término $-\log 5$ y simplemente escribir

$$\int \frac{dx}{\sqrt{x^2+3x-4}} = \log(2x+3+2\sqrt{x^2+3x-4}).$$

• Ahora vamos a resolver la misma integral pero completando el cuadrado. Ponemos

$$x^2+3x-4 = x^2+3x+\frac{9}{4}-\frac{25}{4} = \left(x+\frac{3}{2}\right)^2-\left(\frac{5}{2}\right)^2,$$

hacemos la sustitución

$$u=\frac{2}{5}\left(x+\frac{3}{2}\right),\quad \frac{5}{2}u=\left(x+\frac{3}{2}\right),\quad dx=\frac{5}{2}\,du,\quad \sqrt{x^2+3x-4}=\frac{5}{2}\sqrt{u^2-1},$$

que convierte la integral en

$$\int\frac{dx}{\sqrt{x^2+3x-4}} = \int\frac{du}{\sqrt{u^2-1}}.$$

Ahora ponemos

$$u=\cosh(t)\qquad \sqrt{u^2-1}=\sqrt{\cosh^2(t)-1}=\sqrt{\sinh^2(t)}=\sinh(t),\qquad du=\sinh(t)\,dt$$

si $t\geq 0$ (donde $\sinh(t)\geq 0$) y queda

$$\int\frac{du}{\sqrt{u^2-1}} = \int\frac{\sinh(t)\,dt}{\sinh(t)} = \int dt = t = \operatorname{arc\,cosh}(u) = \operatorname{arc\,cosh}\left(\frac{2x+3}{5}\right).$$

No se parece a la solución anterior, pero de hecho, se tiene

$$\boxed{\operatorname{arc\,cosh}(x) = \log(x+\sqrt{x^2-1})}$$

para $x>1$ (atención a los dominios!), con lo cual sí lo son.

Como puede verse en este ejemplo, el tema de los dominios de definición siempre subyace el cálculo de antiderivadas y cobra mayor importancia para estos ejemplos con raíces. La función $\sqrt{x^2+3x-4}=\sqrt{(x+4)(x-1)}$ está definida como función real solo si $x^2+3x-4\geq 0$, es decir, $x<-4$ ó $x>1$.

• **Ejemplo:** Calcular $\displaystyle\int\frac{dx}{x-\sqrt{x^2+x+1}}$

Aplicamos la primera sustitución de Euler, poniendo

$$t = \sqrt{x^2+x+1} - x$$

y calculando la inversa

$$\begin{aligned}\sqrt{x^2+x+1} = x+t &\implies x^2+x+1 = x^2+2xt+t^2 \implies 1-t^2 = (2t-1)x \\ &\implies x = \frac{1-t^2}{2t-1} \\ &\implies dx = -2\,\frac{(t^2-t+1)}{(2t-1)^2}\,dt\end{aligned}$$

La integral racional que resulta es

$$\int \frac{dx}{x-\sqrt{x^2+x+1}} = \int \frac{1}{(-t)} \cdot -2\,\frac{(t^2-t+1)}{(2t-1)^2}\,dt = \int \frac{2(t^2-t+1)}{t(2t-1)^2}\,dt.$$

El grado del numerador es menor que el grado del denominador así que no necesitamos reducir la función. El denominador ya está factorizado (por comodidad, vamos a dejar las constantes en vez de hacer mónicos todos los factores), por tanto la descomposición en fracciones parciales tiene la forma

$$\frac{a}{t} + \frac{b}{2t-1} + \frac{c}{(2t-1)^2} = \frac{2t^2-2t+2}{t(2t-1)^2}$$

que puesto sobre el denominador común da

$$(4a+2b)t^2 + (-4a-b+c)t + a = 2t^2 - 2t + 2$$

lo cual equivale al sistema de ecuaciones lineales

$$\begin{aligned} a &= 2 \\ -4a-b+c &= -2 \\ 4a+2b &= 2 \end{aligned}$$

cuyas soluciones son

$$\{a \to 2, b \to -3, c \to 3\}.$$

Por tanto la descomposición en fracciones parciales es

$$\frac{2\left(t^2-t+1\right)}{t(2t-1)^2} = -\frac{3}{2t-1} + \frac{3}{(2t-1)^2} + \frac{2}{t}$$

y finalmente la integral es

$$\int \frac{2\left(t^2-t+1\right)}{t(2t-1)^2}\,dt = -\frac{3}{2}\cdot\frac{1}{2t-1} - \frac{3}{2}\log|2t-1| + 2\log|t|$$

con lo cual al sustituir $t = \sqrt{x^2+x+1} - x$ queda la integral original, que se puede simplificar algo despejando raíces de los denominadores:

$$2t-1 = 2\sqrt{x^2+x+1} - (2x+1), \qquad \begin{aligned}\frac{1}{2t-1} &= \frac{2x+1+2\sqrt{x^2+x+1}}{4(x^2+x+1)-(4x^2+4x+1)} \\ &= \frac{2x+1+2\sqrt{x^2+x+1}}{3},\end{aligned}$$

y quitando constantes (el $\frac{1}{3}$ de la expresión anterior) queda

$$\int \frac{dx}{x-\sqrt{x^2+x+1}} = -x - \sqrt{x^2+x+1} - \frac{3}{2}\log\left|2x+1-2\sqrt{x^2+x+1}\right| + 2\log\left|\sqrt{x^2+x+1}-x\right|.$$

5.12.3. Comentarios generales sobre integración

Si repasamos los distintos tipos particulares de integrales que hemos visto, nos damos cuenta que todos acaban reduciéndose a las racionales, que son las aritméticas (operaciones de suma/resta y producto/cociente) y por tanto las más básicas y universales.

Las razones últimas de por qué funcionan las sustituciones explicadas se hallan en el Álgebra y la Geometría y no pueden explicarse aquí. Lo que mejor resume muchos de los conceptos vistos es la siguiente definición.

Una función es **elemental** si puede expresarse como una combinación (finita) de

- potencias y raíces n-ésimas, con $n \in \mathbb{N}$,
- exponenciales y logaritmos,
- funciones trigonométricas y sus inversas (sobre $\mathbb{C}$ son lo mismo que las anteriores)

combinadas mediante las operaciones aritméticas y la composición.

Las funciones elementales abarcan gran parte de las funciones que aparecen en las aplicaciones del Cálculo, pero no todas. Están las llamadas funciones transcendentales superiores que también tienen una gran importancia en estudios más avanzados.

Puede observarse que, debido a las reglas de derivación, las derivadas de funciones elementales son también funciones elementales. Sin embargo, no es así con las integrales.

Existen funciones elementales con antiderivadas no elementales. Algunos ejemplos de este fenómeno (que de hecho es más la regla que la excepción) son

$$\int \sqrt{1+x^4}\,dx, \quad \int e^{-x^2}\,dx, \quad \int \frac{dx}{\log x}, \quad \int \log\log x\,dx, \quad \int \frac{\sin x}{x}\,dx, \quad \int \frac{\cos x}{x}\,dx.$$

6. Integrales Impropias

6.1. Definiciones básicas

En la definición tradicional de la integral se parte siempre de dos datos:

(1) Un intervalo **acotado** que contiene a sus extremos: $[a, b]$.

(2) Una función **acotada** $f(x)$ definida sobre $[a, b]$.

La palabra "impropia" al referirse a una integral significa simplemente que se relajan estas dos condiciones. Resulta que no es tan importante la condición de contener a los extremos; bien podríamos considerar intervalos abiertos (a, b) o semiabiertos $[a, b), (a, b]$ para integrar, manteniendo las otras condiciones, sin cambiar demasiado la teoría. Lo que realmente influye son las condiciones de **acotación.**

Definición (informal) de integral impropia

Una integral $\int_a^b f(x)\,dx$ es impropia si se da al menos una de estas condiciones:

*(1) (a, b) es un **intervalo no acotado**, ó*

*(2) f es una **función no acotada** sobre (a, b).*

Detallemos de modo más formal el concepto. Los datos básicos serán

(1) Un intervalo $X \subseteq \mathbb{R}$ de cualquier tipo con extremos en a, b, no necesariamente incluidos en X. En particular, X puede ser no acotado (se permite $a = -\infty$ ó $b = \infty$).

(2) Una función $f : X \to \mathbb{R}$ continua salvo en un número finito de puntos de X.

Las discontinuidades de la función tienen que ser lo suficientemente "malas" para "estropear" una integral y hacerla impropia en el sentido que buscamos expresar.

Definición de función no acotada en un punto

*Sea $f(x)$ una función definida para $x \approx c$, $x \neq c$. Decimos que $f(x)$ no está acotada en c (o más precisamente, cuando $x \to c$) si existe **alguna** sucesión de puntos $x_n \to c$ con $x_n \neq c$ para todo n, tal que $|f(x_n)| \to \infty$.*

La condición de no estar acotada en c es menos fuerte que $\lim_{x\to c} |f(x)| = \infty$, lo cual equivaldría a que $|f(x)| \to \infty$ a lo largo de **todas** las sucesiones que tienden á c.

• **Ejemplo:** Sea $f(x) = x^{-1}\sin(x^{-1})$ para $x \neq 0$. Escogiendo por ejemplo $x_n = \frac{2}{(2n+1)\pi}$ con $n = 1, 2, 3, \ldots$, se tiene $|f(x_n)| = \frac{(2n+1)\pi}{2} \to \infty$, lo cual muestra que f no está acotada cuando $x \to 0$. Sin embargo, no es cierto que $\lim_{x\to 0} |f(x)| = \infty$, pues de hecho el límite no existe. Por ejemplo, a lo largo de la sucesión $y_n = \frac{1}{n\pi}$ se tiene $f(y_n) = 0$. Visualmente, f es una onda que oscila con oscilaciones de amplitudes que tienden a infinito pero no todos los puntos son llevados a infinito.

Distinguiremos ahora los puntos "malos" cuya presencia hacen "impropia" a una integral.

Definición de punto impropio

*Si $f : X \to \mathbb{R}$ cumple las condiciones anteriores, llamaremos **puntos impropios** a*

- *los puntos $c \in (a, b)$ donde $f(x)$ no está acotada,*
- *aquellos de los extremos a, b finitos ($\neq \pm\infty$) donde $f(x)$ no está acotada,*
- *aquellos de los extremos a, b que sean $\pm\infty$, independiente de si f está o no acotada.*

$\pm\infty$ se consideran automáticamente puntos impropios, sin imponer condición alguna acerca del límite de $f(x)$ cuando x se acerca á $\pm\infty$. Puede ser no acotada o acotada.

El objetivo es dar sentido a la integral $\int_X f(x)\,dx$ bajo estas condiciones, asignándole un valor de manera consistente. Para ello, se emplea la siguiente técnica de análisis.

Subdivisión del intervalo de integración en semiabiertos

Dada una función $f(x)$ en un intervalo X de extremos a, b con colección $\mathcal{P}$ de puntos impropios, la técnica para analizar la integral impropia consiste en

(1) ***intercalar puntos:*** *para cada par **consecutivo** de puntos **impropios**, c y c', se coloca cualquier punto (que ya no será impropio) en medio, $c < m < c'$, y se separa el intervalo abierto entre los dos puntos impropios de la siguiente manera:*

$$(c, c') = (c, m] \cup [m, c').$$

*(2) De este modo el intervalo X menos los puntos impropios queda subdividido en una unión finita de subintervalos **semiabiertos,** es decir, de la forma $(c, \beta]$ ó $[\alpha, c)$, con f continua en cada semiabierto pero*

a) *o bien el extremo $c = \pm\infty$, o*

b) *el extremo c es finito pero $f(x)$ no está acotada cuando x se acerca á c.*

Observar que los puntos impropios quedan descartados en la subdivisión. Como hemos comentado anteriormente, incluir o no un extremo del intervalo de integración resulta irrelevante en cuanto a la noción general de integrabilidad.

• **Ejemplo:** $f(x) = 1/x(x-2)(x-4)(x-6)$ es continua en $\mathbb{R}$ salvo en los puntos $0, 2, 4, 6$ donde no está acotada. Si la analizamos sobre $(0, \infty)$, los puntos impropios serán $0, 2, 4, 6, \infty$.

Colocamos puntos entre puntos impropios consecutivos, por ejemplo

$$0, 1, 2, 3, 4, 5, 6, 7, \infty$$

resultando los ocho semiabiertos $(0, 1]$, $[1, 2)$, $(2, 3]$, $[3, 4)$, $(4, 5]$, $[5, 6)$, $(6, 7]$ y $[7, \infty)$.

• **Ejemplo:** $f(x) = x^{-1}$ es continua en $\mathbb{R}$ salvo en 0.

- Si la analizamos entre 0 y 1 no necesitamos intercalar ningún punto: tendremos un único intervalo semiabierto $(0, 1]$.
- Si la analizamos entre -1 y 1, debemos separar en dos semiabiertos, por ejemplo $[-1, 0)$ y $(0, 1]$.
- Si la analizamos entre -1 y ∞, debemos separar en tres semiabiertos, por ejemplo $[-1, 0)$, $(0, 1]$ y $[1, \infty)$.

En general, n puntos impropios requieren $2n - e$ subintervalos semiabiertos, donde $e = 0, 1, 2$ es el número de extremos del intervalo que son impropios.

El siguiente paso es definir formalmente la noción de convergencia de la integral impropia de una función continua sobre un intervalo semiabierto, asignándole un valor.

Definición de convergencia de una integral impropia en un semiabierto

Sea $f(x)$ continua en el intervalo semiabierto $[a, b)$ (respectivamente $(a, b]$), donde

(1) o bien $b = \infty$ (resp. $a = -\infty$), ó

(2) $f(x)$ no está acotada cuando $x \to b^-$ (resp. $x \to a^+$).

Se dice que la integral impropia de f sobre $[a, b)$ es ***convergente*** *si existe el límite*

$$\lim_{x \to b^-} \int_a^x f \stackrel{\text{def}}{=} \int_a^{b^-} f$$

Respectivamente, para un intervalo $(a, b]$ semiabierto a la izquierda, si existe el límite

$$\lim_{x \to a^+} \int_x^b f \stackrel{\text{def}}{=} \int_{a^+}^b f.$$

En ambos casos el ***valor*** *de la integral impropia se define como el valor del límite correspondiente. Si este límite no existe, se dice que la integral es* ***divergente.*** *También se dice que f es* ***impropiamente integrable*** *como sinónimo de convergente.*

Finalmente, con la noción de convergencia sobre un semiabierto y la técnica de subdivisión señalada antes podemos definir la convergencia de una integral impropia en general.

Definición (estricta) de convergencia de una integral impropia general

Sea $f(x)$ una función sobre un intervalo X, continua salvo en un número finito de puntos. Sea $\mathcal{P}$ el conjunto de puntos impropios. Se dice que la integral impropia de f sobre X es ***convergente*** *si son convergentes* ***todas*** *las integrales impropias*

$$\int_{c^+}^{\beta} f(x)\,dx, \qquad \int_{\alpha}^{c^-} f(x)\,dx$$

resultantes de subdividir X, quitando los puntos impropios, en subintervalos semiabiertos del tipo $(c,\beta]$ y $[\alpha,c)$ donde en cada caso $c \in \mathcal{P}$ es el único punto impropio y está en un extremo.

En este caso, la integral impropia de f sobre X se define como la ***suma de los valores*** *de las integrales impropias en cada subintervalo semiabierto resultado de la subdivisión.*

Para ser convergente la integral impropia completa, deben serlo todas y cada una de las integrales impropias al subdividir. En cuanto una sola sea divergente, la integral total ya se define como divergente.

• **Ejemplo:** Usando la función del ejemplo anterior, $f(x) = 1/x(x-2)(x-4)(x-6)$,si queremos integrarla en $(0,\infty)$, podemos subdividir la integral en

$$\int_0^{\infty} f(x)\,dx = \int_{0^+}^{1} + \int_{1}^{2^-} + \int_{2^+}^{3} + \int_{3}^{4^-} + \int_{4^+}^{5} + \int_{5}^{6^-} + \int_{6^+}^{7} + \int_{7}^{\infty} f(x)\,dx.$$

y mirar a ver si cada trozo es convergente.

• **Ejemplo:** Para $f(x) = 1/(x\log x)$ con $a = 0, b = \infty$, los puntos impropios son $0, 1, \infty$, luego podemos subdividir

$$\int_0^{\infty} \frac{dx}{x\log x} = \int_{0^+}^{1/2} + \int_{1/2}^{1^-} + \int_{1^+}^{e} + \int_{e}^{\infty} \frac{dx}{x\log x}.$$

Independencia de la integral respecto a la subdivisión

El modo en que se subdivide el intervalo de integración en semiabiertos no afecta a la convergencia o divergencia de la integral, ni a su valor en caso de que sea convergente.

Demostración. Es una consecuencia de la aditividad de la integral y las propiedades de los límites. La observación clave es que, para $x \in [a,c)$, si cambiamos el punto base a por $\alpha \in [a,c)$, se tiene $\int_a^x f = \int_a^\alpha f + \int_\alpha^x f$. El término $\int_a^\alpha f$ es constante, por tanto existe $\int_a^{c^-} f = \lim_{x\to c^-} \int_a^x f$ si y sólo si existe $\int_\alpha^{c^-} f = \lim_{x\to c^-} \int_\alpha^x f$, siendo entonces $\int_a^{c^-} f = \int_a^\alpha f + \int_\alpha^{c^-} f$. Similarmente, $\int_{c^+}^a f = \int_{c^+}^\alpha f + \int_\alpha^a f$ si $\alpha \in (c,a]$.

Ahora, si c_k, c_{k+1} son puntos impropios consecutivos e intercalamos dos puntos distintos $a, \alpha \in (c_k, c_{k+1})$, siendo $a < \alpha$, se deduce de lo anterior que existen $\int_{c_k^+}^{a} f$ y $\int_{a}^{c_{k+1}^-} f$ si y sólo si existen $\int_{c_k^+}^{\alpha} f$ y $\int_{\alpha}^{c_{k+1}^-} f$, teniéndose entonces

$$\int_{c_k^+}^{a} f + \int_{a}^{c_{k+1}^-} f = \int_{c_k^+}^{\alpha} f - \int_{a}^{\alpha} f + \int_{a}^{\alpha} f + \int_{\alpha}^{c_{k+1}^-} f = \int_{c_k^+}^{\alpha} f + \int_{\alpha}^{c_{k+1}^-} f.$$

Es decir, tanto la convergencia de cada sumando como el valor de la suma no dependen del punto intercalado entre c_k y c_{k+1}.

Los detalles que faltan para el caso general con puntos impropios $c_1 < c_2 < \ldots < c_m$ (con extremos posiblemente infinitos) se dejan para el lector. □

6.2. Técnicas de cálculo de integrales impropias

Por la manera en la que se define la convergencia, hemos visto que para estudiar las integrales impropias en general basta con saber hacerlo en el caso de un intervalo semiabierto.

Integrales parciales

Sea $f(x)$ continua en el intervalo semiabierto $[a, b)$. La ***integral parcial*** *es la función*

$$F(x) = \int_{a}^{x} f(t)\,dt, \qquad x \in [a, b).$$

Por definición, la convergencia de la integral impropia de f sobre $[a, b)$ equivale a la existencia del límite

$$\lim_{x \to b^-} F(x)$$

y podemos escribir

$$\int_{a}^{b^-} f(x)\,dx = \lim_{x \to b^-} F(x) \stackrel{\text{def}}{=} F(b^-),$$

donde la notación $F(b^-)$ es la habitual para indicar un límite lateral. (ver 2.1.1). Recordamos que se escribe $F(a^+)$ para un límite por la derecha.

Por el Teorema Fundamental del Cálculo, la integral parcial es una antiderivada de f en $[a, b)$, es decir,

- F es continua en $[a, b)$
- F es derivable en (a, b) con $F' = f$

y también por el Teorema Fundamental, cualquier antiderivada sirve tanto para determinar la convergencia como para calcular una integral impropia.

Regla de Barrow para integrales impropias

Sea f continua en $[a,b)$, con antiderivada F en este intervalo. La integral impropia es convergente si y sólo si existe $F(b^-)$, siendo entonces

$$\int_a^{b^-} f(x)\,dx = F(b^-) - F(a).$$

De manera similar, para f continua en $(a,b]$ con antiderivada F, la integral impropia es convergente si y sólo si existe $F(a^+)$, siendo entonces

$$\int_{a^+}^{b} f(x)\,dx = F(b) - F(a^+).$$

Demostración. Sea $x \in [a,b)$. Por la Regla de Barrow usual aplicada a f en el intervalo $[a,x]$, donde es continua y con antiderivada F, se tiene

$$\int_a^x f(t)\,dt = F(x) - F(a) \quad \therefore \quad \exists \lim_{x\to b^-} \int_a^x f(t)\,dt \iff \exists \lim_{x\to b^-} (F(x) - F(a))$$
$$\iff \exists \lim_{x\to b^-} F(x)$$

ya que $F(a)$ es una constante. En caso de converger, entonces se tiene

$$\int_a^{b^-} f(x)\,dx = \lim_{x\to b^-} (F(x) - F(a)) = F(b^-) - F(a).$$

El caso semiabierto a la izquierda es análogo. □

Para una integral impropia donde ambos extremos son puntos impropios, se tiene

$$\int_{a^+}^{b^-} f(x)\,dx = F(b^-) - F(a^+).$$

Esto es consecuencia de las definiciones y la aditividad de la integral: intercalando un punto c en medio, queda

$$\begin{aligned}\int_{a^+}^{b^-} f(x)\,dx &= \int_{a^+}^{c} f(x)\,dx + \int_c^{b^-} f(x)\,dx \\ &= \big(F(b^-) - F(c)\big) + \big(F(c) - F(a^+)\big) = F(b^-) - F(a^+).\end{aligned}$$

La Regla de Barrow impropia significa que una integral impropia (en un semiabierto) se puede calcular como una integral “normal” reemplazando el **valor** de la antiderivada en los extremos con los **límites** (si existen) de la antiderivada en los extremos.

• Ejemplo:

$$\int_0^\infty \frac{dx}{1+x^2} = \lim_{R\to\infty} \int_0^R \frac{dx}{1+x^2} = \lim_{R\to\infty} [\arctan x]_0^R = \lim_{R\to\infty} \arctan R = \frac{\pi}{2}.$$

• Ejemplo:

$$\int_{0^+}^{1} \frac{dx}{\sqrt{x}} = \lim_{x\to 0^+} \int_{x}^{1} \frac{dt}{\sqrt{t}} = \lim_{x\to 0^+} \left[2\sqrt{t}\right]_x^1 = \lim_{x\to 0^+} \left(2 - 2\sqrt{x}\right) = 2.$$

En la práctica uno deja de escribir el límite y se usa una notación simbólica:

$$\int_{0^+}^{1} \frac{dx}{\sqrt{x}} = \left[2\sqrt{x}\right]_{0^+}^1 = 2.$$

evitando usar dos letras, una para la variable de integración y otra para el límite.

• Ejemplo:

$$\int_{0}^{\infty} e^{-x}\, dx = \left[-e^{-t}\right]_0^\infty = -e^{-\infty} + 1 = 1.$$

Cambio de variable en integrales impropias (caso creciente)

Supongamos que $a < b$ y $c < d$, donde posiblemente $b = \infty$ ó $d = \infty$. Sea

- $f : [c, d) \to \mathbb{R}$ *continua.*
- $\varphi : [a, b) \to [c, d)$ *creciente, con derivada continua en (a, b), tal que*
  - $\varphi(a) = c$
  - $\lim\limits_{x\to b^-} \varphi(x) = \varphi(b^-) = d.$

Entonces

$$\int_{c}^{d^-} f(y)\, dy = \int_{a}^{b^-} f(\varphi(x))\, \varphi'(x)\, dx$$

en el sentido que son simultáneamente convergentes al mismo valor, o divergentes.

Demostración. Las hipótesis implican que $\varphi(x)$ es invertible, con inversa creciente $\varphi^{-1} : [c, d) \to [a, b)$ y con derivada continua en (c, d). Para $t \in [a, b)$, φ aplica el intervalo $[a, t]$ en $[c, \varphi(t)]$. Aplicando el Teorema de cambio de variable usual,

$$\int_{a}^{t} f(\varphi(x))\, \varphi'(x)\, dx = \int_{c}^{\varphi(t)} f(y)\, dy$$

y haciendo $t \to b^-$ queda el resultado. □

Dado que $\varphi(x)$ crece a d cuando x crece a b, se puede escribir $\varphi(b^-) = d^-$ y aplicar el resultado como si fuera una integral normal:

$$\int_{c}^{d^-} f(y)\, dy = \int_{\varphi(a)}^{\varphi(b^-)} f(y)\, dy = \int_{a}^{b^-} f(\varphi(x))\, \varphi'(x)\, dx.$$

• **Ejemplo:** Haciendo $y = x^2$

$$\int_0^{1^-} \frac{x\,dx}{\sqrt{1-x^2}} = \frac{1}{2}\int_0^{1^-} \frac{dy}{\sqrt{1-y}} = \Big[-\sqrt{1-y}\Big]_0^{1^-} = 1.$$

• **Ejemplo:** Haciendo $u = \log x$,

$$\int_e^{\infty} \frac{dx}{x\log x} = \int_1^{\infty} \frac{du}{u} = [\log u]_1^{\infty} = \infty.$$

• **Ejemplo:** Haciendo $x = \sin\theta$

$$\int_0^{1^-} \frac{dx}{\sqrt{1-x^2}} = \int_0^{(\pi/2)^-} \frac{\cos\theta\,d\theta}{\sqrt{1-\sin^2\theta}} = \int_0^{(\pi/2)^-} \frac{\cos\theta\,d\theta}{\sqrt{\cos^2\theta}} = \int_0^{\pi/2} d\theta = \frac{\pi}{2}.$$

Como puede comprobarse en el ejemplo anterior, a veces, al cambiar de variable y simplificar, lo que era una integral impropia se convierte en una “normal”.

Cambio de variable en integrales impropias (caso decreciente)

Supongamos que $a < b$ y $c < d$, donde posiblemente $b = \infty$ ó $c = -\infty$. Sea

- $f : (c, d] \to \mathbb{R}$ *continua.*
- $\varphi : [a, b) \to (c, d]$ *decreciente, con derivada continua en (a, b), tal que*
  - $\varphi(a) = d$
  - $\lim_{x\to b^-} \varphi(x) = c$.

Entonces

$$\int_{c^+}^{d} f(y)\,dy = -\int_a^{b^-} f(\varphi(x))\,\varphi'(x)\,dx$$

en el sentido que son simultáneamente convergentes al mismo valor, o divergentes.

Dado que $\varphi(x)$ decrece á c cuando x crece a b, se puede escribir $\varphi(b^-) = c^+$ y aplicar el resultado como si fuera una integral normal:

$$\int_{c^+}^{d} f(y)\,dy = \int_{\varphi(b^-)}^{\varphi(a)} f(y)\,dy = \int_{b^-}^{a} f(\varphi(x))\,\varphi'(x)\,dx = -\int_a^{b^-} f(\varphi(x))\,\varphi'(x)\,dx.$$

El hecho que una φ decreciente invierte el orden habitual, que es de menor a mayor, explica el signo en este caso: para “enderezarlos” hace falta cambiar de signo. De todas formas los pasos siguen siendo los mismos que para una integral no impropia.

Se pueden unificar los dos casos poniendo la derivada en valor absoluto:

$$\int_c^{d^-} f(y)\,dy = \int_a^{b^-} f(\varphi(x))\cdot|\varphi'(x)|\,dx \qquad \int_{c^+}^{d} f(y)\,dy = \int_a^{b^-} f(\varphi(x))\cdot|\varphi'(x)|\,dx$$

siempre son válidas, para φ creciente o decreciente respectivamente.

• **Ejemplo:** Haciendo $y = 1 - x$,

$$\int_0^{1^-} \frac{dy}{\sqrt{1-y}} = \int_1^{0^+} \frac{-dx}{\sqrt{x}} = \int_{0^+}^{1} \frac{dx}{\sqrt{x}} = 2.$$

• **Ejemplo:** Haciendo $u = -\log x$, $x = e^{-u}$

$$\int_0^1 -\sin(\log x)\, dx = \int_\infty^0 -e^{-u}\sin(u)\, du = \int_0^\infty e^{-u}\sin(u)\, du.$$

Integración por partes en integrales impropias

La fórmula de integración por partes sigue siendo válida,

$$\int_a^{b^-} u\, dv = [uv]_a^{b^-} - \int_a^{b^-} v\, du$$

en el sentido de que los tres términos representan límites, y la existencia de dos de ellos implica la del tercero y entonces cumplen la igualdad indicada.

• **Ejemplo:**

$$\int_0^\infty \underbrace{e^{-x}}_{u} \underbrace{\sin x\, dx}_{dv} = -\left[e^{-x}\cos x\right]_0^\infty - \int_0^\infty e^{-x}\cos x\, dx = 1 - \int_0^\infty e^{-x}\cos x\, dx.$$

6.2.1. Las potencias impropias

Potencias en ∞

Dado $s \in \mathbb{R}$,

$$\int_1^\infty \frac{dx}{x^s} = \begin{cases} \dfrac{1}{s-1} & \text{si } s > 1 \\ \infty & \text{si } s \le 1. \end{cases}$$

En particular, una integral $\int_c^\infty \frac{dx}{x^s}$ con $c > 0$ es convergente si y sólo si $s > 1$.

Demostración. Para $s \neq 1$,

$$\int_1^\infty \frac{dx}{x^s} = \left[\frac{x^{-s+1}}{-s+1}\right]_1^\infty = \frac{\infty^{-s+1}}{-s+1} - \frac{1}{-s+1} = \frac{1}{s-1} + \frac{\infty^{-s+1}}{-s+1}$$

y el límite en ∞ existe si y sólo si $-s+1 < 0$, es decir, $s > 1$, en cuyo caso $\infty^{-s+1} = 0$.

Cuando $s = 1$,

$$\int_1^\infty \frac{dx}{x} = [\log x]_1^\infty = \log\infty = \infty.$$

□

Potencias en 0

Dado $s \in \mathbb{R}$,

$$\int_0^1 \frac{dx}{x^s} = \begin{cases} \dfrac{1}{1-s} & \text{si } s < 1 \\ \infty & \text{si } s \geq 1. \end{cases}$$

En particular, una integral $\int_0^c \frac{dx}{x^s}$ *con* $c > 0$ *es convergente si y sólo si* $s < 1$.

Demostración. Se puede ver haciendo el cambio de variable $x = u^{-1}$ y aplicando el resultado en ∞:

$$\int_0^1 \frac{dx}{x^s} = \int_\infty^1 \frac{-u^{-2}\,du}{u^{-s}} = \int_1^\infty \frac{du}{u^{2-s}} = \begin{cases} \dfrac{1}{(2-s)-1} = \dfrac{1}{1-s} & \text{si } 2-s > 1, \text{ o sea } s < 1 \\ \infty & \text{si } 2-s \leq 1, \text{ o sea } s \geq 1. \end{cases}$$

□

Potencias en un punto finito $c \in \mathbb{R}$

Dado $s \in \mathbb{R}$ *y* $-\infty < a < c < b < \infty$, *las integrales*

$$\int_c^b \frac{dx}{(x-c)^s}, \qquad \int_a^c \frac{dx}{(c-x)^s}$$

son convergentes si y sólo si $s < 1$.

Demostración. Hacer los cambios de variable $u = x - c$ y $u = c - x$ respectivamente, aplicando el resultado para el caso $c = 0$. □

Estos resultados sobre potencias son muy útiles para los métodos de comparación que comentaremos más adelante.

6.3. Otros tipos de convergencia para integrales impropias

Sea f continua en $[a, b]$ salvo por el punto impropio $c \in (a, b)$. Según la definición que hemos dado de convergencia para una integral impropia, a la que nos referiremos como estricta, debemos separar el intervalo $[a, b]$ en dos semiabiertos $[a, c)$ y $(c, b]$ y considerar por separado

$$\int_a^b f(t)\,dt = \int_a^{c^-} f(t)\,dt + \int_{c^+}^b f(t)\,dt = \lim_{x \to c^-} \int_a^x f(t)\,dt + \lim_{y \to c^+} \int_y^b f(t)\,dt.$$

La variables x, y en estos límites son independientes. Es decir, no se impone relación alguna entre cómo se acerca la variable x al punto c y cómo lo hace la variable y.

Si existe cada límite con $x \to c^-$, $y \to c^+$ de modo independiente, entonces también existen y coinciden los valores de los límites cuando las dos variables x, y están ligadas, por ejemplo, de modo simétrico poniendo $x = c - \varepsilon$, $y = c + \varepsilon$ con $\varepsilon \to 0^+$.

La condición de existencia de un límite con una ligadura es más débil que la condición estricta de independencia. Por ejemplo, puede que exista el límite simétrico

$$\lim_{\varepsilon\to 0^+}\left(\int_a^{c-\varepsilon}+\int_{c+\varepsilon}^b\right)f(t)\,dt$$

sin que existan por separado los dos límites

$$\lim_{x\to c^-}\int_a^x f(t)\,dt,\qquad \lim_{y\to c^+}\int_y^b f(t)\,dt,$$

que a menudo serán ambos infinitos.

Este tipo de límite simétrico en una integral impropia recibe el nombre de **valor principal** (de Cauchy). Se trata de una definición alternativa capaz de asignar valores a más integrales impropias que la definición estricta. Para distinguirlo de la noción más estricta de convergencia, se suele indicar con un símbolo como v.p. antes de la integral.

Valor principal en un intervalo acotado con un punto impropio interior

*Sean $a,b\in\mathbb{R}$, f continua en $[a,b]$ salvo por el punto impropio $c\in(a,b)$. El **valor principal** de la integral $\int_a^b f$ es el límite simétrico (si existe)*

$$\text{v.p.}\int_a^b f\overset{\text{def}}{=}\lim_{\varepsilon\to 0^+}\left(\int_a^{c-\varepsilon}+\int_{c+\varepsilon}^b\right)f$$

• Ejemplo:

$$\text{v.p.}\int_{-1}^1\frac{dt}{t}=\lim_{\varepsilon\to 0^+}\left(\int_{-1}^{-\varepsilon}+\int_\varepsilon^1\right)\frac{dt}{t}=0$$

pues para todo $\varepsilon>0$,

$$\begin{aligned}\left(\int_{-1}^{-\varepsilon}+\int_\varepsilon^1\right)\frac{dt}{t}&=\left(\Big[\log|t|\Big]_{-1}^{-\varepsilon}+\Big[\log|t|\Big]_\varepsilon^1\right)\\&=(\log\varepsilon-\log 1+\log 1-\log\varepsilon)=0\end{aligned}$$

pero por separado,

$$\lim_{x\to 0^-}\int_{-1}^x\frac{dt}{t}=\lim_{x\to 0^-}\log|x|=-\infty,\qquad \lim_{y\to 0^+}\int_y^1\frac{dt}{t}=\lim_{y\to 0^+}-\log|y|=\infty$$

y por tanto la integral es divergente en sentido estricto.

La convergencia de una integral impropia en sentido estricto implica la convergencia en el sentido de valor principal, asignándoles a ambas integrales el mismo valor, pero la existencia del valor principal no implica la convergencia de la integral impropia; es una condición más débil.

El procedimiento de análisis de una integral de valor principal cuando hay más de un punto impropio se define por subdivisión, al igual que en una integral impropia general, pero en el caso de un valor principal el punto impropio no se coloca en un extremo de un intervalo semiabierto sino dentro de un intervalo donde es el único punto impropio, para poder aplicar la definición anterior.

Valor principal en un intervalo acotado con varios puntos impropios interiores

Sea f continua en el intervalo $[a, b]$ excepto en un número finito de puntos interiores,

$$a < c_1 < c_2 < \cdots < c_m < b.$$

El valor principal de la integral $\int_a^b f$ se define intercalando puntos

$$a = a_0 < c_1 < a_1 < c_2 < a_2 < \cdots < a_{m-1} < c_m < a_m = b$$

y considerando el valor principal aislado en cada uno de los m subintervalos $[a_{k-1}, a_k]$:

$$\text{v.p.} \int_{a_{k-1}}^{a_k} f = \lim_{\varepsilon \to 0^+} \left(\int_{a_{k-1}}^{c_k - \varepsilon} + \int_{c_k + \varepsilon}^{a_k} \right) f.$$

Si todos convergen, decimos que v.p. $\int_a^b f$ converge, y definimos su valor como la suma

$$\text{v.p.} \int_a^b f \stackrel{\text{def}}{=} \text{v.p.} \int_{a_0}^{a_1} f + \text{v.p.} \int_{a_1}^{a_2} f + \cdots + \text{v.p.} \int_{a_{m-1}}^{a_m} f.$$

Se puede comprobar que tanto la existencia de los valores principales como el valor de la suma no dependen de la elección de puntos intercalados.

Si alguno de los sumandos no converge, el valor principal total es divergente.

Para un valor principal en un intervalo acotado, se necesitan tantos subintervalos en la subdivisión como puntos impropios.

• **Ejemplo:** Consideremos la integral de valor principal

$$\text{v.p.} \int_{-2}^{2} \frac{dx}{x(1-x)}.$$

El integrando tiene puntos impropios en 0 y 1 que podemos aislar intercalando $\frac{1}{2}$. Sólo hacen falta dos intervalos, $[-2, \frac{1}{2}]$, que aísla al punto impropio en 0, y $[\frac{1}{2}, 2]$, que aísla al punto impropio en 1. Teniendo en cuenta que

$$\int \frac{dx}{x(1-x)} = \int \left(\frac{1}{x} + \frac{1}{1-x} \right) dx = \log|x| - \log|1-x|$$

calculamos primero (suponiendo $0 < \varepsilon < 1$)

$$\begin{aligned}
&\text{v.p.} \int_{-2}^{\frac{1}{2}} \frac{dx}{x(1-x)} \\
&= \lim_{\varepsilon \to 0^+} \left(\int_{-2}^{-\varepsilon} + \int_{\varepsilon}^{\frac{1}{2}} \right) \frac{dx}{x(1-x)} \\
&= \lim_{\varepsilon \to 0^+} \left(\Big[\log|x| - \log|1-x| \Big]_{-2}^{-\varepsilon} + \Big[\log|x| - \log|1-x| \Big]_{\varepsilon}^{\frac{1}{2}} \right) \\
&= \lim_{\varepsilon \to 0^+} \left(\log\varepsilon - \log|1+\varepsilon| - \log 2 + \log 3 + \log \tfrac{1}{2} - \log \tfrac{1}{2} - \log\varepsilon + \log|1-\varepsilon| \right) \\
&= \lim_{\varepsilon \to 0^+} \left(\log \frac{1-\varepsilon}{1+\varepsilon} - \log 2 + \log 3 \right) \\
&= \log 3 - \log 2.
\end{aligned}$$

Para la segunda parte,

$$\begin{aligned}
&\text{v.p.} \int_{\frac{1}{2}}^{2} \frac{dx}{x(1-x)} \\
&= \lim_{\varepsilon \to 0^+} \left(\int_{\frac{1}{2}}^{1-\varepsilon} + \int_{1+\varepsilon}^{2} \right) \frac{dx}{x(1-x)} \\
&= \lim_{\varepsilon \to 0^+} \left(\Big[\log|x| - \log|1-x| \Big]_{\frac{1}{2}}^{1-\varepsilon} + \Big[\log|x| - \log|1-x| \Big]_{1+\varepsilon}^{2} \right) \\
&= \lim_{\varepsilon \to 0^+} \left(\log|1-\varepsilon| - \log\varepsilon - \log \tfrac{1}{2} + \log \tfrac{1}{2} + \log 2 - \log 1 - \log|1+\varepsilon| + \log\varepsilon \right) \\
&= \lim_{\varepsilon \to 0^+} \left(\log \frac{1-\varepsilon}{1+\varepsilon} + \log 2 \right) \\
&= \log 2.
\end{aligned}$$

Como cada valor principal aislado es convergente, el valor principal total también lo es, y

$$\text{v.p.} \int_{-2}^{2} \frac{dx}{x(1-x)} = \text{v.p.} \int_{-2}^{\frac{1}{2}} \frac{dx}{x(1-x)} + \int_{\frac{1}{2}}^{2} \frac{dx}{x(1-x)} = \log 3 - \log 2 + \log 2 = \log 3.$$

También en este caso, la integral es divergente en sentido estricto. Para verlo hay que analizarla siguiendo el procedimiento correspondiente a la subdivisión en intervalos semiabiertos, por ejemplo $[-2, 0)$, $(0, \frac{1}{2}]$, $[\frac{1}{2}, 1)$ y $(1, 2]$. Entonces

$$\int_{1^+}^{2} \frac{dx}{x(1-x)} = \Big[\log|x| - \log|1-x| \Big]_{1^+}^{2} = \log 2 - \log 1 - \log 1^+ + \log 0^+ = -\infty.$$

De manera similar se puede ver que cada uno de los otros trozos diverge a $\pm\infty$:

$$\int_{-2}^{0^-} \frac{dx}{x(1-x)} = -\infty, \qquad \int_{0^+}^{\frac{1}{2}} \frac{dx}{x(1-x)} = \int_{\frac{1}{2}}^{1^-} \frac{dx}{x(1-x)} = \infty.$$

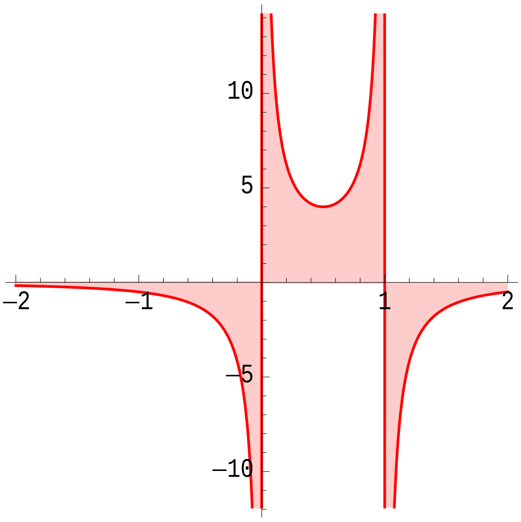

Figura 29 – Gráfica de $\dfrac{1}{x(1-x)}$ entre $x=-2$ y $x=2$.

Si f es continua en $\mathbb{R}$, para estudiar su integral impropia sobre $(-\infty,\infty)$, dividimos en dos partes, por ejemplo

$$\int_{-\infty}^{0} f, \qquad \int_{0}^{\infty} f$$

y exigimos la convergencia de ambas, lo cual equivale a la existencia de los límites

$$\lim_{S\to-\infty}\int_{S}^{0} f, \qquad \lim_{R\to\infty}\int_{0}^{R} f,$$

donde $S\to-\infty$ y $R\to\infty$ de manera independiente. Para definir valores principales en $(-\infty,\infty)$, tenemos en cuenta la simetría, ligando las dos variables poniendo $S=-R$.

Valor principal sobre $(-\infty,\infty)$ sin otros puntos impropios

Si f es continua en $\mathbb{R}$, el valor principal de $\int_{-\infty}^{\infty} f$ es el del límite simétrico (si existe)

$$\text{v.p.}\int_{-\infty}^{\infty} f(x)\,dx \overset{\text{def}}{=} \lim_{R\to\infty}\int_{-R}^{R} f(x)\,dx$$

• Ejemplo:

$$\text{v.p.}\int_{-\infty}^{+\infty} x\,dx = 0 \quad \text{ya que} \quad \int_{-R}^{R} x\,dx = \left[\frac{x^2}{2}\right]_{-R}^{R} = 0 \quad \forall R>0,$$

pero la integral impropia $\displaystyle\int_{-\infty}^{+\infty} x\,dx$ no converge en sentido estricto, pues

$$\int_{0}^{\infty} x\,dx = \infty, \qquad \int_{-\infty}^{0} x\,dx = -\infty.$$

Valor principal de una función impar sobre $\mathbb{R}$

*Para cualquier función continua $f(x)$ **impar** en $\mathbb{R}$ se tiene*

$$\text{v.p.} \int_{-\infty}^{\infty} f(x)\,dx = 0$$

Demostración. Para cualquier $R > 0$, se tiene

$$\int_{-R}^{R} f(x)\,dx = \left(\int_{-R}^{0} + \int_{0}^{R}\right) f(x)\,dx = 0,$$

por simetría alrededor de $x = 0$, es decir, cambiando de variable $x \to -x$:

$$\int_{-R}^{0} f(x)\,dx = \int_{R}^{0} f(-x)(-dx) = \int_{0}^{R} f(-x)\,dx = \int_{0}^{R} -f(x)\,dx = -\int_{0}^{R} f(x)\,dx.$$

$\square$

alternativa. Si f es impar, cualquier antiderivada suya es par. Efectivamente, sea F una antiderivada de f y consideremos $G(x) = F(x) - F(-x)$. Se tiene $G'(x) = F'(x) + F'(-x) = f(x) + f(-x) = f(x) - f(x) = 0$, luego G es constante. Como $G(0) = F(0) - F(0) = 0$, debe ser $G(x) = 0$ para todo x, es decir, $F(-x) = F(x)$. Entonces

$$\int_{-R}^{R} f(x)\,dx = F(R) - F(-R) = 0$$

para todo $R > 0$. $\square$

• **Ejemplo:** Consideremos la integral de valor principal v.p. $\displaystyle\int_{-\infty}^{\infty} \frac{2x+1}{x^2+1}\,dx$. Como

$$\int \frac{2x+1}{x^2+1}\,dx = \int \frac{2x}{x^2+1}\,dx + \int \frac{dx}{x^2+1} = \log(x^2+1) + \arctan x$$

tenemos

$$\begin{aligned} \text{v.p.} \int_{-\infty}^{\infty} \frac{2x+1}{x^2+1}\,dx &= \lim_{R\to\infty} \Big[\log(x^2+1) + \arctan x\Big]_{-R}^{R} \\ &= \lim_{R\to\infty} 2\arctan(R) = 2 \cdot \frac{\pi}{2} = \pi. \end{aligned}$$

La integral es divergente en sentido estricto, pues por ejemplo

$$\int_{0}^{\infty} \frac{2x+1}{x^2+1}\,dx = \Big[\log(x^2+1) + \arctan x\Big]_{0}^{\infty} = \infty + \frac{\pi}{2} - \log 1 - \arctan 0 = \infty.$$

Si ahora añadimos otros puntos impropios aparte de $\pm\infty$, la definición de valor principal se hace juntando todos los conceptos de las definiciones en los casos anteriores.

Valor principal sobre $(-\infty, \infty)$ con otros puntos impropios finitos

Sea f continua en $\mathbb{R}$ excepto en un número finito de puntos,

$$c_1 < c_2 < \cdots < c_m.$$

El valor principal de la integral $\int_a^b f$ se define intercalando puntos

$$-\infty < a_0 < c_1 < a_1 < c_2 < a_2 < \cdots < a_{m-1} < c_m < a_m < \infty$$

y considerando el valor principal aislado en cada uno de los m subintervalos $[a_{k-1}, a_k]$:

$$\text{v.p.} \int_{a_{k-1}}^{a_k} f = \lim_{\varepsilon \to 0^+} \left(\int_{a_{k-1}}^{c_k - \varepsilon} + \int_{c_k + \varepsilon}^{a_k} \right) f$$

junto con el límite simétrico en infinito

$$\lim_{R \to \infty} \left(\int_{-R}^{a_0} + \int_{a_m}^{R} \right) f.$$

Si todos convergen, decimos que v.p. $\int_{-\infty}^{\infty} f$ converge, y definimos su valor como la suma

$$\text{v.p.} \int_{-\infty}^{\infty} f \overset{\text{def}}{=} \text{v.p.} \int_{a_0}^{a_1} f + \text{v.p.} \int_{a_1}^{a_2} f + \cdots + \text{v.p.} \int_{a_{m-1}}^{a_m} f + \lim_{R \to \infty} \left(\int_{-R}^{a_0} + \int_{a_m}^{R} \right) f.$$

Se puede comprobar que tanto la existencia de los valores principales como el valor de la suma no dependen de la elección de puntos intercalados.

• **Ejemplo:** Consideremos la integral de valor principal

$$\text{v.p.} \int_{-\infty}^{\infty} \frac{dx}{x}.$$

Tenemos un punto impropio finito en 0. Podemos intercalar, por ejemplo, -1 y 1. Entonces debemos considerar por separado

$$\text{v.p.} \int_{-1}^{1} \frac{dx}{x} = \lim_{\varepsilon \to 0^+} \left(\int_{-1}^{-\varepsilon} + \int_{\varepsilon}^{1} \right) \frac{dx}{x} = \lim_{\varepsilon \to 0^+} 0 = 0$$

ya que la función es impar. Similarmente, el límite simétrico en infinito también es nulo:

$$\lim_{R \to \infty} \left(\int_{-R}^{-1} + \int_{1}^{R} \right) \frac{dx}{x} = \lim_{R \to \infty} 0 = 0.$$

Por tanto el valor principal es 0.

Cuando sólo hay un único punto impropio finito c, la existencia del valor principal equivale a la del límite doble

$$\lim_{\substack{\varepsilon \to 0^+ \\ R \to \infty}} \left(\int_{-R}^{c - \varepsilon} + \int_{c + \varepsilon}^{R} \right) f.$$

aunque esto no es tan fácil de demostrar sin la noción de límite superior.

• Ejemplo:

$$\text{v.p.} \int_{-\infty}^{\infty} \frac{\cos x}{x}\, dx = \lim_{\substack{\varepsilon \to 0^+ \\ R \to \infty}} \left(\int_{-R}^{-\varepsilon} + \int_{\varepsilon}^{R} \right) \frac{\cos x}{x}\, dx = 0$$

siendo $\pm\infty, 0$ los puntos impropios, pues el integrando es impar.

En el caso anterior de un valor principal sobre $(-\infty, \infty)$ con un punto impropio finito adicional, a veces se emplea una definición distinta:

$$\lim_{\varepsilon \to 0^+} \left(\int_{c-\varepsilon^{-1}}^{c-\varepsilon} + \int_{c+\varepsilon}^{c+\varepsilon^{-1}} \right) f,$$

donde se han ligado los límites de integración de una manera diferente a las que hemos dado. Dichas expresiones suelen provenir de la integración compleja sobre contornos.

Suele haber varias nociones razonables sobre cuál debe ser el valor principal en una situación dada. A menudo la elección depende de la interpretación que se quiera dar a la integral.

Hay aún más situaciones aparte de las que ya hemos listado en las que el concepto de valor principal puede tener sentido. Por ejemplo, si tenemos una función continua en un intervalo acotado, cuyos extremos son ambos puntos impropios, se emplea la siguiente definición de valor principal, siempre basado en la idea de aproximación simétrica.

Valor principal simétrico en un intervalo acotado con extremos impropios

Sea f continua en (a, b) y a, b ambos puntos impropios finitos. Definimos

$$\text{v.p.} \int_a^b f(x)\, dx = \lim_{\varepsilon \to 0^+} \int_{a+\varepsilon}^{b-\varepsilon} f(x)\, dx.$$

• Ejemplo:

$$\text{v.p.} \int_{-1}^{1} \frac{2t}{1-t^2}\, dt = 0 \quad \text{aunque} \quad \int_{-1}^{1} \frac{2t}{1-t^2}\, dt \quad \text{es divergente en sentido estricto.}$$

La función es impar, por tanto la integral es nula sobre cualquier intervalo simétrico $(-R, R)$, y en concreto sobre $(-1+\varepsilon, 1-\varepsilon)$; por tanto el valor principal bajo consideración es nulo.

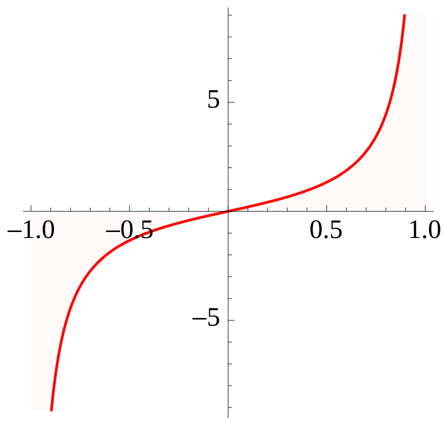

Figura 30 – Gráfica de $\dfrac{2t}{1-t^2}$ para $t \in (-1, 1)$.

Si observamos que $f(t)$ tiene antiderivada $F(t) = -\log(1-t^2)$ en $(-1,1)$ y separamos la integral en $t = 0$, vemos que cada parte es divergente:

$$\lim_{y\to 1^-} \int_0^y \frac{2t}{1-t^2}\,dt = \lim_{y\to 1^-} -\log(1-y^2) = \infty,$$

$$\lim_{x\to -1^+} \int_x^0 \frac{2t}{1-t^2}\,dt = \lim_{x\to -1^+} \log(1-x^2) = -\infty.$$

Otra manera de entender la divergencia es considerar simultáneamente

$$\int_x^y \frac{2t}{1-t^2}\,dt = F(y) - F(x) = -\log(1-y^2) + \log(1-x^2) = \log\frac{1-x^2}{1-y^2}.$$

Para cualquier $\lambda > 0$ fijo se pueden encontrar x, y arbitrariamente cerca de -1^+ y 1^- respectivamente, tales que

$$\frac{1-x^2}{1-y^2} = \lambda.$$

Esto se puede ver poniendo $x = -1+\delta$ con $\delta \approx 0^+$ y resolviendo la ecuación para $y \approx 1^-$:

$$\frac{1-(-1+\delta)^2}{1-y^2} = \lambda \iff \frac{2\delta-\delta^2}{\lambda} = 1-y^2 \iff y = \sqrt{1-\frac{\delta(2-\delta)}{\lambda}}.$$

El límite de la expresión resultante para y cuando $\delta \to 0^+$ es 1, luego escogiendo $\delta > 0$ suficientemente pequeño se obtiene x arbitrariamente cerca de -1^+ é y arbitrariamente cerca de 1^-. Esto significa que dado cualquier λ, existen dos sucesiones x_n, y_n tales que

$$x_n \to -1^+, y_n \to 1^-, \qquad \int_{x_n}^{y_n} \frac{2t}{1-t^2}\,dt = F(y_n) - F(x_n) = \log\lambda \quad \forall n.$$

Es decir, no sólo no existe el límite doble

$$\lim_{\substack{x\to -1^+\\ y\to 1^-}} \int_x^y \frac{2t}{1-t^2}\,dt$$

sino que ligando las variables adecuadamente se puede conseguir que la integral asuma cualquier valor real $\log\lambda$ a medida que $x \to -1^-$, $y \to 1^-$. Haciéndolas simétricas $(x = -y)$ se consigue $\lambda = 1$, por tanto $\log\lambda = 0$, que es el valor principal.

Linealidad del valor principal

Con cualquiera de los distintos tipos de valor principal sobre un intervalo X, si existen los valores principales $\int_X f(x)\,dx$ y $\int_X g(x)\,dx$, entonces también existen los valores principales de las integrales de cualquier múltiplo kf y de la suma $f+g$, siendo

$$\text{v.p.}\int_X kf = k\cdot \text{v.p.}\int_X f, \qquad \text{v.p.}\int_X (f+g) = \text{v.p.}\int_X f + \text{v.p.}\int_X g.$$

Demostración. Cualquier noción de valor principal se define a partir de un límite, por tanto es una consecuencia de la linealidad de los límites. □

• **Ejemplo:** Calculamos la integral de valor principal

$$\text{v.p.}\int_{-2}^{2} \frac{dx}{x(1-x)}$$

que ya hemos estudiado en un ejemplo anterior, pero ahora aprovechando la linealidad:

$$\text{v.p.}\int_{-2}^{2} \frac{dx}{x(1-x)} = \text{v.p.}\int_{-2}^{2}\left(\frac{1}{x}+\frac{1}{1-x}\right)dx = \underbrace{\text{v.p.}\int_{-2}^{2}\frac{dx}{x}}_{\text{impar}} + \text{v.p.}\int_{-2}^{2}\frac{dx}{1-x}$$

$$= \text{v.p.}\int_{-2}^{2}\frac{dx}{1-x} = \lim_{\varepsilon\to 0^+}\left(\int_{-2}^{1-\varepsilon} + \int_{1+\varepsilon}^{2}\right)\frac{dx}{1-x}$$

$$= \lim_{\varepsilon\to 0^+}\Big[-\log|1-x|\Big]_{-2}^{1-\varepsilon} + \Big[-\log|1-x|\Big]_{1+\varepsilon}^{2}$$

$$= \lim_{\varepsilon\to 0^+}\left(-\log\varepsilon + \log 3 - \log 1 + \log\varepsilon\right) = \log 3.$$

No es convergente en sentido estricto pues, por ejemplo

$$\int_{-2}^{0^+}\frac{dx}{x(1-x)} = \left[\log|x| - \log|1-x|\right]_{-2}^{0^+} = -\infty.$$

Se debe prestar especial atención al siguiente fenómeno general: puede ser que exista

$$\text{v.p.}\int_X (f+g)$$

sin que existan por separado

$$\text{v.p.}\int_X f, \quad \text{v.p.}\int_X g.$$

Un caso frecuente es cuando dan una indeterminación de tipo $\infty-\infty$ al sumar.

• **Ejemplo:** Puede comprobarse que

$$\text{v.p.} \int_{-1/2}^{1/2} \frac{x^2}{(x+x^2)^3}\, dx = -\frac{28}{9} - \log 3$$

pero

$$\text{v.p.} \int_{-1/2}^{1/2} \frac{2x+1}{(x+x^2)^3}\, dx = -\infty, \qquad \text{v.p.} \int_{-1/2}^{1/2} \frac{x^2+2x+1}{(x+x^2)^3}\, dx = -\infty,$$

con lo cual la separación

$$\frac{x^2}{(x+x^2)^3} = \frac{x^2+2x+1}{(x+x^2)^3} - \frac{2x+1}{(x+x^2)^3}$$

no da una expresión válida al distribuir la integral de valor principal con la resta.

La siguiente observación puede resultar más sorprendente.

La técnica de cambio de variable no es válida en general para valores principales.

• **Ejemplo:** Se puede comprobar que

$$\text{v.p.} \int_{-\frac{1}{2}}^{\frac{\sqrt{2}-1}{2}} \frac{2x+1}{(x+x^2)^3}\, dx = \lim_{\varepsilon\to 0^+} \left(\int_{-\frac{1}{2}}^{-\varepsilon} + \int_{\varepsilon}^{\frac{\sqrt{2}-1}{2}} \right) \frac{2x+1}{(x+x^2)^3}\, dx. = -\infty$$

Como el numerador es igual a la derivada de la expresión en paréntesis en el denominador, parece adecuado hacer el cambio de variable $u = x + x^2$, quedando

$$\int \frac{2x+1}{(x+x^2)^3} = \int \frac{du}{u^3} = -\frac{1}{2u^2} = -\frac{1}{2(x+x^2)^2}$$

que, además de darnos la antiderivada sencilla, si tenemos en cuenta los límites de integración, vemos que pasan a ser $\pm\frac{1}{4}$ en u. Entonces quedaría la integral de valor principal

$$\text{v.p.} \int_{-1/4}^{1/4} \frac{du}{u^3} = \lim_{\varepsilon\to 0^+} \left(\int_{-1/4}^{-\varepsilon} + \int_{\varepsilon}^{1/4} \right) \frac{du}{u^3} = 0$$

al ser la función impar y la región de integración simétrica respecto al origen.

El resultado ha cambiado de $-\infty$ á 0, aunque se puede comprobar que la sustitución $u = x + x^2$ cumple todas las hipótesis del cambio de variable: es invertible en el intervalo señalado y tanto ella como su inversa tienen derivadas continuas.

El fenómeno que se ha producido es que el cambio de variable puede romper la simetría. El valor principal en x, con punto impropio en $x = 0$, requiere aproximarse desde fuera de un intervalo simétrico $[-\varepsilon, \varepsilon]$ alrededor de 0. Bajo el cambio $u = x + x^2$ este intervalo se transforma en el intervalo $[-\varepsilon + \varepsilon^2, \varepsilon + \varepsilon^2]$, que contiene al 0 si $0 < \varepsilon < 1$, pero no es simétrico. El límite simétrico en u no corresponde al que proviene del límite simétrico en x.

6.4. Criterios de convergencia

6.4.1. Criterios básicos

A veces sólo interesa conocer si una integral impropia es convergente o divergente, bien porque el valor no sea relevante, bien porque sea muy complicado de calcular. Para determinar esto sin necesidad de calcular, existen una serie de criterios.

El criterio de Cauchy

Sea f continua en $[a,b)$. Entonces $\displaystyle\int_a^{b^-} f(x)\,dx$ *es convergente si y sólo si*

$$\lim_{x,y\to b^-}\int_x^y f = 0.$$

Demostración. Se aplica el criterio de Cauchy para límites,

$$\exists \lim_{x\to b^-} F(x) \iff \lim_{x,y\to b^-}|F(y)-F(x)| = \lim_{x,y\to b^-}\left|\int_x^y f\right| = 0$$

□

Intervalo y función acotados

Dados

- $[a,b)$ *un intervalo semiabierto acotado.*
- f *continua en* $[a,b)$ *y acotada:*

$$|f(x)|\le M \quad \forall\ x\in[a,b).$$

  Esto se cumple, por ejemplo, si existe $\lim_{x\to b^-} f(x)$.

Entonces existe $\displaystyle\int_a^{b^-} f = \lim_{x\to b^-}\int_a^x f.$

Demostración. Es consecuencia de la desigualdad triangular y el Criterio de Cauchy:

$$\left|\int_x^y f\right| \le M\,|y-x| \quad \therefore \quad \lim_{x,y\to b^-}\int_x^y f = 0.$$

□

• **Ejemplo:** $\displaystyle\int_{0^+}^1 \frac{\sin x}{x}\,dx$ es convergente, pues $\sin x/x$ es continua en $(0,1]$ y

$$\lim_{x\to 0^+}\frac{\sin x}{x} = \lim_{x\to 0^+}\frac{\cos x}{1} = 1$$

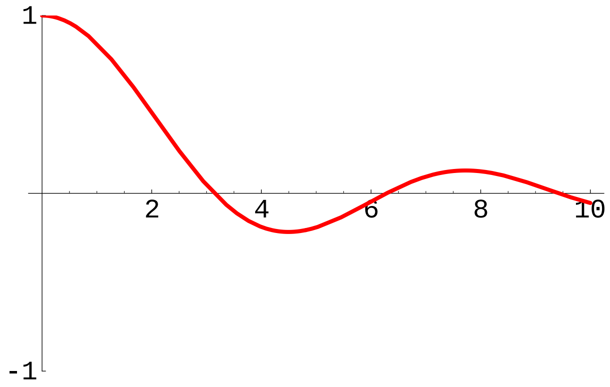

Figura 31 – la función $\sin x/x$.

• **Ejemplo:** La integral

$$\int_{0^+}^{1} \sin\log x\, dx$$

es convergente, pues $\sin\log x$ es continua en $(0,1]$, y $|\sin\log x| \le 1$. Sin embargo, no existe $\lim_{x\to 0^+} \sin\log x$.

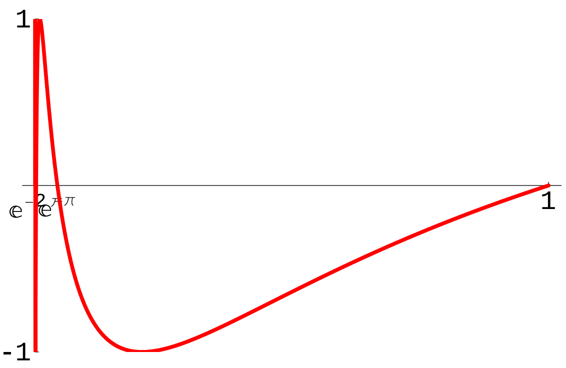

Figura 32 – la función $\sin\log x$.

Este tipo de integral, con integrando continuo y acotado en un intervalo semiabierto acotado, no es realmente una integral impropia, pues no hay nada que no esté acotado. De hecho, definiendo $f(b)$ arbitrariamente siempre se obtiene una función integrable en el sentido habitual. Así pues, son sólo integrales que "parecen impropias pero no lo son."

• **Ejemplo:** $\displaystyle\int_{0^+}^{1} x\,\log x\, dx$ es convergente, pues $x\,\log x$ es continua en $(0,1]$, y

$$\lim_{x\to 0^+} x\,\log x = \lim_{x\to 0^+} \frac{\log x}{x^{-1}} = \lim_{x\to 0^+} \frac{x^{-1}}{-x^{-2}} = \lim_{x\to 0^+} -x = 0.$$

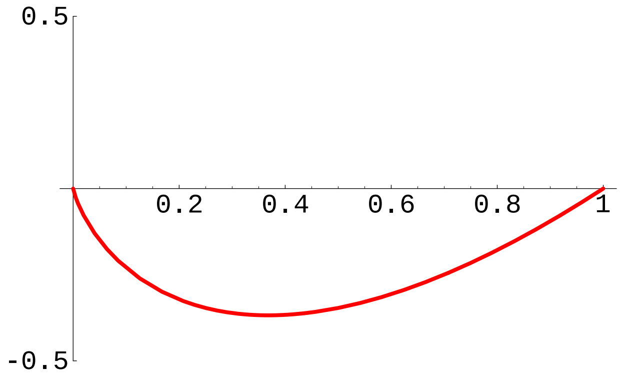

Figura 33 – la función $x \log x$.

El resultado es falso si el intervalo no es acotado aunque la función sea acotada:

$$\int_0^\infty \sin x \, dx \quad \text{diverge.}$$

El resultado es falso si el intervalo es acotado pero la función no está acotada:

$$\int_{0^+}^1 \frac{dx}{x} = \infty.$$

Estas integrales ya son impropias "de verdad."

6.4.2. Convergencia absoluta y Criterios de comparación

Volvemos al caso general de un intervalo semiabierto $I = [a, b)$, posiblemente no acotado.

Integrando positivo de integral acotada

Sea f continua en $[a, b)$ tal que $f \geq 0$. Entonces

$$\int_a^{b^-} f \quad \textit{es convergente si y sólo si} \quad F(x) = \int_a^x f \quad \textit{está acotada en } [a, b).$$

Demostración. Si $f \geq 0$, entonces

$$F(x) = \int_a^x f$$

es creciente:

$$y \geq x \implies F(y) - F(x) = \int_x^y f \geq 0 \implies F(y) \geq F(x).$$

Como F es creciente en $[a, b)$, siempre existe el límite ampliado $F(b^-) \in [0, +\infty]$. Es finito si y sólo si F está acotada. $\square$

Para una función positiva, sólo puede pasar que

$$\int_a^{b^-} f = \infty \qquad \text{ó} \qquad 0 \leq \int_a^{b^-} f < \infty.$$

Es decir, la única posible divergencia es con límite infinito.

Convergencia absoluta

Sea f continua en $[a, b)$. Decimos que una integral impropia

$$\int_a^{b^-} f(x)\, dx$$

*es **absolutamente convergente** si*

$$\int_a^{b^-} |f(x)|\, dx < \infty.$$

*También se dice en este caso que la función f es **absolutamente integrable** en $[a, b)$.*

La convergencia absoluta implica la convergencia

Si $\displaystyle\int_a^{b^-} |f| < \infty$ entonces $\displaystyle\int_a^{b^-} f$ es convergente.

Demostración. Se aplica la desigualdad triangular y el criterio de Cauchy: si $a \leq x < y < b$, y $\int |f| < \infty$,

$$\left| \int_x^y f \right| \leq \int_x^y |f| \to 0 \qquad x, y \to b^-.$$

$\square$

Recordamos que la notación $f(x) = O(g(x))$ cuando $x \to b^-$ significa que hay algún $\delta > 0$ y una constante $C > 0$ tales que

$$\boxed{|f(x)| \leq C|g(x)| \quad \forall\, x \in (b - \delta, b).}$$

y $f(x) = o(g(x))$ cuando $x \to b^-$ significa que dada **cualquier** constante $\varepsilon > 0$, existe $\delta > 0$ tal que

$$\boxed{|f(x)| \leq \varepsilon|g(x)| \quad \forall\, x \in (b - \delta, b).}$$

Si $g(x) \neq 0$, esto equivale a $\lim_{x \to b^-} f(x)/g(x) = 0$.

Esta notación de orden de magnitud es extremadamente útil para comparar funciones unas con otras. Aquí nos interesan las comparaciones entre integrales impropias.

Comparación $f = O(g)$

Sean f, g *continuas en* $[a, b)$ *con* g *absolutamente integrable. Si* $f(x) = O(g(x))$ *cuando* $x \to b^-$*, entonces* f *es absolutamente integrable.*

Demostración. Es consecuencia de la monotonía de la integral y de los límites:

$$|f| \leq C|g| \quad \text{para } b-\delta \leq x < b \implies \int_a^x |f| = \int_a^{b-\delta} |f| + \int_{b-\delta}^x |f| \leq \int_a^{b-\delta} |f| + C\int_{b-\delta}^x |g|.$$

Si g es absolutamente integrable, al hacer $x \to b^-$ vemos que también $\int_a^{b^-} |f| < \infty$. $\square$

Comparación $f = o(g)$

Sean f, g *continuas en* $[a, b)$ *con* g *absolutamente integrable. Si* $f(x) = o(g(x))$ *cuando* $x \to b^-$*, entonces* f *es absolutamente integrable.*

Demostración. Simplemente basta observar que $f = o(g) \implies f = O(g)$. $\square$

Comparación asintótica

Sean f, g *continuas en* $[a, b)$ *con* $g(x) \neq 0$*. Si*

$$\lim_{x \to b^-} \left| \frac{f(x)}{g(x)} \right| = \ell \in (0, \infty)$$

entonces f *es absolutamente integrable en* $[a, b)$ *si y sólo si* g *lo es.*

Demostración. Basta con observar que la hipótesis sobre el límite implica tanto $f = O(g)$ como $g = O(f)$ cuando $x \to b^-$: dado un ε con $0 < \varepsilon < \ell$, existe $\delta > 0$ tal que

$$0 < \ell - \varepsilon < \left| \frac{f(x)}{g(x)} \right| < \ell + \varepsilon \quad \forall x \in [b-\delta, b)$$

por tanto $|f(x)| \leq (\ell + \varepsilon)|g(x)|$ y $|g(x)| \leq \dfrac{1}{\ell - \varepsilon}|f(x)|$ para todo $x \in [b-\delta, b)$.

$\square$

En particular, el caso $\ell = 1$ corresponde a la equivalencia asintótica

$$\boxed{f \sim g \quad (x \to b^-) \iff \lim_{x \to b^-} \frac{f}{g} = 1}$$

En general hablaremos de comparación asintótica cuando se de este caso de límite $\ell \in (0, \infty)$, aunque no sea $\ell = 1$. La multiplicación por una constante no nula no afecta a la convergencia. Es una manera de verificar las condiciones simétricas $f = O(g), g = O(f)$.

6.4.3. Algunas comparaciones útiles

Comparaciones entre potencias en ∞

Sean $c_1, c_2, \ldots, c_m$ con $c_m \neq 0$ y $p_1 < p_2 < \cdots < p_m$. Entonces

$$c_1x^{p_1} + c_2x^{p_2} + \cdots + c_mx^{p_m} \sim c_mx^{p_m} \qquad (x \to \infty)$$

• **Ejemplo:** $\displaystyle\int_1^\infty \frac{dx}{x + x^2 + x^3}$ es convergente por comparación asintótica con x^{-3} $(x \to \infty)$.

• **Ejemplo:** $\displaystyle\int_1^\infty \frac{dx}{\sqrt{x + x^2 + x^3}}$ converge por comparación asintótica con $x^{-3/2}$ $(x \to \infty)$.

• **Ejemplo:** $\displaystyle\int_\pi^\infty \frac{\sin x}{x^{1,00001}}$ converge absolutamente, al ser el integrando $O(x^{-1,00001})$ $(x \to \infty)$.

Comparaciones entre potencias en 0

Sean $c_1, c_2, \ldots, c_m$ con $c_1 \neq 0$ y $p_1 < p_2 < \cdots < p_m$. Entonces

$$c_1x^{p_1} + c_2x^{p_2} + \cdots + c_mx^{p_m} \sim c_1x^{p_1} \qquad (x \to 0^+)$$

• **Ejemplo:** $\displaystyle\int_0^1 \frac{dx}{x + x^2 + x^3}$ diverge por comparación asintótica con x^{-1} $(x \to 0^+)$.

• **Ejemplo:** $\displaystyle\int_0^1 \frac{dx}{\sqrt{x + x^2 + x^3}}$ converge por comparación asintótica con $x^{-1/2}$ $(x \to 0^+)$.

• **Ejemplo:** $\displaystyle\int_0^1 \frac{e^x\,dx}{x^{0,999999}}$ converge por comparación asintótica con $x^{-0,999999}$ $(x \to 0^+)$.

• **Ejemplo:** Si hay más de un punto impropio realizamos varias comparaciones:

$$\int_0^\infty \frac{dx}{\sqrt{x + x^3}}$$

converge por comparación asintótica con $x^{-1/2}$ cuando $x \to 0^+$ y con $x^{-3/2}$ cuando $x \to \infty$.

Un método sencillo para utilizar estos resultados es:

- recordar los resultados en ∞
- obtener los correspondientes en 0 por el cambio de variable $x \to 1/x$
- cambiar a cualquier punto $c \in \mathbb{R}$ mediante $x \to x - c$ ó $x \to c - x$

• **Ejemplo:** $\displaystyle\int_0^2 \frac{dx}{\sqrt{x}|x-1|^{2/3}(2-x)^{5/6}}$ es convergente por comparación asintótica con

- $x^{-1/2}$ cuando $x \to 0^+$
- $|x-1|^{-2/3}$ cuando $x \to 1$
- $(2-x)^{-5/6}$ cuando $x \to 2^-$

• **Ejemplo:** $\displaystyle\int_0^\infty \frac{e^{-x}}{\sqrt[3]{|x^3-x|}} = \int_0^\infty \frac{e^{-x}}{\sqrt[3]{|x(x-1)(x+1)|}}$ converge por

- comparación asintótica con $x^{-1/3}$ cuando $x \to 0^+$
- comparación asintótica con$|x-1|^{-1/3}$ cuando $x \to 1$
- comparación $O(e^{-x})$ cuando $x \to \infty$.

Comparaciones entre potencias, logaritmos y exponenciales en ∞

Para todo $p \in \mathbb{R}, q > 0$,

$$\lim_{x\to\infty} \frac{\log^p x}{x^q} = \lim_{x\to\infty} \frac{x^p}{e^{qx}} = 0.$$

Es decir, $\log^p x = o(x^q)$ y $x^p = o(e^{qx})$ cuando $x \to \infty$.

• **Ejemplo:** $\displaystyle\int_0^\infty x^{10^{10}} e^{-x}\, dx$ es convergente porque el integrando es $o(e^{-x/2})$ cuando $x \to \infty$.

• **Ejemplo:** $\displaystyle\int_0^\infty e^{-x}\frac{\sin x}{x}\, dx$ es absolutamente convergente por comparación asintótica con 1 cuando $x \to 0^+$ y porque el integrando es $O(e^{-x})$ cuando $x \to \infty$.

Comparaciones entre potencias y logaritmos en 0

Para todo $p > 0, q \in \mathbb{R}$,

$$\lim_{x\to 0^+} x^p |\log x|^q = 0.$$

Es decir, $x^p|\log x|^q = o(1)$ (tiende a cero) cuando $x \to 0^+$.

• **Ejemplo:** $\displaystyle\int_0^1 x^p|\log x|^q\, dx$ con $p, q > 0$ es convergente pues el integrando tiende a 0 cuando $x \to 0^+$, luego es una función acotada en un intervalo acotado.

• **Ejemplo:** $\displaystyle\int_0^1 \frac{\log x}{x}\, dx = -\infty$ pues el integrando es negativo y $\dfrac{|\log x|}{x} > \dfrac{1}{x}$ si $0 < x < \frac{1}{e}$.

• **Ejemplo:** $\displaystyle\int_0^1 \frac{|\log x|}{\sqrt{x}}\,dx$ es convergente, pues eligiendo cualquier ε con $0 < \varepsilon < \frac{1}{2}$, entonces

$$\frac{|\log x|}{\sqrt{x}} = \frac{x^{\varepsilon}|\log x|}{x^{\frac{1}{2}+\varepsilon}} = o\left(\frac{1}{x^{\frac{1}{2}+\varepsilon}}\right) \qquad (x \to 0^+)$$

y $\frac{1}{2} + \varepsilon < 1$.

6.4.4. La analogía entre integrales impropias y series infinitas

Para aquellos que conozcan las propiedades básicas de las series infinitas, se hace patente una relación tanto en terminología como en resultados, entre ellas y las integrales impropias. Vamos a resaltar esta analogía.

Integrales impropias	Series infinitas
variable continua x en un intervalo semiabierto $[a,b)$	variable discreta $n \in \mathbb{N}$
límite $x \to b^-$	límite $n \to \infty$
función $f : [a,b) \to \mathbb{R}$ ó $\mathbb{C}$	sucesión $a : \mathbb{N} \to \mathbb{R}$ ó $\mathbb{C}$
integral parcial $F(x) = \displaystyle\int_a^x f$	suma parcial $\displaystyle\sum_{n=1}^{N} a_n = A_N$
convergencia $\displaystyle\int_a^{b^-} f = \lim_{x\to b^-} F(x)$	convergencia $\displaystyle\sum_{n=1}^{\infty} a_n = \lim_{N\to\infty} A_N$
convergencia absoluta $\displaystyle\int_a^{b^-} \lvert f(x)\rvert\,dx < \infty$	convergencia absoluta $\displaystyle\sum_{n=1}^{\infty} \lvert a_n\rvert < \infty$

Los criterios para integrales también son análogos a los de series.

Criterio para integrales	Criterio para series
criterio de Cauchy: $\displaystyle\lim_{x,y\to b^-} \int_x^y f = 0$	criterio de Cauchy: $\displaystyle\lim_{M,N\to\infty} \sum_{n=M}^{N} a_n = 0$
acotación positiva:	acotación positiva:
Si $f \geq 0$, $\displaystyle\int_a^{b^-} f < \infty \iff F$ acotada	Si $a_n \geq 0$, $\displaystyle\sum_{n=1}^{\infty} a_n < \infty \iff A_N$ acotada
convergencia absoluta $\implies$ convergencia	convergencia absoluta $\implies$ convergencia
Si $f = O(g)$, $\int \lvert g\rvert < \infty \implies \int \lvert f\rvert < \infty$	Si $a_n = O(b_n)$, $\sum \lvert b_n\rvert < \infty \implies \sum \lvert a_n\rvert < \infty$
Si $f \sim g$, $\int \lvert g\rvert < \infty \iff \int \lvert f\rvert < \infty$	Si $a_n \sim b_n$, $\sum \lvert b_n\rvert < \infty \iff \sum \lvert a_n\rvert < \infty$

El lector familiarizado con la teoría de series infinitas habrá observado que hay un criterio sencillo para series cuyo análogo no hemos tratado para integrales: si $\sum_{n=1}^{\infty} a_n$ es convergente, su término general tiende á cero: $\lim_n a_n = 0$. La demostración es muy sencilla: si A es el valor de la suma y $A_n = a_1 + \cdots + a_n$ es la suma parcial n-ésima, se tiene por definición de convergencia que $\lim_n A_n = A$, luego

$$\lim_n a_n = \lim_n (A_n - A_{n-1}) = A - A = 0.$$

El resultado análogo para integrales sería:

"Si $\int_a^{b^-} f(x)\,dx$ es convergente, entonces $\lim_{x\to b^-} f(x) = 0$."

Esto es falso en general, principalmente por dos razones:

- Si $b < \infty$, suponiendo que es un punto impropio, de existir el límite de f, por definición debería ser $\lim_{x\to b^-} f(x) = \infty$ y no 0. Por ejemplo, $f(x) = 1/\sqrt{x}$ en $(0, 1]$.
- Aunque sea $b = \infty$, el límite $\lim_{x\to\infty} f(x)$ no tiene por qué existir. Un ejemplo clásico es una función positiva $f(x)$ sobre $[0, \infty)$ cuya gráfica consta de una sucesión T_n de triángulos disjuntos de modo que la suma de sus áreas sea convergente pero sus alturas tiendan a infinito. Por ejemplo, tomando como base de T_n un segmento de longitud 4^{-n} contenido en el intervalo $[n, n+1]$ y como altura uno de longitud 2^{n+1}. Entonces el área de T_n será 2^{-n}, y $\int_0^{\infty} f(x)\,dx = \sum_{n=1}^{\infty} 2^{-n} = 1$, pero $\lim_{x\to\infty} f(x)$ no existe.

Un resultado análogo correcto es el siguiente aunque no es demasiado útil como criterio.

Criterio del Límite para integrales impropias sobre intervalos no acotados

Sea $f(x)$ continua en $[a, \infty)$ (respectivamente $(-\infty, b]$), con integral $\int_a^{\infty} f(x)\,dx$ convergente (resp. $\int_{-\infty}^{b} f(x)\,dx$). Si existe el límite $\lim_{x\to\infty} f(x) = \ell$ (resp. $\lim_{x\to-\infty} f(x) = \ell$), entonces $\ell = 0$.

Demostración. Por simetría supondremos el caso $[a, \infty)$. Damos dos demostraciones. Sea $F(x) = \int_a^x f(t)\,dt$, de modo que existe $I = \lim_{x\to\infty} F(x) = \int_a^{\infty} f(t)\,dt$. Por el Teorema del Valor Medio, para $x \geq a$,

$$F(x+1) - F(x) = \int_x^{x+1} f(t)\,dt = f(c_x)$$

para algún $c_x \in [x, x+1]$. Haciendo $x \to \infty$, también $c_x \to \infty$ y $x + 1 \to \infty$, luego

$$\ell = \lim_{x\to\infty} f(c_x) = \lim_{x\to\infty} (F(x+1) - F(x)) = I - I = 0.$$

Alternativamente, si $\ell \neq 0$, podemos suponer que $\ell > 0$ (si no, cambiamos f por $-f$ y aplicamos el mismo razonamiento). Como $\lim_{x\to\infty} f(x) = \ell > 0$, existe $R \geq a$ tal que $f(x) \geq \frac{\ell}{2} > 0$ si $x \geq a$. Pero entonces, para $x \geq R$, se tiene

$$F(x) = \int_a^x f = \int_a^R f + \int_R^x f \geq \int_a^R f + \int_R^x \frac{\ell}{2} = F(R) + \frac{\ell}{2}(x - R),$$

por tanto $\lim_{x\to\infty} F(x) = \infty$, lo cual contradice que $\int_a^{\infty} f$ sea convergente; luego $\ell = 0$. $\square$

6.4.5. Comparaciones entre integrales y series

La analogía entre integrales impropias y series sugiere que hay una relación entre ambas cosas y que quizá pueda aprovecharse para obtener resultados en un ámbito y pasarlos al otro. Vamos a estudiar una manera de obtener una serie a partir de una integral y una integral a partir de una serie.

Partición numerable

Dado un intervalo semiabierto, $[a,b)$, una partición numerable es una sucesión infinita

$$a = c_0 < c_1 < c_2 < \cdots < c_n < \cdots < b, \quad \text{tal que} \quad \lim_{n\to\infty} c_n = b.$$

Tal sucesión divide al intervalo $[a,b)$ en una unión numerable de subintervalos

$$[a,b) = [c_0,c_1] \cup [c_1,c_2] \cup \cdots \cup [c_{n-1},c_n] \cup \cdots$$

Aditividad numerable respecto a una partición

Sea f continua en $[a,b)$ y

$$a = c_0 < c_1 < c_2 < \cdots < c_n < \cdots < b, \quad \lim c_n = b,$$

una partición numerable de $[a,b)$. Consideremos los dos límites

$$\int_a^{b^-} f(x)\,dx, \qquad \sum_{n=0}^{\infty} \int_{c_n}^{c_{n+1}} f(x)\,dx.$$

- *La convergencia de la integral siempre implica la de la serie.*
- *Si $f(x)$ no cambia de signo en cada intervalo $[c_n, c_{n+1}]$, entonces la convergencia de la serie implica la de la integral.*

En ambos casos, la integral y la serie convergen al mismo valor.

Demostración. Denotamos respectivamente a la integral parcial, al término general de la serie, y a las sumas parciales de la serie, por

$$F(u) = \int_a^u f(x)\,dx, \qquad I_n = \int_{c_n}^{c_{n+1}} f(x)\,dx, \qquad S_N = I_0 + I_1 + I_2 + \cdots + I_N.$$

Por aditividad (finita) de la integral, las sumas parciales de la serie son integrales parciales:

$$S_N = \left(\int_{c_0}^{c_1} + \int_{c_1}^{c_2} + \cdots + \int_{c_{N-1}}^{c_N} f(x)\,dx \right) = \int_a^{c_N} f(x)\,dx = F(c_N).$$

Comparemos entonces lo que significan la convergencia de la serie y de la integral.

- La convergencia de la serie equivale a la existencia del límite $S = \lim_{N\to\infty} F(c_N)$.
- La convergencia de la integral equivale a la existencia del límite $I = \lim_{u\to b^-} F(u)$.

Entonces

- La convergencia de la serie equivale a la existencia del límite discreto de las integrales parciales $F(u)$ a lo largo de la sucesión concreta $u = c_N$ que tiende á b por la izquierda.
- La convergencia de la integral equivale a la existencia del límite continuo cuando $u \to b^-$ de las integrales parciales. A su vez, esto es equivalente a la existencia e igualdad de todos los límites discretos a lo largo de cualquier sucesión que tiende á b por la izquierda.

Por tanto, la convergencia de la integral (cualquier sucesión) es una condición más fuerte que la de la serie (una sucesión particular). Esto es lo que dice el primer enunciado. Específicamente, si converge la integral, entonces

$$\sum_{n=0}^{\infty} \int_{c_n}^{c_{n+1}} f(x)\,dx = \lim_{N\to\infty} S_N = \lim_{N\to\infty} F(c_N) = \lim_{u\to b^-} F(u) = \int_a^{b^-} f(x)\,dx,$$

es decir, también la serie converge al mismo valor.

• En cambio, el segundo enunciado va en la dirección contraria: de la convergencia a lo largo de la sucesión particular $u = c_N$ queremos deducir la convergencia continua cuando $u \to b^-$. En general esto será falso sin alguna hipótesis adicional. Veamos por qué la que hemos dado, que $f(x)$ no cambie de signo en cada intervalo $[c_{n-1}, c_n]$, es suficiente.

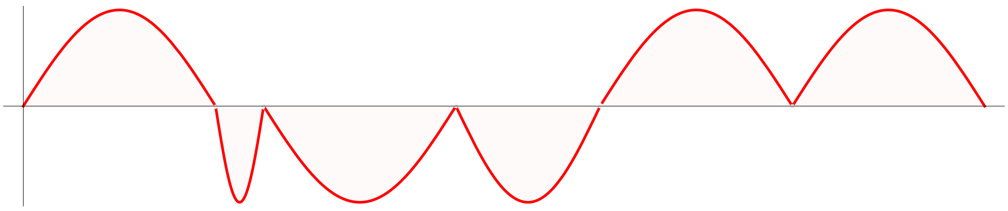

Figura 34 – Ejemplo de función que no cambia de signo en los intervalos determinados por una sucesión.

Como c_N crece á b, dado $u \in [a, b)$, existe un único entero $N(u)$, que depende de u, tal que

$$c_{N(u)} \le u < c_{N(u)+1} \quad \therefore \quad u \to b^- \iff N(u) \to \infty.$$

La integral parcial $F(u)$ es una suma parcial de la serie más un resto que depende de u:

$$F(u) = \int_a^u f(x)\,dx = \left(\int_a^{c_{N(u)}} + \int_{c_{N(u)}}^u\right) f(x)\,dx = S_{N(u)} + R(u), \quad R(u) = \int_{c_{N(u)}}^u f(x)\,dx.$$

Si la serie converge á S, es decir, $\lim_{N\to\infty} S_N = S$, entonces también $\lim_{u\to b^-} S_{N(u)} = S$ y basta con demostrar que $\lim_{u\to b^-} R(u) = 0$ para concluir que entonces la integral también converge á S:

$$\lim_{u\to b^-} \int_a^u f(x)\,dx = \lim_{u\to b^-} (S_{N(u)} + R(u)) = \lim_{u\to b^-} S_{N(u)} + \lim_{u\to b^-} R(u) = S + 0 = S.$$

Para demostrar que el resto $R(u)$ tiende á cero, la hipótesis que f no cambia de signo en cada intervalo $[c_N, c_{N+1}]$ es fundamental, pues asegura que la desigualdad triangular en ese intervalo es igualdad. Es decir, si una función f integrable en un intervalo $[\alpha, \beta]$ no cambia de signo en él, entonces cumple

$$\int_\alpha^\beta |f(x)|\,dx = \left|\int_\alpha^\beta f(x)\,dx\right|$$

ya que:

- Si $f \geq 0$ en $[\alpha, \beta]$, entonces (por monotonía) $\int_\alpha^\beta f(x)\,dx \geq 0$, luego efectivamente
$$\left|\int_\alpha^\beta f(x)\,dx\right| = \int_\alpha^\beta f(x)\,dx = \int_\alpha^\beta |f(x)|\,dx.$$
- Si $f \leq 0$ en $[\alpha, \beta]$, entonces (también por monotonía) $\int_\alpha^\beta f(x)\,dx \leq 0$, luego
$$\left|\int_\alpha^\beta f(x)\,dx\right| = -\int_\alpha^\beta f(x)\,dx = \int_\alpha^\beta -f(x)\,dx = \int_\alpha^\beta |f(x)|\,dx.$$

Entonces, volviendo a la demostración, como $c_{N(u)} \leq u < c_{N(u)+1}$ y $|f| \geq 0$,

$$\begin{aligned}|R(u)| = \left|\int_{c_{N(u)}}^u f(x)\,dx\right| = \int_{c_{N(u)}}^u |f(x)|\,dx &\leq \int_{c_{N(u)}}^{c_{N(u)+1}} |f(x)|\,dx \\ &= \left|\int_{c_{N(u)}}^{c_{N(u)+1}} f(x)\,dx\right| = |I_{N(u)}|.\end{aligned}$$

Si la serie es convergente, su término general I_N tiende á 0, por tanto $R(u)$ también. □

Si se tiene una fórmula explícita para una antiderivada $F(x)$ de $f(x)$, la serie de integrales es siempre **telescópica:**

$$\sum_{n=0}^{\infty} \int_{c_n}^{c_{n+1}} f(x)\,dx = \sum_{n=0}^{\infty} \big(F(c_{n+1}) - F(c_n)\big)$$

luego, como método, no sirve tanto para hallar el valor de la integral, sino para demostrar su convergencia o divergencia mediante el análisis de la serie.

Aditividad numerable para funciones positivas

Sea f continua en $[a, b)$, con $f \geq 0$, y $\{c_n\}_{n=0}^\infty$ una partición infinita numerable de $[a, b)$, es decir, $a = c_0 < \ldots < c_{n-1} < c_n < b$ con $\lim_n c_n = b$. Entonces

$$\int_a^{b^-} f(x)\,dx = \sum_{n=0}^{\infty} \int_{c_n}^{c_{n+1}} f(x)\,dx,$$

en el sentido de que o bien ambas convergen al mismo valor, o ambas divergen a ∞.

(de la aditividad para funciones positivas). Esto es consecuencia inmediata del resultado anterior, pues si $f \geq 0$, en particular no cambia de signo en ningún intervalo. Entonces, la convergencia de la integral equivale a la de la serie. Además, al tratarse de una integral y una serie positivas, la divergencia sólo puede ser á ∞. □

Aditividad numerable y convergencia absoluta

Sea f es continua en $[a, b)$ y $\{c_n\}_{n=0}^{\infty}$ una partición infinita numerable de $[a, b)$, es decir, $a = c_0 < \ldots < c_{n-1} < c_n < b$ con $\lim_n c_n = b$. Si

$$\sum_{n=0}^{\infty} \int_{c_n}^{c_{n+1}} |f(x)| \, dx < \infty$$

entonces

$$\int_a^{b^-} f(x) \, dx = \sum_{n=0}^{\infty} \int_{c_n}^{c_{n+1}} f(x) \, dx,$$

siendo la convergencia de ambas absoluta y al mismo valor.

Demostración. El resultado anterior para funciones positivas implica que

$$\int_a^{b^-} |f(x)| \, dx = \sum_{n=0}^{\infty} \int_{c_n}^{c_{n+1}} |f(x)| \, dx < \infty$$

es decir, la integral

$$\int_a^{b^-} f(x) \, dx$$

converge absolutamente. En particular, converge. Como la integral es convergente, la serie

$$\sum_{n=0}^{\infty} \int_{c_n}^{c_{n+1}} f(x) \, dx$$

también lo es, y ambas convergen al mismo valor. Además, como

$$\sum_{n=0}^{\infty} \left| \int_{c_n}^{c_{n+1}} f(x) \, dx \right| \leq \sum_{n=0}^{\infty} \int_{c_n}^{c_{n+1}} |f(x)| \, dx < \infty,$$

la serie también converge absolutamente. □

• **Ejemplo:** $\displaystyle\int_0^{\infty} \sin x \, dx$ es divergente, pues la serie obtenida mediante $c_n = n\pi$ es

$$\sum_{n=0}^{\infty} \int_{(n-1)\pi}^{n\pi} \sin x \, dx = \sum_{n=0}^{\infty} (-1)^{n+1} \cdot 2,$$

que es divergente al no tender a 0 su término general. Observar que con $c_n = 2\pi n$,

$$\int_{2\pi(n-1)}^{2\pi n} \sin x \, dx = 0$$

y por tanto
$$\sum_{n=1}^{\infty} \int_{2\pi(n-1)}^{2\pi n} \sin x \, dx = \sum_{n=1}^{\infty} 0 = 0,$$
pero esto no implica la convergencia de la integral pues $\sin x$ no mantiene un signo constante en $(2\pi(n-1), 2\pi n)$.

• **Ejemplo:**
$$\int_0^{\infty} \left| \frac{\sin x}{x} \right| dx = \infty$$
pues eligiendo $c_n = \pi n$, para $n \geq 1$,
$$\int_{\pi(n-1)}^{\pi n} \left| \frac{\sin x}{x} \right| dx \geq \frac{1}{\pi n} \int_{\pi(n-1)}^{\pi n} |\sin x| \, dx = \frac{2}{\pi n} \quad \text{pero} \quad \sum_{n=1}^{\infty} \frac{1}{n} = \infty.$$
Sin embargo, la integral
$$\int_0^{\infty} \frac{\sin x}{x} \, dx$$
es convergente, pues en 0 no es impropia ya que el integrando es asintóticamente equivalente á 1 y, separándola en
$$\int_0^{\infty} \frac{\sin x}{x} \, dx = \int_0^{\pi/2} \frac{\sin x}{x} \, dx + \int_{\pi/2}^{\infty} \frac{\sin x}{x} \, dx,$$
la primera integral es convergente al ser un integrando acotado en un intervalo acotado, y podemos integrar la segunda por partes con $u = x^{-1}$, $dv = \sin x \, dx$, resultando
$$\int_{\pi/2}^{\infty} \frac{\sin x}{x} \, dx = \left[-\frac{\cos x}{x} \right]_{\pi/2}^{\infty} - \int_{\pi/2}^{\infty} \frac{\cos x}{x^2} \, dx = -\int_{\pi/2}^{\infty} \frac{\cos x}{x^2} \, dx,$$
y esta última integral es absolutamente convergente por ser el integrando $O(x^{-2})$ cuando $x \to \infty$.

Esta integral proporciona un ejemplo de integral convergente pero no absolutamente convergente. En estos casos se dice que tenemos convergencia condicional.

A continuación damos una manera de asociar a una serie una integral, y de usar el comportamiento de la integral para determinar el de la serie.

El criterio integral de Cauchy

Sea $a \in \mathbb{Z}$, $f(x)$ continua en $[a, \infty)$, positiva y decreciente. Entonces
$$\int_a^{\infty} f(x) \, dx \qquad \sum_{n=a}^{\infty} f(n)$$
son ambas convergentes o ambas divergentes a ∞.

No es verdad que la integral y la suma en el criterio integral de Cauchy tengan el mismo valor en caso de ser ambas convergentes.

Demostración. Sean

$$I_N = \int_a^N f(x)\,dx, \qquad S_N = \sum_{n=a}^{N} f(n).$$

Entonces, como f es positiva y decreciente,

$$\begin{aligned} n \le x \le n+1 &\implies f(n) \ge f(x) \ge f(n+1) \\ &\implies \int_n^{n+1} f(n)\,dx \ge \int_n^{n+1} f(x)\,dx \ge \int_n^{n+1} f(n+1)\,dx \\ &\iff f(n) \ge \int_n^{n+1} f(x)\,dx \ge f(n+1). \end{aligned}$$

Sumando de $n = a$ hasta $n = N-1$ queda

$$S_{N-1} = \sum_{n=a}^{N-1} f(n) \ge \int_a^N f = I_N \ge \sum_{n=a+1}^{N} f(n) = S_N - f(a).$$

Como $f \ge 0$, tanto la integral como la suma son de términos positivos, luego la convergencia equivale a la acotación. Las desigualdades anteriores muestran inmediatamente que S_N está acotada si y sólo si I_N está acotada. □

• **Ejemplo:** La serie armónica $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n}$ es divergente por comparación con la integral

$$\int_1^{\infty} \frac{dx}{x} = \infty$$

• **Ejemplo:** La serie zeta $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^s}$ es convergente si y sólo si $s > 1$, por comparación con

$$\int_1^{\infty} \frac{dx}{x^s} = \frac{1}{s-1} \quad (s > 1).$$

Un problema famoso, el **problema de Basilea,** es demostrar que $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$. Observar que la integral correspondiente vale 1.

• **Ejemplo:** La serie $\displaystyle\sum_{n=2}^{\infty} \frac{1}{n \log n}$ es divergente por comparación con

$$\int_2^{\infty} \frac{dx}{x \log x} = [\log \log x]_2^{\infty} = \infty.$$

• **Ejemplo:** La serie $\displaystyle\sum_{n=0}^{\infty} \frac{1}{n^2+1}$ es convergente por comparación con

$$\int_0^{\infty} \frac{dx}{x^2+1} = \frac{\pi}{2}.$$

De hecho,

$$\sum_{n=0}^{\infty} \frac{1}{n^2+1} = \frac{1 + \pi \coth \pi}{2}.$$

6.4.6. Criterios de Dirichlet y de Abel

La analogía entre integrales impropias y series va aún más lejos, a otros criterios más sofisticados. De ellos, resaltaremos dos que pueden aplicarse en determinadas situaciones cuando no haya una comparación directa más clara.

El criterio de Dirichlet

Sean f, g continuas en $[a, b)$ tales que

- *Las integrales parciales $G(x) = \displaystyle\int_a^x g$ están acotadas en $[a, b)$, o sea, existe una constante M tal que $|G(x)| \leq M$ para todo $x \in [a, b)$.*
- *$f(x)$ es monótona en $[a, b)$, con $\lim_{x\to b^-} f(x) = 0$.*

Entonces la integral $\displaystyle\int_a^{b^-} f(x)g(x)\, dx$ es convergente.

Para series, el criterio de Dirichlet es el siguiente:

Sean dos sucesiones a_n, b_n tales que

- Las sumas parciales $B_N = b_1 + b_2 + \cdots + b_N$ están acotadas, o sea, existe una constante M tal que $|B_N| \leq M$ para todo N.
- a_n es positiva y decreciente con $\lim_{n\to\infty} a_n = 0$.

Entonces la serie $\sum_{n=1}^{\infty} a_n b_n$ es convergente.

Demostración. Supondremos para la demostración que f es derivable con derivada continua (no hace falta pero simplifica el razonamiento considerablemente). Consideramos las colas $\int_x^y f(t)g(t)\, dt$, siendo $a \leq x < y < b$. Aplicaremos el Criterio de Cauchy, demostrando que $\lim_{x,y\to b^-} \int_x^y f(t)g(t)\, dt = 0$. Integramos la cola por partes:

$$\int_x^y \underbrace{f(t)}_{u}\, \underbrace{g(t)\, dt}_{dv} = \Big[f(t)G(t)\Big]_{t=x}^{t=y} - \int_x^y f'(t)G(t)\, dt.$$

Observar que hemos usado el Teorema Fundamental para deducir que $v = G$ es una antiderivada de g. Ahora, para el primer término tenemos

$$\lim_{x,y\to b^-} (f(y)G(y) - f(x)G(x)) = 0,$$

ya que se trata de límites de la forma $0 \cdot$ acotado. En cuanto al segundo, aplicando la desigualdad triangular para integrales, estimamos

$$\left|\int_x^y f'(t)G(t)\, dt\right| \leq \int_x^y |f'(t)| \cdot |G(t)|\, dt \leq M \int_x^y |f'(t)|\, dt$$

y, como suponemos f derivable y monótona, el signo $\varepsilon = \pm 1$ de f' es constante en $[a,b)$, por tanto, observando que $|f'| = \varepsilon f'$, en la integral tenemos

$$\int_x^y |f'(t)|\,dt = \varepsilon \int_x^y f'(t)\,dt = \varepsilon(f(x) - f(y)) \to 0,$$

cuando $x, y \to b^-$, ya que $f(x), f(y) \to 0$. Por tanto las colas tienden a cero. □

• **Ejemplo:** La integral y suma análoga

$$\int_0^\infty \frac{\sin x}{x}\,dx, \qquad \sum_{n=0}^{\infty} \frac{\sin(n)}{n},$$

son convergentes, por el criterio de Dirichlet. El integrando tiene límite 1 cuando $x \to 0$ luego 0 no es un punto impropio y declaramos $\sin 0/0 = 1$ por continuidad. Basta con analizar lo que ocurre cuando $x \to \infty$:

- $f(x) = \dfrac{1}{x}$ es positiva y decreciente a 0 cuando $x \to \infty$.
- $g(x) = \sin x$ tiene integrales (respectivamente, sumas parciales) acotadas, pues

$$\left| \int_0^x \sin t\,dt \right| = |1 - \cos x| \le 2 \qquad \forall\, x \ge 0.$$

Los valores son

$$\int_0^\infty \frac{\sin x}{x}\,dx = \frac{\pi}{2}, \qquad \sum_{n=0}^{\infty} \frac{\sin(n)}{n} = \frac{\pi - 1}{2}.$$

La primera se puede demostrar mediante integración paramétrica (§7) y la segunda fórmula con series de Fourier (§9), pues

$$\sum_{n=1}^{\infty} \frac{\sin(nt)}{n} = \frac{\pi - t}{2} \qquad (0 < t < 2\pi).$$

Sin embargo, la introducción de un múltiplo positivo de la variable no afecta a la integral:

$$\int_0^\infty \frac{\sin tx}{x}\,dx = \int_{0\cdot t}^{\infty\cdot t} \frac{\sin x}{(x/t)}\frac{dx}{t} = \int_0^\infty \frac{\sin x}{x}\,dx \quad (t > 0).$$

El criterio de Abel

Sean f, g continuas en $[a,b)$ tales que

- *La integral $\displaystyle\int_a^{b^-} f(x)\,dx$ es convergente.*
- *$g(x)$ es monótona y acotada en $[a,b)$.*

Entonces la integral $\displaystyle\int_a^{b^-} f(x)g(x)\,dx$ es convergente.

Para series, el criterio de Abel es el siguiente:

Sean dos sucesiones a_n, b_n tales que

- $\sum_{n=1}^{\infty} a_n$ es convergente.
- b_n es monótona y acotada.

Entonces la serie $\sum_{n=1}^{\infty} a_n b_n$ es convergente.

Demostración. Al igual que en la demostración del Criterio de Dirichlet, para simplificar la demostración supondremos que uno de los factores, en este caso g, es derivable con derivada continua. Aplicamos el Criterio de Cauchy, demostrando que las colas tienden a cero: $\lim_{x,y\to b^-} \int_x^y f(t)g(t)\,dt = 0$.

Denotamos la integral parcial por $F(x) = \int_a^x f(t)\,dt$. Hacemos algunas observaciones.

- Sabemos que $F' = f$ por el Teorema Fundamental.
- La convergencia de $\int_a^{b-} f$ significa que existe el límite (finito) $F(b^-) \stackrel{\text{def}}{=} \lim_{x\to b^-} F(x)$.
- Como g es monótona y acotada, existe el límite (finito) $g(b^-) \stackrel{\text{def}}{=} \lim_{x\to b^-} g(x)$.

Integrando por partes la cola de la integral,

$$\int_x^y \underbrace{g(t)}_{u}\,\underbrace{f(t)\,dt}_{dv} = \Big[g(t)F(t)\Big]_{t=x}^{t=y} - \int_x^y g'(t)F(t)\,dt.$$

Al hacer límite en el primer término queda

$$\lim_{x,y\to b^-} (g(y)F(y) - g(x)F(x)) = g(b^-)\cdot F(b^-) - g(b^-)\cdot F(b^-) = 0.$$

En cuanto al segundo, aplicando la desigualdad triangular para integrales, estimamos

$$\left|\int_x^y g'(t)F(t)\,dt\right| \le \int_x^y |g'(t)|\cdot|F(t)|\,dt \le \max_{t\in[x,y]} |F(t)|\cdot \int_x^y |g'(t)|\,dt$$

y, como g es derivable y monótona, debe ser o bien $g'(t) > 0$ para todo $t \in [a,b)$ (creciente) o bien $g'(t) < 0$ para todo $t \in [a,b)$ (decreciente), con lo cual $|g'| = \varepsilon g'$, siendo ε constante con $\varepsilon = \pm 1$, y entonces

$$\int_x^y |g'(t)|\,dt = \varepsilon \int_x^y g'(t)\,dt = \varepsilon(g(y) - g(x)),$$

que tiende á $\varepsilon \cdot (g(b^-) - g(b^-)) = 0$ cuando $x, y \to b^-$. Por tanto

$$\lim_{x,y\to b^-} \max_{t\in[x,y]} |F(t)| \cdot \int_x^y |g'(t)|\,dt = |F(b^-)|\cdot 0 = 0$$

y las colas tienden a cero, como queríamos demostrar. □

El criterio de Abel también se puede deducir del criterio de Dirichlet. Basta con reemplazar g por $\widetilde{g}(x) \stackrel{\text{def}}{=} g(x) - g(b^-)$, que es monótona con límite nulo en b^-. Entonces $f, \widetilde{g}$ cumplen las hipótesis del criterio de Dirichlet, y como suponemos que $\int_a^{b^-} f$ es convergente, entonces $\int_a^{b^-} f\widetilde{g} = \int_a^{b^-} fg - g(b^-)\int_a^{b^-} f$, luego el criterio de Dirichlet implica que $\int_a^{b^-} fg$ es convergente.

• **Ejemplo:** Sabemos que $\int_0^\infty \frac{\sin x}{x}\,dx$ es convergente, por tanto

$$\int_0^\infty \frac{\sin x}{x} \arctan x \, dx$$

es convergente, pues $g(x) = \arctan x$ crece de 0 á $\pi/2$ cuando $x \in [0, \infty)$. Esto también se puede demostrar con el criterio de Dirichlet, pues $\sin x$ tiene integrales parciales acotadas y $\arctan x/x$ es decreciente a 0.

Es difícil encontrar entre los ejemplos habituales, uno al que sea aplicable el Criterio de Abel pero no el de Dirichlet. De hecho, como hemos comentado, la única diferencia es el límite del factor monótono, el cual si no es nulo siempre se puede restar y aplicar el criterio de Dirichlet. Si ya es nulo, el criterio de Abel es directamente un caso particular del criterio de Dirichlet. En todo caso, estos criterios deben reservarse para cuando no sean concluyentes los anteriores más sencillos.

7. Integrales Paramétricas

7.1. Definiciones y ejemplos básicos

Integral paramétrica

Una ***integral paramétrica*** *es una integral donde tanto el integrando como los extremos del intervalo de integración dependen de una o más variables, llamadas* ***parámetros.*** *Es decir, son integrales de la forma*

$$I(t) = \int_{a(t)}^{b(t)} f(t,x)\,dx$$

donde $t = (t_1, t_2, \ldots, t_p)$, *siendo* $t_k \in \mathbb{R}$ *ó* $\mathbb{C}$, *los parámetros individuales.*

Vamos a limitarnos en general al caso de un parámetro. Esto no resta demasiada generalidad, pues las operaciones habituales (límite, derivada, integral) proceden por pasos, operando sobre una variable a la vez mientras las demás se mantienen constantes.

Una integral paramétrica define una función de los parámetros. Por tanto estamos tratando con **funciones definidas por integrales.**

• **Ejemplo:** las integrales parciales $F(t) = \int_0^t f(x)\,dx$ son integrales paramétricas.

El Teorema Fundamental dice que esta función es una antiderivada de f: $F'(x) = f(x)$ para $x \in [0,t]$ (las derivadas en los extremos son laterales).

La integral en una integral paramétrica puede ser una **integral impropia.**

• **Ejemplo:** $I(t) = \int_0^{\infty} e^{-tx}\,dx$ converge si $t > 0$, siendo entonces $I(t) = 1/t$.

Para una función definida por una integral paramétrica, interesa conocer sus propiedades básicas: límites, continuidad, derivabilidad, integrabilidad.

• **Ejemplo:** $I(t) = \int_0^{\infty} \frac{dx}{1+t^2x^2}$, donde $f(t,x) = \frac{1}{1+t^2x^2}$, converge para $t \neq 0$, siendo

$$I(t) = \int_0^{\infty} \frac{dx}{1+t^2x^2} = \left[\frac{1}{t}\arctan(tx)\right]_0^{\infty} = \frac{\pi}{2|t|}$$

que es derivable si $t \neq 0$. Sin embargo, el integrando $f(t,x)$ está definido para todo t, incluyendo $t = 0$, y es derivable para todo t. Hay que estudiar en detalle estas cuestiones.

7.2. Extremos variables, integrando fijo

Vamos a estudiar primero el caso particular de integrales paramétricas de la forma

$$\boxed{I(t) = \int_{a(t)}^{b(t)} f(x)\,dx}$$

donde $a(t), b(t)$ son funciones de un parámetro real t. Aquí suponemos que

- $a(t), b(t)$ son funciones definidas para t en un intervalo $\mathbb{T} \subseteq \mathbb{R}$
- Existe un intervalo abierto J tal que $[a(t), b(t)] \subseteq J$ para todo $t \in \mathbb{T}$.
- $f(x)$ es continua para $x \in J$.

Estas condiciones garantizan que $I(t) = \displaystyle\int_{a(t)}^{b(t)} f(x)\,dx$ está definida para $t \in \mathbb{T}$.

Ahora, sea F cualquier antiderivada de f en J (su existencia está dada por el Teorema Fundamental del Cálculo). Entonces, por la regla de Barrow,

$$\boxed{I(t) = \int_{a(t)}^{b(t)} f(x)\,dx = F(b(t)) - F(a(t))}$$

con lo cual podemos asegurar que

- Si $a(t), b(t)$ son continuas, también lo es $I(t)$.
- Si $a(t), b(t)$ son derivables, entonces $I(t)$ es derivable, y se tiene

$$\begin{aligned} I'(t) &= F'(b(t)) \cdot b'(t) - F'(a(t)) \cdot a'(t) \\ &= f(b(t)) \cdot b'(t) - f(a(t)) \cdot a'(t). \end{aligned}$$

Resaltamos la fórmula obtenida para la derivada.

Derivada de una integral con límites variables e integrando fijo

$$\frac{d}{dt} \int_{a(t)}^{b(t)} f(x)\,dx = f\big(b(t)\big) \cdot b'(t) - f\big(a(t)\big) \cdot a'(t)$$

• **Ejemplo:** Sea $I(t) = \displaystyle\int_{2t}^{3t} \sin x\,dx$. Entonces $I'(t) = 3\sin(3t) - 2\sin(2t)$.

• **Ejemplo:** Sea $I(t) = \displaystyle\int_{t^2}^{t^3} \log x\,dx$, siendo $t > 0$. Entonces

$$I'(t) = \log(t^3)(3t^2) - \log(t^2)(2t) = 9t^2 \log t - 4t \log t = t(9t-4)\log t.$$

En estos ejemplos podríamos calcular explícitamente la antiderivada y después derivar, por ejemplo, en el anterior,

$$\begin{aligned} I(t) = \Big[x\log x - x\Big]_{x=t^2}^{x=t^3} &= t^3\log t^3 - t^3 - t^2\log t^2 + t^2 \\ &= 3t^3\log t - 2t^2\log t - t^3 + t^2, \end{aligned}$$

por tanto

$$\begin{aligned} I'(t) &= 9t^2\log t + 3t^3\cdot t^{-1} - 4t\log t - 2t^2\cdot t^{-1} - 3t^2 + 2t \\ &= 9t^2\log t + 3t^2 - 4t\log t - 2t - 3t^2 + 2t \\ &= (9t^2 - 4t)\,\log t \\ &= t(9t-4)\,\log t. \end{aligned}$$

Llegamos al mismo resultado, pero por un camino más largo.

En los siguientes ejemplos, los integrandos no tienen antiderivada elemental (expresable como combinación de potencias, exponenciales o trigonométricas), así que esta vía se cierra.

• **Ejemplo:** Sea $I(t) = \displaystyle\int_{2t}^{3t} \frac{\sin x}{x}\,dx$, siendo $t > 0$. Entonces

$$I'(t) = \frac{\sin 3t}{3t}\cdot 3 - \frac{\sin 2t}{2t}\cdot 2 = \frac{\sin 3t - \sin 2t}{t}.$$

• **Ejemplo:** Sea $I(t) = \displaystyle\int_{\sqrt{t}}^{2\sqrt{t}} e^{-x^2}\,dx$, siendo $t > 0$. Entonces

$$I'(t) = e^{-4t}\cdot\frac{1}{\sqrt{t}} - e^{-t}\cdot\frac{1}{2\sqrt{t}} = \frac{2e^{-4t} - e^{-t}}{2\sqrt{t}}.$$

• **Ejemplo:** Sea $I(t) = \displaystyle\int_{t}^{2t} \frac{x}{\log x}\,dx$, siendo $t > 1$. Entonces

$$I'(t) = \frac{4t}{\log(2t)} - \frac{t}{\log t}.$$

• **Ejemplo:** Sea $I(t) = \displaystyle\int_{1}^{\log t} \frac{e^x}{x}\,dx$, siendo $t > 1$. Entonces

$$I'(t) = \frac{e^{\log t}}{\log t}\cdot\frac{1}{t} = \frac{1}{\log t}.$$

Para derivar una integral con integrando fijo y límites variables, se emplea el Teorema Fundamental del Cálculo y la Regla de la Cadena, evitando así tener que determinar una antiderivada explícita. Este es realmente el camino más corto y preferible.

7.2.1. Fórmulas de derivación de integrales paramétricas generales

Consideremos una integral paramétrica donde los límites de integración y el parámetro son variables independientes:

$$F(r,s,t) = \int_r^s f(t,x)\,dx.$$

Por el Teorema Fundamental del Cálculo, se tiene

- $\dfrac{\partial F}{\partial s} = f(t,s)$
- $\dfrac{\partial F}{\partial r} = -f(t,r)$

y, si es legítimo el intercambio de la derivada respecto al parámetro con la integral,

- $\dfrac{\partial F}{\partial t} = \displaystyle\int_r^s \frac{\partial f}{\partial t}(t,x)\,dx$

Con estas tres fórmulas y las reglas del Cálculo Diferencial se pueden obtener las derivadas de integrales donde las variables están ligadas entre sí. Por ejemplo,

$$F(a(t),b(t),t) = \int_{a(t)}^{b(t)} f(t,x)\,dx$$

tiene derivada

$$\begin{aligned}\frac{d}{dt}F(a(t),b(t),t) &= \frac{\partial F}{\partial r}(a(t),b(t),t)\cdot a'(t) + \frac{\partial F}{\partial s}(a(t),b(t),t)\cdot b'(t) + \frac{\partial F}{\partial t}(a(t),b(t),t)\\ &= -f(t,a(t))\cdot a'(t) + f(t,b(t))\cdot b'(t) + \int_{a(t)}^{b(t)} \frac{\partial f}{\partial t}(t,x)\,dx.\end{aligned}$$

Esta fórmula se conoce como fórmula de Leibniz.

Este tipo de función es de bastante importancia en Física. Más que memorizar fórmulas como la anterior, es importante poder seguir y reproducir los pasos para obtenerla, pudiendo modificarla para otras situaciones.

Damos una versión más compacta y general.

La fórmula de Leibniz

Si los límites de integración y el integrando son funciones de t,

$$\frac{d}{dt}\int_r^s f(t,x)\,dx = f(t,s)\cdot\frac{\partial s}{\partial t} - f(t,r)\cdot\frac{\partial r}{\partial t} + \int_r^s \frac{\partial f}{\partial t}(t,x)\,dx.$$

7.3. Extremos fijos, integrando paramétrico

En esta sección vamos a considerar integrales paramétricas de la forma

$$\boxed{I(t) = \int_a^b f(t,x)\,dx}$$

donde los extremos a, b son fijos y el integrando depende de x y del parámetro $t \in \mathbb{T}$. La integral puede ser impropia, de hecho lo es en la mayoría de casos interesantes. Por ello llamaremos $\mathbb{X}$ al intervalo de integración y sus extremos a, b pueden ser $\pm\infty$.

Para abarcar todos los casos, llamaremos integrable a una función que lo es en el sentido ordinario o cuya integral es convergente en el sentido impropio. Para que la integral tenga sentido, tenemos que $f(t,x)$ es integrable en $x \in \mathbb{X}$ para cada $t \in \mathbb{T}$ fijo.

7.3.1. Intercambio del límite con la integral

Convergencia puntual

Decimos que $f(t,x)$ converge puntualmente a una función $f_0(x)$ cuando $t \to t_0$ si

$$\lim_{t\to t_0} f(t,x) = f_0(x)$$

para cada $x \in \mathbb{X}$ fijo. Aquí t_0 puede ser también un punto en infinito, $\pm\infty$.

En el caso particular donde t_0 es finito y $f_0(x) = f(t_0,x)$, es decir

$$\lim_{t\to t_0} f(t,x) = f(t_0,x),$$

esto significa que $f(t,x)$ es continua en t_0 como función de t para cada $x \in \mathbb{X}$ fijo.

Si $f(t,x)$ es una función continua de t en t_0, entonces la continuidad en t_0 de la integral equivale a un intercambio de límite con la integral:

$$\lim_{t\to t_0} I(t) = I(t_0) \iff \lim_{t\to t_0} \int_a^b f(t,x)\,dx = \int_a^b f(t_0,x)\,dx = \int_a^b \lim_{t\to t_0} f(t,x)\,dx.$$

No siempre es cierto que este cambio en el orden de operaciones sea válido. Vamos a estudiar una condición general que lo garantiza.

Acotación Uniforme o Dominación

Decimos que $f(t,x)$, definida para $t \in \mathbb{T}$ y $x \in \mathbb{X}$, está **dominada** *por la función $g(x)$ si*

$$|f(t,x)| \le g(x) \quad \forall\, t \in \mathbb{T}, x \in \mathbb{X}, \qquad \int_{\mathbb{X}} g(x)\,dx < \infty.$$

es decir, si $|f(t,x)|$ está uniformemente acotada en el parámetro (o sea, independientemente del parámetro), por una función de integral finita.

Condición para intercambio de límite con integral

Sea $f(t,x) : \mathbb{T} \times \mathbb{X} \to \mathbb{R}, \mathbb{C}$ y supongamos que $t_0 \in \mathbb{T} \cup \{\pm\infty\}$ es tal que:

(1) Para cada $x \in \mathbb{X}$, existe el límite puntual $\lim_{t \to t_0} f(t,x) \stackrel{\text{def}}{=} f_0(x)$.

(2) $f(t,x)$ está dominada (es decir, existe $g(x) \geq 0$ con $|f(t,x)| \leq g(x)$ para cada $t \in \mathbb{T}, x \in \mathbb{X}$ y $\int_{\mathbb{X}} g(x)\,dx < \infty$).

Entonces $f(t,x)$ es integrable en $x \in \mathbb{X}$ para cada $t \in \mathbb{T}$ fijo, $f_0(x)$ es integrable, y la integral paramétrica tiene límite cuando $t \to t_0$ igual a la integral del límite:

$$\lim_{t \to t_0} \int_{\mathbb{X}} f(t,x)\,dx = \int_{\mathbb{X}} \lim_{t \to t_0} f(t,x)\,dx = \int_{\mathbb{X}} f_0(x).$$

Demostración. La demostración rigurosa de este hecho queda fuera del alcance de la teoría expuesta aquí. □

Continuidad de la integral paramétrica

Sea $f(t,x) : \mathbb{T} \times \mathbb{X} \to \mathbb{R}, \mathbb{C}$ tal que:

(1) $t \to f(t,x)$ es continua en $t \in \mathbb{T}$ para cada $x \in \mathbb{X}$ fijo.

(2) $f(t,x)$ está dominada (es decir, existe $g(x) \geq 0$ con $|f(t,x)| \leq g(x)$ para cada $t \in \mathbb{T}, x \in \mathbb{X}$ y $\int_{\mathbb{X}} g(x)\,dx < \infty$).

Entonces $f(t,x)$ es integrable en $x \in \mathbb{X}$ para cada $t \in \mathbb{T}$ fijo, y la integral paramétrica

$$I(t) = \int_{\mathbb{X}} f(t,x)\,dx$$

es continua para $t \in \mathbb{T}$, es decir,

$$\lim_{t \to t_0} \int_{\mathbb{X}} f(t,x)\,dx = \int_{\mathbb{X}} \lim_{t \to t_0} f(t,x)\,dx = \int_{\mathbb{X}} f(t_0,x)\,dx \qquad (t_0 \in \mathbb{T}).$$

Demostración. Es una consecuencia directa del resultado anterior, pues dado cualquier $t_0 \in \mathbb{T}$, existe el límite $f_0(x) = \lim_{t \to t_0} f(t,x) = f(t_0,x)$. □

La condición de dominación de $f(t,x)$ por una función $g(x) \geq 0$ implica por comparación que de hecho $f(t,x)$ es **absolutamente integrable:**

$$\int_{\mathbb{X}} |f(t,x)|\,dx \leq \int_{\mathbb{X}} g(x)\,dx < \infty.$$

La dominación es una condición suficiente pero no necesaria. Cuando $\int f(t,x)\,dx$ converge pero no absolutamente, se necesitan hipótesis alternativas que permitan el intercambio de límite con integral, pero son más complicadas de enunciar.

• **Ejemplo:** Para $t, x \in \mathbb{R}$ consideremos

$$f(t,x) = \frac{t}{1+t^2x^2}, \qquad I(t) = \int_0^\infty f(t,x)\,dx.$$

Claramente $f(t,x)$ es continua simultáneamente en t, x e infinitamente diferenciable. Por cálculo directo de la antiderivada, tenemos

$$I(t) = \int_0^\infty \frac{t\,dx}{1+t^2x^2} = \Big[\arctan(tx)\Big]_{x=-\infty}^{x=\infty} = \begin{cases} \arctan(\infty) - \arctan(-\infty) = \pi & \text{si } t > 0, \\ 0 & \text{si } t = 0, \\ \arctan(-\infty) - \arctan(\infty) = -\pi & \text{si } t < 0, \end{cases}$$

es decir, $I(t) = \pi \operatorname{sgn} t$ donde $\operatorname{sgn} t$ es la función signo. Evidentemente no es continua en 0, aunque el integrando lo sea. Vamos a ver que la función $f(t,x)$ no está dominada por ninguna función integrable $g(x)$. Si se tiene

$$\left|\frac{t}{1+t^2x^2}\right| \le g(x) \quad \forall\, t, x \in \mathbb{R}$$

entonces en particular, poniendo $t = x^{-1}$ si $x \ne 0$, queda $\frac{1}{2}|x^{-1}| \le g(x)$. Pero entonces

$$\infty = \tfrac{1}{2}\int_{-\infty}^\infty \frac{dx}{|x|} \le \int_{-\infty}^\infty g(x)\,dx$$

o sea, $g(x)$ no tiene integral finita.

• **Ejemplo:** Para $t > 0$, $x \in [0, \infty)$ consideremos

$$f(t,x) = e^{-tx}, \qquad I(t) = \int_0^\infty f(t,x)\,dx.$$

Se tiene

$$I(t) = \int_0^\infty e^{-tx}\,dx = \left[-\frac{e^{-tx}}{t}\right]_{x=0}^{x=\infty} = \frac{1}{t}.$$

La continuidad de $I(t)$ para $t > 0$ se puede explicar con los resultados teóricos expuestos, sin calcularla, de la siguiente manera: la acotación obvia $e^{-tx} \le 1$ no sirve pues $\int_0^\infty 1\,dx = \infty$. Usamos la siguiente técnica de separación: fijamos $a > 0$ y consideramos $t \ge a$. Entonces

$$|f(t,x)| = e^{-tx} \le e^{-ax} \overset{\text{def}}{=} g(x) \quad \forall\, t \ge a, x \ge 0$$

y $g(x)$ es integrable, pues $\int_0^\infty g(x)\,dx = a^{-1} < \infty$. Por tanto $I(t)$ es continua para $t \ge a$. Como $a > 0$ era arbitrario, se concluye que $I(t)$ es continua para $t > 0$.

En el ejemplo anterior vemos un caso del fenómeno de dominación local: $|f(t,x)| \le g_a(x)$ para t solamente en un cierto entorno del punto a, no necesariamente en todo el espacio de parámetros $\mathbb{T}$. Al variar a puede cambiar la función $g_a(x)$, pero esto también es suficiente para resolver cuestiones de límites, continuidad o derivabilidad, que son locales (dependen solo de lo que ocurre en un entorno del punto donde se estudian).

• **Ejemplo:** Para $t > 0$, $x \in [0, \infty)$ consideremos

$$f(t,x) = \frac{\sin x}{x} e^{-tx}, \qquad I(t) = \int_0^\infty f(t,x)\,dx = \int_0^\infty \frac{\sin x}{x} e^{-tx}\,dx.$$

Como $\sin x \sim x$ cuando $x \to 0$, no es impropia en 0. La integral es absolutamente convergente para $t > 0$ pero solo condicionalmente convergente para $t = 0$ (ver este ejemplo).

La antiderivada de $f(t,x)$ respecto á x no es elemental, así que debemos analizar la integral indirectamente. Se puede demostrar fácilmente que

$$\left|\frac{\sin x}{x}\right| \leq 1, \quad \therefore \quad |f(t,x)| \leq \frac{|\sin x|}{x} e^{-tx} \leq e^{-tx}$$

pero en $t = 0$ no sirve para dominar pues $\sin x/x$ no es absolutamente integrable. Usando la técnica de separación, para $t \geq a > 0$ con a fijo,

$$|f(t,x)| \leq e^{-tx} \leq e^{-ax} \stackrel{\text{def}}{=} g(x) \quad \forall\, t \geq a, x \geq 0$$

y $\int_0^\infty g(x)\,dx < \infty$, por tanto $I(t)$ es continua para $t > 0$.

De hecho, $I(t)$ también es continua en 0, pero esto ya no es tan sencillo pues requiere otro método para demostrarlo.

Condiciones sencillas para la continuidad de la integral paramétrica

Bajo las siguientes hipótesis:

- *Si los intervalos $\mathbb{X} = [a,b]$ y $\mathbb{T}$ son compactos.*
- *Si la función $f(t,x)$ es continua como función de las dos variables (t,x).*

Entonces la integral paramétrica

$$I(t) = \int_a^b f(t,x)\,dx$$

está definida para $t \in \mathbb{T}$ y es continua.

• **Ejemplo:**

$$\int_0^1 e^{tx^6} \sin(t^3 x)\,dx, \quad \int_a^b \log(1 + t^2 x^2)\,dx, \quad \int_0^1 \frac{dx}{\sqrt{1 + t^2 x^4}}$$

son todas funciones continuas de $t \in \mathbb{R}$.

7.3.2. Intercambio de la derivada con la integral

Derivabilidad de la integral paramétrica

Sea $f(t,x)$ integrable en $x \in \mathbb{X}$ para cada $t \in \mathbb{T}$ fijo, de modo que está definida la función

$$I(t) = \int_{\mathbb{X}} f(t,x)\, dx.$$

Supongamos que $f(t,x)$ es derivable en t para cada $x \in \mathbb{X}$ fijo, es decir, que existe la derivada parcial $\partial f/\partial t$ para $t \in \mathbb{T}, x \in \mathbb{X}$. Si

$$\left|\frac{\partial f}{\partial t}(t,x)\right| \leq g(x), \qquad \int_{\mathbb{X}} g(x)\, dx < \infty,$$

(la derivada parcial está dominada por g), entonces $\partial f/\partial t$ es integrable para cada x fijo, y $I(t)$ es derivable para $t \in \mathbb{T}$, con

$$\frac{\partial I}{\partial t} = \frac{\partial}{\partial t}\int_{\mathbb{X}} f(t,x)\, dx = \int_{\mathbb{X}} \frac{\partial f}{\partial t}(t,x)\, dx.$$

• **Ejemplo:** Para $t > 0$, $x \in [0,\infty)$ consideremos otra vez

$$f(t,x) = \frac{\sin x}{x} e^{-tx}, \qquad I(t) = \int_0^\infty f(t,x)\, dx = \int_0^\infty \frac{\sin x}{x} e^{-tx}\, dx.$$

Calculamos $\dfrac{\partial f}{\partial t}(t,x) = -e^{-tx}\sin x$. Usando la técnica de separación, para $t \geq a > 0$

$$\left|\frac{\partial f}{\partial t}(t,x)\right| \leq e^{-tx} \leq e^{-ax} \stackrel{\text{def}}{=} g(x) \quad \forall\, t \geq a, x \geq 0$$

y $\int_0^\infty g(x)\, dx < \infty$, por tanto $I(t)$ es derivable para $t > 0$ y se tiene

$$I'(t) = \int_0^\infty e^{-tx} \sin x\, dx.$$

No es difícil calcular el valor de esta última integral (se puede integrar por partes dos veces o emplear la fórmula de Euler de variable compleja). Resulta que

$$I'(t) = -\frac{1}{1+t^2},$$

luego $I(t) = K - \arctan t$ para alguna constante K que se puede determinar intercambiando el límite $t \to \infty$ con la integral, justificado por la acotación uniforme $|f(t,x)| \leq e^{-ax}$ para $t \geq a > 0$. Queda por un lado $I(\infty) = K - \arctan(\infty) = K - \frac{\pi}{2}$, y por otro

$$I(\infty) = \lim_{t\to\infty}\int_0^\infty \frac{\sin x}{x} e^{-tx}\, dx = \int_0^\infty \lim_{t\to\infty} \frac{\sin x}{x} e^{-tx}\, dx = \int_0^\infty 0\, dx = 0,$$

por tanto $K = \frac{\pi}{2}$ y $I(t) = \frac{\pi}{2} - \arctan t$ para $t > 0$.

Como hemos comentado anteriormente, $I(t)$ es de hecho continua en 0 aunque esto no es consecuencia directa de los resultados expuestos. Entonces

$$\int_0^\infty \frac{\sin x}{x}\,dx = I(0) = \lim_{t\to 0^+} I(t) = \lim_{t\to 0^+}\left(\frac{\pi}{2} - \arctan t\right) = \frac{\pi}{2}.$$

Como puede constatarse, estas técnicas de cálculo son más avanzadas que las que veníamos empleando para calcular integrales ordinarias o incluso integrales impropias. Sin embargo, son técnicas muy comunes.

Condiciones sencillas para la derivabilidad de la integral paramétrica

Bajo las siguientes hipótesis:

- *Si los intervalos $\mathbb{X} = [a,b]$ y $\mathbb{T}$ son compactos.*
- *Si $f(t,x)$ es continua como función de las dos variables (t,x)*
- *Si existe $\dfrac{\partial f}{\partial t}(t,x)$ y es continua como función de las dos variables.*

Entonces la integral paramétrica

$$I(t) = \int_a^b f(t,x)\,dx$$

está definida para $t \in \mathbb{T}$ y es derivable, con

$$I'(t) = \int_a^b \frac{\partial f}{\partial t}(t,x)\,dx.$$

• **Ejemplo:** Consideremos

$$f(t,x) = e^{-tx}, \qquad I(t) = \int_0^1 f(t,x)\,dx = \int_0^1 e^{-tx}\,dx, \qquad \frac{\partial f}{\partial t} = -xe^{-tx}.$$

Se cumplen las condiciones sencillas con $\mathbb{X} = [0,1]$ y con $\mathbb{T}$ cualquier intervalo compacto, por ejemplo $\mathbb{T} = [-1,1]$. Debe ser entonces

$$I'(t) = -\int_0^1 xe^{-tx}\,dx.$$

En particular $I(0) = \int_0^1 1\,dx = 1$ y $I'(0) = -\int_0^1 x\,dx = -\frac{1}{2}$. Curiosamente, es más difícil ver esto integrando directamente:

$$I(t) = \int_0^1 e^{-tx}\,dx = \left[-\frac{e^{-tx}}{t}\right]_{x=0}^{x=1} = \frac{1-e^{-t}}{t} \qquad (t \neq 0).$$

Efectivamente $I(t)$ es continua también en 0:

$$\lim_{t\to 0} I(t) = \lim_{t\to 0} \frac{1-e^{-t}}{t}. = \lim_{t\to 0} \frac{e^{-t}}{1} = 1 = I(0),$$

y se puede comprobar que

$$I'(0) = \lim_{t \to 0} \frac{I(t) - I(0)}{t} = \lim_{t \to 0} \frac{\dfrac{1 - e^{-t}}{t} - 1}{t} = \lim_{t \to 0} \frac{1 - t - e^{-t}}{t^2} = -\frac{1}{2}.$$

7.3.3. Intercambio de la integral con la integral

Igual que hemos tomado límites y derivado respecto al parámetro en una integral, se puede integrar respecto al parámetro. El teorema que lo justifica es de gran importancia.

Teorema de Fubini para integrales paramétricas

Sea $f(t,x)$ absolutamente integrable en $x \in \mathbb{X}$ para cada $t \in \mathbb{T}$ fijo, es decir,

$$I^*(t) \stackrel{\text{def}}{=} \int_{\mathbb{X}} |f(t,x)|\, dx < \infty \quad \forall\, t \in \mathbb{T}$$

y supongamos que $I^*(t)$ es integrable sobre $t \in \mathbb{T}$. Como $I^*(t) \geq 0$, esto equivale a

$$\int_{\mathbb{T}} I^*(t)\, dt = \int_{\mathbb{T}} \left(\int_{\mathbb{X}} |f(t,x)|\, dx \right) dt < \infty.$$

Entonces

- *Para $t \in \mathbb{T}$ fijo, $x \to f(t,x)$ es integrable sobre $x \in \mathbb{X}$. Por tanto, está definida*

  $$t \to \int_{\mathbb{X}} f(t,x)\, dx, \quad t \in \mathbb{T}.$$

  Además, esta función de t es integrable sobre $t \in \mathbb{T}$, es decir,

  $$\int_{\mathbb{T}} \left(\int_{\mathbb{X}} f(t,x)\, dx \right) dt$$

  *es convergente (se denomina **integral iterada**).*

- *Para $x \in \mathbb{X}$ fijo, $t \to f(t,x)$ es integrable sobre $t \in \mathbb{T}$. Por tanto, está definida*

  $$x \to \int_{\mathbb{X}} f(t,x)\, dt, \quad x \in \mathbb{X}.$$

  Además, esta función de x es integrable sobre $x \in \mathbb{X}$, es decir,

  $$\int_{\mathbb{X}} \left(\int_{\mathbb{T}} f(t,x)\, dt \right) dx$$

  es convergente.

- *Ambas integrales iteradas son iguales, es decir, se puede intercambiar la integral en t con la integral en x:*

  $$\int_{\mathbb{T}} \int_{\mathbb{X}} f(t,x)\, dx\, dt = \int_{\mathbb{X}} \int_{\mathbb{T}} f(t,x)\, dt\, dx.$$

Si $f(t,x) \geq 0$, siempre se puede intercambiar el orden de integración, aunque el resultado del intercambio puede ser la ecuación trivial $\infty = \infty$.

• **Ejemplo:** Sea $f(t,x) = e^{-tx}$. A partir de

$$\int_0^\infty e^{-tx}\,dx = \left[-\frac{e^{-tx}}{t}\right]_{x=0}^{x=\infty} = \frac{1}{t} \qquad (t>0)$$

dados $0<a<b$, consideramos la integral doble

$$\int_a^b \int_0^\infty e^{-tx}\,dx\,dt = \int_a^b \frac{dt}{t} = \log b - \log a = \log\frac{b}{a} < \infty,$$

y como $f(t,x) = e^{-tx} > 0$, esto verifica las condiciones del Teorema de Fubini. Por tanto se puede intercambiar el orden de integración, dando

$$\int_0^\infty \int_a^b e^{-tx}\,dt\,dx = \int_0^\infty \left[-\frac{e^{-tx}}{x}\right]_{t=a}^{t=b} = \int_0^\infty \frac{e^{-ax}-e^{-bx}}{x}\,dx.$$

Intercambiar el orden ha llevado entonces a la fórmula

$$\int_0^\infty \frac{e^{-ax}-e^{-bx}}{x}\,dx = \log\frac{b}{a} \quad (0<a<b).$$

Condiciones sencillas para el Teorema de Fubini

Si $f(t,x)$ continua como función de dos variables en un rectángulo compacto $(t,x) \in [\alpha,\beta]\times[a,b]$, entonces

$$\int_\alpha^\beta \int_a^b f(t,x)\,dx\,dt = \int_a^b \int_\alpha^\beta f(t,x)\,dt\,dx.$$

Variables separadas en el teorema de Fubini

Si $f(x)$ es continua en $[a,b]$ y $g(t)$ es continua en $[\alpha,\beta]$, entonces

$$\int_\alpha^\beta \int_a^b f(x)g(t)\,dx\,dt = \int_a^b \int_\alpha^\beta f(x)g(t)\,dt\,dx = \left(\int_a^b f(x)\,dx\right)\left(\int_\alpha^\beta g(t)\,dt\right).$$

Demostración. Cuando se integra respecto á x, la expresión $g(t)$ es una constante y por tanto se puede sacar fuera de la integral:

$$\int_a^b f(x)g(t)\,dx = g(t)\int_a^b f(x)\,dx$$

y ahora la integral que ha quedado es una constante, y por tanto se puede sacar fuera al integrar respecto á t, quedando el producto de las dos integrales. □

El Teorema de Fubini en un rectángulo compacto es la herramienta básica para desarrollar la teoría de integración en dos variables. Se puede generalizar el teorema a un número finito cualquiera de variables. Como aquí se trata fundamentalmente con funciones de una variable, dejaremos el tema.

7.4. La Función Gamma

7.4.1. Definición y propiedades básicas

Definición de $\Gamma(s)$

*La **función gamma** está definida para $s > 0$ por la integral paramétrica*

$$\Gamma(s) = \int_0^\infty x^{s-1} e^{-x}\, dx.$$

Se comprueba que efectivamente la integral es convergente para $s > 0$, pues

- $x^{s-1}e^{-s} \sim x^{s-1}$ cuando $x \to 0^+$.
- $x^{s-1}e^{-x} = o(e^{-x/2})$ cuando $x \to \infty$, pues $x^{s-1}e^{-x/2} = o(1)$. para cualquier s.

La ecuación funcional de Γ

Para $s > 0$ se tiene $\Gamma(s+1) = s\Gamma(s)$

Demostración. Integrando por partes,

$$\Gamma(s+1) = \int_0^\infty \underbrace{x^s}_{u}\, \underbrace{e^{-x}\, dx}_{dv} = \left[-x^s\, e^{-x}\right]_0^\infty + \int_0^\infty s x^{s-1}\, e^{-x}\, dx$$

$$= 0 + s \int_0^\infty x^{s-1}\, e^{-x}\, dx = s\,\Gamma(s).$$

Hemos usado que $x^s e^{-x} = o(1)$ cuando $x \to \infty$. □

Demostración alternativa. Usando el Teorema Fundamental,

$$0 = \left[x^s e^{-x}\right]_0^\infty = \int_0^\infty \frac{d}{dx}\left(x^s e^{-x}\right) dx = \int_0^\infty \left(s x^{s-1} e^{-x} - x^s e^{-x}\right) dx = s\Gamma(s) - \Gamma(s+1).$$

□

Cálculo recursivo de $\Gamma(s)$

Para cualquier $n = 1, 2, 3, \ldots$, se tiene

$$\Gamma(s) = (s-1)(s-2)\cdots(s-n)\,\Gamma(s-n) \quad (n < s \le n+1)$$

Por tanto los valores de $\Gamma(s)$ para $s > 0$ están determinados por sus valores en $(0, 1]$.

Demostración. Usando la ecuación funcional, por inducción vemos que

$$\Gamma(s) = (s-1)\,\Gamma(s-1) = (s-1)(s-2)\,\Gamma(s-2) = \cdots$$
$$\cdots = (s-1)(s-2)\cdots(s-n)\,\Gamma(s-n).$$

□

Prolongación de $\Gamma(s)$ para $s \leq 0$

Para $s \leq 0$, se define

$$\Gamma(s) = \begin{cases} \dfrac{1}{s(s+1)\cdots(s+n-1)}\,\Gamma(s+n) & \text{si } -n < s < -n+1 \text{ con } n = 1,2,3,\dots \\ \infty & \text{si } s = 0,-1,-2,-3,\dots \end{cases}$$

Se trata simplemente de sustituir $s \to s+n$ en la anterior relación para ponerla en la forma

$$\Gamma(s+n) = (s+n-1)(s+n-2)\cdots(s+1)s\,\Gamma(s)$$

y por tanto

$$\Gamma(s) = \frac{1}{s(s+1)\cdots(s+n-1)}\,\Gamma(s+n).$$

Si $n \in \mathbb{N}$ entonces

$$-n < s < -n+1 \implies 0 < s+n < 1$$

y entonces $\Gamma(s+n)$ está definida por la integral.

Como $\Gamma(0) = \int_0^\infty x^{-1}e^{-x}\,dx = \infty$, para que se siga cumpliendo la ecuación funcional nos vemos obligados a definir

$$\Gamma(0) = \Gamma(-1) = \Gamma(-2) = \Gamma(-3) = \cdots = \infty.$$

La fórmula de Reflexión

Para todo $s \notin \mathbb{Z}$, se tiene

$$\Gamma(s)\,\Gamma(1-s) = \frac{\pi}{\sin(\pi s)}$$

Demostración. La demostración de la fórmula de reflexión queda fuera del alcance de la teoría desarrollada aquí. □

Si $s = 0,-1,-2,-3,\dots$ se tiene $\Gamma(s) = \infty$, mientras que en $1-s = 1,2,3,\dots$ es el factorial, y viceversa si $s = 1,2,3,\dots$ entonces $1-s = 0,-1,-2,-3,\dots$. Por otra parte en este caso $\sin(\pi s) = 0$. Si interpretamos $c\cdot\infty = \infty$ y $c/0 = \infty$ para una constante $c > 0$, la fórmula de reflexión sigue siendo válida para $s \in \mathbb{Z}$.

7.4.2. Valores especiales de Γ

Interpolación del Factorial

Para $n = 0, 1, 2, 3, \cdots$, *se tiene* $\Gamma(n+1) = n!$

Demostración.

$$\Gamma(1) = \int_0^\infty e^{-x}\, dx = \left[-e^{-x}\right]_0^\infty = 1$$

y por la ecuación funcional

$$\Gamma(n+1) = n(n-1)(n-2)\cdots(1)\Gamma(1) = n!.$$

□

Es tradicional la traslación por 1 en la fórmula, aunque igualmente válido sería

$$\Gamma(n) = (n-1)! \qquad (n \in \mathbb{N}).$$

Esta traslación se debe al $s-1$ en la definición de $\Gamma(s)$.

El valor de $\Gamma(\frac{1}{2})$

Se tiene

$$\Gamma(\tfrac{1}{2}) = \int_0^\infty x^{-1/2}\, e^{-x}\, dx = \int_{-\infty}^\infty e^{-x^2}\, dx = \sqrt{\pi}.$$

Demostración. Cambiando de variable en la definición con $x = u^2$, $dx = 2u\, du$ se tiene

$$\Gamma(\tfrac{1}{2}) = \int_0^\infty u^{-1}\, e^{-u^2}\, 2u\, du = 2\int_0^\infty e^{-u^2}\, du = \int_{-\infty}^\infty e^{-u^2}\, du.$$

El truco para evaluar $\Gamma(\frac{1}{2})$ consiste en convertir la integral en doble, aplicar el Teorema de Fubini y cambiar a coordenadas polares:

$$\begin{aligned}
\Gamma(\tfrac{1}{2})^2 &= \left(\int_{-\infty}^\infty e^{-u^2}\, du\right)^2 = \int_{-\infty}^\infty e^{-x^2}\, dx \int_{-\infty}^\infty e^{-y^2}\, dy = \int_{-\infty}^\infty \int_{-\infty}^\infty e^{-x^2}\, e^{-y^2}\, dx\, dy \\
&= \iint_{\mathbb{R}^2} e^{-(x^2+y^2)}\, dx\, dy \\
&= \int_0^{2\pi} \int_0^\infty e^{-r^2}\, r\, dr\, d\theta = 2\pi \int_0^\infty e^{-r^2}\, r\, dr = \pi\left[-e^{-r^2}\right]_0^\infty = \pi
\end{aligned}$$

por tanto, como claramente $\Gamma(s) > 0$ para $s > 0$, queda $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. □

"Un matemático es alguien para quien $\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}$ es tan obvio como que dos más dos son cuatro."

– Lord Kelvin

La integral parcial

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-u^2}\,du$$

se conoce como la **función de error de Gauss** y el integrando

$$e^{-x^2}$$

es una **campana de Gauss** o simplemente **Gaussiana.** Es fundamental en la Probabilidad y Estadística, donde se la conoce también como **distribución normal.**

$\Gamma(s)$ en los medio enteros positivos

Para $n = 0, 1, 2, 3, \ldots$ se tiene

$$\Gamma\left(n + \frac{1}{2}\right) = \frac{(2n-1)!!}{2^n}\sqrt{\pi} = \frac{(2n)!}{2^{2n}\,n!}\sqrt{\pi},$$

donde $m!!$ es el ***doble factorial,*** *definido para $m \in \mathbb{N}$ por*

$$m!! = \begin{cases} m(m-2)(m-4)\cdots(3)(1) & \text{si } m \text{ es impar} \\ m(m-2)(m-4)\cdots(4)(2) & \text{si } m \text{ es par.} \end{cases}$$

Demostración. Usando la ecuación funcional iterada

$$\Gamma(s) = (s-1)(s-2)\cdots(s-n)\,\Gamma(s-n)$$

se tiene

$$\begin{aligned}
\Gamma\left(n+\frac{1}{2}\right) &= \left(n-\frac{1}{2}\right)\left(n-\frac{3}{2}\right)\cdots\frac{1}{2}\,\Gamma\left(\frac{1}{2}\right) \\
&= \frac{(2n-1)(2n-3)\cdots 5\cdot 3\cdot 1}{2^n}\sqrt{\pi} = \frac{(2n-1)!!}{2^n}\sqrt{\pi} \\
&= \frac{(2n)(2n-1)(2n-2)(2n-3)\cdots 5\cdot 4\cdot 3\cdot 2\cdot 1}{2^n(2n)(2n-2)(2n-4)\cdots 4\cdot 2}\sqrt{\pi} \\
&= \frac{(2n)!}{2^n\,(2\cdot n)\,(2\cdot(n-1))\cdots(2\cdot 3)\cdot(2\cdot 2)\cdot(2\cdot 1)}\sqrt{\pi} \\
&= \frac{(2n)!}{2^n\cdot 2^n\,n!}\sqrt{\pi} = \frac{(2n)!}{2^{2n}\,n!}\sqrt{\pi}.
\end{aligned}$$

□

• **Ejemplo:** $\Gamma\left(\frac{3}{2}\right) = \frac{1}{2}\Gamma\left(\frac{1}{2}\right) = \frac{\sqrt{\pi}}{2}$.

• **Ejemplo:** $\Gamma\left(\frac{5}{2}\right) = \frac{3}{2}\cdot\frac{1}{2}\cdot\Gamma\left(\frac{1}{2}\right) = \frac{3\sqrt{\pi}}{4}$.

• **Ejemplo:** $\Gamma\left(\frac{7}{2}\right) = \frac{5}{2}\cdot\frac{3}{2}\cdot\frac{1}{2}\cdot\Gamma\left(\frac{1}{2}\right) = \frac{15\sqrt{\pi}}{8}$.

$\Gamma(s)$ en los medio enteros negativos

Para $n = 1, 2, 3, \ldots$ se tiene

$$\Gamma\left(-n+\frac{1}{2}\right) = \frac{(-1)^n 2^{2n}\, n!\sqrt{\pi}}{(2n)!}.$$

Demostración. Para enteros negativos, usamos

$$\Gamma(s) = \frac{1}{s(s+1)\cdots(s+n-1)}\Gamma(s+n).$$

Entonces

$$\begin{aligned}
\Gamma\left(-n+\frac{1}{2}\right) &= \frac{1}{\left(-n+\frac{1}{2}\right)\left(-n+\frac{3}{2}\right)\cdots\left(-\frac{1}{2}\right)}\Gamma\left(\frac{1}{2}\right) = \frac{2^n\sqrt{\pi}}{(-2n+1)(-2n+3)\cdots(-3)(-1)} \\
&= \frac{(-1)^n 2^n\sqrt{\pi}}{(2n-1)(2n-3)\cdots(3)(1)} = \frac{(-1)^n 2^n\,(2n)(2n-2)\cdots(4)(2)\sqrt{\pi}}{(2n)!} \\
&= \frac{(-1)^n 2^n\, 2^n n!\sqrt{\pi}}{(2n)!} = \frac{(-1)^n 2^{2n}\, n!\sqrt{\pi}}{(2n)!}.
\end{aligned}$$

□

• **Ejemplo:** $\Gamma\left(-\frac{1}{2}\right) = \frac{\sqrt{\pi}}{\left(-\frac{1}{2}\right)} = -2\sqrt{\pi}$.

• **Ejemplo:** $\Gamma\left(-\frac{3}{2}\right) = \frac{\sqrt{\pi}}{\left(-\frac{3}{2}\right)\left(-\frac{1}{2}\right)} = \frac{4\sqrt{\pi}}{3}$.

• **Ejemplo:** $\Gamma\left(-\frac{5}{2}\right) = \frac{\sqrt{\pi}}{\left(-\frac{5}{2}\right)\left(-\frac{3}{2}\right)\left(-\frac{1}{2}\right)} = -\frac{8}{15}\sqrt{\pi}$.

Otros valores. No se conocen fórmulas cerradas para $\Gamma(1/k)$ para enteros $k > 2$. Se sabe que son números transcendentes para $k = 3, 4$. Si se conocen diversas relaciones entre estos valores.

7.4.3. La Función Beta

Definición de la función beta

*La **función beta** o **integral beta** se define por*

$$B(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1}\,dx \qquad (p,q>0)$$

Para verificar que es convergente para $p, q > 0$, en 0 se compara con

$$x^{p-1}(1-x)^{q-1} \sim x^{p-1}$$

que converge si y sólo si $p-1 > -1$, o sea para $p > 0$. Haciendo el cambio $x \to 1-x$ se ve que en 1 converge si y sólo si $q > 0$.

Simetría de la función beta

$$B(p,q) = B(q,p).$$

Demostración. Si cambiamos de variable con $x \to 1-x$ queda

$$B(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1}\,dx = \int_0^1 (1-x)^{p-1}x^{q-1}\,dx = B(q,p).$$

□

Integral trigonométrica para la función beta

La función beta es expresable como la integral trigonométrica

$$B(p,q) = 2\int_0^{\pi/2} (\sin\theta)^{2p-1}(\cos\theta)^{2q-1}\,d\theta \qquad (p,q>0).$$

Demostración. Cambiando de variable con

$$\left\{\begin{aligned} x &= \sin^2\theta \\ 1-x &= \cos^2\theta \end{aligned}\right\} \qquad dx = 2\sin\theta\cos\theta\,d\theta, \qquad \theta\in[0,\pi/2]$$

queda

$$\begin{aligned} B(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1}\,dx &= 2\int_0^{\pi/2} (\sin\theta)^{2p-2}(\cos\theta)^{2q-2}\,\sin\theta\cos\theta\,d\theta \\ &= 2\int_0^{\pi/2} (\sin\theta)^{2p-1}(\cos\theta)^{2q-1}\,d\theta. \end{aligned}$$

□

Relación entre las funciones Beta y Gamma

$$B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)} \qquad (p,q>0).$$

Demostración. Usamos la misma técnica empleada para calcular $\Gamma\left(\frac{1}{2}\right)$.

Primero, en la definición de $\Gamma(s)$ ponemos $x = u^2$:

$$\Gamma(s) = \int_0^\infty x^{s-1}\, e^{-x}\, dx = \int_0^\infty u^{2s-2}\, e^{-u^2}\, 2u\, du = 2\int_0^\infty u^{2s-1}\, e^{-u^2}\, du.$$

Esta fórmula es útil por sí misma, así que la resaltamos:

$$\boxed{\Gamma(s) = 2\int_0^\infty u^{2s-1}\, e^{-u^2}\, du \qquad (s>0).}$$

Por el Teorema de Fubini,

$$\begin{aligned}\Gamma(p)\Gamma(q) &= 4\int_0^\infty \int_0^\infty u^{2p-1}\, v^{2q-1}\, e^{-u^2}\, e^{-v^2}\, du\, dv \\ &= 4\iint_{u,v>0} u^{2p-1}\, v^{2q-1}\, e^{-(u^2+v^2)}\, du\, dv\end{aligned}$$

Finalmente, cambiando a coordenadas polares poniendo $u = r\sin\theta, v = r\cos\theta$, queda

$$\begin{aligned}\Gamma(p)\Gamma(q) &= 4\int_0^{\pi/2} \int_0^\infty r^{2p-1}\, r^{2q-1}\, (\sin\theta)^{2p-1}(\cos\theta)^{2q-1}\, e^{-r^2}\, r\, dr\, d\theta \\ &= 2\int_0^\infty r^{2p+2q-1}\, e^{-r^2}\, dr \cdot 2\int_0^{\pi/2} (\sin\theta)^{2p-1}(\cos\theta)^{2q-1}\, d\theta \\ &= \Gamma(p+q)\, B(p,q).\end{aligned}$$

□

Integrales trigonométricas sobre $[0,\pi/2]$

Cambiando $r = 2p-1, s = 2q-1$, o sea $p = \frac{r+1}{2}, q = \frac{s+1}{2}$, en la fórmula trigonométrica para la integral beta y expresando beta en términos de gamma, queda

$$\int_0^{\pi/2} \sin^r\theta\cos^s\theta\, d\theta = \frac{1}{2}\frac{\Gamma(\frac{r+1}{2})\Gamma(\frac{s+1}{2})}{\Gamma(\frac{r+s+2}{2})} \qquad (r,s>-1).$$

En particular, poniendo $r = 0$ y $s = 0$,

$$\int_0^{\pi/2} (\sin\theta)^r\, d\theta = \int_0^{\pi/2} (\cos\theta)^r\, d\theta = \frac{\sqrt{\pi}}{2}\frac{\Gamma(\frac{r+1}{2})}{\Gamma(\frac{r+2}{2})} \qquad (r>-1).$$

• **Ejemplo:**

$$B(\tfrac{1}{2}, \tfrac{1}{2}) = 2\int_0^{\pi/2} d\theta = \pi = \frac{\Gamma(\frac{1}{2})^2}{\Gamma(1)} = \Gamma(\tfrac{1}{2})^2$$

• **Ejemplo:** Esto permite calcular recursivamente usando la ecuación funcional

$$\begin{aligned}\int_0^{\pi/2} \sin^7\theta\cos^9\theta\, d\theta &= \frac{1}{2}\cdot\frac{\Gamma(4)\Gamma(5)}{\Gamma(9)} = \frac{1}{2}\cdot\frac{3!4!}{8!}\\ &= \frac{1}{2}\cdot\frac{3!}{8\cdot 7\cdot 6\cdot 5}\\ &= \frac{1}{2}\cdot\frac{1}{8\cdot 7\cdot 5} = \frac{1}{560}\end{aligned}$$

• **Ejemplo:** Con potencias pares salen los valores en los medio enteros

$$\begin{aligned}\int_0^{\pi/2} \sin^4\theta\cos^5\theta\, d\theta &= \frac{1}{2}\cdot\frac{\Gamma(\frac{5}{2})\Gamma(3)}{\Gamma(\frac{11}{2})}\\ &= \frac{1}{2}\cdot\frac{\left(\frac{3}{2}\cdot\frac{1}{2}\cdot\sqrt{\pi}\right)\cdot 2}{\left(\frac{9}{2}\cdot\frac{7}{2}\cdot\frac{5}{2}\cdot\frac{3}{2}\cdot\frac{1}{2}\cdot\sqrt{\pi}\right)}\\ &= \frac{3\cdot 2^3}{(9\cdot 7\cdot 5\cdot 3)} = \frac{8}{315}\end{aligned}$$

• **Ejemplo:** Potencias naturales del seno:

$$\int_0^{\pi/2} \sin^7\theta\, d\theta = \frac{\sqrt{\pi}}{2}\frac{\Gamma(4)}{\Gamma(\frac{9}{2})} = \frac{\sqrt{\pi}}{2}\frac{3!}{\frac{7}{2}\cdot\frac{5}{2}\cdot\frac{3}{2}\cdot\frac{1}{2}\cdot\sqrt{\pi}} = \frac{16}{35}$$

• **Ejemplo:** Potencias reales del seno:

$$\int_0^{\pi/2} \frac{d\theta}{\sqrt{\sin\theta}} = \frac{\sqrt{\pi}}{2}\frac{\Gamma(\frac{1}{4})}{\Gamma(\frac{3}{4})},$$

o

$$\int_0^{\pi/2} \sqrt{\sin\theta}\, d\theta = \frac{\sqrt{\pi}}{2}\frac{\Gamma(\frac{3}{4})}{\Gamma(\frac{5}{4})}$$

• **Ejemplo:** Cambiando r por $-r$ queda

$$\int_0^{\pi} \frac{dx}{(\sin x)^r} = \sqrt{\pi}\,\frac{\Gamma(\frac{1-r}{2})}{\Gamma(1-\frac{r}{2})} \qquad (r<1).$$

En muchos casos, las integrales de seno y coseno sobre otros intervalos como $[0,\pi]$ y $[0,2\pi]$ son múltiplos de las integrales sobre $\left[0,\frac{\pi}{2}\right]$ por simetría y periodicidad, recordando que

$$\left\{\begin{aligned}\sin\left(\frac{\pi}{2}-\theta\right) &= \cos\theta,\\ \cos\left(\frac{\pi}{2}-\theta\right) &= \sin\theta.\end{aligned}\right\} \qquad \left\{\begin{aligned}\sin(\pi-\theta) &= \sin\theta,\\ \cos(\pi-\theta) &= -\cos\theta.\end{aligned}\right\}$$

8. Ecuaciones Diferenciales

Recordemos el método de derivación implícita. Cuando se tiene una ecuación, como

$$x^2 + y^2 = 1,$$

y sabemos que en esta ecuación y es función de x, entonces aún sin despejarla explícitamente, podemos derivar respecto á x, obteniendo

$$x + yy' = 0, \quad \therefore \quad y' = -\frac{x}{y}.$$

El tipo de ecuación que resulta, donde se encuentran la variable x, la variable y que depende de x (o sea, que es función de x), y la derivada y' de y, es un ejemplo de ecuación diferencial. También es una ecuación diferencial la que resulta de derivar más veces:

$$1 + (y')^2 + yy'' = 0.$$

Definición de Ecuación Diferencial Ordinaria (E.D.O.)

Una ecuación diferencial ordinaria es una ecuación del tipo

$$F(x, y, y', y'', \ldots, y^{(n)}) = 0$$

que relaciona una sola variable independiente x y una función o variable dependiente de ella, y, junto con sus derivadas hasta un cierto orden n. Se dice entonces que la ecuación es de ***orden** n.*

• **Ejemplo:** La ecuación $x + yy' = 0$ es de orden 1.

• **Ejemplo:** La ecuación $1 + (y')^2 + yy'' = 0$ es de orden 2.

Decimos que una ecuación diferencial está expresada en forma **explícita** si podemos escribirla como

$$y^{(n)} = f(x, y, y', y'', \ldots, y^{(n-1)}),$$

es decir, si la derivada más alta de y (el orden de la ecuación) se puede expresar en función de la variable x y las derivadas de orden más bajo de y. De lo contrario diremos que la ecuación es **implícita.**

• **Ejemplo:** $y' = -\frac{x}{y}$ está en forma explícita.

• **Ejemplo:** $x\sin(y') + (y')^2e^x = 0$ es implícita.

Resolver una ecuación diferencial es poder expresar la variable dependiente o función y en términos sólo de la variable independiente x.

Dicho formalmente, una **solución** de la E.D. es una función $y = f(x)$, definida en algún intervalo abierto I y derivable hasta el orden de la ecuación, tal que al sustituir la expresión $y = f(x)$ en dicha ecuación, ésta se cumple.

• **Ejemplo:** Poniendo $y = \sqrt{1 - x^2}$ para $x \in I = (-1, 1)$, vemos que esta y resuelve la ecuación $y' = -x/y$, puesto que efectivamente

$$y' = \frac{d}{dx}(1 - x^2)^{1/2} = \frac{1}{2}(1 - x^2)^{-1/2} \cdot (-2x) = -\frac{x}{\sqrt{1 - x^2}} = -\frac{x}{y}.$$

Integrar, en el sentido de hallar antiderivadas, es un caso particular de resolución de E.D.O. En efecto, el problema de hallar

$$\int f(x)\, dx = F(x) + C$$

se puede plantear como la resolución de la E.D. explícita

$$y' = f(x)$$

siendo $y = F(x) + C$ la solución general. Así pues, se puede pensar que las técnicas de resolución de E.D. son generalizaciones de las técnicas de integración. Sin embargo, no es cierto en general que toda E.D. se pueda resolver haciendo sucesivamente integrales. Puede ser más complicado.

Se dice que una ecuación diferencial es **integrable por cuadraturas** si se puede resolver expresando su solución como una serie de integrales sucesivas.

Al igual que se exponen los diversos métodos de integración y las técnicas de reducción de unas integrales a otras, se hace lo mismo con las ecuaciones diferenciales.

Las ecuaciones que vamos a estudiar van a ser de momento todas explícitas de primer orden, es decir, de la forma

$$y' = f(x, y).$$

Empezamos con un tipo sencillo de E.D.O. cuyo método de solución consiste esencialmente en calcular dos antiderivadas.

8.1. Variables separadas

E.D.O. de variables separadas

Decimos que una E.D.O. es de ***variables separadas*** *si es de la forma*

$$y' = f(x)g(y)$$

Para resolver una ecuación de variables separadas, observamos que la notación diferencial sugiere las manipulaciones

$$\frac{dy}{dx} = f(x)g(y) \quad \therefore \quad \frac{dy}{g(y)} = f(x)\,dx,$$

donde en esta última ecuación, como el nombre indica, las variables están separadas, x a un lado, y al otro.

Si integramos ambos lados de la ecuación separada, tratando x, y como variables, quedará una ecuación entre sus antiderivadas, con la introducción de una constante:

$$G_*(y) = F(x) + C, \quad G_*'(y) = \frac{1}{g(y)}, \quad F'(x) = f(x).$$

Suponiendo que $\tilde{G}$ es invertible (en algún intervalo), podemos despejar, dejando

$$y = G_*^{-1}(F(x) + C).$$

Por supuesto estos paso hay que justificarlos; por ejemplo, debe ser $g(y) \neq 0$ si vamos a integrar $1/g$, y debe tener sentido la composición.

Método de resolución de una E.D. de variables separadas

Sea

$$y' = f(x)\,g(y)$$

siendo f, g continuas, $g(y) \neq 0$. Entonces la solución general de la ecuación se obtiene manipulando la notación diferencial para separar las variables e integrando:

$$\frac{dy}{dx} = f(x)g(y) \iff \frac{dy}{g(y)} = f(x)\,dx \iff \int \frac{dy}{g(y)} = \int f(x)\,dx + C$$

intentando finalmente expresar y explícitamente como función de x.

• **Ejemplo:** Resolvamos la ecuación que provenía del círculo como ecuación implícita:

$$y' = -\frac{x}{y} \iff \frac{dy}{dx} = -\frac{x}{y}$$
$$\iff y\,dy = -x\,dx \iff \frac{y^2}{2} = -\frac{x^2}{2} + C \iff y^2 = C - x^2$$

[12] donde C representa la frase "cualquier constante," de modo que es legítimo escribir cosas como $2C = C$... no tiene un valor concreto.

Si especificamos **condiciones iniciales** tales como $y(0) = 1$, entonces la constante queda determinada: $C = 1$.

• **Ejemplo:** $y' + y\sin(x) = 0$ es de esta forma:

$$\frac{dy}{dx} = -y\sin(x) \iff \frac{dy}{y} = -\sin(x)\,dx \iff \int \frac{dy}{y} = -\int \sin(x)\,dx + C$$
$$\iff \log|y| = \cos(x) + C \iff |y| = e^{\cos(x)+C} = e^{\cos(x)} \cdot e^{C}$$
$$\iff y = \pm K e^{\cos(x)}$$

siendo $K = e^C$ una constante que en principio es positiva, pero como tenemos $\pm K$ de hecho puede tener cualquier signo. Es más, incluso tomando $K = 0$, dando $y = 0$, es una solución trivial de la ecuación, aunque no haya resultado de la integración. Por tanto la solución general es $y = Ke^{\cos(x)}$ con $K \in \mathbb{R}$.

• En esta exposición, que está orientada a técnicas más que a teoría, no nos vamos a preocupar de justificar con rigor, la existencia y/o la unicidad de las soluciones de las ecuaciones diferenciales que vayamos estudiando. Nos limitaremos a indicar los métodos de resolución y a comentar que se puede verificar sustituyendo en la ecuación que el método ha llevado de hecho a una solución.

Al igual que ocurre con los métodos de integración, lo importante es conocer y saber emplear el **método** de resolución, separando las variables e integrando, que aprender la fórmula de memoria.

A continuación vamos estudiando otros tipos de ecuaciones diferenciales que se pueden reducir a una de variables separadas utilizando diversas técnicas.

Ecuación de la forma

$$y' = f(ax + by)$$

donde $a, b \neq 0$ (si no ya es de variables separadas). Se resuelve cambiando

$$z = ax + by,$$

que la convierte en una de variables separadas.

En efecto, escribiendo

$$z = ax + by, \quad z' = a + by', \quad y' = \frac{z' - a}{b},$$

la ecuación se cambia por

$$\frac{z' - a}{b} = f(z), \quad z' = a + bf(z),$$

12

que en notación diferencial es

$$\frac{dz}{dx} = a + bf(z) \quad \therefore \quad \frac{dz}{a+bf(z)} = dx,$$

luego se resuelve integrando:

$$G(z) = \int \frac{dz}{a+bf(z)} = x + C,$$

lo cual es una ecuación implícita para z y por tanto para y. Suponiendo que se puede invertir G, tendremos $z = ax + by = G^{-1}(x+C)$, lo cual nos dará y en función de x.

• **Ejemplo:** $y' - e^{x+y} = -1$. Poniendo $z = x + y$ se tiene $z' = 1 + y' = e^z$, por tanto $e^{-z}dz = dx$ y $-e^{-z} = x+C$, por tanto $e^{-z} = -x-C$ y finalmente $-z = -x-y = \log(-x-C)$, $y = -x - \log(-x-C)$.

8.2. Homogéneas

Ecuaciones homogéneas

Una ecuación diferencial de la forma

$$y' = f\left(\frac{y}{x}\right)$$

se puede resolver haciendo el cambio de variable $z = \frac{y}{x}$, es decir, $y = xz$, lo cual la transforma en una de variables separadas.

Poniendo $y = xz$, se tiene

$$y' = z + xz',$$

la ecuación se transforma en

$$z + xz' = f(z) \iff xz' = f(z) - z \iff x\frac{dz}{dx} = f(z) - z$$
$$\iff \frac{dz}{f(z)-z} = \frac{dx}{x} \iff G(z) \stackrel{\text{def}}{=} \int \frac{dz}{f(z)-z} = \log|x| + C$$

y si podemos invertir G, tendremos

$$z = \frac{y}{x} = G^{-1}(\log|x| + C) \quad \therefore \quad y = x\,G^{-1}(\log|x| + C).$$

• **Ejemplo:** $y' = \dfrac{2xy - y^2}{x^2}$ es homogénea pues el lado derecho es

$$2\,\frac{y}{x} - \left(\frac{y}{x}\right)^2 = 2z - z^2, \quad z = \frac{y}{x}.$$

Queda

$$y' = z + xz' = 2z - z^2 \iff x\frac{dz}{dx} = z - z^2 \iff \frac{dz}{z-z^2} = \frac{dx}{x}.$$

Descomponiendo en fracciones parciales,

$$\frac{1}{z-z^2} = \frac{1}{z(1-z)} = \frac{1}{z} + \frac{1}{1-z}$$

se obtiene

$$\log z - \log(1-z) = \log x + C$$

o sea

$$\log \frac{z}{1-z} = \log Kx$$

donde $C = \log K$, siendo $K > 0$ ya que C es cualquier constante real. Entonces

$$\frac{z}{1-z} = \frac{\dfrac{y}{x}}{1-\dfrac{y}{x}} = Kx$$

y simplificando

$$\begin{aligned} \frac{y}{x} = \left(1-\frac{y}{x}\right)Kx = Kx - Ky \iff & y = Kx^2 - Kxy \\ \iff & y(1+Kx) = Kx^2 \\ \iff & y = \frac{Kx^2}{Kx+1} \end{aligned}$$

y si tenemos en cuenta que $K > 0$ es cualquier constante, podemos dividir numerador y denominador por K para dejar la solución general como

$$y = \frac{x^2}{x+K} \quad (K > 0).$$

Ahora, si pusiéramos también $K = 0$, nos daría $y = x$, que también es una solución de la ecuación original, e incluso poniendo $K = \infty$ daría $y = 0$, que es otra solución. Esto es bastante típico de los métodos de solución. Otra manera de obtener estas nuevas soluciones era darse cuenta que al dividir por $z - z^2$ estábamos suponiendo que no era nulo. Tomando $z = 0, 1$ que anulan este factor, saldrían estas otras soluciones a partir de la relación $y = xz$.

8.3. Exactas

Sea $f(x, y)$ una función de dos variables x, y. Las **derivadas parciales** de f se obtienen derivando respecto a una de las variables considerando la otra como constante:

(1) $\dfrac{\partial f}{\partial x}$ ó f_x ó $\partial_x f$: derivar respecto á x considerando y constante.

(2) $\dfrac{\partial f}{\partial y}$ ó f_y ó $\partial_y f$: derivar respecto á y considerando x constante.

• **Ejemplo:** Si $f(x, y) = xy + x^2 + 2y^2$, entonces $f_x = y + 2x$, $f_y = x + 4y$.

• **Ejemplo:** Si $f(x,y) = \sin(x)\cos(y)$, entonces $f_x = \cos(x)\cos(y)$, $f_y = -\sin(x)\sin(y)$.

• **Ejemplo:** Si $f(x,y) = e^x \sin(y)$ entonces $f_x = e^x \sin(y)$, $f_y = e^x \cos(y)$.

• **Ejemplo:** Si $f(x,y) = \log(1+xy)$ entonces $f_x = \dfrac{y}{1+xy}$, $f_y = \dfrac{x}{1+xy}$.

Ecuación diferencial exacta

Una ecuación diferencial es exacta si se puede escribir en la forma

$$P(x,y)\,dx + Q(x,y)\,dy = 0 \quad \text{con} \quad \frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}.$$

Lema de Poincaré

Supongamos que P, Q son funciones derivables con derivadas parciales continuas en un dominio estrellado E del plano, lo cual significa que existe un "punto base" $b \in E$ tal que dado cualquier otro punto $a \in E$, el segmento que une a con b se halla dentro de E. Entonces la condición

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$$

equivale a la existencia de una función $f(x,y)$ definida en E tal que

$$P = \frac{\partial f}{\partial x}, \quad Q = \frac{\partial f}{\partial y}.$$

Tal f es única salvo por la suma de una constante y se denomina función potencial.

El Cálculo Diferencial en varias variables nos proporciona entonces un método de solución para una E.D. exacta.

Solución de una ecuación diferencial exacta

Sea $Pdx + Qdy = 0$ una ecuación diferencial exacta en un dominio estrellado del plano. Sea f tal que $f_x = P$, $f_y = Q$ (potencial). Entonces la solución general de la ecuación viene dada en forma implícita por

$$f(x,y) = C$$

donde C es cualquier constante.

Por tanto se trata de encontrar el potencial f dadas P, Q con $P_y = Q_x$. Para ello, se puede integrar una de las ecuaciones, por ejemplo,

$$\frac{\partial f}{\partial x} = P$$

respecto a la variable de derivación, x, considerando la otra, y, constante. Se obtiene

$$f(x,y) = \int P(x,y)\,dx + C(y)$$

donde $C(y)$ es la "constante" de integración, pero que en realidad es una función de y, ya que ésta se estaba considerando como constante, es decir, con un valor fijo. Para determinar f, derivamos la ecuación anterior respecto á y:

$$Q(x,y) = \frac{\partial f}{\partial y} = \frac{\partial}{\partial y}\int P(x,y)\,dx + C'(y)$$

con lo cual queda

$$\boxed{C'(y) = Q(x,y) - \frac{\partial}{\partial y}\int P(x,y)\,dx}$$

En esta expresión el lado derecho parece que será también función de x, pero en realidad no es así. Para ver que no depende de x, basta con ver que es constante respecto á x, lo cual puede mostrarse derivando respecto á x y comprobando que el resultado es nulo:

$$\frac{\partial Q}{\partial x} - \frac{\partial}{\partial x}\frac{\partial}{\partial y}\int P(x,y)\,dx = \frac{\partial Q}{\partial x} - \frac{\partial}{\partial y}\frac{\partial}{\partial x}\int P(x,y)\,dx = \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 0.$$

Hemos usado otro resultado del Cálculo Diferencial en varias variables: la igualdad de las derivadas parciales cruzadas, además del Teorema Fundamental del Cálculo en la variable x. En todo caso, la expresión para $C'(y)$ es función solamente de y, por tanto se puede integrar respecto á y para obtener el valor de $C(y)$, y por tanto de f, salvo una constante "de verdad," es decir, independiente de x y de y.

• También se puede partir de la otra ecuación,

$$\frac{\partial f}{\partial y} = Q,$$

integrando respecto á y, procediendo de la misma manera pero intercambiando los papeles de x, y. Así, se obtendrá

$$f(x,y) = \int Q(x,y)\,dy + C(x)$$

y para hallar $C(x)$ derivamos respecto á x, obteniendo

$$P(x,y) = \frac{\partial f}{\partial x} = \frac{\partial}{\partial x}\int Q(x,y)\,dy + C'(x)$$

por tanto

$$\boxed{C'(x) = P(x,y) - \frac{\partial}{\partial x}\int Q(x,y)\,dy}$$

y la expresión a la derecha no depende de y, con lo cual, integrando respecto á x, se obtiene $C(x)$, y por tanto f, salvo una constante.

• **Ejemplo:** Resolvemos $3y + e^x + (3x + \cos y)y' = 0$. En la forma diferencial, es

$$(3y + e^x)\,dx + (3x + \cos y)\,dy = 0$$

y se comprueba que es exacta:

$$P(x,y) = 3y + e^x, \quad Q(x,y) = 3x + \cos y, \quad P_y = 3 = Q_x.$$

Vamos a hallar el potencial integrando respecto á x: $f_x = P$, luego

$$f(x,y) = \int (3y + e^x)\,dx + C(y) = 3xy + e^x + C(y).$$

Ahora, derivando respecto á y, como $f_y = Q$, se obtiene

$$3x + \cos y = 3x + C'(y), \quad \therefore \quad C'(y) = \cos y \quad \therefore \quad C(y) = \sin y + C$$

y queda finalmente

$$f(x,y) = 3xy + e^x + \sin y + C,$$

luego las solución, en forma implícita, de la E.D. es

$$3xy + e^x + \sin y = C.$$

Usamos la misma letra C para designar el concepto de una constante arbitraria. Esto no debe llevar a la confusión de pensar que C denota un valor concreto. Por ejemplo,

$$f(x,y) + C = C'$$

siendo C, C' constantes arbitrarias, significa lo mismo que simplemente

$$f(x,y) = C.$$

Incluso pueden llegar a verse a veces expresiones tales como

$$f(x,y) + C = C$$

donde ambas "C" denotan constantes arbitrarias pero en principio distintas! Este abuso de notación es parecido a lo que ocurre con la O-notación (2.8). Otro ejemplo se da cuando vemos una expresión como $K = e^C$. Esto es una constante arbitraria pero positiva ya que $e^x > 0$.

8.4. Lineales de primer orden

Ecuación diferencial lineal de primer orden

Una ecuación diferencial es lineal de primer orden si es de la forma

$$y' + a(x)y = b(x).$$

Existen varios procedimientos para resolver este tipo de ecuaciones. Daremos uno que funciona bastante bien en general.

La **ecuación homogénea asociada** a una E.D. lineal de primer orden

$$y' + a(x)y = b(x)$$

es la ecuación

$$y' + a(x)y = 0.$$

Aunque se usa la misma palabra homogénea, esto no es el mismo concepto que el de ecuación diferencial homogénea que vimos antes. Es sencillamente un caso de uso múltiple de la misma terminología.

• Una ecuación lineal homogénea (en el sentido actual) es de variables separadas:

$$y' + a(x)y = 0 \iff \frac{dy}{dx} = -a(x)y \iff \frac{dy}{y} = -a(x)\,dx$$

y por tanto resoluble "por cuadraturas," en otras palabras, integrando ambos lados. Se obtiene

$$\log y = -\int a(x)\,dx + C = -A(x) + C$$

donde $A'(x) = a(x)$ y C es una constante arbitraria. Entonces

$$y = Ce^{-A(x)}$$

donde $C > 0$. Para resolver la ecuación general no homogénea

$$y' + a(x)y = b(x)$$

se utiliza el **método de variación de constantes.** Imaginamos que C es una función de x y escribimos, acorde con esto

$$\boxed{y = C(x)\,e^{-A(x)}}$$

Ahora, derivando

$$\begin{aligned} y' &= C'(x)e^{-A(x)} - C(x)A'(x)e^{-A(x)} \\ &= C'(x)e^{-A(x)} - A'(x)\,y \\ &= C'(x)e^{-A(x)} - a(x)\,y, \end{aligned}$$

por tanto debe ser

$$y' + a(x)y = b(x) = C'(x)e^{-A(x)},$$

o sea,

$$C'(x) = b(x)e^{A(x)},$$

ecuación que en principio se puede integrar respecto á x para obtener $C(x)$ y por tanto una solución y a la ecuación general.

• **Ejemplo:** Resolvemos $2xy' - 3y = 4x^2$. Es lineal de primer orden, con

$$y' - \frac{3}{2x}\,y = 2x, \quad a(x) = -\frac{3}{2x}, \quad b(x) = 2x.$$

La ecuación homogénea asociada es

$$\begin{aligned} y' - \frac{3}{2x}\,y = 0 \iff & \frac{dy}{y} = \frac{3}{2x}\,dx \\ \iff & \log y = \frac{3}{2}\log x + C = \log x^{3/2} + C \\ \iff & y = Cx^{3/2}. \end{aligned}$$

Ahora variamos la constante:

$$y = C(x)\,x^{3/2}$$

y ponemos la ecuación original, confirmando que se cancelará $C(x)$ quedando solamente su derivada $C'(x)$:

$$\begin{aligned} y' - \frac{3}{2x}y = 2x \iff & C'(x)x^{3/2} + C(x)\cdot\frac{3}{2}x^{1/2} - \frac{3}{2x}C(x)x^{3/2} = 2x \\ \iff & C'(x)x^{3/2} = 2x \\ \iff & C'(x) = 2x^{-1/2} \\ \iff & C(x) = 4x^{1/2} + C, \end{aligned}$$

quedando finalmente como solución general

$$y = (4x^{1/2} + C)\,x^{3/2} = 4x^2 + Cx^{3/2}.$$

9. Series de Fourier [13]

9.1. Introducción

El Análisis de Fourier o Análisis Armónico fue desarrollado por J. Fourier al estudiar la ecuación del calor, que es la ecuación diferencial que describe la propagación del calor en un medio. Descubrió que se podían describir las soluciones de la ecuación en términos de series infinitas iguales a combinaciones de las funciones $\sin(nx)$ y $\cos(nx)$, y conjeturó que "cualquier" función podría expresarse en términos de dichas series.

En general, los fenómenos de onda tales como la vibración acústica, las interferencias en óptica, y en general todo tipo de oscilaciones, también corresponden matemáticamente a las combinaciones de estas funciones trigonométricas. Dentro de las matemáticas, el Análisis de Fourier tiene aplicaciones importantes a la teoría de números, ecuaciones diferenciales, combinatoria, álgebra y geometría, por mencionar solo algunas.

Una aplicación del Análisis Armónico de gran relevancia actual es en la Teoría de Señales, y en particular el procesamiento de señales. Esto se refleja en situaciones cotidianas como la transmisión digital inalámbrica, hablar por teléfono, trabajar con imágenes digitales, o manipular el ecualizador en un sistema de audio.

La palabra armónico proviene de la Música, donde, dado un tono que vibra a una cierta frecuencia fundamental, se utiliza para designar los tonos cuyas frecuencias son múltiplos enteros de la frecuencia fundamental. Desde Pitágoras se conoce que el oído humano interpreta los múltiplos enteros sencillos como tonos agradables, en armonía. Por extensión, la noción de armónico superior en el Análisis de Fourier designa a los términos $\sin(nx)$, $\cos(nx)$ donde $n \geq 2$ (segundos, terceros, cuartos armónicos, etc.)

Como indican los ejemplos anteriores, una señal puede tener una naturaleza física de onda de muchos tipos, por ejemplo, eléctrica, de radio, acústica, óptica, etc.

La idea central en la que se basa el Analisis de Fourier es que una señal general es la superposición (combinación lineal posiblemente infinita) de los armónicos básicos o "puros" $\sin(nx)$ y $\cos(nx)$.

9.2. Conceptos básicos generales

9.2.1. Terminología de la Teoría de Señales

A menudo en las aplicaciones, la variable en las funciones representa el tiempo. Entonces una función periódica modela un fenómeno que se repite en el tiempo.

[13] Algunas secciones están adaptadas de unos apuntes de la U. de la Rioja.

Señales

Llamaremos **señal** *(periódica) u* **onda** *a una función periódica $v(t)$, donde habitualmente la variable t denota tiempo.*

También se puede hablar de señal no periódica, pero ese tipo no va a entrar dentro de nuestro análisis, por lo cual a partir de ahora "señal" querrá decir "señal periódica."

• **Ejemplo:** El "tono puro" $v(t) = \sin(t)$ es una onda.

• **Ejemplo:** Las ondas cuadradas son ondas, en apariencia distintas a la onda seno.

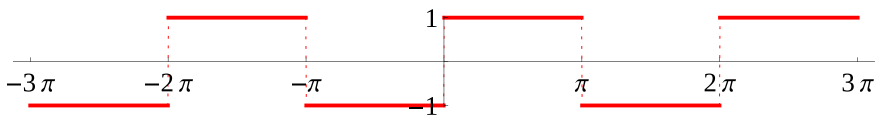

Figura 35 – Una onda cuadrada.

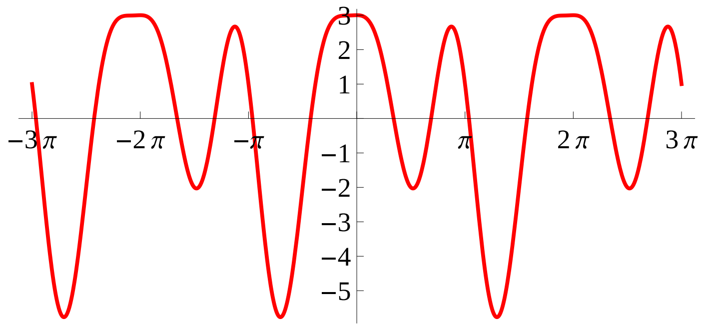

Figura 36 – Una señal periódica algo más complicada.

Período y ciclo

El **periodo** *de una señal es el tiempo* mínimo *T que ha de transcurrir hasta que la señal se repita: $v(t+T) = v(t)$.*

Un **ciclo** *completo de la señal es un intervalo de tiempo de longitud igual al periodo T.*

• **Ejemplo:** La señal $v(t) = \sin(t)$ tiene periodo 2π, la proporción entre la longitud de una circunferencia y su radio. Esta constante aparecerá en multitud de fórmulas.

• **Ejemplo:** Las señales $v(t) = \sin(2\pi t), \cos(2\pi t)$ tienen periodo 1, ya que, por ejemplo,

$$v(t+1) = \sin\big(2\pi(t+1)\big) = \sin(2\pi t + 2\pi) = \sin(2\pi t) = v(t)$$

y similarmente con el coseno. Muchas veces se utilizan estas modificaciones normalizadas de las funciones trigonométricas en vez de las originales.

• **Ejemplo:** Dado cualquier $T > 0$, las señales

$$v(t) = \sin\left(\frac{2\pi}{T}t\right), \qquad v(t) = \cos\left(\frac{2\pi}{T}t\right),$$

tiene periodo T, pues por ejemplo

$$v(t+T) = \sin\left(\frac{2\pi}{T}(t+T)\right) = \sin\left(\frac{2\pi}{T}t + 2\pi\right) = \sin\left(\frac{2\pi}{T}t\right) = v(t)$$

y similarmente para el coseno.

Frecuencia

*La **frecuencia** de la señal es el número de ciclos que se repiten por unidad de tiempo.*

Matemáticamente, la frecuencia es el recíproco del periodo:

$$\boxed{f = 1/T}$$

ya que por ejemplo, tomando el segundo como unidad de tiempo, el periodo es el número de segundos por ciclo, y la frecuencia el número de ciclos por segundo.

Físicamente, la frecuencia es el número de ciclos completos de la onda que pasan por un punto dado por unidad de tiempo. La unidad de frecuencia es el hercio (Hz), correspondiente a un ciclo por segundo; también son habituales el kilohercio (KHz), igual a mil ciclos por segundo y el megahercio (MHz), igual un millón de ciclos por segundo.

• **Ejemplo:** Midiendo en segundos, una onda de periodo $T = 1/2$ se reproduce después de $1/2$ segundo. En un segundo, se reproduce dos veces, por tanto su frecuencia es 2. Una onda de periodo $1/3$ tiene frecuencia 3 pues en un segundo se reproduce tres veces, etc.

• **Ejemplo:** Para señales acústicas, percibimos los tonos de alta frecuencia como sonidos agudos y los de baja frecuencia como sonidos graves.

Frecuencia Angular

*La **frecuencia angular** de la señal es el número de ciclos por 2π unidades de tiempo. Equivalentemente, es el número de ciclos,* medidos en radianes, *por unidad de tiempo.*

Entonces, la relación entre frecuencia y frecuencia angular es

$$\boxed{\omega = 2\pi f}$$

por una sencilla regla de tres, ya que la frecuencia f es el número de ciclos por una unidad de tiempo (habitualmente el segundo) y ahora contamos ciclos por 2π unidades de tiempo. Por otra parte, como $f = 1/T$,

$$\boxed{\omega = \frac{2\pi}{T}}$$

La razón más profunda para esta normalización (cambio de constante) se debe a que las funciones seno y coseno son las funciones trigonométricas circulares: como cumplen

$$\sin^2 t + \cos^2 t = 1,$$

el punto $p = (x, y) = (\cos t, \sin t)$ recorre la circunferencia unidad (centro en el origen y radio 1), ya que la ecuación de esta circunferencia es

$$x^2 + y^2 = 1$$

(puntos del plano cuya distancia euclidea al origen es 1). Se dice que parametrizan la circunferencia. La propia palabra ciclo significa círculo, y el hecho de aparecer la constante 2π se debe a que la longitud de la circunferencia unidad – el ciclo – mide precisamente 2π.

Entonces, si la unidad de ciclo en radianes es 2π, la frecuencia angular también es, como hemos dicho en la definición, el número de ciclos en radianes por unidad de tiempo.

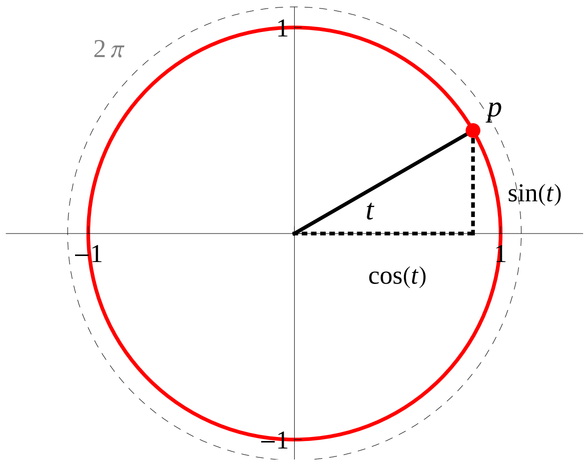

Figura 37 – Las funciones sin y cos parametrizan la circunferencia unidad.

Amplitud

*Por **amplitud** de una señal se entienden varios a números, relacionados entre sí:*

- *La amplitud pico es el máximo de su valor absoluto.*
- *La amplitud pico a pico es la diferencia entre su máximo y su mínimo.*
- *La semiamplitud es la mitad de la amplitud pico a pico.*

• **Ejemplo:** En una señal acústica, la amplitud mide la intensidad del tono, es decir, a mayor amplitud, el tono se percibe a mayor volumen.

• **Ejemplo:** Las señales eléctricas se pueden representar como voltaje en función del tiempo. La amplitud pico es el voltaje máximo. Físicamente, el voltaje es el concepto análogo para una corriente eléctrica al concepto de presión en un fluido.

• **Ejemplo:** Las ondas sísmicas relacionadas con los terremotos se miden con un sismógrafo; la frecuencia se reflejará en la velocidad del movimiento de la aguja del sismógrafo y la amplitud en la altura del trazo que deja sobre el papel.

9.2.2. Ondas sinusoidales

La onda sinusoidal en el tiempo

La onda sinusoidal general es la función

$$s(t) = A\sin(\omega t + \phi)$$

- La frecuencia angular de $s(t)$ es ω.
- La frecuencia de $s(t)$ es $f = \omega/2\pi$.
- El periodo de $s(t)$ es $T = 1/f = 2\pi/\omega$.
- La amplitud pico de $s(t)$ es A.
- La amplitud pico a pico de $s(t)$ es $2A$.
- La semiamplitud de $s(t)$ es A.

Fase de la onda sinusoidal

*El número ϕ (medido en radianes) en una onda sinusoidal $s(t) = A\sin(\omega t + \phi)$, que representa el desplazamiento en el tiempo respecto a la onda seno pura, se llama **fase** de la onda sinusoidal.*

• **Ejemplo:** La función coseno es la función seno cambiada de fase por $\pi/2$:

$$\sin\left(t + \frac{\pi}{2}\right) = \sin t\cos\frac{\pi}{2} + \cos t\sin\frac{\pi}{2} = \sin t \cdot 0 + \cos t \cdot 1 = \cos t.$$

• **Ejemplo:** Si la señal es acústica, hay un debate acerca de si el oído humano es capaz de percibir diferencias en la fase entre dos ondas cuyas frecuencias y amplitudes sean iguales. En principio uno pensaría que no es posible pues se trata de la misma onda desplazada temporalmente, pero algunos experimentos indican lo contrario.

9.2.3. Otros tipos de ondas: clasificaciones generales

Señales digitales y analógicas

*Una señal $v(t)$ es **digital** si como función, $v(t)$ toma solo un número finito de valores, por ejemplo, 0 y 1. Es **analógica** si toma un número infinito de valores.*

Muchas ondas correspondientes a tipos de señales habituales reciben su nombre de la forma que tienen.

• **Ejemplo:** Las ondas cuadradas son digitales.

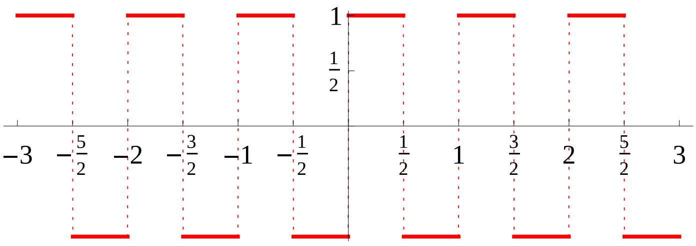

Figura 38 – Onda cuadrada.

• **Ejemplo:** Las ondas triangulares y las ondas dientes de sierra son analógicas.

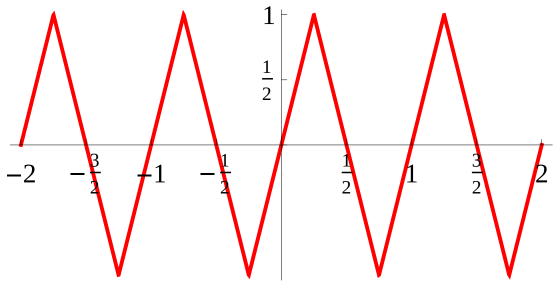

Figura 39 – Onda triangular.

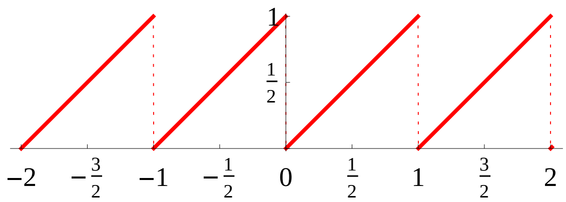

Figura 40 – Onda dientes de sierra.

Más adelante estudiaremos en detalle las ecuaciones de estas ondas, cuando encontremos sus series de Fourier.

Señales continuas y discretas

*Una señal es **continua** si la función que la describe está definida en cualquier instante en un intervalo de tiempo. Es **discreta** si sólo está definida en instantes $t_0 < t_1 < t_2 < \dots$ de tiempo separados por una distancia mínima $\delta > 0$.*

Aquí solamente vamos a estudiar señales continuas. Las señales discretas se estudian con la transformada de Fourier discreta (DFT por sus siglas en inglés) y tiene gran importancia en aplicaciones como la teoría de códigos o el procesamiento de imágenes.

9.3. Conceptos básicos matemáticos

Desde el punto de vista matemático, la idea central del Análisis Armónico, que una señal general es la superposición de armónicos básicos, corresponde a una teoría de la aproximación mediante el cual unas funciones más o menos arbitrarias se pueden aproximar por otras funciones, más sencillas. Esto sirve tanto para entender el comportamiento teórico de las funciones como para en la práctica, por ejemplo, poder calcular sus valores con un cierto grado de precisión, o para representarlas gráficamente a una resolución dada.

Un ejemplo de aproximación es el desarrollo de Taylor estudiado en la sección 4.12, donde se utilizan polinomios, es decir, combinaciones lineales de las funciones potencias enteras no negativas $e_n(x) = x^n$ para aproximar a las funciones suficientemente derivables. El requisito de existencia de muchas derivadas presenta una cierta desventaja al restringir el tipo de funciones que se pueden aproximar de esta manera.

En el Análisis de Fourier, se estudia la aproximación por los armónicos básicos

$$\sin(nx), \quad n = 1, 2, \ldots, \quad \text{y} \quad \cos(nx), \quad n = 0, 2, \ldots .$$

Estas funciones jugarán el papel que las potencias hacen en el caso del desarrollo de Taylor. Sin embargo, veremos que una de las ventajas de usar funciones trigonométricas es que el resultado de aproximación correspondiente es aplicable a más funciones que el desarrollo de Taylor. En particular, no se requiere que las funciones que se van a aproximar tengan derivadas más allá de la primera.

Armónicos básicos

A la colección de funciones $\{\sin(nx), n = 1, 2, 3, \ldots\} \cup \{\cos(nx), n = 0, 1, 2, 3, \ldots\}$ *la llamaremos la* ***base estándar*** *de funciones trigonométricas y sus elementos* ***armónicos básicos.*** *A menudo omitiremos la palabra "básico."*

Las funciones $\sin(x)$ y $\cos(x)$ se consideran los armónicos fundamentales. El periodo fundamental es $T = 2\pi$, por tanto la frecuencia es $f = 1/2\pi$ y la frecuencia angular $\omega = 1$. Su amplitud es $A = 1$.

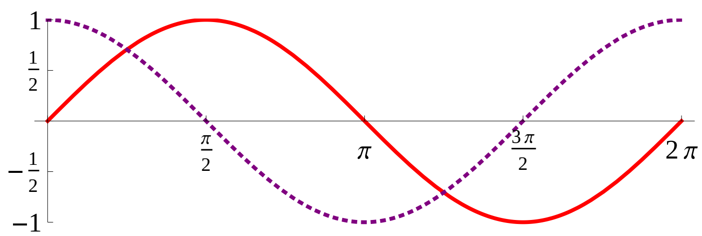

Figura 41 – Los armónicos fundamentales $\sin(x), \cos(x)$.

Los armónicos n-ésimos son $\sin(nx)$, $\cos(nx)$. Para $n \geq 2$ son los llamados armónicos superiores. Su periodo es $T = 2\pi/n$, la frecuencia $f = n/2\pi$, la frecuencia angular $\omega = n$ y la amplitud $A = 1$.

Recordamos que el coseno es igual al seno cambiado de fase por $\pi/2$, con lo cual todos los armónicos son sinusoidales.

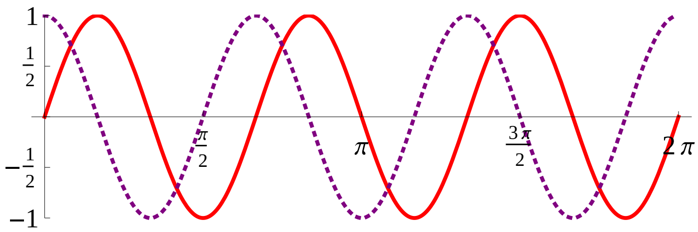

Figura 42 – Los terceros armónicos $\sin(3x), \cos(3x)$.

La función constante $1 = \cos(0 \cdot x)$ $(n = 0)$ se considera como un caso aparte.

• **Ejemplo:** En la teoría de los circuitos eléctricos, donde la onda representa la fluctuación del voltaje respecto al tiempo, la onda constante corresponde a la corriente directa donde el voltaje no cambia, y los demás armónicos serían ejemplos de corriente alterna, donde hay un cambio alternante a una cierta frecuencia (50 Hz en la mayor parte del mundo, exceptuando los continentes americanos y algún otro territorio, donde es 60 Hz).

• **Ejemplo:** Al combinar linealmente un número finito de armónicos obtendremos ondas más complejas, como por ejemplo $v(t) = \sin(t) + 3\sin(2t) + 2\sin(3t)$.

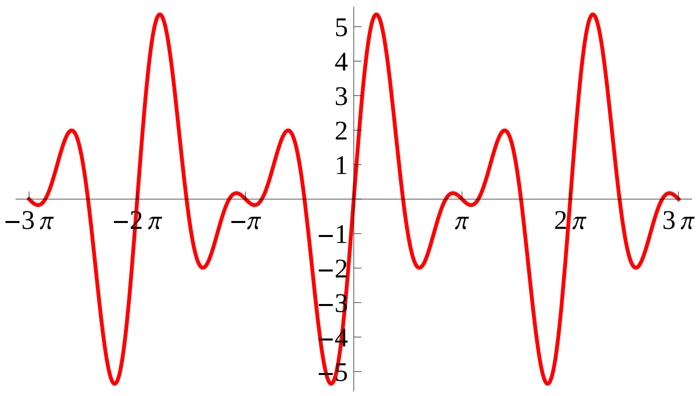

Figura 43 – La onda $\sin(t) + 3\sin(2t) + 2\sin(3t)$

Matemáticamente, la representación de una función mediante armónicos cobra la forma de una serie que en la mayoría de los casos es infinita, llamada serie de Fourier. El avance que realizó Fourier fue precisamente intuir que al permitir combinar un número infinito de armónicos, se pueden conseguir funciones casi arbitrarias.

Serie de Fourier (real)

Se llama **serie de Fourier** *a una serie infinita igual a una combinación lineal de las funciones de la base trigonométrica estándar,*

$$\begin{aligned} &c + a_1 \cos(x) + a_2 \cos(2x) + a_3 \cos(3x) + \dots \\ &+ b_1 \sin(x) + b_2 \sin(2x) + b_3 \sin(3x) + \dots \end{aligned}$$

Por conveniencia, es habitual denotar el término constante c como $a_0/2$: entonces, empleando la notación para series infinitas de funciones, escribimos

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

Dada una función $\upsilon(x)$, buscaremos una serie de Fourier que en algún sentido la representa. La interpretación más común de esta idea es que la función $\upsilon(x)$ sea la suma de la serie de Fourier en algún dominio X, es decir, que sea el límite de sus sumas parciales,

$$\upsilon(x) = \lim_{N\to\infty} S_N(x), \qquad S_N(x) \stackrel{\text{def}}{=} \frac{a_0}{2} + \sum_{n=1}^{N} \big(a_n \cos(nx) + b_n \sin(nx)\big), \quad x \in X.$$

En este caso escribimos simplemente

$$\upsilon(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \big(a_n \cos(nx) + b_n \sin(nx)\big), \quad x \in X$$

Habitualmente el dominio será los números reales $\mathbb{R}$ o un intervalo. Muchas veces basta con un número relativo pequeño de sumandos N para que la serie de Fourier empiece ya a adoptar la forma de la onda $\upsilon(x)$.

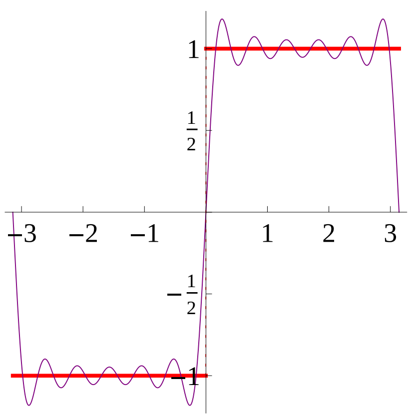

Figura 44 – Aproximación a una onda cuadrada por su serie de Fourier (6 sumandos).

Dado que las funciones $\sin(nx)$ y $\cos(nx)$ tienen periodo $2\pi/n$, todas ellas juntas comparten el periodo común 2π, por tanto el tipo de función $\upsilon(x)$ representable por una serie de Fourier, es decir, combinando armónicos, también deberá ser periódica de periodo 2π:

$$\upsilon(x + 2\pi) = \upsilon(x).$$

Es decir, la periodicidad es una condición necesaria para que una función esté representada por una serie de Fourier.

En general, podemos considerar funciones periódicas con cualquier periodo simplemente efectuando un cambio de variable en las fórmulas, como veremos más adelante.

9.4. Los coeficientes de la serie de Fourier

Supongamos entonces que $\upsilon : \mathbb{R} \to \mathbb{R}$ es una función periódica de periodo 2π y vamos a estudiar cómo se puede representar mediante una serie de Fourier. Buscamos entender cómo, dada una señal periódica, podemos dar con una combinación, posiblemente infinita, de armónicos que, al superponerlos, adoptarán la forma de esta señal.

El siguiente resultado nos dice que para una función, existe a lo sumo una serie de Fourier que la representa, y nos dice cuál tiene que ser esa serie.

Unicidad de la serie de Fourier y fórmulas para los coeficientes

Sea $\upsilon : \mathbb{R} \to \mathbb{R}$ periódica de periodo 2π. Si υ es representable por una serie de Fourier en el intervalo $[-\pi, \pi]$,

$$\upsilon(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \big(a_n \cos(nx) + b_n \sin(nx)\big), \quad x \in [-\pi, \pi],$$

entonces los coeficientes a_n, b_n están dados por las fórmulas

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} \upsilon(x)\, dx, \qquad a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} \upsilon(x) \cos(nx)\, dx, \quad n = 0, 1, 2, 3, \ldots$$

y

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} \upsilon(x) \sin(nx)\, dx, \quad n = 1, 2, 3, \ldots.$$

En particular, la serie de Fourier que representa á υ es única.

El intervalo de longitud 2π que se usa para estudiar la onda no es esencial. Hemos enunciado el resultado para $[-\pi, \pi]$ pero sirve cualquier intervalo de la forma $[c, c + 2\pi]$. Físicamente, el instante c en el que se empieza a recibir la señal periódica no afecta a sus características básicas.

Demostración de las fórmulas para los coeficientes. La demostración se basa en unas relaciones fundamentales que cumplen los armónicos y que vamos a señalar como un resultado aparte:

Relaciones de ortogonalidad entre los armónicos

$$\text{(a)} \quad \int_{-\pi}^{\pi} \sin(nx)\,dx = \int_{-\pi}^{\pi} \cos(nx)\,dx = 0 \qquad \forall\, n = 1, 2, 3, \ldots;$$

$$\text{(b)} \quad \int_{-\pi}^{\pi} \cos(nx)\cos(mx)\,dx = \begin{cases} 0 & si\ n \neq m, \\ \pi & si\ n = m, \end{cases} \qquad \forall\, n, m = 1, 2, 3, \ldots;$$

$$\text{(c)} \quad \int_{-\pi}^{\pi} \sin(nx)\sin(mx)\,dx = \begin{cases} 0 & si\ n \neq m, \\ \pi & si\ n = m, \end{cases} \qquad \forall\, n, m = 1, 2, 3, \ldots;$$

$$\text{(d)} \quad \int_{-\pi}^{\pi} \sin(nx)\cos(mx)\,dx = 0 \qquad \forall\, n, m = 1, 2, 3, \ldots$$

Demostración de las relaciones de ortogonalidad. Esto es un ejercicio en técnicas básicas de integración. La denominación “relación de ortogonalidad” viene de una interpretación geométrica de los resultados que consiste en entender las funciones como elementos de un espacio vectorial dotado de una manera de medir ángulos. □

Retomando la demostración principal, supongamos entonces dada una función $\upsilon : \mathbb{R} \to \mathbb{R}$ de periodo 2π y que

$$\upsilon(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \big(a_n \cos(nx) + b_n \sin(nx)\big), \quad x \in [-\pi, \pi] \tag{$*$}$$

• Cálculo de a_0. Vamos a integrar los dos miembros de la igualdad $(*)$ entre $-\pi$ y π. Dando por hecho que se puede intercambiar la suma infinita con la integral, tendríamos, usando las relaciones de ortogonalidad (a) y (b),

$$\begin{aligned} \int_{-\pi}^{\pi} \upsilon(x)\,dx &= \frac{a_0}{2}\int_{-\pi}^{\pi} dx + \sum_{n=1}^{\infty} a_n \int_{-\pi}^{\pi} \cos(nx)\,dx + \sum_{n=1}^{\infty} b_n \int_{-\pi}^{\pi} \sin(nx)\,dx \\ &= \frac{a_0}{2} \cdot 2\pi + 0 + 0 = a_0\pi \end{aligned}$$

y por tanto obtenemos la fórmula

$$\boxed{a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} v(x)\,dx}$$

• Cálculo de a_n. Multipliquemos los dos miembros de la igualdad $(*)$ por $\cos mx$ e integremos entre $-\pi$ y π. De nuevo suponiendo que se puede intercambiar la suma infinita con la integral, o sea, que la integral de una combinación lineal infinita es la combinación lineal correspondiente de las integrales, tenemos

$$\begin{aligned}\int_{-\pi}^{\pi} f(x)\cos(mx)\,dx &= \frac{a_0}{2}\int_{-\pi}^{\pi}\cos(mx)\,dx + \sum_{n=1}^{\infty} a_n \int_{-\pi}^{\pi}\cos(nx)\cos(mx)\,dx \\ &\quad + \sum_{n=1}^{\infty} b_n \int_{-\pi}^{\pi}\sin(nx)\cos(mx)\,dx \\ &= 0 + a_m\pi + 0 = a_m\pi,\end{aligned}$$

ya que la única integral que no se anula es $\int_{-\pi}^{\pi} a_n\cos(nx)\cos(mx)\,dx$ cuando $n=m$. De aquí se obtiene el valor de a_m dividiendo por π. Así, si en lugar de m volvemos a usar n, hemos obtenido

$$\boxed{a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(nx)\,dx, \quad n = 1, 2, \dots}$$

Obsérvese además que, sustituyendo $n=0$ en esta fórmula, como $\cos(nx)=1$, esta nueva fórmula corresponde exactamente con la que aparecía al calcular a_0. En consecuencia, la fórmula para calcular a_n no sólo es válida para $n = 1, 2, \dots$, sino que se puede aplicar para el rango completo $n = 0, 1, 2, \dots$. Precisamente para que esto ocurra es por lo que se usa $a_0/2$ en lugar de a_0 al escribir la serie de Fourier de una función f, tal como hemos hecho en $(*)$. De este modo se unifican fórmulas, al menor coste de tener que recordar el factor $1/2$.

• Cálculo de b_n. Multipliquemos ahora los dos miembros de la igualdad $(*)$ por $\sin mx$ e integremos entre $-\pi$ y π. Como antes, obtenemos

$$\begin{aligned}\int_{-\pi}^{\pi} f(x)\sin(mx)\,dx &= \frac{a_0}{2}\int_{-\pi}^{\pi}\sin(mx)\,dx + \sum_{n=1}^{\infty}\int_{-\pi}^{\pi} a_n\cos(nx)\sin(mx)\,dx \\ &\quad + \sum_{n=1}^{\infty}\int_{-\pi}^{\pi} b_n\sin(nx)\sin(mx)\,dx = b_m\pi,\end{aligned}$$

pues la única integral que no se anula es $\int_{-\pi}^{\pi} b_n\sin(nx)\sin(mx)\,dx$ con $n=m$. Empleando de nuevo n en lugar de m, obtenemos

$$\boxed{b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(nx)\,dx, \quad n = 1, 2, \dots}$$

□

Vamos a mencionar un caso particular de interés.

Polinomios Trigonométricos

Un ***polinomio trigonométrico*** *es una serie de Fourier finita, es decir, una combinación lineal finita de armónicos*

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

La serie de Fourier de un polinomio trigonométrico es él mismo

Sea $v(x)$ un polinomio trigonométrico. Entonces la serie de Fourier de $v(x)$ es su propia expresión única como combinación lineal de armónicos.

Demostración. La expresión del polinomio trigonométrico como combinación lineal de armónicos es una serie de Fourier. Como ésta es única, debe ser precisamente esa. □

En el lenguaje del Álgebra Lineal, los armónicos básicos son linealmente independientes.

9.5. Convergencia de la serie de Fourier

Dada una función periódica $v(x)$, hemos visto cuáles deberían ser los coeficientes a_n y b_n para que se pueda dar la igualdad

$$v(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \big(a_n \cos(nx) + b_n \sin(nx)\big), \quad x \in [-\pi, \pi] \tag{$*$}$$

Había que suponer que es legítimo el intercambio de una suma infinita con una integral, algo que no hemos justificado rigurosamente. Aún así, lo que hemos obtenido es una condición necesaria, que no tiene por qué ser suficiente.

Es decir, nada garantiza, desde el punto de vista lógico, que, aún tomando los coeficientes a_n, b_n como dicen las fórmulas de la sección 9.4 que tienen que ser para la función $v(x)$, vaya a cumplirse la representación $(*)$, recordando que "representar" quiere decir:

- Que la serie de Fourier sea convergente para $x \in [-\pi, \pi]$,
- Que, siendo convergente, la suma de la serie sea igual a la función $v(x)$

En el siglo XIX se descubrieron ejemplos de series de Fourier que no convergen, y es relativamente fácil encontrar funciones cuya serie de Fourier no converge a ellas en un punto dado x del intervalo. Este problema de convergencia es bastante complicado, pero afortunadamente, también en el siglo XIX, Dirichlet encontró un criterio para las funciones que, de cumplirlo, asegura una respuesta afirmativa a la cuestión de la validez de la representación de una función por su serie de Fourier. Además, es un criterio que abarca una amplia categoría de funciones, incluyendo las que surgen en las aplicaciones de la teoría.

Continuidad a trozos (con discontinuidades de salto)

*Una función $\upsilon : [a, b] \to \mathbb{R}$ es **continua a trozos** si existe una partición*

$$\wp : a = x_0 < x_1 < x_2 < \cdots < x_n = b$$

tal que

- *υ es continua en cada subintervalo (x_{k-1}, x_k), $k = 1, \ldots, n$.*
- *los límites laterales $\upsilon(x_k^+)$ y $\upsilon(x_k^-)$ existen y son finitos $\forall\, k$.*

Nota: en $x_0 = a$ sólo pedimos la existencia de $\upsilon(a^+)$ y en $x_n = b$, la de $\upsilon(b^-)$.

Brevemente, esto significa que υ es continua salvo en un número finito de puntos, donde tiene discontinuidades de salto. En un contexto más general, también se llaman continuas a trozos a las funciones acotadas y con un número finito de discontinuidades, del tipo que sean. Aquí estamos expresamente reduciendo al caso de discontinuidades de salto, simplemente diciendo "continua a trozos" e implícitamente incluyendo la condición sobre las discontinuidades.

El criterio de convergencia para series de Fourier de Dirichlet requiere algo más de la función para que ésta se pueda representar por su serie de Fourier.

Función de clase C^1 a trozos

*Una función $\upsilon : [a, b] \to \mathbb{R}$ es **de clase C^1 a trozos** si existe una partición*

$$\wp : a = x_0 < x_1 < x_2 < \cdots < x_n = b$$

tal que

- *υ es derivable en cada subintervalo (x_{k-1}, x_k), $k = 1, \ldots, n$ (en particular también es continua en estos subintervalos).*
- *los límites laterales de la función $\upsilon(x_k^+)$ y $\upsilon(x_k^-)$ existen y son finitos $\forall\, k$.*
- *los límites laterales de la derivada $\upsilon'(x_k^+)$ y $\upsilon'(x_k^-)$ existen y son finitos $\forall\, k$.*

En los extremos a, b sólo pedimos los límites en a^+, b^-.

Abreviando, se suele decir simplemente que

"υ es C^1 a trozos significa que υ y υ' [existen y] son continuas a trozos."

Esto es un tanto impreciso desde el rigor lógico, pues antes de poder hablar de la derivada hay que especificar que existe salvo en un número finito de puntos, como hemos hecho en la definición formal. Aunque a menudo se omite la declaración de su existencia, la imprecisión se resuelve entendiendo que la frase "a trozos" se refiere tanto a la existencia como a la continuidad de la derivada.

En general, el concepto se extiende a cualquier orden de derivada.

Se dice que una función v es **de clase C^p a trozos,** con $0 \le p \le \infty$, si $v^{(k)}$ existen y son continuas a trozos para $k = 1, 2, \ldots, p$."

• **Ejemplo:** Las funciones de clase C^1 son por definición las derivables con derivada continua. En particular, también son de clase C^1 a trozos, pues "solo hay un trozo".

La partición que se toma en la definición de C^1 a trozos incluye puntos que pueden ser

- discontinuidades de v, que tienen que ser saltos de la función,
- puntos donde no existe la derivada v', que tienen que ser saltos de la derivada.

Un punto en esta partición puede que sea de continuidad para v pero no para v'. En cambio, en una discontinuidad de v no puede existir la derivada v', por tanto para estas funciones, estos puntos tienen que ser también discontinuidades de salto de v'.

• **Ejemplo:** Una onda sinusoidal es de clase C^∞ sin más, desde luego también "a trozos," pero basta con un trozo. Igualmente, un polinomio trigonométrico es de clase C^∞. En general, el índice p en una función de clase C^p es una medida del grado de suavidad de la onda.

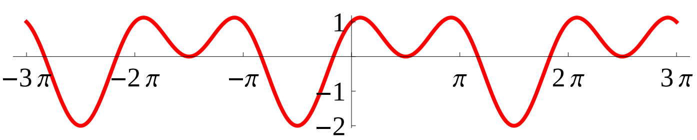

Figura 45 – Onda correspondiente a un polinomio trigonométrico.

• **Ejemplo:** Una onda "dientes de sierra" es continua y derivable con derivada continua (de hecho, constante) excepto en los puntos de salto de la propia función. En estos puntos la derivada tiene límites laterales, en particular, son saltos de la derivada.

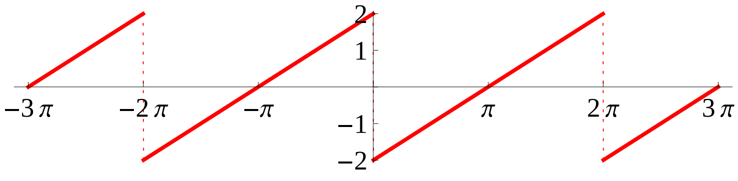

Figura 46 – Una onda dientes de sierra.

• **Ejemplo:** Una onda cuadrada es continua y derivable, con derivada continua (de hecho, nula) excepto en los puntos de salto de la propia función. En estos puntos la derivada tiene límites laterales nulos, en particular, son "saltos" de la derivada, aunque la altura del salto es 0.

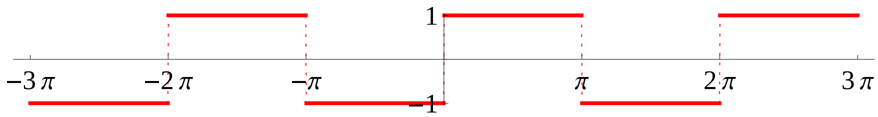

Figura 47 – Una onda cuadrada.

• **Ejemplo:** Una onda triangular es continua en todo punto, incluyendo los picos, y es derivable con derivada continua (de hecho, constante) entre los picos. En los picos la derivada no existe, pero sí tiene límites laterales, es decir, presenta discontinuidades de salto.

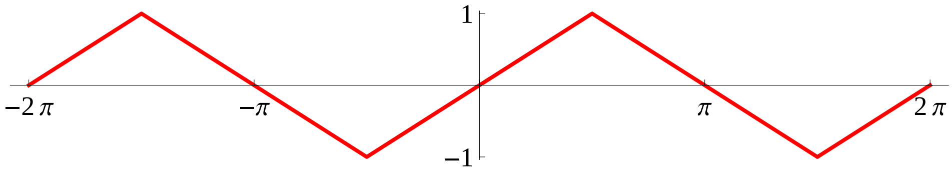

Figura 48 – Una onda triangular.

Con estas definiciones podemos enunciar el teorema principal de convergencia.

Teorema de Dirichlet acerca de la convergencia puntual de series de Fourier

Si υ es una función periódica de periodo 2π que es de clase C^1 a trozos en el intervalo $[-\pi, \pi]$, entonces la serie de Fourier

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

con coeficientes a_n, b_n dados por

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx)\, dx, \quad n = 0, 1, 2, \ldots,$$
$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx)\, dx, \quad n = 1, 2, \ldots,$$

es convergente a

(a) $\upsilon(x)$ *si x es un punto de continuidad;*

(b) $\dfrac{\upsilon(x^+) + \upsilon(x^-)}{2}$ *si x es un punto de discontinuidad de υ.*

En particular, si υ es una función continua, y derivable con derivada continua a trozos, entonces el caso (a) del Teorema de Dirichlet dice que la función está representada por su serie de Fourier, o sea, es cierta la fórmula

$$v(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \big(a_n \cos(nx) + b_n \sin(nx)\big), \quad x \in [-\pi, \pi]$$

Incluso se sigue escribiendo la fórmula de la misma manera, como una igualdad, en el caso (b), es decir, también en los puntos en los que v tiene un salto, y por tanto la serie converge al valor medio de sus límites laterales,

$$\frac{v(x^+) + v(x^-)}{2},$$

aunque no es necesariamente igual al valor $f(x)$.

Como ya mencionamos en la sección 9.4, los coeficientes de Fourier de una función periódica v de periodo 2π también se pueden calcular integrando sobre cualquier otro intervalo de longitud 2π, es decir, sobre cualquier $[c, c + 2\pi]$ con $c \in \mathbb{R}$. Asimismo, el Teorema de Dirichlet es igualmente válido cambiando $[-\pi, \pi]$ por este otro intervalo.

9.6. Extensión por periodicidad

Con el Teorema de Dirichlet ya se tiene un resultado aplicable a muchas funciones que se puede usar para desarrollarlas en series de Fourier. La única "pega" es que necesitamos que las funciones sean periódicas de periodo 2π. Que el periodo sea 2π no es importante. Veremos más adelante que los resultados se adaptan de manera fácil a cualquier otro periodo. Más fundamental es el hecho que la señal sea periódica.

A continuación vamos a dar un procedimiento que, dada una función $v(x)$ definida en el intervalo $[-\pi, \pi]$ (o en general, cualquier intervalo de longitud 2π), ésta se puede extender o prolongar a una función definida en todo $\mathbb{R}$, periódica de periodo 2π. De este modo se generan de manera fácil señales periódicas.

Para evitar una contradicción en la definición de valores en los extremos, se puede utilizar un intervalo semiabierto. Con ello basta con definir el valor en uno de los extremos.

La extensión periódica de una función en $[-\pi, \pi)$

*Sea $v : [-\pi, \pi) \to \mathbb{R}$. La **extensión periódica** de v á $\mathbb{R}$ es la función $V(x)$ definida por*

$$v_{\text{per}}(x) = v(x - 2\pi n) \quad \textit{si} \quad -\pi + 2\pi n \le x < \pi + 2\pi n \quad \textit{con} \quad n \in \mathbb{Z}.$$

En particular, $v_{\text{per}}(x) = v(x)$ para $x \in [-\pi, \pi)$ (el caso $n = 0$).

Geométricamente, la extensión periódica consiste en trasladar por múltiplos enteros de 2π copias de la gráfica de la función original v sobre el intervalo $[-\pi, \pi)$.

• **Ejemplo:** Consideremos la función $\upsilon(x) = x$ sobre $[-\pi, \pi)$. Su extensión periódica se muestra en la figura

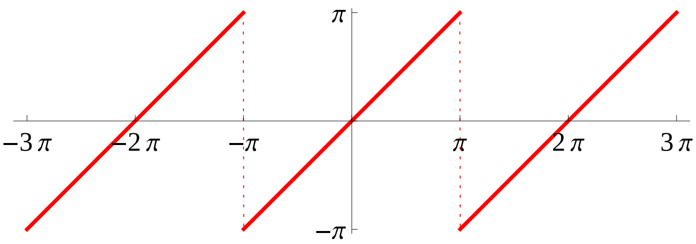

Figura 49 – Parte de la gráfica de la extensión periódica de $\upsilon(x) = x$ sobre el intervalo $[-\pi, \pi)$.

Por abuso de notación, a menudo utilizaremos la misma letra para designar tanto a la función original como a la función extendida por periodicidad.

En general, se puede realizar esta operación para cualquier función definida en un intervalo semiabierto finito de cualquier longitud T.

Extensión periódica general

Sea $\upsilon : [c, c+T)$. La extensión periódica de υ á $\mathbb{R}$ es la función $\upsilon_{\text{per}}(x)$ definida por

$$\upsilon_{\text{per}}(x) = \upsilon(x - nT) \quad \textit{si} \quad c + nT \leq x < c + (n+1)T \quad \textit{con} \quad n \in \mathbb{Z}.$$

En particular, $\upsilon_{\text{per}}(x) = \upsilon(x)$ para $x \in [c, c+T)$ (el caso $n = 0$).

Para dar una fórmula para la extensión periódica en términos sólo de x, la recta $\mathbb{R}$ se subdivide en la unión disjunta de los intervalos

$$I_n = \big[c + nT, c + (n+1)T\big) : \quad n \in \mathbb{Z}$$

de longitud T:

$$\begin{aligned}\mathbb{R} = \cdots \cup [c - 2T, c - T) \cup [c - T, c) \cup {\color{green}[c, c+T)} \\ \cup [c + T, c + 2T) \cup [c + 2T, c + 3T) \cup \ldots.\end{aligned}$$

Se trata simplemente de ubicar un punto x en el intervalo I_N de esta forma que le corresponde, y luego, restando nT, se tiene $x - nT \in [c, c+T)$, el intervalo de definición de υ, por tanto se puede evaluar $\upsilon(x - nT)$.

La condición $x \in I_n$ equivale a

$$c + nT \leq x < c + (n+1)T \iff nT \leq x - c < (n+1)T$$
$$\iff n \leq \frac{x-c}{T} < n+1 \iff \left\lfloor \frac{x-c}{T} \right\rfloor = n$$

donde

$$\boxed{\lfloor x \rfloor}$$

es la función suelo, que redondea un número real x al entero más próximo y menor que él. Por ejemplo, $\lfloor \pi \rfloor = 3$.

Geométricamente, la extensión periódica general consiste en trasladar por múltiplos enteros del periodo la gráfica de la función original υ sobre el intervalo $[c, c+T)$.

También se puede partir de una función definida en un intervalo semiabierto a la izquierda, $(c, c+T]$. En este caso si numeramos $I_n = (c + (n-1)T, c + nT]$, la condición $x \in I_n$ equivale a

$$c + (n-1)T < x \leq c + nT \iff (n-1)T < x - c \leq nT$$
$$\iff n - 1 < \frac{x-c}{T} \leq n \iff \left\lceil \frac{x-c}{T} \right\rceil = n$$

donde

$$\boxed{\lceil x \rceil}$$

es la función techo que redondea un número real x al entero más próximo y mayor que él. Por ejemplo, $\lceil \pi \rceil = 4$.

Como puede observarse en el ejemplo dado de la función $\upsilon(x) = x$ sobre $[-\pi, \pi)$, la extensión periódica produce una función $\upsilon_{\text{per}}(x)$ de clase C^1 a trozos, con discontinuidades de salto en los múltiplos impares de π (figura 49). Se tiene

$$\upsilon_{\text{per}}\big((2n+1)\pi^-\big) = \pi, \qquad \upsilon_{\text{per}}\big((2n+1)\pi^+\big) = -\pi,$$

y por tanto en esos puntos, por el Teorema de Dirichlet, la serie de Fourier converge a

$$\frac{\upsilon_{\text{per}}\big((2n+1)\pi^+\big) + \upsilon_{\text{per}}\big((2n+1)\pi^-\big)}{2} = 0.$$

El valor en esos puntos es, sin embargo, el valor $\upsilon_{\text{per}}((2n+1)\pi) = -\pi$, ya que hemos empezado con $\upsilon(x) = x$ sobre $[-\pi, \pi)$, incluyendo el extremo izquierdo π y no el derecho $-\pi$. De haber extendido $\upsilon(x) = x$ sobre $(-\pi, \pi]$, tendríamos el valor $\upsilon_{\text{per}}((2n+1)\pi) = \upsilon(\pi) = \pi$.

Este tipo de fenómeno plantea el siguiente problema:

¿Cómo se debe definir el valor de una función en un punto de salto?

La respuesta es que realmente no tiene gran importancia, por dos motivos:

- Los coeficientes de Fourier vienen dados por integrales, y cambiar una función en un número finito de puntos no cambia ni su integrabilidad ni el valor de su integral.
- Se defina como se defina el valor en un punto de salto c, la serie de Fourier sigue convergiendo al valor medio de los límites laterales,

$$\frac{f(c^+) + f(c^-)}{2},$$

  y esta media es independiente de lo que sea $f(c)$. Recordamos que esto es parte fundamental de la definición de límite: lo que ocurra en el propio punto no afecta al valor del límite.

Por las razones expuestas, es habitual al calcular series de Fourier

- No definir el valor en un punto de salto, o
- Tener dos valores contradictorios en un punto de salto como resultado de pegar dos gráficas (aunque esto debería solucionarse con mayor esmero que lo primero).

También es habitual decir que extendemos una función por periodicidad de un intervalo cerrado, como $[-\pi, \pi]$, aunque si los valores en los extremos no son iguales, esto conduce al segundo problema.

Resumimos algunos datos fáciles de verificar sobre la extensión periódica.

Propiedades de la extensión periódica

Sea $v : [c, c+T) \to \mathbb{R}$ y v_{per} su extensión periódica á $\mathbb{R}$.

- *Si v es continua a trozos en $[c, c+T)$, entonces v_{per} lo es en $\mathbb{R}$.*
- *Si v es de clase C^1 a trozos en $[c, c+T)$, entonces v_{per} lo es en $\mathbb{R}$.*

Si v es continua en el intervalo cerrado $[c, c+T]$, entonces v_{per} es continua en $\mathbb{R}$ si y sólo si

$$v(c) = v(c+T).$$

Si no se cumple esta condición, la función extendida v_{per} no está bien definida en los puntos $c+nT$, $n \in \mathbb{Z}$ (hay que decidir qué valor asignarle) y en todo caso va a tener saltos en estos puntos.

9.6.1. Algunos ejemplos de series de Fourier

La extensión periódica proporciona muchos ejemplos de señales periódicas a las que son aplicables el Teorema de Dirichlet y por tanto, están representadas por su serie de Fourier, teniendo en cuenta el fenómeno de convergencia al valor medio en los saltos.

• **Ejemplo:** Vamos a hallar la serie de Fourier de la función υ que, en $[-\pi, \pi)$, se define como

$$\upsilon(x) = \begin{cases} x + \pi, & \text{si } -\pi \le x < 0, \\ 0, & \text{si } 0 \le x < \pi, \end{cases}$$

extendida por periodicidad a una función $\upsilon_{\text{per}}(x)$ en todo $\mathbb{R}$, periódica de periodo 2π. Esta función tiene la siguiente gráfica:

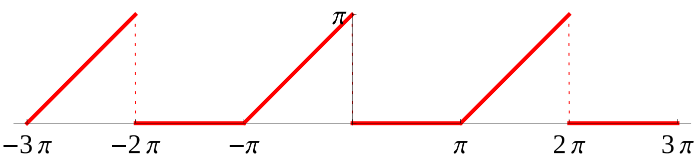

Figura 50

Calculemos los coeficientes de Fourier de υ. Comencemos con a_0:

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} \upsilon(x)\, dx = \frac{1}{\pi}\int_{-\pi}^{0} (x+\pi)\, dx = \frac{1}{\pi}\left(\frac{x^2}{2} + \pi x\right)\Big|_{-\pi}^{0} = \frac{1}{\pi}\cdot\frac{\pi^2}{2} = \frac{\pi}{2}.$$

$$\boxed{a_0 = \frac{\pi}{2}}$$

Para calcular a_n con $n \ge 1$, integramos por partes con

$$\begin{aligned} u &= x + \pi & dv &= \cos(nx)\, dx \\ du &= dx & v &= \frac{\sin(nx)}{n}. \end{aligned}$$

De este modo,

$$\begin{aligned} a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} \upsilon(x)\cos(nx)\, dx = \frac{1}{\pi}\int_{-\pi}^{0} (x+\pi)\cos(nx)\, dx \\ &= \frac{1}{\pi}\left((x+\pi)\frac{\sin(nx)}{n}\Big|_{-\pi}^{0} - \frac{1}{n}\int_{-\pi}^{0} \sin(nx)\, dx \right) \\ &= \frac{1}{\pi}\left(\pi \cdot \frac{\sin(0)}{n} - 0 \cdot \frac{\sin(-n\pi)}{n} - \frac{1}{n}\cdot\left[-\frac{\cos(nx)}{n}\right]_{-\pi}^{0} \right) \\ &= \frac{1}{\pi n^2}\Big[\cos(nx)\Big]_{-\pi}^{0} \\ &= \frac{1}{\pi n^2}(1 - \cos(-n\pi)) = \frac{1}{\pi n^2}(1 - \cos(n\pi)) \end{aligned}$$

y como $\cos(n\pi) = (-1)^n$ para $n \in \mathbb{Z}$,

$$= \frac{1}{\pi n^2}(1 - (-1)^n)$$

con lo cual

$$\boxed{a_n = \begin{cases} 0 & \text{si } n \text{ es par,} \\ \dfrac{2}{\pi n^2} & \text{si } n \text{ es impar.} \end{cases} \qquad n = 1, 2, 3, \ldots}$$

Dado que los subíndices pares dan coeficientes nulos, podemos reescribir este resultado como

$$\boxed{a_{2n-1} = \frac{2}{\pi(2n-1)^2}, \qquad a_{2n} = 0, \qquad n = 1, 2, 3, \ldots}$$

Para calcular b_n con $n \geq 1$, integramos por partes con

$$\begin{aligned} u &= x + \pi & dv &= \sin(nx)\,dx \\ du &= dx & v &= -\frac{\cos(nx)}{n}. \end{aligned}$$

De este modo, obtenemos

$$\begin{aligned} b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} \upsilon(x)\sin(nx)\,dx = \frac{1}{\pi}\int_{-\pi}^{0}(x+\pi)\sin(nx)\,dx \\ &= \frac{1}{\pi}\left((x+\pi)\frac{-\cos(nx)}{n}\bigg|_{-\pi}^{0} + \frac{1}{n}\int_{-\pi}^{0}\cos(nx)\,dx \right) \\ &= \frac{1}{\pi}\left(\pi\cdot\frac{-\cos(0)}{n} - 0\cdot\frac{-\cos(-n\pi)}{n} + \frac{1}{n}\cdot\left[\frac{\sin(nx)}{n}\right]_{-\pi}^{0} \right) \end{aligned}$$

y como $\sin(n\pi) = 0$ para todo $n \in \mathbb{Z}$,

$$= \frac{1}{\pi}\cdot\pi\cdot\frac{-1}{n} = -\frac{1}{n}$$

$$\boxed{b_n = -\frac{1}{n}, \qquad n = 1, 2, 3, \ldots}$$

En consecuencia, por el Teorema de Dirichlet,

$$\boxed{\upsilon(x) = \frac{\pi}{4} + \sum_{n=1}^{\infty}\frac{2\cos((2n-1)x)}{\pi(2n-1)^2} - \sum_{n=1}^{\infty}\frac{\sin(nx)}{n} = \begin{cases} x+\pi, & \text{si } -\pi \leq x < 0, \\ 0, & \text{si } 0 \leq x < \pi, \end{cases}}$$

en los puntos de continuidad de υ.

En los puntos de salto de υ la serie converge al valor medio de los límites laterales. Dado que $\upsilon(-\pi) = \upsilon(\pi) = 0$, el único salto de υ es en 0, y éste se reproduce a saltos de $\upsilon_{\text{per}}(x)$ en todo $x = 2\pi n$ con $n \in \mathbb{Z}$. En estos puntos

$$\frac{\upsilon(2n\pi^+) + \upsilon(2n\pi^-)}{2} = \frac{\upsilon(0^+) + \upsilon(0^-)}{2} = \frac{0+\pi}{2} = \frac{\pi}{2}.$$

Valores particulares. Si hacemos $x = 0$, entonces queda

$$\frac{\pi}{4} + \sum_{n=1}^{\infty} \frac{2}{\pi(2n-1)^2} - 0 = \frac{v(0^+) + v(0^-)}{2} = \frac{\pi}{2}$$

y de aquí es inmediato llegar a

$$\boxed{\sum_{n=1}^{\infty} \frac{1}{(2n-1)^2} = \frac{\pi^2}{8}}$$

Es interesante darse cuenta que el uso de series de Fourier resulta muy adecuado para calcular (de manera exacta) la suma de algunas series (no sólo para saber que la serie en cuestión es convergente, que en general es mucho más sencillo).

Precisamente del valor hallado se deduce otro muy conocido: descomponiendo en sumandos pares e impares, podemos escribir

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \sum_{n=1}^{\infty} \frac{1}{(2n)^2} + \sum_{n=1}^{\infty} \frac{1}{(2n-1)^2} = \frac{1}{4}\sum_{n=1}^{\infty} \frac{1}{n^2} + \frac{\pi^2}{8}$$

y despejando, obtenemos

$$\boxed{\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}}$$

Hallar la suma de esta serie es conocido como el Problema de Basilea y tiene una historia larga en la que intervinieron muchos de los grandes matemáticos de los siglos pasados.

Existen fórmulas similares para cualquier serie de la forma $\sum_{n=1}^{\infty} 1/n^{2k}$ con k un entero positivo; sin embargo, no se conoce el valor exacto de $\sum_{n=1}^{\infty} 1/n^{2k+1}$ para ningún entero positivo k. Este es un problema abierto desde hace varios siglos en la Teoría de Números.

• **Ejemplo:** Consideremos la función

$$\boxed{v(x) = x \quad \text{sobre} \quad (-\pi, \pi)}$$

extendida por periodicidad á $\mathbb{R}$, dejando de momento sin definir el valor en los múltiplos impares de π, y denotando por $v_{\text{per}}(x)$ a la función extendida.

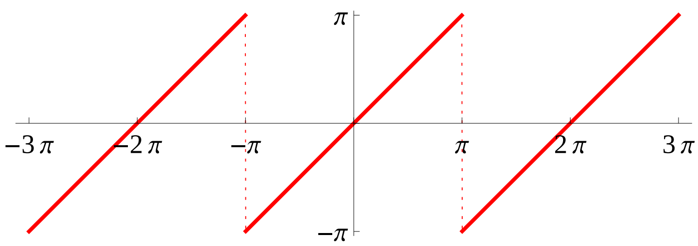

Figura 51 – La extensión periódica de la función identidad en $[-\pi, \pi]$.

Calculamos

$$\boxed{a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} x\cos(nx)\,dx = 0, \quad n = 0, 1, 2, \dots}$$

al ser $x\cos(nx)$ una función impar en el intervalo $(-\pi, \pi)$. Por otra parte, integrando por partes,

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} \underbrace{x}_{u}\,\underbrace{\sin(nx)\,dx}_{dv}, \qquad \left\{\begin{matrix} u = x & dv = \sin(nx)\,dx \\ du = dx & v = -\dfrac{1}{n}\cos(nx) \end{matrix}\right\}$$

$$= \frac{1}{\pi}\left(-\frac{1}{n}\Big[x\cos(nx)\Big]_{-\pi}^{\pi} + \frac{1}{n}\int_{-\pi}^{\pi}\cos(nx)\,dx\right)$$

$$= \frac{1}{\pi}\left(-\frac{1}{n}\Big[\pi\cos(n\pi) - (-\pi)\cos(-n\pi)\Big] + \frac{1}{n}\cdot\frac{1}{n}\Big[\sin(nx)\Big]_{-\pi}^{\pi}\right)$$

$$= \frac{1}{\pi}\left(-\frac{1}{n}\cdot 2\pi\cos(n\pi) + \frac{1}{n^2}\cdot\Big[\sin(nx)\Big]_{-\pi}^{\pi}\right)$$

y como $\cos(n\pi) = (-1)^n$ y $\sin(n\pi) = 0$ para $n \in \mathbb{Z}$,

$$= \frac{1}{\pi}\left(-\frac{2\pi}{n}(-1)^n + \frac{1}{n^2}\cdot 0\right)$$

$$= (-1)^{n+1}\frac{2}{n}$$

$$\boxed{b_n = (-1)^{n+1}\frac{2}{n} \quad n = 1, 2, 3, \dots}$$

y por el Teorema de Dirichlet,

$$\boxed{x = 2\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}\sin(nx), \quad -\pi < x < \pi}$$

En los extremos $\pm\pi$, el teorema predice que la serie converge al valor medio de los límites laterales, que es 0. Efectivamente, sustituyendo $x = \pm\pi$ de hecho da la serie nula.

Valores particulares. Si hacemos $x = \frac{\pi}{2}$, teniendo en cuenta que

$$\sin\left(\frac{n\pi}{2}\right) = \begin{cases} 0 & \text{si } n \text{ es par} \\ (-1)^{(n-1)/2} & \text{si } n \text{ es impar} \end{cases}$$

queda, cambiando n por $2n-1$ para eliminar los coeficientes nulos de índice par,

$$\frac{\pi}{2} = 2\sum_{n=1}^{\infty}\frac{(-1)^{2n}}{2n-1}(-1)^{(2n-2)/2} = 2\sum_{n=1}^{\infty}\frac{(-1)^{(n-1)}}{2n-1}$$

y despejando,

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

una serie famosa conocida como la serie de Leibniz.

9.7. Paridad

En el ejemplo de la extensión periódica de $v(x) = x$ vimos que los coeficientes de Fourier correspondientes a los cosenos se anulaban porque la función $v(x)$ era impar. Si esto es un principio general, será útil porque nos ahorrará la mitad de las cuentas para calcular los coeficientes de Fourier.

Surge por tanto la pregunta de cómo se refleja la paridad en la serie de Fourier. Recordamos la definición de la paridad para funciones:

$$\begin{array}{ll} v(-x) = v(x) & \text{función par} \\ v(-x) = -v(x) & \text{función impar} \end{array}$$

La terminología proviene de lo que ocurre para las potencias $v(x) = x^n$ con $n \in \mathbb{N}$:

$$(-1)^n = \begin{cases} 1 & \text{si } n \text{ es par: } n = 0, 2, 4, 6, \dots \\ -1 & \text{si } n \text{ es impar: } n = 1, 3, 5, 7, \dots \end{cases}$$

por tanto x^n es par (como función) cuando n es par (como número) e impar (como función) cuando n es impar (como número).

Igulamente, las leyes de paridad para números corresponden a las reglas de signos:

$$\begin{array}{ll} \text{par} \times \text{par} = \text{par} & + \cdot + = + \\ \text{impar} \times \text{impar} = \text{par} & - \cdot - = + \\ \text{par} \times \text{impar} = \text{impar} & + \cdot - = - \\ \text{impar} \times \text{par} = \text{impar} & - \cdot + = - \end{array}$$

y se reflejan en las reglas de paridad para funciones:

- el producto de dos funciones pares o dos impares es par
- el producto de una función par y otra impar es impar

En lo que respecta a los armónicos,

- $\cos(nt)$ es par
- $\sin(nt)$ es impar

Geométricamente, la paridad se manifiesta como simetría:

- Las funciones pares tienen gráfica simétrica respecto al eje y
- Las funciones impares tienen gráfica simétrica respecto al origen

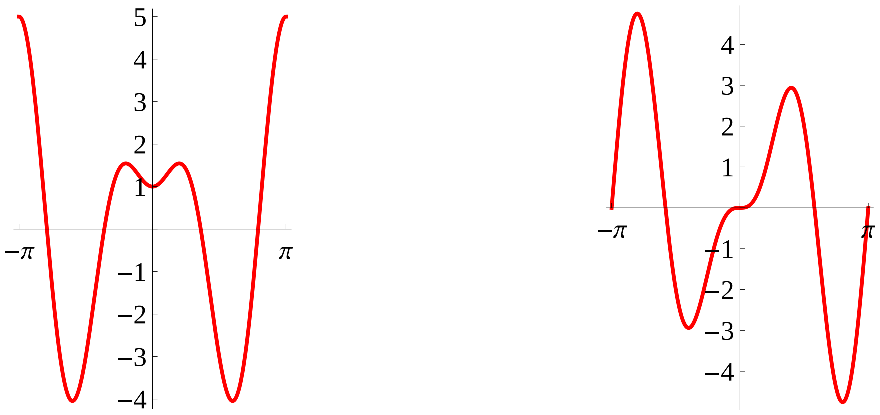

Figura 52 – Onda par (izquierda) e impar (derecha).

La interpretación geométrica de la paridad nos lleva a asegurar que

- Si υ es impar, $\displaystyle\int_{-T}^{T} \upsilon(t)\,dt = 0.$

  O sea, la integral sobre un intervalo simétrico respecto al origen es nula.

- Si υ es par, $\displaystyle\int_{-T}^{T} \upsilon(t)\,dt = 2\int_{0}^{T} \upsilon(t)\,dt.$

  O sea, la integral sobre un intervalo simétrico respecto al origen es el doble que sobre la mitad del intervalo (también es cierto con la mitad $[-T, 0]$).

- Sea ahora $\upsilon : [-\pi, \pi]$ extendida por periodicidad con periodo 2π.

Si $\upsilon(t)$ es par entonces sus coeficientes de Fourier de los **senos** son nulos:

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} \upsilon(t)\sin(nt)\,dt = 0, \quad n = 1, 2, 3, \ldots$$

porque la función $\upsilon(t)\sin(nt)$ es impar al ser producto de una función par (υ) con una impar ($\sin(nt)$). Por otra parte, $\upsilon(t)\cos(nt)$ es par al ser par por par, y esto se puede aprovechar para dividir por la mitad el intervalo de integración en la fórmula para los coeficientes de los cosenos:

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} \upsilon(t)\,\cos(nt)\,dt = \frac{2}{\pi}\int_{0}^{\pi} \upsilon(t)\,\cos(nt)\,dt.$$

Por tanto,

Serie de Fourier de una función par

La serie de Fourier de una función par consta sólo de ***cosenos***

$$\upsilon \text{ par}: \qquad \upsilon(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n\cos(nt), \qquad a_n = \frac{2}{\pi}\int_{0}^{\pi} \upsilon(t)\,\cos(nt)\,dt,$$

Si $\upsilon(t)$ es impar entonces sus coeficientes de Fourier de los **cosenos** son nulos:

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} \upsilon(t)\cos(nt)\,dt = 0, \qquad n = 0, 2, 3, \ldots$$

porque la función $\upsilon(t)\cos(nt)$ es impar al ser producto de una función impar (υ) con una par ($\cos(nt)$). Por otra parte, $\upsilon(t)\sin(nt)$ es par al ser impar por impar, y esto se puede aprovechar para dividir por la mitad el intervalo de integración en la fórmula para los coeficientes de los senos:

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} \upsilon(t)\,\sin(nt)\,dt = \frac{2}{\pi}\int_{0}^{\pi} \upsilon(t)\,\sin(nt)\,dt.$$

Por tanto,

Serie de Fourier de una función impar

La serie de Fourier de una función impar consta sólo de ***senos***

$$\upsilon \text{ impar}: \qquad \upsilon(t) = \sum_{n=1}^{\infty} b_n \sin(nt), \qquad b_n = \frac{2}{\pi}\int_{0}^{\pi} \upsilon(t)\,\sin(nt)\,dt,$$

Merece la pena caer en la cuenta de que todo esto que hemos obtenido es bastante natural pues, cuando υ es par, su serie de Fourier sólo tiene funciones pares (los cosenos); y, cuando υ es impar, sólo funciones impares (los senos).

• **Ejemplo:** Consideremos la función par

$$\boxed{f(x) = \pi^2 - x^2, \quad x \in [-\pi, \pi]}$$

extendida á $\mathbb{R}$ con periodo 2π

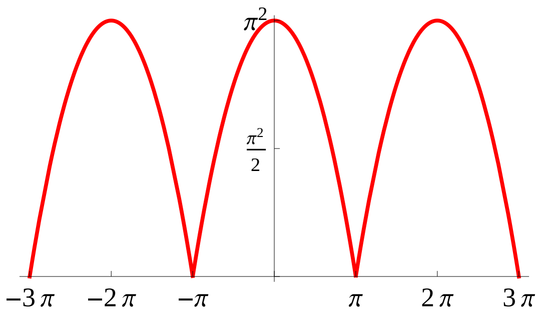

Figura 53 – La onda $\pi^2 - x^2$

Veamos que admite el siguiente desarrollo en serie de Fourier de cosenos:

$$\pi^2 - x^2 = \frac{2\pi^2}{3} + 4\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^2}\cos(nx), \quad x \in [-\pi, \pi].$$

Primero calculamos

$$a_0 = \frac{2}{\pi}\int_0^\pi (\pi^2 - x^2)\,dx = \frac{2}{\pi}\left(\pi^3 - \frac{\pi^3}{3}\right) = \frac{2}{\pi}\cdot\frac{2}{3}\pi^3 = \frac{4\pi^2}{3}$$

y después, para $n \geq 1$, teniendo en cuenta la relación de ortogonalidad de $\cos(nx)$ con 1,

$$\begin{aligned} a_n &= \frac{2}{\pi}\int_0^\pi (\pi^2 - x^2)\cos(nx)\,dx \\ &= 2\pi\int_0^\pi \cos(nx)\,dx - \frac{2}{\pi}\int_0^\pi x^2\cos(nx)\,dx \\ &= -\frac{2}{\pi}\int_0^\pi x^2\cos(nx)\,dx \end{aligned}$$

Integramos por partes dos veces, primero con

$$\begin{aligned} u &= x^2 & dv &= \cos(nx) \\ du &= 2x\,dx & v &= \frac{1}{n}\sin(nx) \end{aligned} \quad \therefore \quad \int x^2\cos(nx) = \frac{1}{n}x^2\sin(nx) - \frac{2}{n}\int x\sin(nx)\,dx$$

y volviendo a integrar por partes, pero ahora con

$$\begin{aligned} u &= x & dv &= \sin(nx) \\ du &= dx & v &= -\frac{1}{n}\cos(nx) \end{aligned} \quad \therefore \quad \begin{aligned} \int x\sin(nx)\,dx &= -\frac{1}{n}x\cos(nx) + \frac{1}{n}\int\cos(nx)\,dx \\ &= -\frac{1}{n}x\cos(nx) + \frac{1}{n^2}\sin(nx) \end{aligned}$$

por tanto

$$\boxed{\int x^2\cos(nx)\,dx = \frac{1}{n}x^2\sin(nx) + \frac{2}{n^2}x\cos(nx) - \frac{2}{n^3}\sin(nx)}$$

y finalmente obtenemos

$$\begin{aligned} a_n = -\frac{2}{\pi}\int_0^\pi x^2\cos(nx)\,dx &= -\frac{2}{\pi}\left[\frac{1}{n}x^2\sin(nx) + \frac{2}{n^2}x\cos(nx) - \frac{2}{n^3}\sin(nx)\right]_0^\pi \\ &= -\frac{2}{\pi}\left(0 + \frac{2}{n^2}\pi(-1)^n - 0\right) = 4\frac{(-1)^{n+1}}{n^2} \end{aligned}$$

con lo cual queda probada la fórmula que habíamos puesto para la serie de Fourier.

9.7.1. Otras simetrías

Si al dibujar la gráfica de una función, vemos que ésta tiene simetría respecto a un eje que no es el eje y ó respecto a un punto p que no es el origen, entonces la función no será par o impar según las definiciones anteriores. Aún así, trasladando el eje o el punto para que coincidan con el eje y o el origen, respectivamente, obtendremos funciones pares o impares. Entonces podemos calcular la serie de Fourier de la nueva función y, deshaciendo la traslación, obtener la serie de la función original.

Sea u la función

$$u(x) = \begin{cases} 0 & \text{si } -\pi < x < 0 \\ 1/2 & \text{si } 0 < x < \pi \end{cases}$$

que se extiende por periodicidad á $\mathbb{R}$. Su gráfica es como en la figura 54.

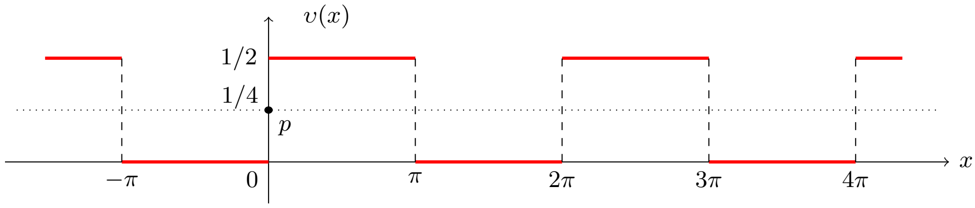

Figura 54 – Onda cuadrada con simetría alrededor del punto $p = (0, \frac{1}{4})$.

La función u no es ni par ni impar, pero se ve fácilmente que su gráfica es simétrica alrededor del punto $p = (0, \frac{1}{4})$. Trasladando el eje horizontal hacia abajo por $1/4$, movemos este punto al origen y obtenemos la función impar

$$\upsilon(x) = u(x) - \frac{1}{4} = \begin{cases} -1/4, & \text{si } -\pi < x < 0, \\ 1/4, & \text{si } 0 < x < \pi, \end{cases}$$

que se extiende a todo $\mathbb{R}$ con periodicidad 2π.

Cálculo de la serie de Fourier de υ. Como υ es impar, los coeficientes a_n de los cosenos son todos nulos, y los coeficientes del seno son

$$\begin{aligned} b_n &= \frac{2}{\pi}\int_0^\pi \upsilon(x)\sin(nx)\,dx = \frac{2}{\pi}\int_0^\pi \frac{1}{4}\sin(nx)\,dx = \frac{1}{2\pi}\int_0^\pi \sin(nx)\,dx \\ &= \frac{1}{2\pi}\cdot\left[-\frac{1}{n}\cos(nx)\right]_0^\pi = \frac{1-\cos(n\pi)}{2\pi n} \\ &= \frac{1-(-1)^n}{2\pi n} = \begin{cases} 0, & \text{si } n \text{ es par,} \\ \dfrac{1}{\pi n}, & \text{si } n \text{ es impar.} \end{cases} \end{aligned}$$

Escribiendo $2n-1$ en lugar de n para quedarnos sólo con los coeficientes impares, tendremos

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

Como consecuencia, la serie de Fourier de $u(x) = 1/4 + \upsilon(x)$ será

$$u(x) = \frac{1}{4} + \frac{1}{\pi}\sum_{n=1}^\infty \frac{\sin((2n-1)x)}{2n-1} = \begin{cases} 0 & \text{si } -\pi < x < 0 \\ 1/2 & \text{si } 0 < x < \pi \end{cases}$$

Valores particulares. Haciendo $x = \frac{\pi}{2}$ vuelve a quedar la serie de Leibniz para $\frac{\pi}{4}$.

Obsérvese que el término constante $a_0/2 = 1/4 = \frac{1}{\pi}\int_{-\pi}^{\pi} v(t)\,dt$ que nos está apareciendo ahora es el que representa el desplazamiento vertical de la forma de onda de la función $u(x)$; no tiene ninguna influencia en la variable independiente x. En cambio, si hubiera un desplazamiento horizontal, sí se reflejaría en la variable (ver la sección 1.6.4).

9.8. Series de Fourier con periodos arbitrarios

Supongamos que una señal $v : \mathbb{R} \to \mathbb{R}$ tiene periodo T. Entonces

$$\boxed{u(t) = v\left(\frac{T}{2\pi}t\right)}$$

tiene periodo 2π:

$$u(t+2\pi) = v\left(\frac{T}{2\pi}(t+2\pi)\right) = v\left(\frac{T}{2\pi}t + T\right) = v\left(\frac{T}{2\pi}t\right) = u(t).$$

El cambio de variable lineal que transforma v en u también transforma el intervalo $[-\frac{T}{2}, \frac{T}{2}]$ en $[-\pi, \pi]$. Además, conserva la propiedad de ser de clase C^1 que aparece como hipótesis en el Teorema de Dirichlet. Suponiendo entonces que v es de clase C^1 a trozos en $[-\frac{T}{2}, \frac{T}{2}]$, entonces u lo es en $[-\pi, \pi]$ y por tanto está representada por su serie de Fourier:

$$u(t) = v\left(\frac{T}{2\pi}t\right) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\big(a_n\cos(nt) + b_n\sin(nt)\big), \qquad t \in [-\pi, \pi]$$

Deshaciendo el cambio con $t \to \frac{2\pi}{T}t = \omega t$ (la frecuencia angular) queda

$$\boxed{v(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\big(a_n\cos(n\omega t) + b_n\sin(n\omega t)\big)}$$

como expresión general de serie de Fourier con cualquier periodo.

También es deseable modificar las fórmulas para los coeficientes de Fourier para expresarlas como integrales cobre el intervalo simétrico $[-\frac{T}{2}, \frac{T}{2}]$. Esto también se hace aplicando el mismo cambio de variable $t \to \frac{2\pi}{T}t = \omega t$ a las fórmulas originales. Por ejemplo,

$$\begin{aligned} a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} u(t)\,\cos(nt)\,dt &= \frac{1}{\pi}\int_{-\pi}^{\pi} v\left(\frac{T}{2\pi}t\right)\,\cos(nt)\,dt \\ &= \frac{1}{\pi}\int_{-T/2}^{T/2} v(t)\,\frac{2\pi}{T}\,\cos\left(\frac{2\pi n}{T}t\right)\,dt \\ &= \frac{2}{T}\int_{-T/2}^{T/2} v(t)\,\cos(n\omega t)\,dt \end{aligned}$$

y similarmente para los coeficientes b_n de los senos. Resumimos los resultados.

La serie de Fourier con un periodo arbitario

Sea $v : \mathbb{R} \to \mathbb{R}$ periódica de periodo T, con frecuencia angular correspondiente $\omega = \dfrac{2\pi}{T}$. La serie de Fourier de v es

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} \big(a_n \cos(n\omega t) + b_n \sin(n\omega t)\big)$$

con los coeficientes dados por

$$a_n = \frac{2}{T}\int_{-T/2}^{T/2} v(t)\, \cos(n\omega t)\, dt, \qquad b_n = \frac{2}{T}\int_{-T/2}^{T/2} v(t)\, \sin(n\omega t)\, dt.$$

Puede comprobarse inmediatamente que para el periodo $T = 2\pi$, cuya frecuencia angular es $\omega = 1$, estas fórmulas nuevas nos devuelven las originales.

• **Ejemplo:** En un ejemplo anterior hemos hallado la serie de Fourier de la extensión 2π-periódica de la función $I(x) = x$ sobre $[-\pi, \pi]$:

$$\boxed{x = 2\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sin(nx), \quad -\pi < x < \pi}$$

Cambiando de variable por $x \to \dfrac{2\pi}{T} x = \omega x$, queda

$$\omega x = 2\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sin(n\omega x), \quad -\pi < \frac{2\pi}{T} x < \pi$$

o sea

$$\boxed{x = 2\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n\omega} \sin(n\omega x), \quad -\frac{T}{2} < x < \frac{T}{2}}$$

y esta es la serie de Fourier de la extensión T-periódica de $I(x) = x$ sobre $[-T/2, T/2]$ para cualquier periodo T.

Con este procedimiento sencillo de cambio de variable, vemos que no se pierde realmente generalidad por suponer que usamos un periodo normalizado. Habitualmente se utilizan $T = 2\pi$, con frecuencia angular $\omega = 1$, ó $T = 1$, con frecuencia angular 2π.

• **Ejemplo:** Sea $\{x\}$ la función parte fraccional, dada por

$$\{x\} = x - \lfloor x \rfloor$$

es decir, quitando a un número el entero $n = \lfloor x \rfloor$ más proximo y menor o igual que él, o sea, redondeando hacia abajo al entero más próximo.

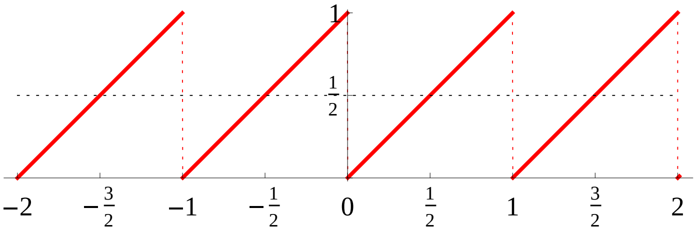

Figura 55 – La función parte fraccional $\{x\}$.

$\{x\}$ es periódica de periodo $T = 1$, ya que al sumar 1 a un número x, sólo cambia su parte entera pero no su parte fraccional. Como puede apreciarse en la figura 55, la función $\{x\}$ no es ni par ni impar, pero sí tiene simetría alrededor del punto $(0, 1/2)$. Entonces, desplazándola por $1/2$, o sea, considerando

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

se consigue una función impar, cuya serie de Fourier constará por tanto solamente de senos.

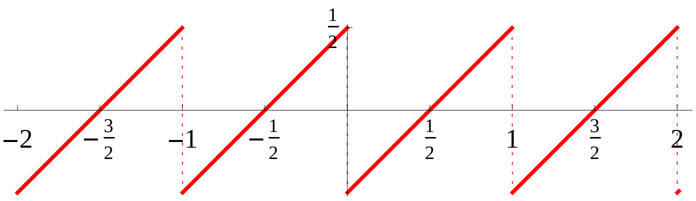

Figura 56 – La función $\{x\} - \frac{1}{2}$

Calculamos los coeficientes de los senos también aprovechando la simetría. Como la frecuencia angular $\omega = \frac{2\pi}{T} = 2\pi$, tenemos

$$\begin{aligned} b_n &= \frac{1}{T/2} \int_{-T/2}^{T/2} f(x) \sin(n\omega x)\, dx \\ &= 2 \int_{-1/2}^{1/2} \left(\{x\} - \frac{1}{2} \right) \sin(2\pi n x)\, dx \\ &= 2 \cdot 2 \int_{0}^{1/2} \left(\{x\} - \frac{1}{2} \right) \sin(2\pi n x)\, dx \end{aligned}$$

y como $\{x\} = x$ para $0 \leq x < 1$ (la parte entera es nula),

$$= 4 \int_0^{1/2} \left(x - \frac{1}{2} \right) \sin(2\pi n x)\, dx$$

ahora, integrando por partes con $u = x - \frac{1}{2}$, $du = dx$, $dv = \sin(2\pi n x)$, $v = -\frac{\cos(2\pi n x)}{2\pi n}$,

$$= 4 \left(\left[-\left(x - \frac{1}{2} \right) \frac{\cos(2\pi n x)}{2\pi n} \right]_0^{1/2} + \frac{1}{2\pi n} \int_0^{1/2} \cos(2\pi n x)\, dx \right)$$

$$= 4\left(-\frac{1}{2}\frac{\cos(0)}{2\pi n} + \frac{1}{(2\pi n)^2}\Big[\sin(2\pi n x)\Big]_0^{1/2}\right)$$
$$= 4\left(-\frac{1}{4\pi n} + \frac{1}{(2\pi n)^2}\Big[\sin(\pi n) - \sin(0)\Big]\right)$$
$$= 4\left(-\frac{1}{4\pi n} + \frac{1}{(2\pi n)^2}\Big[\sin(\pi n) - \sin(0)\Big]\right) = -\frac{1}{\pi n}.$$

Por tanto queda

$$\boxed{\{x\} - \frac{1}{2} = -\frac{1}{\pi}\sum_{n=1}^{\infty}\frac{\sin(2\pi n x)}{n}}$$

y esta ecuación es correcta para todo $x \in \mathbb{R}$, al menos en el sentido del Teorema de Dirichlet, pues la función $\{x\}$ es periódica (no hablamos de una extensión periódica aquí sino de la propia función).

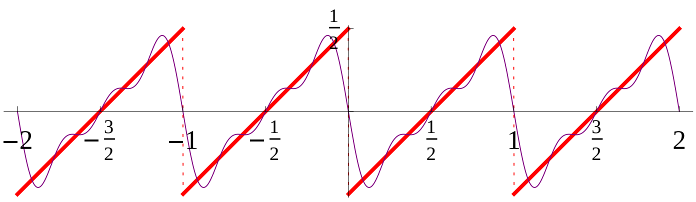

Figura 57 – La función $\{x\} - \frac{1}{2}$ y su aproximación de Fourier hasta los terceros armónicos.

Los saltos de la función $\{x\} - \frac{1}{2}$ están en los x enteros, donde los límites laterales son $\pm\frac{1}{2}$, por tanto su valor medio es 0, y es a 0 a lo que tiende la serie de Fourier en los enteros, tal como señala el Teorema de Dirichlet. Sin embargo, la función $\{x\}$ vale 0 en los enteros, y por tanto $\{x\} - \frac{1}{2}$ vale $-\frac{1}{2}$ en los enteros, pero como ya comentamos en su momento, aún así es habitual usar el símbolo de igualdad en la representación de la función por la serie, entendiendo que existen estos matices.

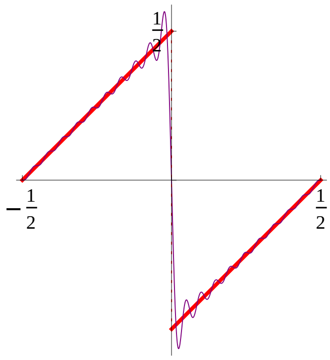

Figura 58 – La función $\{x\}-\frac{1}{2}$ y los primeros 20 sumandos de su serie de Fourier. La oscilación en los saltos no es un artefacto gráfico: es una oscilación real conocida como **fenómeno de Gibbs.**

9.9. Resultados teóricos

Linealidad de la serie de Fourier

La operación de formar la serie de Fourier de una función es lineal. Es decir, si f, F son C^1 a trozos, con series de Fourier

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

$$g(x) = \frac{A_0}{2} + \sum_{n=1}^{\infty}(A_n \cos(n\omega t) + B_n \sin(n\omega t)),$$

entonces su suma también es C^1 a trozos, con serie de Fourier obtenida sumando las series de Fourier de f, F y reuniendo términos:

$$f(x) + g(x) = \frac{a_0 + A_0}{2} + \sum_{n=1}^{\infty}\big((a_n + A_n)\cos(n\omega t) + (b_n + B_n)\sin(n\omega t)\big)$$

Equivalentemente, los coeficientes de Fourier de la suma son la suma de los coeficientes de Fourier correspondientes de cada sumando.

Similarmente, los coeficientes de Fourier de un múltiplo kf se obtienen multiplicando los coeficientes de f por k.

Demostración. Esto es consecuencia de las propiedades básicas de las series infinitas. □

9.9.1. La identidad de Parseval

La identidad de Parseval

Sea υ una función periódica de periodo T y frecuencia angular $\omega = \dfrac{2\pi}{T}$ que es C^1 a trozos. Entonces

$$\frac{2}{T}\int_{-T/2}^{T/2} |f(x)|^2\, dx = \frac{a_0^2}{2} + \sum_{n=1}^{\infty}\big(a_n^2 + b_n^2\big)$$

Demostración. Por el Teorema de Dirichlet,

$$\upsilon(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos(n\omega t) + b_n \sin(n\omega t)).$$

Multipliquemos esta igualdad por $\upsilon(t)$:

$$\upsilon(t)^2 = \frac{a_0}{2}\upsilon(t) + \sum_{n=1}^{\infty}(a_n\upsilon(t)\ \cos(n\omega t) + b_n\upsilon(t)\ \sin(n\omega t)).$$

Ahora, integremos entre $-T/2$ y $T/2$. Suponiendo cierto que en este caso la integral se puede intercambiar con la suma infinita, obtenemos

$$\begin{aligned}
&\int_{-T/2}^{T/2} v(t)^2\,dt \\
&= \frac{a_0}{2}\int_{-T/2}^{T/2} v(t)\,dt + \sum_{n=1}^{\infty} a_n \int_{-T/2}^{T/2} v(t)\,\cos(n\omega t)\,dt + b_n \int_{-T/2}^{T/2} v(t)\,\sin(n\omega t)\,dt
\end{aligned}$$

reconociendo en estas integrales las fórmulas para los coeficientes salvo el factor $2/T$,

$$\begin{aligned}
&= \frac{a_0}{2}\cdot\frac{T}{2}\cdot a_0 + \sum_{n=1}^{\infty}\left(a_n\cdot\frac{T}{2}\cdot a_n + b_n\cdot\frac{T}{2}\cdot b_n\right) \\
&= \frac{T}{2}\cdot\left(\frac{a_0^2}{2} + \sum_{n=1}^{\infty}(a_n^2+b_n^2)\right)
\end{aligned}$$

y al multiplicar por $2/T$ queda la identidad de Parseval. □

El razonamiento es válido aunque en las discontinuidades de salto la serie de Fourier no tiene por qué converger al valor de la función, pues como ya comentamos antes, cambiar los valores de las funciones en un número finito de puntos no cambia los valores de sus integrales.

• **Ejemplo:** Aplicando la identidad de Parseval a la serie de Fourier para la extensión 2π-periódica de la identidad hallada antes,

$$\boxed{x = 2\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}\sin(nx), \qquad -\pi < x < \pi}$$

nos da $a_n = 0$, $b_n = 2(-1)^n/n$, por tanto

$$\frac{1}{\pi}\int_{-\pi}^{\pi} x^2\,dx = \frac{1}{\pi}\left[\frac{x^3}{3}\right]_{-\pi}^{\pi} = \frac{1}{\pi}\cdot\frac{2\pi^3}{3} = 4\sum_{n=1}^{\infty}\frac{1}{n^2} \iff \sum_{n=1}^{\infty}\frac{1}{n^2} = \frac{\pi^2}{6}.$$

9.9.2. El Teorema de Multiplicación

El Teorema de Multiplicación

Sean u, v periódicas de periodo T y frecuencia angular $\omega = \dfrac{2\pi}{T}$, ambas de clase C^1 a trozos y con series de Fourier

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

$$F(x) = \frac{A_0}{2} + \sum_{n=1}^{\infty}(A_n \cos(n\omega t) + B_n \sin(n\omega t)).$$

Entonces

$$\frac{1}{\ell}\int_{-\ell}^{\ell} F(x)f(x)\,dx = \frac{A_0 a_0}{2} + \sum_{n=1}^{\infty}(A_n a_n + B_n b_n)$$

Demostración. Sumando y restando una serie de otra, se obtiene

$$F(x) + f(x) = \frac{A_0 + a_0}{2} + \sum_{n=1}^{\infty}\big((A_n + a_n)\cos(n\omega t) + (B_n + b_n)\sin(n\omega t)\big),$$

$$F(x) - f(x) = \frac{A_0 - a_0}{2} + \sum_{n=1}^{\infty}\big((A_n - a_n)\cos(n\omega t) + (B_n - b_n)\sin(n\omega t)\big).$$

Aplicando la identidad de Parseval a estos dos nuevos desarrollos de Fourier, obtenemos

$$\frac{1}{T/2}\int_{-T/2}^{T/2}(F(x) + f(x))^2\,dx = \frac{(A_0 + a_0)^2}{2} + \sum_{n=1}^{\infty}\big((A_n + a_n)^2 + (B_n + b_n)^2\big),$$

$$\frac{1}{T/2}\int_{-T/2}^{T/2}(F(x) - f(x))^2\,dx = \frac{(A_0 - a_0)^2}{2} + \sum_{n=1}^{\infty}\big((A_n - a_n)^2 + (B_n - b_n)^2\big).$$

Finalmente, restando estas dos expresiones, llegamos a

$$\frac{1}{T/2}\int_{-T/2}^{T/2}4F(x)f(x)\,dx = \frac{4A_0 a_0}{2} + \sum_{n=1}^{\infty}(4A_n a_n + 4B_n b_n).$$

Dividiendo por 4 queda demostrado el teorema. □

El caso $f = F$ en el Teorema de Multiplicación es la identidad de Parseval. Dado que hemos utilizado la identidad de Parseval para demostrar el Teorema de Multiplicación, desde el punto de vista lógico, los dos resultados son equivalentes (cada uno es consecuencia del otro).

9.10. Derivación de series de Fourier

Si $f : \mathbb{R} \to \mathbb{R}$ es una función T-periódica derivable, entonces su derivada también es T-periódica. Esto es una consecuencia inmediata de la regla de la Cadena:

$$f(t+T) = f(t) \implies f'(t+T) = f'(T)$$

Por tanto, si f y f' están representadas por series de Fourier, cabe esperar alguna relación entre ambas series. Supongamos que

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

Si la serie fuera finita, como la derivada es lineal, podríamos derivar término a término, es decir, intercambiar la operación de derivada con la suma. Para la serie infinita, esto daría

$$f'(t) = \sum_{n=1}^{\infty} \big(-n\omega a_n \sin(n\omega t) + n\omega b_n \cos(n\omega t)\big)$$

Desafortunadamente, esto no siempre va a funcionar, y por varias razones, como veremos a continuación.

Por supuesto, primero hay que suponer que f, f' satisfacen las condiciones adecuadas para estar representadas por series de Fourier. Aún así, hay otro problema que resaltamos.

La derivación término a término de una serie de Fourier convergente no da lugar necesariamente a una serie convergente.

Sin más condiciones, no va a ser legítimo intercambiar la derivada con la suma infinita.

• **Ejemplo:** En un ejemplo anterior hemos hallado el desarrollo de Fourier de la extensión 2π-periódica de la función identidad $I(x) = x$ para $x \in [-\pi, \pi]$. Resultaba

$$I_{\text{per}}(x) = 2\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sin(nx) = 2\left(\sin(x) - \frac{\sin(2x)}{2} + \frac{\sin(3x)}{3} - \frac{\sin(4x)}{4} + \cdots\right)$$

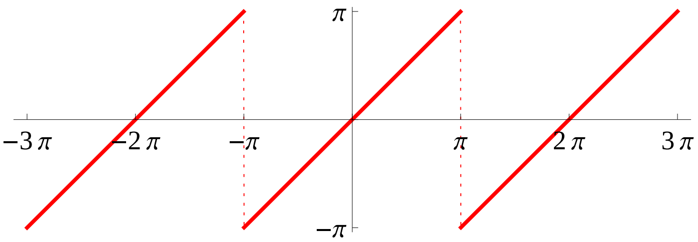

Figura 59 – La extensión 2π-periódica de la identidad.

La función tiene saltos en los múltiplos enteros de π, donde la serie converge al valor medio de los límites laterales, que es 0. En esos puntos I_{per} no es derivable, pero sí existen los límites laterales de la derivada. En los demás puntos, $I'_{\text{per}}(x) = 1$, por tanto, el enunciado resultante al derivar término a término es

$$1 \overset{?}{=} 2\sum_{n=1}^{\infty}(-1)^{n+1}\cos(nx) = 2\big(\cos(x) - \cos(2x) + \cos(3x) - \cos(4x) + \cdots\big)$$

pero esta serie no es convergente para ningún $x \in \mathbb{R}$, pues el término general $\cos(nx)$ no tiende á cero para ningún x.

Divergencia de $\cos(nx)$

No hay ningún $x \in \mathbb{R}$ tal que $\lim_{n\to\infty}\cos(nx) = 0$

Demostración. Empleamos el siguiente "truco:" si fuera $\lim_{n\to\infty}\cos(nx) = 0$, también sería cierto, cambiando de variable por $n \to 2n$, que $\lim_{n\to\infty}\cos(2nx) = 0$, pero como

$$1 + \cos(2nx) = 2\cos^2(nx),$$

haciendo $n \to \infty$ quedaría $0 = 1$! Esta contradicción prueba que no puede darse nunca el límite nulo. [14] □

Problema: utilizando el mismo truco pero con las dos identidades $\cos(2x) = 2\cos^2 x - 1$ y $\cos(3x) = 4\cos^3(x) - 3\cos(x)$, demostrar que el único valor posible para

$$\lim_{n\to\infty}\cos(nx) = \ell$$

es $\ell = 1$, que se da por ejemplo cuando x es un múltiplo entero de 2π. Es más difícil demostrar que sólo se da en este caso.

En el ejemplo anterior,

- La onda es del tipo dientes de sierra.
- La función que la representa es I_{per}, la extensión 2π-periódica de la función identidad $I(x) = x$ en $[-\pi, \pi]$.
- Esta función es de clase C^∞ a trozos, con saltos en los múltiplos enteros de π.

[14] La Teoría de Números da una explicación más profunda. Se puede demostrar que si $x \in \mathbb{R}$ es un múltiplo irracional de π, entonces dado cualquier $\xi \in \mathbb{R}$, hay alguna sucesión creciente de enteros positivos $n_1 < n_2 < \ldots < n_k < \ldots$ tales que $\lim_{k\to\infty}\cos(n_k x) = \cos(\xi)$. En este caso la sucesión $\{\cos(nx) : n \in \mathbb{N}\}$ no sólo no tiene un límite concreto como 0, sino que de hecho oscila, aproximándose a cualquier valor entre -1 y $+1$ infinitas veces. En cambio, si $x \in \mathbb{R}$ es un múltiplo racional de π, la sucesión $\cos(nx)$ asume periódicamente un número finito de valores. Sólo da un valor fijo, 1, cuando x es un múltiplo entero de 2π.

Dado que partiendo de una función tan "buena" como la identidad, ya nos falla la derivación término a término, está claro que va a ser necesaria alguna condición adicional para que se cumpla.

Pensando ahora en requisitos "técnicos," para que tengamos la representación de la derivada f' por su serie de Fourier (que no va a ser necesariamente la obtenida al derivar término a término la de f),

$$\boxed{f'(t) = \frac{\alpha_0}{2} + \sum_{n=1}^{\infty} \big(\alpha_n \cos(n\omega t) + \beta_n \sin(n\omega t)\big)}$$

por el Teorema de Dirichlet bastaría con que f' y f'' sean continuas a trozos o, lo que es lo mismo, que f sea de clase C^2 a trozos.

Sin embargo, el ejemplo anterior muestra que con esto no puede ser suficiente para garantizar que la serie de f' sea de hecho la obtenida derivando término a término la de f. Hace falta algo más. Resulta que la continuidad de la propia f, en todos los puntos, o sea, sin el calificativo "a trozos," es suficiente.

Teorema de derivación de series de Fourier

Sea $f : \mathbb{R} \to \mathbb{R}$ periódica tal que

- *f es de clase C^2 a trozos.*
- *f es continua en todo punto.*

Entonces la serie de Fourier de f' se obtiene por derivación término a término de la serie de Fourier de f.

Entonces, el ejemplo anterior se explicaría – al menos en parte – porque la extensión periódica de la identidad no es continua en todo punto (es una onda dientes de sierra).

9.10.1. Derivación de la serie de Fourier de una extensión periódica

En la práctica, se construyen funciones periódicas por extensión periódica de una función en un intervalo de longitud igual al periodo, como hemos hecho con la función identidad.

Entonces, teniendo en cuenta el Teorema de Derivación para series de Fourier, debemos analizar qué significa la hipótesis de continuidad en todo punto cuando se emplea este método.

Sea pues f una función de clase C^2 a trozos en el intervalo $[-T/2, T/2]$. Entonces su extensión T-periódica f_{per} a todo $\mathbb{R}$ es también C^2 a trozos. Aquí no hay ningún problema.

Sin embargo,

La extensión periódica f_{per} de una función incluso C^∞ a trozos, no es necesariamente continua en todo punto, como muestra el ejemplo de $f = I$, la identidad, en $[-\pi, \pi]$.

Por tanto, la pregunta clave es

¿Cuándo es continua en todo punto la extensión periódica de una función?

Claramente una condición necesaria es que f sea continua en el intervalo $[-T/2, T/2]$. En ese caso, es inmediato ver que f_{per} será continua salvo posiblemente en los puntos de la forma

$$(T/2) + nT, \quad n \in \mathbb{Z},$$

donde podría haber saltos. Para evitar esto, vemos que el requisito adicional para la continuidad de f_{per} en todo punto es el siguiente.

Para que la extensión periódica f_{per} de una función f en $[-T/2, T/2]$ sea continua, es necesario y suficiente que f sea continua en $[-T/2, T/2]$ y que se cumpla

$$f(-T/2) = f(T/2),$$

siendo este valor común igual también al valor de f_{per} en todos los puntos $(T/2) + nT$ con $n \in \mathbb{Z}$. De lo contrario, f_{per} tendrá saltos en estos puntos.

Geométricamente, la condición $f(-T/2) = f(T/2)$ dice que al trasladar por el periodo T, podemos “pegar” copias de la gráfica de f de manera continua, sin que haya saltos.

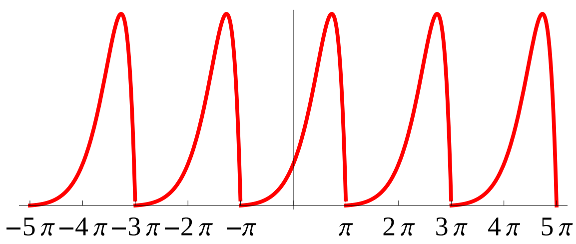

Figura 60 – Una función continua f en $[-\pi, \pi]$ cuya extensión periódica es continua en todo punto puesto que $f(-\pi) = f(\pi)$.

Con esta aclaración, podemos enunciar el Teorema de Derivación para funciones construidas por extensión periódica:

Teorema de derivación de series de Fourier para extensiones periódicas

Sea $f : [-T/2, T/2] \to \mathbb{R}$ *tal que*

- f *es de clase* C^2 *a trozos;*
- f *es continua en todo punto de* $[-T/2, T/2]$*;*
- $f(-T/2) = f(T/2)$.

Sea f_{per} *la extensión* T*-periódica de* f*. Entonces la serie de Fourier de la derivada* f'_{per} *se obtiene por derivación término a término de la serie de Fourier de* f_{per}*.*

• **Ejemplo:** Consideremos la función par (estudiada en un ejemplo anterior)

$$f(x) = \pi^2 - x^2, \quad x \in [-\pi, \pi]$$

extendida á $\mathbb{R}$ con periodo 2π

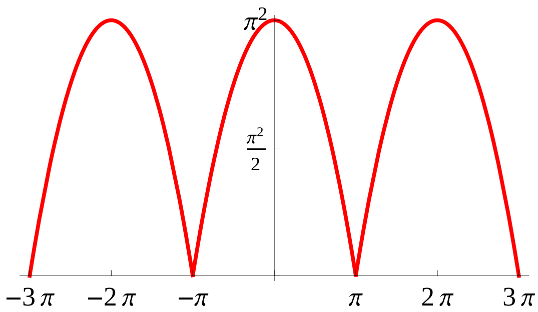

Figura 61 – La onda $\pi^2 - x^2$

y que, como vimos, admite el siguiente desarrollo en serie de Fourier de cosenos:

$$\pi^2 - x^2 = \frac{2\pi^2}{3} + 4\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^2} \cos(nx), \quad x \in [-\pi, \pi].$$

La función f

- es de clase C^∞ a trozos, en particular de clase C^2 a trozos, por ser un polinomio;
- es continua en todo punto, ya que cumple $f(-\pi) = f(\pi) = 0$.

Entonces, por el Teorema de Derivación, la serie de Fourier de f' se obtiene por derivación término a término de la serie de Fourier de f,

$$-2x = -4\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^2}\, n \sin(nx), \quad x \in [-\pi, \pi].$$

Así pues, simplificando recuperamos la serie de Fourier de la función x en $[-\pi, \pi]$:

$$x = 2\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sin(nx), \quad x \in [-\pi, \pi]$$

como también habíamos obtenido por cálculo directo de sus coeficientes. También cabe recordar que los extremos $\pm\pi$ son puntos de salto de la función x extendida por periodicidad, y por tanto en ellos la serie converge al valor medio de los límites laterales, que es 0. Es decir, aunque la función f extendida por periodicidad es continua en todo punto, su derivada no tiene por qué serlo.

Acabamos anotando un resultado de conmutación.

Derivación y extensión periódica

Sea $f : [c, c+T) \to \mathbb{R}$ derivable a trozos. Considerando la extensión T-periódica desde $[c, c+T)$ á $\mathbb{R}$,

$$(f')_{\text{per}} = (f_{\text{per}})'$$

es decir, la derivada de la extensión periódica es la extensión periódica de la derivada.

Demostración. Sea $x \in \mathbb{R}$. Existe un único $n \in \mathbb{Z}$ tal que $x - nT \in [c, c+T)$. Por definición,

$$f_{\text{per}}(x) = f(x - nT), \qquad (f')_{\text{per}}(x) = f'(x - nT)$$

Derivando la primera, queda la igualdad: $(f_{\text{per}})'(x) = f'(x - nT) = (f')_{\text{per}}(x)$. □

9.11. Integración de series de Fourier

9.11.1. Integración término a término de un desarrollo de Fourier

La integración término a término de series de Fourier se puede hacer con requisitos mucho más débiles que para la derivación.

Teorema de integración de series de Fourier

Sea f una función continua a trozos en $[-T/2, T/2]$ y cuya extensión T-periódica esté representada por su serie de Fourier:

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

Entonces es válida para todo $x \in \mathbb{R}$ la igualdad

$$\int_{-T/2}^{x} f(t)\, dt = \frac{a_0}{2} \int_{-T/2}^{x} dt + \sum_{n=1}^{\infty} \left(a_n \int_{-T/2}^{x} \cos(n\omega t) dt + b_n \int_{-T/2}^{x} \sin(n\omega t) dt \right),$$

obtenida integrando término a término la representación.

Nota: *este nuevo desarrollo en general* ***no*** *será una serie de Fourier para la integral.*

Aquí la "pega" principal es que al integrar el término constante, queda

$$\frac{a_0}{2}\int_{-T/2}^{x} dt = \frac{a_0}{2}\left(x + \frac{T}{2}\right)$$

pero el nuevo término con x no es un armónico básico.

• **Ejemplo:** La serie de Fourier de $f(t) = 1$ en cualquier intervalo $[-T/2, T/2]$ es

$$1 = 1 + \sum_{n=1}^{\infty}\big(0\cos(n\omega t) + 0\sin(n\omega t)\big).$$

Al integrar hasta x, es decir, hallar una antiderivada, nos va a quedar simplemente

$$F(x) = \int_{-T/2}^{x} dt = x + \frac{T}{2}$$

(la integral del término constante) pero ésta no es una serie de Fourier.

El obstáculo fundamental a poder obtener una serie de Fourier al integrar se puede apreciar en este ejemplo: si bien la función constante $f = 1$ es periódica, la antiderivada F no lo es. Vamos a analizar este fenómeno.

9.11.2. Periodicidad de la antiderivada

Sea $f : \mathbb{R} \to \mathbb{R}$ una función T-periódica y supongamos que f_{per} está representada por su serie de Fourier (será el caso si por ejemplo es C^1 a trozos):

$$\boxed{f_{\text{per}}(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n\cos(n\omega t) + b_n\sin(n\omega t)), \quad t \in \mathbb{R}}$$

Sea ahora

$$F(x) = \int_{-T/2}^{x} f(t)\,dt$$

que por el Teorema Fundamental del Cálculo, es una antiderivada de f en todo $\mathbb{R}$; es decir

$$F' = f.$$

Además, cualquier antiderivada es simplemente esta F más una constante, con lo cual es inmediato que F es periódica si y sólo si lo es cualquier antiderivada.

La pregunta principal que tenemos es

> ¿Cuándo es periódica la antiderivada de una función periódica?

La respuesta la da el siguiente resultado. Esencialmente, el obstáculo era únicamente el que habíamos señalado antes: el término a_0 en la serie de Fourier de f, que al integrar introduce un término con x. Si este término es nulo, desaparece el problema.

Criterio de periodicidad de la antiderivada

Sea $f : \mathbb{R} \to \mathbb{R}$ una función T-periódica continua a trozos, con serie de Fourier

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} \big(a_n \cos(n\omega t) + b_n \sin(n\omega t)\big),$$

y sea F una antiderivada de f. Las siguientes condiciones son equivalentes:

(1) $F(x)$ *es T-periódica, o sea,* $F(x+T) = F(x)$.

(2) $F(-T/2) = F(T/2)$.

(3) $a_0 = 0$, *o sea,* $\displaystyle\int_{-T/2}^{T/2} f(t)\,dt = 0$.

Para cualquier f periódica continua a trozos, la función $F(x) - \dfrac{a_0}{2}x$ es periódica.

Demostración. Demostraremos la equivalencia de las condiciones cerrando el círculo de implicaciones (1) $\implies$ (2) $\implies$ (3) $\implies$ (1):

Una antiderivada cualquiera de f tiene la forma

$$\boxed{F(x) = \int_{-T/2}^{x} f(t)\,dt + C}$$

para alguna constante C.

- (1) $\implies$ (2) : Si F es T-periódica, en particular $F(T/2) = F(-T/2+T) = F(-T/2)$.
- (2) $\implies$ (3) : Si $F(-T/2) = F(T/2)$, entonces por la Regla de Barrow

  $$\boxed{\int_{-T/2}^{T/2} f(t)\,dt = F(T/2) - F(-T/2) = 0}$$

  y como $a_0 = \dfrac{2}{T}\displaystyle\int_{-T/2}^{T/2} f(t)\,dt$, esto significa que $a_0 = 0$.
- (3) $\implies$ (1) : Supongamos que $\displaystyle\int_{-T/2}^{T/2} f(t)\,dt = 0$. Entonces

  $$\begin{aligned} F(x+T) = \int_{-T/2}^{x+T} f(t)\,dt + C &= \int_{-T/2}^{T/2} f(t)\,dt + \int_{T/2}^{x+T} f(t)\,dt + C \\ &= \int_{T/2}^{x+T} f(t)\,dt + C; \end{aligned}$$

  cambiando de variable con $t \to t+T$,

  $$= \int_{-T/2}^{x} f(t+T)\,d(t+T) + C;$$

como $f(t)$ es T-periódica,

$$= \int_{-T/2}^{x} f(t)\,dt + C = F(x)$$

por tanto $F(x)$ es T-periódica.

• Esto demuestra la equivalencia. Para f periódica en general, consideramos

$$G(x) = F(x) - \frac{a_0}{2}x, \quad \therefore \quad G'(x) = F'(x) - \frac{a_0}{2}x = f(x) - \frac{a_0}{2} \stackrel{\text{def}}{=} g(x).$$

La función $g(x)$ es claramente periódica y continua a trozos si f lo es, y tiene término constante en su serie de Fourier nulo ya que ésta es

$$\boxed{g(t) = f(t) - \frac{a_0}{2} = \sum_{n=1}^{\infty}\big(a_n\cos(n\omega t) + b_n\sin(n\omega t)\big).}$$

Por tanto, el criterio ya demostrado implica que su antiderivada G es periódica. □

• **Ejemplo:** La función $f(x) = \sin x$ es periódica e impar, por tanto su coeficiente constante a_0 es nulo y, efectivamente, su antiderivada $F(x) = -\cos x$ es periódica.

• **Ejemplo:** La función $f(x) = |\sin x|$ en $[-\pi,\pi]$ es periódica y par. Su coeficiente constante es

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} |\sin x|\,dx = \frac{4}{\pi}$$

y en consecuencia su antiderivada

$$F(x) = \int_{-\pi}^{x} |\sin t|\,dt$$

no es periódica. Por ejemplo, en el intervalo $[-\pi,\pi]$ se tiene

$$F(x) = \int_{-\pi}^{x} -\sin t\,dt = 1 + \cos x, \quad -\pi \le x \le 0$$

y

$$F(x) = \int_{-\pi}^{x} |\sin t|\,dt = \int_{-\pi}^{0} |\sin t|\,dt + \int_{0}^{x} \sin t\,dt$$

$$= 2 + 1 - \cos x = 3 - \cos x, \quad 0 \le x \le \pi.$$

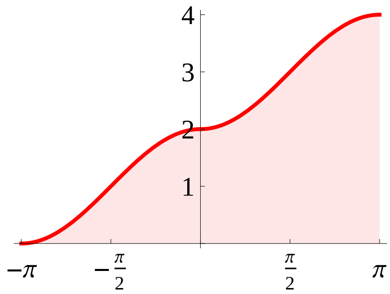

Figura 62 – Antiderivada de $|\sin x|$ en $[-\pi,\pi]$.

En este caso, la función

$$\boxed{F(x) - \frac{a_0}{2}x = \int_{-\pi}^{x} |\sin t|\, dt - \frac{2}{\pi}x}$$

es periódica.

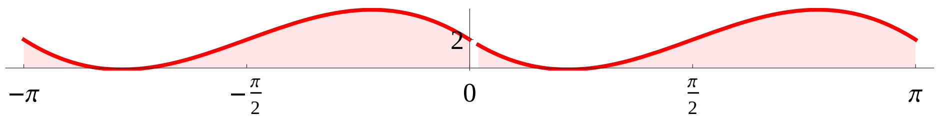

Figura 63 – La antiderivada de $|\sin x|$ quitando el término lineal $\frac{2}{\pi}x$.

• **Ejemplo:** Sea $I(x) = x$ y $I_{\text{per}}(x)$ su extensión 2π-periódica desde $[-\pi, \pi]$ á $\mathbb{R}$. Entonces $a_0 = 0$ pues I es impar. Por tanto cualquier antiderivada $F(x)$ de la extensión periódica I_{per} es periódica. Por ejemplo, la antiderivada que hemos usado en la demostración, que para este ejemplo es

$$\boxed{F(x) = \int_{-\pi}^{x} I_{\text{per}}(t)\, dt}$$

Es importante entender que se integra la extensión periódica I_{per}, no la función I.

Ahora, si restringimos x al intervalo $[-\pi, \pi]$ desde el cual se extiende, y sobre el cual por definición $I_{\text{per}} = I$, se obtiene una fórmula válida en ese intervalo:

$$F(x) = \int_{-\pi}^{x} t\, dt = \left[\frac{t^2}{2}\right]_{-\pi}^{x} = \frac{x^2 - \pi^2}{2}, \quad x \in [-\pi, \pi].$$

La antiderivada efectivamente satisface $F(-\pi) = F(\pi)$ (con valor común 0). Para evaluar $F(x)$ en cualquier $x \in \mathbb{R}$, dada su fórmula para $x \in [-\pi, \pi]$, debemos encontrar $n \in \mathbb{Z}$ tal que $x - 2\pi n \in [-\pi, \pi]$ y entonces $F(x) = F(x - 2\pi n)$.

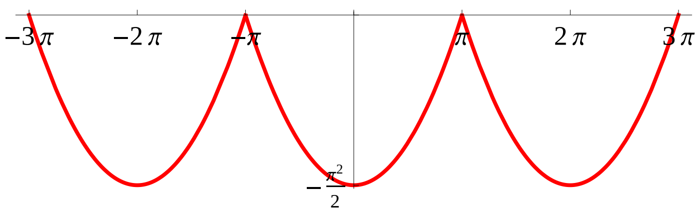

Figura 64 – La integral de la extensión 2π-periódica de la identidad.

En (9.11.6) estudiaremos la antiderivada de una extensión periódica en más detalle.

9.11.3. Cálculo de la serie de Fourier de una antiderivada periódica

En esta sección vamos a suponer que $f(t)$ es una función T-periódica continua a trozos con término constante $a_0 = 0$. Suponemos que f está representada por su serie de Fourier, que entonces será de la forma

$$\boxed{f(x) = \sum_{n=1}^{\infty} \big(a_n \cos(n\omega x) + b_n \sin(n\omega x)\big)}$$

En estas circunstancias, como hemos visto en la sección 9.11.2, cualquier antiderivada F de f será periódica, y por tanto tendrá una serie de Fourier, que queremos calcular.

En estas circunstancias, la manera más sencilla de proceder es considerando integrales indefinidas. Es decir, integrando término a término pero sin límites de integración.

La serie de Fourier de una antiderivada periódica

Sea $f : \mathbb{R} \to \mathbb{R}$ una función T-periódica continua a trozos y representada por una serie de Fourier sin término constante:

$$f(x) = \sum_{n=1}^{\infty} \big(a_n \cos(n\omega x) + b_n \sin(n\omega x)\big).$$

Entonces la serie de Fourier de una antiderivada F de f se obtiene haciendo la integral indefinida término a término:

$$F(x) = \sum_{n=1}^{\infty} \left(\frac{a_n}{n\omega} \sin(n\omega x) - \frac{b_n}{n\omega} \cos(n\omega x) \right) + C$$

donde C es una constante arbitraria.

Demostración. Utilizamos la antiderivada concreta

$$F(x) = \int_{-T/2}^{x} f(t)\, dt.$$

Por el Teorema de Integración, la serie de Fourier de f,

$$f(t) = \sum_{n=1}^{\infty} \big(a_n \cos(n\omega t) + b_n \sin(n\omega t)\big)$$

se puede integrar término a término entre $-T/2$ y x. Queda

$$\begin{aligned} F(x) &= \sum_{n=1}^{\infty} \left(a_n \int_{-T/2}^{x} \cos(n\omega t)\, dt + b_n \int_{-T/2}^{x} \sin(n\omega t)\, dt \right) \\ &= \sum_{n=1}^{\infty} \left(a_n \left[\frac{1}{n\omega} \sin(n\omega t) \right]_{-T/2}^{x} + b_n \left[-\frac{1}{n\omega} \cos(n\omega t) \right]_{-T/2}^{x} \right) \end{aligned}$$

y como $n\omega(-T/2) = -n\pi$,

$$= \sum_{n=1}^{\infty} \left(\frac{a_n}{n\omega} \sin(n\omega x) - \frac{b_n}{n\omega} \Big(\cos(n\omega x) - (-1)^n \Big) \right)$$

$$= \sum_{n=1}^{\infty} \left(\frac{a_n}{n\omega} \sin(n\omega x) - \frac{b_n}{n\omega} \cos(n\omega x) \right) + K$$

donde nos ha quedado para esta antiderivada concreta, la constante

$$\boxed{K = \sum_{n=1}^{\infty} \frac{(-1)^n b_n}{n\omega}}$$

pero como sabemos, al cambiar de antiderivada, cambiará el valor de la constante. □

Hay varias maneras de obtener la constante de integración de la antiderivada F:

La constante C está dada por la fórmula general para el coeficiente de Fourier:

$$C = \frac{A_0}{2}, \qquad A_0 = \frac{2}{T} \int_{-T/2}^{T/2} F(x)\, dx.$$

C también está determinada dando un valor concreto á x en el desarrollo de Fourier integrado, pero esto requiere saber sumar la serie en ese punto.

• **Ejemplo:** Partimos de la serie de Fourier de la extensión 2π-periódica de la función identidad $I(x) = x$ en $[-\pi, \pi]$, vista antes

$$\boxed{I_{\text{per}}(x) = 2 \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sin(nx), \qquad x \in \mathbb{R}}$$

Al ser impar la función $f = I$, tiene $a_n = 0$ para todo $n \geq 0$. Entonces, tomando la integral indefinida término a término, queda

$$F(x) \stackrel{\text{def}}{=} \int I_{\text{per}}(x)\, dx = 2 \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \cdot -\frac{1}{n} \cos(nx) + K,$$

$$= 2 \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} \cos(nx) + K.$$

Si restringimos x al intervalo $[-\pi, \pi]$ entonces $I_{\text{per}}(x) = I(x) = x$, y tomamos la antiderivada específica que hemos usado en la demostración:

$$\boxed{F(x) = \int_{-\pi}^{x} I_{\text{per}}(t)\, dt = \int_{-\pi}^{x} t\, dt = \frac{x^2 - \pi^2}{2}}$$

quedará

$$\frac{x^2 - \pi^2}{2} = 2 \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} \cos(nx) + K, \quad x \in [-\pi, \pi],$$

y se puede determinar la constante calculando $K = A_0/2$:

$$A_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{x^2-\pi^2}{2}\,dx = \frac{1}{\pi}\left[\frac{x^3}{6} - \frac{\pi^2 x}{2}\right]_{-\pi}^{\pi} = \frac{1}{\pi}\cdot 2\cdot\left[\frac{\pi^3}{6} - \frac{\pi^3}{2}\right] = \frac{2}{\pi}\cdot -\frac{\pi^3}{3} = -\frac{2\pi^2}{3}$$

con lo cual queda

$$\frac{x^2-\pi^2}{2} = -\frac{\pi^2}{3} + 2\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}\cos(nx), \quad x\in[-\pi,\pi],$$

que al despejar nos da

$$\boxed{x^2 = \frac{\pi^2}{3} + 4\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}\cos(nx), \quad x\in[-\pi,\pi]}$$

Nota 1: Para calcular K para la antiderivada específica F también podíamos haber usado la fórmula

$$K = \sum_{n=1}^{\infty}\frac{(-1)^n b_n}{n\omega} = \sum_{n=1}^{\infty}\frac{(-1)^n\, 2(-1)^{n+1}}{n\cdot n} = -2\sum_{n=1}^{\infty}\frac{1}{n^2} = -2\cdot\frac{\pi^2}{6} = -\frac{\pi^2}{3},$$

pero este método requiere que conozcamos el valor de la suma infinita, y el primero no.

Nota 2: podemos tomar cualquier antiderivada $F(x)$. En vez de la anterior, si usamos

$$\boxed{F(x) = \frac{x^2}{2}}$$

(que por otra parte es la más “lógica”) quedará ahora

$$\frac{x^2}{2} = 2\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}\cos(nx) + K, \quad x\in[-\pi,\pi],$$

y al igual que antes, determinamos la constante calculando

$$K = \frac{A_0}{2} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{x^2}{2}\,dx = \frac{1}{2\pi}\cdot 2\int_{0}^{\pi}\frac{x^2}{2}\,dx = \frac{1}{2\pi}\cdot\frac{\pi^3}{3} = \frac{\pi^2}{6}$$

con lo cual queda

$$\frac{x^2}{2} = \frac{\pi^2}{6} + 2\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}\cos(nx), \quad x\in[-\pi,\pi],$$

que tras multiplicar por 2 lleva al mismo desarrollo de antes.

Nota 3: Debe contrastarse el método empleado aquí para obtener la serie de Fourier de x^2 con este otro donde se empleaban directamente las fórmulas para los coeficientes.

9.11.4. Desarrollo de una antiderivada en general

En las secciones anteriores, hemos visto que el único obstáculo a poder obtener de manera rápida una serie de Fourier a partir de otra mediante integración indefinida, es que al integrar la serie de Fourier de una función periódica f, aparece un término con x debido al término constante $a_0/2$. En particular, si $a_0 = 0$, entonces no aparece tal término y no hay obstáculo. Así, obtuvimos el resultado principal para antiderivadas periódicas.

Ahora, vamos a ver cómo se puede proceder cuando $a_0 \neq 0$.

Para una función T-periódica $f : \mathbb{R} \to \mathbb{R}$ continua a trozos y representada por una serie de Fourier general

$$\boxed{f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \big(a_n \cos(n\omega t) + b_n \sin(n\omega t)\big), \quad t \in \mathbb{R}}$$

sabemos (sección 9.11.2) que si F es una antiderivada de f, entonces

$$G(x) = F(x) - \frac{a_0}{2} x$$

es periódica, y es la antiderivada de la función

$$g(x) = f(x) - \frac{a_0}{2} = \sum_{n=1}^{\infty} \big(a_n \cos(n\omega t) + b_n \sin(n\omega t)\big)$$

por tanto, el resultado de la sección anterior (9.11.3) es aplicable para calcular la serie de Fourier de G. Se tiene entonces

$$\boxed{F(x) - \frac{a_0}{2} x = \sum_{n=1}^{\infty} \left(\frac{a_n}{n\omega} \sin(n\omega x) - \frac{b_n}{n\omega} \cos(n\omega x) \right) + C, \quad x \in \mathbb{R}}$$

Sin embargo, como ya comentamos en su momento, el término con x no forma parte de una serie de Fourier. Por eso lo hemos puesto del lado de la antiderivada $F(x)$.

Qué más podemos decir en este caso? La antiderivada no es periódica, así que no puede estar representada por una serie de Fourier, que sí lo es.

Para poder avanzar, debemos tener en cuenta que lo habitual es construir funciones periódicas por extensión, y entender qué ocurre en ese caso.

9.11.5. Serie de Fourier de la extensión periódica de una antiderivada [15]

Vamos a suponer que tenemos una función f dada en el intervalo "fundamental" de periodicidad $[-T/2, T/2]$ (incluimos ambos extremos por comodidad). Restringiéndonos a este intervalo, podemos hablar de una antiderivada F de f en él.

[15] Esta sección se puede omitir en una lectura rápida

Después, podemos extender ambas funciones, tanto f como su antiderivada F, del intervalo fundamental a todo $\mathbb{R}$. O sea, primero tomamos una antiderivada en el intervalo fundamental $[-T/2, T/2]$, y luego la extendemos por periodicidad a todo $\mathbb{R}$.

Entonces, investigamos qué relación tienen

- la serie de Fourier de la extensión T-periódica f_{per} de una función f en $[-T/2, T/2]$;
- la serie de Fourier de la extensión T-periódica F_{per} de una antiderivada F de f en $[-T/2, T/2]$.

Consideramos la antiderivada particular F de f en $[-T/2, T/2]$ dada por

$$F(x) = \int_{-T/2}^{x} f(t)\,dt, \quad x \in [-T/2, T/2].$$

Para esta F se tiene

$$F(-T/2) = 0, \quad F(T/2) = \int_{-T/2}^{T/2} f(t)\,dt = \frac{T}{2}a_0$$

Con suponer solamente que f es continua a trozos, entonces por el Teorema Fundamental, F será continua, y como $F' = f$, de hecho es C^1 a trozos. En particular si f es C^1 a trozos, F será continua y C^2 a trozos.

Supongamos que la extensión periódica f_{per} está representada por su serie de Fourier

$$f_{\text{per}}(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\big(a_n \cos(n\omega t) + b_n \sin(n\omega t)\big), \quad t \in \mathbb{R}$$

Denotemos a los coeficientes de Fourier de la extensión periódica de F por letras mayúsculas:

$$F_{\text{per}}(x) = \frac{A_0}{2} + \sum_{n=1}^{\infty}\Big(A_n \cos(n\omega x) + B_n \sin(n\omega x)\Big), \quad x \in \mathbb{R}$$

• Vamos a calcular los coeficientes A_n, B_n integrando la serie de Fourier de f entre $-T/2$ y x término a término y reordenando términos, como ya hicimos cuando $a_0 = 0$ en la sección 9.11.3.

Integrando término a término, queda, para $x \in [-T/2, T/2]$,

$$\begin{aligned} F(x) &= \int_{-T/2}^{x} f(t)\,dt \\ &= \frac{a_0}{2}\left(x + \frac{T}{2}\right) + \sum_{n=1}^{\infty}\left(a_n \int_{-T/2}^{x} \cos(n\omega t)\,dt + b_n \int_{-T/2}^{x} \sin(n\omega t)\,dt\right) \end{aligned}$$

$$= \frac{a_0}{2}x + \frac{a_0 T}{4} + \sum_{n=1}^{\infty}\left(a_n\left[\frac{\sin(n\omega t)}{n\omega}\right]_{-T/2}^{x} + b_n\left[-\frac{\cos(n\omega t)}{n\omega}\right]_{-T/2}^{x}\right)$$

y como $n\omega(-T/2) = -n\pi$,

$$= \frac{a_0}{2}x + \frac{a_0 T}{4} + \sum_{n=1}^{\infty}\left(a_n \cdot \frac{\sin(n\omega x)}{n\omega} + b_n \cdot \frac{-\cos(n\omega x) + (-1)^n}{n\omega}\right)$$

$$= \frac{a_0}{2}x + \frac{a_0 T}{4} + \sum_{n=1}^{\infty}\left(\frac{a_n}{n\omega}\sin(n\omega x) - \frac{b_n}{n\omega}\cos(n\omega x)\right) + \sum_{n=1}^{\infty}\frac{(-1)^n b_n}{n\omega}$$

Nos encontramos con el problema del término con x, que no pertenece a una serie de Fourier. Para resolverlo, desarrollamos x en serie de Fourier, usando el resultado que obtuvimos antes

$$\boxed{x = 2\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n\omega}\sin(n\omega x), \qquad -\frac{T}{2} < x < \frac{T}{2}}$$

y sustituimos en lo anterior para obtener

$$F(x) = \frac{a_0 T}{4} + \sum_{n=1}^{\infty}\frac{(-1)^n b_n}{n\omega}$$

$$+ \frac{a_0}{2} \cdot 2\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n\omega}\sin(n\omega x) + \sum_{n=1}^{\infty}\left(\frac{a_n}{n\omega}\sin(n\omega x) - \frac{b_n}{n\omega}\cos(n\omega x)\right)$$

$$= \frac{a_0 T}{4} + \sum_{n=1}^{\infty}\frac{(-1)^n b_n}{n\omega} + \sum_{n=1}^{\infty}\left(\frac{(-1)^{n+1}a_0 + a_n}{n\omega}\sin(n\omega x) - \frac{b_n}{n\omega}\cos(n\omega x)\right).$$

Esta última sí es una serie de Fourier, y en ella identificamos entonces el término constante

$$\frac{A_0}{2} = \frac{a_0 T}{4} + \sum_{n=1}^{\infty}\frac{(-1)^n b_n}{n\omega}$$

por tanto

$$\boxed{A_0 = \frac{a_0 T}{2} + 2\sum_{n=1}^{\infty}\frac{(-1)^n b_n}{n\omega}}$$

y para los otros coeficientes obtenemos las expresiones:

$$\boxed{A_n = -\frac{b_n}{n\omega}, \qquad B_n = \frac{(-1)^{n+1}a_0 + a_n}{n\omega}, \qquad n \geq 1}$$

• Para contrastar métodos, ahora vamos a calcular los coeficientes A_n, B_n mediante las fórmulas generales.

• Cálculo de A_0: Integramos por partes usando $u = F(x)$, $dv = dx$, luego $du = f(x)\,dx$, $v = x$:

$$A_0 = \frac{2}{T}\int_{-T/2}^{T/2} F(x)\,dx = \frac{2}{T}\left(\Big[xF(x)\Big]_{-T/2}^{T/2} - \int_{-T/2}^{T/2} xf(x)\,dx\right)$$

$$= \frac{2}{T}\left(\frac{T}{2}F\left(\frac{T}{2}\right) - \int_{-T/2}^{T/2} xf(x)\,dx\right)$$

$$= \frac{T}{2}a_0 - \frac{2}{T}\int_{-T/2}^{T/2} xf(x)\,dx$$

Lo resaltamos:

$$\boxed{A_0 = \frac{T}{2}a_0 - \frac{2}{T}\int_{-T/2}^{T/2} xf(x)\,dx}$$

• Cálculo de A_n con $n \geq 1$: Integramos por partes, usando $u = F(x)$, $dv = \cos(n\omega x)$, luego $du = f(x)\,dx$, $v = \sin(n\omega x)/n\omega$:

$$A_n = \frac{2}{T}\int_{-T/2}^{T/2} F(x)\cos(n\omega x)\,dx$$

$$= \frac{2}{T}\left[\frac{1}{n\omega}F(x)\sin(n\omega x)\right]_{-T/2}^{T/2} - \frac{2}{T}\cdot\frac{1}{n\omega}\int_{-T/2}^{T/2} f(x)\sin(n\omega x)\,dx = -\frac{b_n}{n\omega}$$

pues $n\omega \cdot \pm\frac{T}{2} = \pm n \cdot \frac{2\pi}{T}\cdot\frac{T}{2} = \pm n\pi$. Lo resaltamos:

$$\boxed{A_n = -\frac{b_n}{n\omega}, \quad n \geq 1}$$

• Cálculo de B_n con $n \geq 1$: Integramos por partes, usando $u = F(x)$, $dv = \sin(n\omega x)$, luego $du = f(x)\,dx$, $v = -\cos(n\omega x)/n\omega$:

$$B_n = \frac{2}{T}\int_{-T/2}^{T/2} F(x)\sin(n\omega x)\,dx$$

$$= \frac{2}{T}\left[-\frac{1}{n\omega}F(x)\cos(n\omega x)\right]_{-T/2}^{T/2} + \frac{2}{T}\cdot\frac{1}{n\omega}\int_{-T/2}^{T/2} f(x)\cos(n\omega x)\,dx$$

$$= \frac{2}{T}\cdot -\frac{1}{n\omega}\cdot F\left(\frac{T}{2}\right)\cdot\cos(n\pi) + \frac{a_n}{n\omega}$$

(ya que $F(-T/2) = 0$)

$$= \frac{2}{T}\cdot -\frac{1}{n\omega}\cdot\frac{T}{2}\,a_0\cdot(-1)^n + \frac{a_n}{n\omega}$$

$$= \frac{(-1)^{n+1}a_0 + a_n}{n\omega}$$

Lo resaltamos:

$$\boxed{B_n = \frac{(-1)^{n+1}a_0 + a_n}{n\omega}, \quad n \geq 1}$$

• Mediante estos procedimientos hemos obtenido dos expresiones equivalentes para el término constante, que unidas a la fórmula habitual son tres:

$$\boxed{A_0 = \frac{2}{T}\int_{-T/2}^{T/2} F(x)\,dx = \frac{a_0 T}{2} - \frac{2}{T}\int_{-T/2}^{T/2} x f(x)\,dx = \frac{a_0 T}{2} + 2\sum_{n=1}^{\infty}\frac{(-1)^n b_n}{n\omega}}$$

Al igualar las dos últimas y cancelar el término común, queda la igualdad

$$-\frac{2}{T}\int_{-T/2}^{T/2} x f(x)\,dx = 2\sum_{n=1}^{\infty}\frac{(-1)^n b_n}{n\omega} \iff \frac{1}{T}\int_{-T/2}^{T/2} x f(x)\,dx = \sum_{n=1}^{\infty}\frac{(-1)^{n+1} b_n}{n\omega}$$

que también se podría verificar independientemente, integrando término a término la serie de Fourier de $f(x)$ multiplicada por x.

• Nota: Hemos estado trabajando con la antiderivada concreta $F(x) = \int_{-T/2}^{x} f(t)\,dt$. Como cambiar de antiderivada significa cambiar de constante, las fórmulas para el término constante A_0 que dependían de usar esta F y no otra dejan de ser válidas para cualquier F en general. La que siempre es válida es la fórmula general para el coeficiente de Fourier.

Resumimos los resultados obtenidos.

Coeficientes de Fourier de la extensión periódica de una antiderivada

Sea f una función en $[-T/2, T/2]$, con antiderivada F en este intervalo. Si f es continua a trozos y su extensión T-periódica está representada por la serie de Fourier

$$f_{\text{per}}(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\big(a_n\cos(n\omega x) + b_n\sin(n\omega x)\big)$$

entonces F es continua, de clase C^1 a trozos y su extensión T-periódica está representada por la serie de Fourier

$$F_{\text{per}}(x) = \frac{A_0}{2} + \sum_{n=1}^{\infty}\big(A_n\cos(n\omega x) + B_n\sin(n\omega x)\big),$$

donde los coeficientes A_n, B_n están dados por las fórmulas

$$A_n = -\frac{b_n}{n\omega}, \qquad B_n = \frac{(-1)^{n+1}a_0 + a_n}{n\omega}, \quad n \ge 1$$

y para $n = 0$, si la antiderivada es $F(x) = \displaystyle\int_{-T/2}^{x} f(t)\,dt$,

$$A_0 = \frac{2}{T}\int_{-T/2}^{T/2} F(x)\,dx = \frac{a_0 T}{2} - \frac{2}{T}\int_{-T/2}^{T/2} x f(x)\,dx = \frac{a_0 T}{2} + 2\sum_{n=1}^{\infty}\frac{(-1)^n b_n}{n\omega}$$

pero si es otra antiderivada, solo la primera expresión sigue siendo válida en general.

Como caso particular ($a_0 = 0$) se obtienen de nuevo los resultados de la sección 9.11.3.

9.11.6. Relación entre la extensión periódica de una antiderivada y una antiderivada de la extensión periódica [16]

A diferencia de lo que ocurre con la derivación y la extensión, la extensión periódica de la antiderivada no es necesariamente una antiderivada de la extensión periódica. Es decir, no tiene por qué ser

$$F'_{\text{per}} = f_{\text{per}}$$

Para resumir los resultados y resaltar las operaciones relevantes, vamos a usar la siguiente notación:

- $\int(*)$: operación de tomar una antiderivada;
- $(*)_{\text{per}}$: operación de extensión T-periódica;
- $F = \int f$: una antiderivada de f en $[c, c+T)$
- $F_{\text{per}} = \left(\int f\right)_{\text{per}}$: extensión T-periódica de una antiderivada
- $\Phi = \int f_{\text{per}}$: una antiderivada de la extensión periódica

Como $\Phi' = f_{\text{per}}$ sobre todo $\mathbb{R}$ (donde sea derivable, claro), en particular, al restringir á $[c, c+T)$ se tiene

$$\Phi'(x) = f_{\text{per}}(x) = f(x), \quad x \in [c, c+T),$$

es decir, Φ también es una antiderivada de f en $[c, c+T)$, y por tanto,

$$\boxed{K \overset{\text{def}}{=} \Phi - F}$$

es constante sobre $[c, c+T)$. Incluso podríamos hacer K nula eligiendo la antiderivada Φ de f_{per} para que coincida con la antiderivada dada F de f sobre $[c, c+T)$.

Por la Regla de Barrow y la periodicidad

$$\Phi(x+T) - \Phi(x) = \int_x^{x+T} f_{\text{per}}(t)\,dt = \int_c^{c+T} f_{\text{per}}(t)\,dt = \int_c^{c+T} f(t)\,dt = \frac{T}{2}a_0$$

(recordamos que se puede calcular sobre cualquier intervalo de longitud T) y entonces, cambiando de variable

$$\begin{aligned} \Phi(x+T) - \Phi(x) &= \frac{T}{2}a_0, \\ \Phi(x+2T) - \Phi(x+T) &= \frac{T}{2}a_0, \end{aligned}$$

[16] Esta es una sección suplementaria.

$$\vdots$$

$$\Phi(x+nT)-\Phi(x+(n-1)T)=\frac{T}{2}a_0,$$

y sumando, queda

$$\boxed{\Phi(x+nT)=\Phi(x)+n\cdot\frac{T}{2}a_0,\quad n\in\mathbb{Z}}$$

Dado $x\in\mathbb{R}$ existe $n\in\mathbb{Z}$ tal que $x-nT\in[c,c+T)$. Entonces

$$\Phi(x)=\Phi(x-nT)+n\cdot\frac{T}{2}a_0=F(x-nT)+n\cdot\frac{T}{2}a_0+K.$$

Por otra parte, la extensión periódica F_{per} de la antiderivada F de f sobre $[c,c+T)$ es

$$F_{\text{per}}(x)=F(x-nT)$$

con lo cual, en general,

$$\Phi(x)=F_{\text{per}}(x)+n\cdot\frac{T}{2}a_0+K$$

Recordamos que el entero n tal que $x=nT\in[-T/2,T/2)$ está determinado por

$$c\le x-nT<c+T\iff nT\le x-c<(n+1)T\iff n=\left\lfloor\frac{x-c}{T}\right\rfloor$$

Entonces la relación general es

$$\boxed{\left(\int f_{\text{per}}\right)(x)=\left(\int f\right)_{\text{per}}(x)+\left\lfloor\frac{x-c}{T}\right\rfloor\cdot\frac{T}{2}a_0+K}$$

Cuando $a_0=0$, conmutan las dos operaciones, salvo constante:

$$\boxed{\left(\int f_{\text{per}}\right)(x)=\left(\int f\right)_{\text{per}}(x)+K}$$

En general, sólo coincidirán (salvo constante) sobre el intervalo $[c,c+T)$.

9.12. La serie de Fourier Compleja

9.12.1. Funciones de variable real con valores complejos

Debido a la fórmula

$$\boxed{e^{ix}=\cos(x)+i\sin(x)\qquad\textbf{(fórmula de Euler)}}$$

al utilizar números complejos, la Trigonometría (real) se convierte en manipulación de la función exponencial compleja, con sólo aplicar las fórmulas que van en sentido contrario,

$$\cos(x) = \operatorname{Re}\big(e^{ix}\big) = \frac{e^{ix} + e^{-ix}}{2}, \qquad \sin(x) = \operatorname{Im}\big(e^{ix}\big) = \frac{e^{ix} - e^{-ix}}{2i}, \quad x \in \mathbb{R}$$

A la luz de estas fórmulas fundamentales, se pueden reconsiderar los resultados principales de las secciones anteriores con el objetivo no solo de simplificar algunas fórmulas, sino de entenderlas desde otra perspectiva.

Consideramos funciones f de una variable real t pero con **valores** $f(t)$ que pueden ser números complejos, es decir, $f : [a, b] \subseteq \mathbb{R} \to \mathbb{C}$. Una función $f : \mathbb{R} \to \mathbb{C}$ se descompone en su parte real y su parte imaginaria, al escribir sus valores $f(t)$ en forma compleja binómica:

$$f(t) = u(t) + iv(t), \quad u, v : [a, b] \subseteq \mathbb{R} \to \mathbb{T} \qquad u(t) = \operatorname{Re} f(t), \quad v(t) = \operatorname{Im} f(t)$$

La definición de función periódica es la misma que para funciones reales: $f : \mathbb{R} \to \mathbb{C}$ tiene periodo T si $f(t + T) = f(t) \quad \forall t \in \mathbb{R}$. Es inmediato comprobar que f es periódica de periodo T si y sólo si sus partes real e imaginaria lo son.

El hecho fundamental que une la trigonometría con la exponencial compleja es que

la exponencial compleja tiene periodo complejo $2\pi i$:

Es decir,

$$\exp(z + 2\pi i) = \exp(z)$$

y de hecho, en general, esta es la única manera que dos exponenciales pueden ser iguales:

$$\exp(z) = \exp(w) \iff \exists n \in \mathbb{Z} : z = w + 2\pi i n$$

con lo cual, la exponencial cambiada de variable por $z \to iz$, es decir, la función

$$\exp(iz) = e^{iz} = \cos(z) + i \sin(z),$$

tiene periodo 2π, unificando en una sola expresión a las dos funciones trigonométricas básicas seno y coseno.

9.12.2. Relaciones de ortogonalidad para exponenciales complejas

El análogo complejo de las relaciones de ortogonalidad es el siguiente teorema.

Relaciones de ortogonalidad complejas

Sea $c \in \mathbb{R}$ un punto fijado y $k \in \mathbb{Z}$. Entonces

$$\frac{1}{2\pi}\int_c^{c+2\pi} e^{ikx}\,dx = \begin{cases} 0 & k \neq 0 \\ 1 & k = 0 \end{cases}$$

y por tanto, si $m, n \in \mathbb{Z}$, entonces

$$\frac{1}{2\pi}\int_c^{c+2\pi} e^{inx}e^{-imx}\,dx = \begin{cases} 0 & n \neq m \\ 1 & n = m \end{cases}$$

Demostración. Para $k = 0$ la integral es

$$\frac{1}{2\pi}\int_c^{c+2\pi} e^0\,dx = \frac{1}{2\pi}\int_c^{c+2\pi} dx = \frac{1}{2\pi}(c + 2\pi - c) = \frac{1}{2\pi}\cdot 2\pi = 1,$$

y para $k \neq 0$, se tiene

$$\frac{1}{2\pi}\int_c^{c+2\pi} e^{ikx} = \left[\frac{e^{ikx}}{2\pi ik}\right]_{x=c}^{x=c+2\pi} = \frac{e^{ik(c+2\pi)} - e^{ikc}}{2\pi ik} = \frac{e^{ikc+2\pi ik} - e^{ikc}}{2\pi ik} = 0$$

ya que exp tiene periodo $2\pi i$. El segundo resultado es consecuencia del primero poniendo $k = n - m$, ya que $e^{inx}e^{-imx} = e^{i(n-m)x}$. □

La terminología geométrica de ortogonalidad viene motivada por la siguiente construcción. En general, para dos funciones $f, g : [a, b] \subseteq \mathbb{R} \to \mathbb{C}$ se define su producto interior por

$$\langle f, g\rangle \overset{\text{def}}{=} \int_a^b f(x)\overline{g(x)}\,dx \qquad \textbf{(producto interior de funciones)}$$

donde la barra denota la conjugación compleja: dada en forma binomial por $z = a + ib \to \overline{z} = a - ib$. También es habitual dividir la integral por la longitud del intervalo, tomando el valor medio:

$$\langle f, g\rangle \overset{\text{def}}{=} \frac{1}{b-a}\int_a^b f(x)\overline{g(x)}\,dx \qquad \textbf{(producto interior normalizado)}$$

Se demuestra entonces que este producto interior funciona de manera análoga al producto interior que se usa habitualmente en Geometría, para vectores:

$$\langle \vec{x}, \vec{y}\rangle = x_1y_1 + x_2y_2 + \cdots + x_ny_n, \qquad \begin{array}{l} \vec{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \\ \vec{y} = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n \end{array}$$

y más en general, con vectores de números complejos,

$$\langle \vec{x}, \vec{y}\rangle = x_1\overline{y_1} + x_2\overline{y_2} + \cdots + x_n\overline{y_n}, \qquad \begin{array}{l} \vec{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{C}^n \\ \vec{y} = (y_1, y_2, \ldots, y_n) \in \mathbb{C}^n \end{array}$$

En particular, se define el concepto de la "longitud" o norma de una función mediante

$$\|f\| = \sqrt{\langle f, f\rangle} = \left(\int_a^b f(x)\overline{f(x)}\,dx\right)^{1/2} = \left(\int_a^b |f(x)|^2\,dx\right)^{1/2} \qquad \textbf{(norma)}$$

donde $|z| = \sqrt{a^2 + b^2}$ para $z = a + ib$ es el módulo complejo. Al igual que para vectores, se cumple la

Desigualdad de Cauchy-Schwarz: $|\langle f, g\rangle| \leq \|f\|\,\|g\|$

con igualdad si y sólo si existe $\lambda \in \mathbb{R}$ con $g = \lambda f$. Esta desigualdad sirve para medir ángulos entre vectores – en este caso, funciones – declarando que si θ es el ángulo entre f, g entonces

$$\cos\theta = \frac{\langle f, g\rangle}{\|f\|\,\|g\|}$$

En particular, dos funciones f, g son ortogonales cuando

$$\langle f, g\rangle = 0, \quad \text{escrito como} \quad f \perp g \qquad \textbf{(ortogonalidad)}$$

Entonces, si definimos

$$e_n(x) \stackrel{\text{def}}{=} e^{inx} \qquad \text{entonces} \qquad \overline{e^{imx}} = e^{-imx}$$

respecto al producto interior normalizado dado por

$$\langle f, g\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx$$

las llamadas relaciones de ortogonalidad son precisamente eso:

$$e_n \perp e_m \quad (n \neq m), \qquad \langle e_n, e_m\rangle = \begin{cases} 0 & n \neq m \\ 1 & n = m \end{cases}$$

que, en la terminología del Álgebra Lineal, se expresa diciendo que

La colección de funciones $e_n(x) = e^{inx}$ es **ortonormal**

9.12.3. La serie y los coeficientes de Fourier complejos

Serie y coeficientes de Fourier complejos

Sea $f : \mathbb{R} \to \mathbb{C}$ una función periódica de periodo 2π. La ***serie de Fourier (compleja)*** *de f es la serie infinita doble*

$$\sum_{n=-\infty}^{\infty} c_n e^{inx}$$

donde

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\,dx, \quad n \in \mathbb{Z}$$

Observar que los signos de la exponencial en la serie de Fourier y en la fórmula para el coeficiente son opuestos.

Al igual que en el caso real, un sencillo cambio de variable sirve para desarrollar las funciones con un periodo T cualquiera en serie de Fourier. Si f es T-periódica, entonces

$$\varphi(x) = f\left(\frac{T}{2\pi}x\right) = f\left(\frac{x}{\omega}\right), \quad \omega = \frac{2\pi}{T}$$

es 2π-periódica, y por tanto, si $\varphi(x)$ tiene serie de Fourier compleja

$$\varphi(x) = f\left(\frac{x}{\omega}\right) = \sum_{n=-\infty}^{\infty} c_n e^{inx}$$

entonces

$$\boxed{f(x) = \sum_{n=-\infty}^{\infty} c_n e^{in\omega x}}$$

es su serie de Fourier T-periódica, y la fórmula para los coeficientes también se obtiene por cambio de variable:

$$\begin{aligned} c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} \varphi(x)e^{-inx}\,dx &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f\left(\frac{x}{\omega}\right)e^{-inx}\,dx \\ &= \frac{1}{2\pi}\int_{-\pi/\omega}^{\pi/\omega} f(x)e^{-in\omega x}\,d(\omega x) \\ &= \frac{\omega}{2\pi}\int_{-\pi/\omega}^{\pi/\omega} f(x)e^{-in\omega x}\,dx \\ &= \frac{1}{T}\int_{-T/2}^{T/2} f(x)e^{-in\omega x}\,dx \end{aligned}$$

Resaltamos estos resultados

Serie y coeficientes de Fourier complejos con periodo arbitario

Sea $f : \mathbb{R} \to \mathbb{C}$ una función T-periódica. La serie de Fourier compleja es

$$\sum_{n=-\infty}^{\infty} c_n e^{in\omega x} \quad \text{donde} \quad \omega = \frac{2\pi}{T} \quad \text{con} \quad c_n = \frac{1}{T}\int_{-T/2}^{T/2} f(x)e^{-in\omega x}\,dx$$

9.12.4. Correspondencia entre la serie de Fourier compleja y la trigonométrica

•: la serie trigonométrica a partir de la compleja por agrupación: si agrupamos los coeficientes por pares $(n, -n)$ con $n = 1, 2, 3, \ldots$ y usamos la fórmula de Euler,

$$e^{in\omega x} = \cos(n\omega x) + i\sin(n\omega x),$$

conseguimos que la serie de Fourier compleja, que es una serie infinita doble (el sumatorio va de $-\infty$ á ∞), se convierta en una serie simple (de 0 á ∞):

$$\begin{aligned}
\sum_{n=-\infty}^{\infty} c_n e^{in\omega x} &= c_0 + \sum_{n=1}^{\infty} \left(c_n e^{in\omega x} + c_{-n} e^{-in\omega x}\right) \\
&= c_0 + \sum_{n=1}^{\infty} (c_n \cos(n\omega x) + c_n i \sin(n\omega x) + c_{-n} \cos(n\omega x) - c_{-n} i \sin(n\omega x)) \\
&= c_0 + \sum_{n=1}^{\infty} \Big(\big(c_n + c_{-n}\big) \cos(n\omega x) + i \big(c_n - c_{-n}\big) \sin(n\omega x) \Big) \\
&= c_0 + \sum_{n=1}^{\infty} \Big(a_n \cos(n\omega x) + b_n \sin(n\omega x) \Big)
\end{aligned}$$

siendo

$$\boxed{a_n = c_n + c_{-n} \qquad b_n = i(c_n - c_{-n}) \qquad (n \geq 1)}$$

Vamos a ver que estos a_n, b_n son los coeficientes de Fourier reales vistos cuando expresamos una onda en términos de los armónicos con seno y coseno.

Primero, para $n \geq 1$,

$$\begin{aligned}
a_n = c_n + c_{-n} &= \frac{1}{T} \int_{-T/2}^{T/2} f(x) \big(e^{-in\omega x} + e^{in\omega x}\big)\, dx \\
&= \frac{1}{2(T/2)} \int_{-T/2}^{T/2} f(x) \cdot 2 \cos(n\omega x)\, dx \\
&= \frac{1}{T/2} \int_{-T/2}^{T/2} f(x) \cos(n\omega x)\, dx
\end{aligned}$$

y similarmente

$$\begin{aligned}
b_n = i(c_n - c_{-n}) &= \frac{i}{T} \int_{-T/2}^{T/2} f(x) \big(e^{-in\omega x} - e^{in\omega x}\big)\, dx \\
&= \frac{i}{2(T/2)} \int_{-T/2}^{T/2} f(x) \cdot 2i \cdot (- \sin(n\omega x))\, dx \\
&= \frac{1}{T/2} \int_{-T/2}^{T/2} f(x) \sin(n\omega x)\, dx.
\end{aligned}$$

En cuanto al término constante, recordando que se escribe como $a_0/2$:

$$\frac{a_0}{2} = c_0 = \frac{1}{T} \int_{-T/2}^{T/2} f(x)\, dx \quad \therefore \quad a_0 = \frac{1}{T/2} \int_{-T/2}^{T/2} f(x)\, dx$$

y vemos entonces otra razón para escribirlo de esta manera: para seguir siendo válida la relación $a_n = c_n + c_{-n}$ cuando $n = 0$.

• A partir de las relaciones obtenidas

$$a_n = c_n + c_{-n} \qquad b_n = i(c_n - c_{-n}) \qquad (n \geq 1)$$

se puede también expresar c_n en términos de a_n y b_n:

$$a_n - ib_n = c_n + c_{-n} + c_n - c_{-n} = 2c_n$$
$$a_n + ib_n = c_n + c_{-n} - c_n + c_{-n} = 2c_{-n}$$

por tanto

$$\boxed{c_n = \frac{a_n - ib_n}{2}, \quad c_{-n} = \frac{a_n + ib_n}{2}, \quad (n \geq 1)}$$

Resumimos los resultados.

Relaciones entre los coeficientes de Fourier reales y complejos

Los coeficientes de Fourier reales a_n, b_n y los complejos c_n están relacionados por

$$a_n = c_n + c_{-n} \quad (n \geq 0) \qquad b_n = i(c_n - c_{-n}) \quad (n \geq 1)$$

y recíprocamente

$$c_0 = \frac{a_0}{2}, \quad c_n = \frac{a_n - ib_n}{2}, \quad c_{-n} = \frac{a_n + ib_n}{2}, \quad (n \geq 1),$$

•: la serie trigonométrica a partir de la compleja desarrollando los coeficientes: otra manera de encontrar las relaciones entre los coeficientes es usando la fórmula de Euler

$$e^{-in\omega x} = \cos(n\omega x) - i\sin(n\omega x)$$

y aplicándola a la del coeficiente; queda

$$\begin{aligned} c_n = \frac{1}{T}\int_{-T/2}^{T/2} f(x)e^{-in\omega x}\,dx &= \frac{1}{T}\int_{-T/2}^{T/2} f(x)(\cos(n\omega x) - i\sin(n\omega x))\,dx \\ &= \frac{1}{T}\int_{-T/2}^{T/2} f(x)\cos(n\omega x)\,dx - i\,\frac{1}{T}\int_{-T/2}^{T/2} f(x)\sin(n\omega x)\,dx \\ &= \frac{a_n - ib_n}{2} \end{aligned}$$

donde

$$\boxed{a_n = \frac{2}{T}\int_{-T/2}^{T/2} f(x)\cos(n\omega x)\,dx, \qquad b_n = \frac{2}{T}\int_{-T/2}^{T/2} f(x)\sin(n\omega x)\,dx}$$

son los coeficientes de la serie de Fourier trigonómetrica (atención al factor π, no 2π).

Es importante observar varias cosas:

- Las expresiones a_n, b_n están definidas para todo $n \in \mathbb{Z}$, no solo para $n \geq 0$. Esto no lo necesitábamos antes para estudiar la serie trigonométrica, pero ahora refleja el hecho de que la serie de Fourier compleja es una serie infinita doble.
- Teniendo en cuenta la paridad de seno y coseno,

$$a_{-n} = a_n, \qquad b_{-n} = -b_n, \qquad c_{-n} = \frac{a_{-n} - ib_{-n}}{2} = \frac{a_n + ib_n}{2}$$

  por tanto

$$\boxed{c_n = \frac{a_n - ib_n}{2} \qquad c_{-n} = \frac{a_n + ib_n}{2}}$$

  y entonces, sumando y restando queda

$$c_n + c_{-n} = a_n, \qquad c_n - c_{-n} = -ib_n$$

  o equivalentemente

$$\boxed{a_n = c_n + c_{-n}, \qquad b_n = i(c_n - c_{-n})}$$

9.12.5. Relaciones entre los coeficientes de Fourier para funciones reales

Es importante entender varias cosas acerca de las fórmulas que relacionan los coeficientes de Fourier de la serie compleja con exponenciales y la serie trigonométrica con senos y cosenos:

- Estas fórmulas son válidas para funciones $f : \mathbb{R} \to \mathbb{C}$, es decir, con valores complejos.
- En particular, son válidas para funciones $f : \mathbb{R} \to \mathbb{R}$, es decir, con valores reales.
- En general los coeficientes a_n, b_n y c_n son todos números complejos.
- En el caso de $f : \mathbb{R} \to \mathbb{R}$ (con valores reales), a_n y b_n son reales, pero c_n sigue siendo en general complejo.

Entonces, las expresiones

$$c_n = \frac{a_n - ib_n}{2}, \qquad c_{-n} = \frac{a_n + ib_n}{2} \qquad (n \geq 1)$$

son las formas binómicas de los coeficientes cuando f es real pero no en general. Si f es real, entonces estas fórmulas implican que los coeficientes c_n y c_{-n} son conjugados:

$$\boxed{c_{-n} = \overline{c_n} \qquad \text{si la función es real}}$$

Entonces, en este caso, el coeficiente en forma polar

$$c_n = r_n e^{i\theta_n} \quad n \in \mathbb{Z}$$

viene dado por

$$r_n = r_{-n} = |c_n| = |c_{-n}| = \frac{1}{2}\sqrt{a_n^2 + b_n^2} \qquad (n \geq 1)$$

y

$$\theta_n = \arctan\left(-\frac{b_n}{a_n}\right), \qquad \theta_{-n} = \arctan\left(\frac{b_n}{a_n}\right) = -\theta_n \qquad (n \geq 1)$$

Resumimos estas observaciones.

Forma polar del coeficiente complejo de Fourier para funciones reales

Si $f : \mathbb{R} \to \mathbb{R}$ es una función real periódica de periodo 2π, con coeficentes de Fourier complejos c_n y reales a_n, b_n, entonces c_n y c_{-n} son conjugados:

$$c_{-n} = \overline{c_n}$$

y por tanto sus formas polares $c_n = r_n e^{i\theta_n}$ satisfacen

$$r_n = r_{-n} = \frac{1}{2}\sqrt{a_n^2 + b_n^2}, \qquad \theta_n = \arctan\left(-\frac{b_n}{a_n}\right), \quad \theta_{-n} = -\theta_n \qquad (n \geq 1)$$

Finalmente, observamos que en el caso de una función real, al agrupar por pares $(n, -n)$ los términos en la serie de Fourier compleja obtenemos, en forma polar

$$\begin{aligned} c_n e^{in\omega x} + c_{-n} e^{-in\omega x} &= r_n e^{i\theta_n} e^{in\omega x} + r_n e^{-i\theta_n} e^{-in\omega x} \\ &= r_n \left(e^{i(n\omega x + \theta_n)} + e^{-i(n\omega x + \theta_n)}\right) \\ &= 2r_n \cos(n\omega x + \theta_n) \end{aligned}$$

es decir, hemos demostrado el siguiente resultado.

La serie de Fourier de una función real en forma polar

Si $f : \mathbb{R} \to \mathbb{R}$ es una función real periódica de periodo 2π, su serie de Fourier tiene la forma

$$c_0 + 2\sum_{n=1}^{\infty} r_n \cos(n\omega x + \theta_n), \qquad r_n = \frac{1}{2}\sqrt{a_n^2 + b_n^2}, \quad \theta_n = \arctan\left(-\frac{b_n}{a_n}\right).$$

Observemos que r_n es la **amplitud** y θ_n es la **fase** (respecto a la onda coseno).

El espectro en amplitud y fase

*El **espectro en amplitud** es la función $n \to r_n$ (amplitud en función de la frecuencia) y el **espectro en fase** es la función $n \to \theta_n$ (fase en función de la frecuencia).*

9.12.6. Convergencia de la serie de Fourier compleja

Una vez vistas estas relaciones, es inmediato que, en cuanto a la convergencia, el Teorema de Dirichlet sigue siendo válido para la convergencia de la serie de Fourier compleja:

Teorema de Dirichlet Complejo

Sea $f : \mathbb{R} \to \mathbb{C}$ de clase C^1 a trozos (esto quiere decir que lo son su parte real y su parte imaginaria). Entonces

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}$$

donde la serie converge a $f(x)$ en un punto de continuidad y al valor medio

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

en un punto de salto.

• **Ejemplo:** Vamos a encontrar el desarrollo de Fourier de la extensión 2π-periódica de la identidad $I(x) = x$ en $[-\pi, \pi]$ utilizando la fórmula compleja. Primero,

$$c_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} x\, dx = 0$$

Para $n \geq 1$, integramos por partes con

$$u = x \qquad du = dx$$

$$dv = e^{-inx} \qquad v = \frac{-e^{-inx}}{in}$$

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} x e^{-inx}\, dx = \frac{1}{2\pi} \left(\left[-\frac{x e^{-inx}}{in} \right]_{-\pi}^{\pi} + \frac{1}{in} \int_{-\pi}^{\pi} e^{-inx}\, dx \right)$$

y, como la segunda integral es nula por las relaciones de ortogonalidad complejas,

$$= \frac{1}{2\pi} \left(-\frac{\pi e^{-in\pi} + \pi e^{in\pi}}{in} + 0 \right)$$

$$= \frac{1}{2\pi} \left(-\frac{\pi}{in} (e^{-in\pi} + e^{in\pi}) \right)$$

$$= \frac{1}{2\pi} \left(-\frac{2\pi}{in} \cos(n\pi) \right)$$

$$= \frac{1}{2\pi} \cdot -\frac{2\pi}{in} (-1)^n = \frac{(-1)^{n+1}}{in}$$

Por tanto

$$\boxed{I_{\text{per}}(x) = \sum_{n=-\infty}^{\infty} \frac{(-1)^{n+1}}{in} e^{inx}}$$

Para deducir la serie real, separamos el término constante y agrupamos la compleja por pares $(-n, n)$:

$$\sum_{n=-\infty}^{\infty} \frac{(-1)^{n+1}}{in} e^{inx} = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{in} e^{inx} + \frac{(-1)^{-n+1}}{-in} e^{-inx}$$

y como $(-1)^n = (-1)^{-n}$,

$$= \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{in} e^{inx} + \frac{(-1)^{n+1}}{-in} e^{-inx}$$

$$= \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{in} \left(e^{inx} - e^{-inx}\right)$$

$$= \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{in} \cdot 2\, i \sin(nx)$$

$$= 2 \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sin(nx)$$

que, como puede comprobarse, coincide con la hallada anteriormente.