Cálculo
para matemáticos

Tomo I

Fundamentos del cálculo de una variable

A. Peña

TEORÍA, EJERCICIOS Y PERSPECTIVA

Matemática²biertA

MATEMATICAABIERTA.CL

Cálculo para matemáticos. Tomo I

Fundamentos del cálculo de una variable

A. Peña · Matemática Abierta

Veinte capítulos y ochocientos ejercicios con soluciones desarrolladas.

Esta edición reproduce el manuscrito canónico corregido el 16 de octubre de 2026.

Licencia: GNU Free Documentation License, versión 1.3 o posterior (GFDL-1.3-or-later), conforme a la edición publicada en Matemática Abierta.

`https://www.gnu.org/licenses/fdl-1.3.html`

Índice general

Capítulo 1. Los números reales: axiomas de cuerpo, orden y completitud 1
1.1. Axiomas de cuerpo y estructura de orden 1
1.2. Valor absoluto, distancia y desigualdades 34
1.3. Cotas, máximos, mínimos, supremos e ínfimos 42
1.4. Por qué los racionales no bastan: el ejemplo de Rudin 52
1.5. Completitud: la propiedad que falta en $\mathbb{Q}$ 66
1.6. Completitud en acción: existencia de raíces 81
1.7. La propiedad arquimediana 87
1.8. Entre dos reales siempre hay más números 93
1.9. Intervalos encajados y bisección 97
1.10. Laboratorio de completitud 103
1.11. Ejercicios y soluciones 109

Capítulo 2. Funciones reales: estructura, composición, inversas y gráficas 132
Convenciones conjuntistas indispensables 132
Convenciones conjuntistas indispensables 132
2.1. Una fórmula no es todavía una función 133
2.2. Imágenes y preimágenes 139
2.3. Inyectividad, sobreyectividad y biyectividad 146
2.4. Composición e identidad 152
2.5. Funciones inversas: cuándo se puede deshacer una función 157
2.6. El álgebra de imágenes y preimágenes 163
2.7. Gráficas: una representación, no la definición completa 169
2.8. Funciones por tramos y transformaciones de gráficas 175
2.9. Monotonía e inversas sobre intervalos 182
2.10. Laboratorio de estructura funcional 188
2.11. Ejercicios y soluciones 197

Capítulo 3. Sucesiones y la primera noción rigurosa de límite 218
3.1. Una sucesión es una función, pero el límite vive en la cola 218
3.2. Tan cerca como queramos, desde algún momento: la definición ε–N 225
3.3. Qué fuerza tiene converger: unicidad, colas y acotación 230
3.4. El álgebra de límites: cuándo podemos calcular sin volver a ε 238
3.5. Orden y teorema del sándwich 246
3.6. Un pequeño repertorio de límites fundamentales 252
3.7. Convergencia de sucesiones monótonas y acotadas 259
3.8. Sucesiones recursivas: fabricar y controlar un límite 266
3.9. Cuando no hay límite real: límites infinitos y divergencia 272
3.10. Laboratorio de límites secuenciales 277
3.11. Ejercicios y soluciones 285

Capítulo 4. Límites de funciones 306
4.1. Acercarse a un punto: entornos, acumulación y puntos aislados 306
4.2. De la cola al entorno perforado: la definición ε–δ 313
4.3. Qué fuerza tiene un límite: unicidad y estabilidad local 321
4.4. El criterio secuencial: dos lenguajes para el mismo acercamiento 328
4.5. Álgebra de límites funcionales y control local 334
4.6. Orden, sándwich y oscilación local 344
4.7. Acercarse desde un lado: límites laterales 352
4.8. Cuando la salida o la entrada escapan: límites infinitos y límites en infinito 359
4.9. Componer límites sin esconder la hipótesis delicada 367
4.10. Laboratorio de límite funcional 372
4.11. Ejercicios y soluciones 380

Capítulo 5. Continuidad en la recta: intervalos, compacidad y teoremas fundamentales 407
5.1. Continuidad en un punto: controlar también el centro 408
5.2. Criterio secuencial, álgebra y composición de funciones continuas 416
5.3. Cómo falla la continuidad: reparación, salto y oscilación 424
5.4. De lo local a atravesar valores: teorema del valor intermedio 430
5.5. Bisección: convertir existencia en localización 437
5.6. Recubrimientos abiertos y compactitud de un intervalo cerrado 445
5.7. Los extremos sí se alcanzan: teorema de los valores extremos 456
5.8. Una sola δ para todo el intervalo: continuidad uniforme 464
5.9. Continuidad + inyectividad en un intervalo: monotonía e inversa continua 472
5.10. Laboratorio de continuidad: del control local al global 477
5.11. Ejercicios y soluciones 485

Capítulo 6. La derivada y la aproximación lineal local 506
6.1. De la secante a la tasa instantánea 507
6.2. Derivada en un punto: definición, unicidad y criterio secuencial 512
6.3. Calcular desde la definición y aprender a fallar 518
6.4. Derivadas laterales, extremos, esquinas y pendientes verticales 525
6.5. Diferenciabilidad implica continuidad 530
6.6. La recta tangente: una definición local, no una figura 535
6.7. Diferenciabilidad como aproximación lineal local 539
6.8. Por qué la aproximación de primer orden es única 544
6.9. La derivada como función y la localidad del concepto 548
6.10. Laboratorio: reconstruir el pensamiento diferencial 553
6.11. Ejercicios y soluciones 561

Capítulo 7. Álgebra de derivadas y regla de la cadena 583
7.1. Por qué las reglas de derivación necesitan demostración 583
7.2. La derivación respeta combinaciones lineales 586
7.3. La regla del producto: separar una variación en dos 589
7.4. Recíprocos y cocientes: la hipótesis que impide dividir por cero 592
7.5. Potencias naturales y derivación de polinomios 596
7.6. Potencias enteras negativas y funciones racionales 599
7.7. La regla de la cadena: cuando el incremento interior puede anularse 602
7.8. Composiciones repetidas sin perder los puntos de evaluación 606
7.9. Lo que las reglas no afirman 610
7.10. Laboratorio: reconstruir, reparar y combinar 613

7.11. Ejercicios y soluciones 619

Capítulo 8. Derivación de funciones elementales, inversas e implícitas 635
8.1. La fórmula no crea su función 635
8.2. Inversa: existencia y continuidad primero 638
8.3. Derivada de una función inversa 641
8.4. Raíces de cualquier índice: construir antes de derivar 646
8.5. Potencias racionales y funciones algebraicas explícitas 649
8.6. Puntos excepcionales y funciones por tramos 653
8.7. Una inversa no se deduce de una derivada no nula aislada 656
8.8. Relaciones polinómicas y ramas 659
8.9. Derivación implícita polinómica sin circularidad 663
8.10. Laboratorio P3: leer una curva antes de derivarla 666
8.11. Ejercicios y soluciones 671

Capítulo 9. Teoremas de Rolle y del valor medio 684
9.1. De una pendiente local a una afirmación global 684
9.2. Extremos y puntos interiores 687
9.3. Fermat: condición necesaria de extremo interior 690
9.4. El teorema de Rolle 693
9.5. Auditoría de las hipótesis de Rolle 695
9.6. El teorema del valor medio 698
9.7. Qué afirma realmente el valor medio 700
9.8. Auditoría de las hipótesis del valor medio 703
9.9. Estimaciones mediante la derivada 705
9.10. Derivadas iguales y laboratorio de demostraciones 709
9.11. Ejercicios y soluciones 712

Capítulo 10. Monotonía, extremos, convexidad y forma de las gráficas 725
10.1. Del incremento medio al sentido del cambio 725
10.2. Funciones crecientes y decrecientes: definiciones y cuantificadores 727
10.3. Criterios de monotonía a partir de la derivada 729
10.4. Puntos críticos y cambios de signo 732
10.5. Extremos absolutos y extremos del dominio 737
10.6. Convexidad como propiedad de las cuerdas 741
10.7. Pendientes secantes, tangentes y primera derivada 746
10.8. Segunda derivada y pruebas de curvatura 750
10.9. Puntos de inflexión y errores de diagnóstico 755
10.10. Reconstrucción de gráficas y laboratorio de pruebas 759
10.11. Ejercicios y soluciones 765

Capítulo 11. Derivadas superiores y fórmula de Taylor con resto 776
11.1. De la aproximación lineal a la aproximación polinómica 776
11.2. Derivadas de orden superior: iteración, dominios y notación 780
11.3. Álgebra de derivadas superiores 782
11.4. El polinomio de Taylor y su unicidad 784
11.5. Un lema de Rolle iterado 786
11.6. Fórmula de Taylor con resto de Lagrange 788
11.7. Cotas del resto y control cuantitativo del error 790
11.8. Orden de contacto y residuo normalizado 792

11.9. Hipótesis, fallos y frontera con las series infinitas 793
11.10. Laboratorio de demostraciones y aproximaciones finitas 795
11.11. Ejercicios y soluciones 797

Capítulo 12. Aproximación, método de Newton y problemas de optimización 811
12.1. Aproximar no es adivinar: valor, error y certificado 811
12.2. Aislar una raíz antes de aproximarla 814
12.3. El paso de Newton desde la recta tangente 817
12.4. La identidad del error de Newton 822
12.5. Convergencia local y régimen cuadrático 828
12.6. Certificados a posteriori y criterios de parada 835
12.7. Cuándo Newton falla o engaña 843
12.8. De un problema verbal a una función objetivo 850
12.9. Optimización global en intervalos y dominios no compactos 858
12.10. Newton sobre la derivada y laboratorio integrador 864
12.11. Ejercicios y soluciones 870

Capítulo 13. Del área y las sumas a la integral 892
13.1. Del área geométrica al problema analítico 892
13.2. Particiones, subintervalos y refinamientos 897
13.3. Funciones escalonadas y regiones por rectángulos 903
13.4. La integral elemental de una función escalonada 908
13.5. Álgebra, orden y aditividad de la integral escalonada 913
13.6. Cotas por abajo y por arriba en una partición 918
13.7. Refinar mejora las cotas 923
13.8. Particiones etiquetadas y sumas de Riemann 928
13.9. La brecha de una partición y la oscilación 934
13.10. Laboratorio: de rectángulos a un candidato único 939
13.11. Ejercicios y soluciones 946

Capítulo 14. Integral de Riemann: definición, integrabilidad y propiedades 973
14.1. De todas las particiones a dos extremos globales 973
14.2. Definición de integrabilidad de Riemann 978
14.3. El criterio de Darboux: brechas arbitrariamente pequeñas 982
14.4. Sumas etiquetadas: equivalencia con la definición por malla 986
14.5. Compatibilidad con la integral elemental y cambios finitos 992
14.6. Las funciones continuas son integrables 997
14.7. Las funciones monótonas son integrables 1001
14.8. Álgebra y orden de las funciones integrables 1005
14.9. Restricción, aditividad por intervalos y orientación 1010
14.10. Laboratorio de integrabilidad: qué basta y qué no 1016
14.11. Ejercicios y soluciones 1021

Capítulo 15. Teoremas de valor medio para integrales 1045
15.1. De una integral a una altura media 1045
15.2. El promedio queda entre las cotas de la función 1046
15.3. Positividad estricta y comparación estricta 1048
15.4. Teorema del valor medio para integrales 1051
15.5. Consecuencias del valor medio 1053
15.6. Promedios ponderados 1054

15.7. Teorema del valor medio ponderado 1056
15.8. Por qué las hipótesis importan 1058
15.9. Estimaciones y lectura cuantitativa del promedio 1061
15.10. Laboratorio: escoger el teorema y reparar argumentos 1063
15.11. Ejercicios y soluciones 1067

Capítulo 16. Teorema fundamental del cálculo 1086
16.1. De un intervalo fijo a un extremo variable 1086
16.2. La función de acumulación es continua antes de ser derivable 1090
16.3. Derivar una acumulación en un punto de continuidad 1093
16.4. Primera forma global del teorema fundamental 1097
16.5. Primitivas: existencia y unicidad hasta una constante 1099
16.6. Newton–Leibniz: de una derivada a una integral 1103
16.7. Evaluar integrales cuando una primitiva es conocida 1108
16.8. Extremos variables y regla de la cadena 1111
16.9. Por qué las hipótesis importan 1115
16.10. Laboratorio: escoger la dirección correcta y reparar argumentos 1121
16.11. Ejercicios y soluciones 1126

Capítulo 17. Logaritmo, exponencial y funciones relacionadas desde el cálculo 1149
17.1. Por qué construir una función nueva 1149
17.2. El logaritmo natural como función de acumulación 1151
17.3. La ecuación funcional del logaritmo 1155
17.4. Geometría y rango del logaritmo 1160
17.5. La exponencial como función inversa 1164
17.6. Ley exponencial y notación e^x 1169
17.7. Potencias reales y logaritmos de base arbitraria 1174
17.8. Derivadas de logaritmos, exponenciales y potencias 1181
17.9. Funciones hiperbólicas y sus inversas elementales 1187
17.10. Laboratorio integrador y auditoría de dominios 1195
17.11. Ejercicios y soluciones 1203

Capítulo 18. Técnicas de integración 1225
18.1. Del teorema fundamental al problema inverso 1225
18.2. Cambio de variable: invertir la regla de la cadena 1229
18.3. Integración por partes: invertir la regla del producto 1236
18.4. Repetir con propósito: partes iteradas y fórmulas de reducción 1242
18.5. Potencias trigonométricas: reescribir antes de sustituir 1248
18.6. Productos trigonométricos de frecuencias distintas: convertir productos en sumas 1254
18.7. Funciones racionales I: dividir y descomponer en factores lineales simples 1259
18.8. Funciones racionales II: factores lineales repetidos 1265
18.9. Funciones racionales III: cuadráticos irreducibles y la aparición de arctan 1270
18.10. Funciones racionales IV: cuadráticos irreducibles repetidos y reducción finita 1275
18.11. Estrategia global: reconocer estructura, elegir representación y verificar 1281
Ejercicios y soluciones 1285
Ejercicios y soluciones 1285

Capítulo 19. Aplicaciones geométricas y cuantitativas de la integral 1315
19.1. De una cantidad local a una integral: aditividad y densidad 1315

19.2. Área entre gráficas: diferencia, signo y valor absoluto 1317
19.3. Volumen por secciones: cortar antes de integrar 1320
19.4. Sólidos de revolución I: discos y arandelas 1322
19.5. Sólidos de revolución II: cascarones y elección de representación 1325
19.6. Longitud de la gráfica: de polígonos a $\int \sqrt{1+(f')^2}$ 1328
19.7. Densidad lineal y cantidad total 1331
19.8. Promedio ponderado, primer momento y centro de masa de una varilla 1333
19.9. Trabajo de una fuerza variable 1336
19.10. Cambio neto desde una tasa conocida 1339
19.11. Laboratorio de modelización: elegir cortes, densidades y controles 1343
Ejercicios y soluciones 1346
Ejercicios y soluciones 1346

Capítulo 20. Ecuaciones diferenciales elementales y síntesis Newton–Leibniz 1393
20.1. Solución sobre un intervalo y dato inicial 1393
20.2. Reconstruir una función desde una tasa conocida 1395
20.3. Cuando la tasa es proporcional al estado 1397
20.4. Ecuaciones lineales y factor integrante 1400
20.5. Separación de variables y rama inversa 1401
20.6. Equilibrios y límites de la unicidad 1404
20.7. Del balance a la ecuación: enfriamiento y mezcla 1406
20.8. Restricción, prolongación e intervalo maximal 1408
20.9. Signo de la tasa y lectura cualitativa: la logística 1410
20.10. Síntesis Newton–Leibniz: del dato a la identidad integral 1411
20.11. Laboratorio de reconstrucción y cierre del tomo 1414
Ejercicios y soluciones 1416
Ejercicios y soluciones 1416

Bibliografía 1425

Capítulo 1

Los números reales: axiomas de cuerpo, orden y completitud

En los cálculos elementales usamos los números reales con la familiaridad de quien conoce el terreno: sumamos, comparamos y escribimos raíces confiando en que cada operación permitida dará un número y en que los números que necesitamos estarán disponibles. El cálculo nos obligará a examinar esa confianza. Afirmar que una sucesión tiene un límite o que una función continua alcanza cierto valor exige una razón para asegurar la existencia de ese número.

La imagen de una recta sin agujeros sugiere lo que buscamos, pero todavía debemos darle un significado que pueda utilizarse en una demostración. Los racionales servirán para reconocer la dificultad: permiten las operaciones aritméticas habituales y admiten un orden compatible con ellas, aunque dejan sin solución algunas ecuaciones muy sencillas. La comparación entre ambos sistemas nos lleva a preguntar

¿qué posee $\mathbb{R}$ que no posee $\mathbb{Q}$ y que hace posible el cálculo?

Tomaremos las propiedades de los números reales como punto de partida axiomático. Las operaciones y el orden describirán primero una estructura que también poseen los racionales; la **completitud** añadirá la condición de existencia que esa estructura deja abierta. Construir los reales a partir de objetos más elementales es otra manera de justificar estas propiedades, propia del estudio de sus fundamentos.

La ecuación que pondrá a prueba esa estructura es conocida desde la aritmética escolar, $x^2 = 2$, y, sin embargo, su resolución exige distinguir dos cuestiones: qué operaciones podemos hacer con un número y qué garantiza que exista un número con la propiedad buscada. Para formular la segunda cuestión necesitaremos las nociones de cota y supremo; la completitud permitirá después demostrar la existencia de la raíz real.

Como este es el primer capítulo, fijamos una convención mínima de lectura. La escritura $x \in A$ significa que x pertenece al conjunto A; $A \subseteq B$ indica que todo elemento de A pertenece a B. Las expresiones $\forall x \in A$ y $\exists x \in A$ se leen, respectivamente, «para todo elemento de A» y «existe al menos uno». El símbolo $\exists!$ exigirá, cuando aparezca, existencia **y** unicidad. Estas convenciones sirven para leer los axiomas y las pruebas; las operaciones con conjuntos necesarias para estudiar funciones se presentarán en el capítulo siguiente.

1.1. Axiomas de cuerpo y estructura de orden

Partir de las operaciones conocidas: ¿qué estructura necesitamos? Partamos de las operaciones que ya conocemos: sumar, multiplicar, restar y dividir por elementos no nulos. Los racionales poseen esas operaciones; los reales también. Queremos identificar las propiedades esenciales de esa estructura compartida y demostrar sus consecuencias antes de investigar qué propiedad adicional hace falta para resolver determinados problemas de existencia.

La cadena $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$ puede leerse como una sucesión de extensiones. Cada sistema incorpora al anterior y permite resolver problemas que antes no tenían solución. Pasar de $\mathbb{N}$ a $\mathbb{Z}$ permite restar sin abandonar el sistema; pasar de $\mathbb{Z}$ a $\mathbb{Q}$ permite dividir por enteros no nulos;

pasar de $\mathbb{Q}$ a $\mathbb{R}$ deberá resolver un problema diferente, relacionado con ciertos puntos frontera. Las reglas **algebraicas** permiten operar y despejar incógnitas; las reglas de **orden** permiten comparar los resultados y trabajar con desigualdades. Al reunir ambas obtendremos la estructura de **cuerpo ordenado**, que comparten racionales y reales.

La elección de axiomas recoge las propiedades de la estructura que queremos estudiar. Existen construcciones explícitas de $\mathbb{R}$, mediante cortes de Dedekind o sucesiones de Cauchy de racionales, en las que esas propiedades se demuestran a partir de la construcción. Al tomarlas como punto de partida podremos concentrarnos en las consecuencias que necesita el cálculo.

La parte algebraica: qué significa ser un cuerpo. Consideremos las operaciones $+$ y $\cdot$. Reordenar una suma, sacar factor común o cancelar un término son movimientos familiares, pero cada uno utiliza propiedades determinadas. Si escogemos unas pocas como axiomas y demostramos las demás a partir de ellas, podremos reconocer qué autoriza cada cálculo:

$$\boxed{\text{axioma} \longrightarrow \text{consecuencia demostrada} \longrightarrow \text{regla algebraica de uso cotidiano.}}$$

Una regla demostrada quedará disponible para los cálculos posteriores. Conocer su procedencia nos permitirá reconocer cuándo podemos usarla y cuándo falta alguna de sus hipótesis.

Dos operaciones antes que una lista de reglas. Sea F un conjunto no vacío. Por una **operación binaria** en F entenderemos aquí una regla que, dados dos elementos $a, b \in F$, produce un único elemento de F. Así, decir que en F tenemos una suma y un producto significa que, para cada par $a, b \in F$, existen resultados determinados

$$a + b \in F, \qquad ab \in F.$$

Al hablar de operaciones exigimos que cada par de entradas tenga un único resultado y que ese resultado siga perteneciendo a F. La segunda exigencia recibe el nombre de **clausura**. Por ejemplo, si $F = \mathbb{Z}$, la suma y el producto son operaciones internas, pero la división no lo es: $1/2 \notin \mathbb{Z}$. La estructura que vamos a introducir se formula, por tanto, en términos de dos operaciones primitivas; la resta y la división aparecerán después como operaciones **derivadas**.

Definición 1.1. **Cuerpo.** Un **cuerpo** es un conjunto no vacío F provisto de dos operaciones binarias, suma y producto, y de dos elementos distinguidos $0, 1 \in F$, con $0 \neq 1$, tales que para todos $a, b, c \in F$ se cumplen los axiomas siguientes.

Axiomas aditivos

- **(C1) Asociatividad de la suma**
$$a + (b + c) = (a + b) + c.$$
- **(C2) Existencia de neutro aditivo**
$$a + 0 = a.$$
- **(C3) Existencia de inverso aditivo:** para cada $a \in F$ existe al menos un elemento de F que, provisionalmente, denotaremos por $-a$, tal que
$$a + (-a) = 0.$$
Enseguida demostraremos que ese elemento es único; solo entonces la notación $-a$ quedará inequívocamente determinada.
- **(C4) Conmutatividad de la suma**
$$a + b = b + a.$$

Axiomas multiplicativos

- **(C5) Asociatividad del producto**
$$a(bc) = (ab)c.$$
- **(C6) Existencia de neutro multiplicativo**
$$a1 = a.$$
- **(C7) Existencia de inverso multiplicativo:** para cada $a \neq 0$ existe al menos un elemento de F que, provisionalmente, denotaremos por a^{-1}, tal que
$$aa^{-1} = 1.$$
También aquí demostraremos enseguida la unicidad del elemento y justificaremos definitivamente la notación.
- **(C8) Conmutatividad del producto**
$$ab = ba.$$

Axioma que enlaza ambas operaciones

- **(C9) Distributividad**
$$a(b + c) = ab + ac.$$

En este libro asumiremos que $\mathbb{R}$ posee esta estructura de cuerpo. Todavía no hemos añadido ningún axioma acerca del orden ni, mucho menos, acerca de la completitud.

Conviene advertir algo sobre la lista. Por ejemplo, $0a = 0$
no aparece entre los axiomas. Tampoco aparecen

$$-(-a) = a, \qquad (-a)(-b) = ab,$$

ni las reglas para cancelar términos, ni la regla del producto nulo, ni la multiplicación cruzada de fracciones. Todas esas afirmaciones tendrán que salir de un conjunto mucho más pequeño de supuestos.

¿Por qué podemos sustituir en una demostración?
A lo largo de una demostración diremos a menudo cosas como «sustituyendo $m = 2k$», «tomando $c = -a$» o «reemplazando 0 por $a + c$». Conviene distinguir dos operaciones lógicas diferentes que el lenguaje corriente suele llamar simplemente *sustituir*.
1. Sustituir iguales por iguales. Si sabemos que $x = y$, entonces x e y designan el mismo objeto. Por ello podemos reemplazar uno por el otro dentro de cualquier expresión bien definida. Así, por ejemplo,

$$x = y \implies x + c = y + c$$

y

$$x = y \implies x^2 = y^2.$$

Más generalmente, si $E(t)$ es una expresión bien definida,

$$x = y \implies E(x) = E(y).$$

Y si $P(t)$ es una afirmación acerca de t, sustituir x por y no cambia su valor de verdad. Esta es la **sustituibilidad de la igualdad**: no constituye un nuevo axioma de cuerpo, sino una

regla del lenguaje lógico con igualdad en el que formulamos los axiomas. Como la igualdad es simétrica, podemos usarla en cualquiera de los dos sentidos. Por eso, si $m = 2k$ y sabemos que $m^2 = 2n^2$, podemos escribir $(2k)^2 = 2n^2$.

2. Particularizar una afirmación universal. Si una propiedad ha sido establecida para **todo** elemento de un dominio, podemos aplicarla a cualquier elemento admisible de ese dominio. Por ejemplo, la distributividad afirma que, para cualesquiera $a, b, c \in F$, $a(b + c) = ab + ac$.

Podemos particularizar esta identidad tomando $c = -b$: $a\big(b + (-b)\big) = ab + a(-b)$. Aquí no hemos reemplazado dos objetos porque sean iguales: hemos escogido un caso particular de una afirmación universal. Estas dos reglas explican buena parte de las «sustituciones» que aparecerán en las pruebas. Pero no autorizan reemplazos arbitrarios: debe existir una igualdad previa o una afirmación universal que justifique el paso, y deben respetarse todas las hipótesis bajo las cuales la expresión está definida. Por ejemplo, una identidad demostrada solo para $a \neq 0$ no puede particularizarse tomando $a = 0$. En una prueba especialmente delicada conviene hacerse esta **pregunta de control**: ¿estoy sustituyendo iguales por iguales o particularizando una afirmación universal?

Lo mismo vale para encadenar igualdades por transitividad: también es una regla lógica previa a los axiomas particulares del cuerpo.

Existencia no significa todavía unicidad. Los axiomas (C2), (C3), (C6) y (C7) garantizan la **existencia** de neutros e inversos. Para hablar de cada uno de ellos sin ambigüedad necesitamos saber también que es único. Bastará comparar dos candidatos y utilizar las propiedades que ambos deben satisfacer para demostrar que coinciden.

Proposición 1.1. **Unicidad de neutros e inversos.** En todo cuerpo F se cumplen las afirmaciones siguientes.

1. El neutro aditivo es único.
2. El neutro multiplicativo es único.
3. Para cada $a \in F$, el inverso aditivo de a es único.
4. Para cada $a \in F$ con $a \neq 0$, el inverso multiplicativo de a es único.

En el caso de los neutros, la operación entre los dos candidatos permitirá usar la propiedad de cada uno. Para los inversos, introduciremos el elemento del que ambos son inversos y reagruparemos hasta pasar de un candidato al otro. Las igualdades siguientes muestran cómo funciona esa comparación.

Demostración.

Supongamos que 0 y $\widetilde{0}$ son dos neutros aditivos. Como $\widetilde{0}$ es neutro, $0 + \widetilde{0} = 0$. Por conmutatividad, $0 + \widetilde{0} = \widetilde{0} + 0$, y, como 0 también es neutro, $\widetilde{0} + 0 = \widetilde{0}$.

Por transitividad, $0 = \widetilde{0}$.

Así, el neutro aditivo es único. La prueba para el neutro multiplicativo es idéntica en estructura. Si 1 y $\widetilde{1}$ son dos neutros multiplicativos, entonces $1\widetilde{1} = 1$ porque $\widetilde{1}$ es neutro, mientras que $1\widetilde{1} = \widetilde{1}1 = \widetilde{1}$ por conmutatividad y porque 1 es neutro. Por tanto, $1 = \widetilde{1}$.

Consideremos ahora los inversos aditivos. Sea $a \in F$ y supongamos que b y c satisfacen

$$a + b = 0, \qquad a + c = 0.$$

Queremos demostrar $b = c$. Utilizando únicamente los axiomas ya disponibles,

$$\begin{aligned} b &= b + 0 \\ &= b + (a + c) \\ &= (b + a) + c \\ &= (a + b) + c \\ &= 0 + c \\ &= c. \end{aligned}$$

Luego el inverso aditivo de a es único. Por último, sea $a \neq 0$ y supongamos que b y c son dos inversos multiplicativos de a:

$$ab = 1, \qquad ac = 1.$$

Entonces

$$\begin{aligned} b &= b1 \\ &= b(ac) \\ &= (ba)c \\ &= (ab)c \\ &= 1c \\ &= c. \end{aligned}$$

Por tanto, el inverso multiplicativo también es único. ■

Después de la prueba
Ahora la notación

$$-a \qquad \text{y} \qquad a^{-1}$$

está mejor justificada.
Los axiomas garantizan **existencia** y la proposición garantiza **unicidad**. Por eso tiene sentido hablar de *el* inverso aditivo de a y, cuando $a \neq 0$, de *el* inverso multiplicativo de a.

Por qué multiplicar por cero da cero. La distributividad permite obtener una regla que usamos desde la aritmética elemental: $a0 = 0$. La afirmación resulta tan familiar que es fácil olvidar que no figura entre (C1)–(C9).

Ejemplo 1.1. **Demostrar desde los axiomas que** $a0 = 0$**.** Sea $a \in F$. Entonces $a0 = 0$.

Demostración. Como 0 es neutro aditivo, $0 + 0 = 0$.

Multiplicando ambos miembros por a, $a(0 + 0) = a0$.

Por distributividad, $a0 + a0 = a0$.

Ahora sumamos el inverso aditivo de $a0$ en ambos miembros: $(a0 + a0) + (-a0) = a0 + (-a0)$. Usando asociatividad e inverso aditivo, $a0 + \big(a0 + (-a0)\big) = 0$, de modo que $a0 + 0 = 0$.

Finalmente, por el neutro aditivo, $a0 = 0$.

■

Podemos auditar la prueba en una tabla.

Paso	Qué lo autoriza
$0+0=0$	(C2), con $a=0$
$a(0+0)=a0$	sustitución en una igualdad
$a0+a0=a0$	(C9), distributividad
sumar $-a0$	(C3) y sustitución
reagrupar	(C1)
$a0+(-a0)=0$	(C3)
$a0+0=a0$	(C2)

La demostración ilustra una regla general de lectura: una cadena de símbolos puede ser correcta y, sin embargo, no explicar por sí sola **por qué** cada paso es legal.

De los axiomas a las reglas de signos y de inversos. Una vez demostrada la unicidad, muchas identidades se obtienen mostrando que cierto candidato satisface la propiedad que caracteriza a un inverso.

PROPOSICIÓN 1.2. **Consecuencias algebraicas elementales.** En todo cuerpo F, para cualesquiera $a, b \in F$, se cumplen:

1. $$-0 = 0;$$
2. $$-(-a) = a;$$
3. $$-(a+b) = (-a) + (-b);$$
4. $$(-a)b = -(ab), \qquad a(-b) = -(ab),$$

   y por tanto
   $$(-a)(-b) = ab.$$

Además, si $a \neq 0$, entonces

$$a^{-1} \neq 0, \qquad (a^{-1})^{-1} = a.$$

Si $a \neq 0$ y $b \neq 0$, entonces

$$ab \neq 0$$

y

$$(ab)^{-1} = a^{-1}b^{-1}.$$

Demostración.

Como $0+0=0$, el propio 0 es un inverso aditivo de 0. Por la unicidad demostrada en Proposición 1.1, $-0=0$. Para la segunda afirmación, observemos que $(-a)+a = a+(-a) = 0$. Así, a es un inverso aditivo de $-a$. Por unicidad, $-(-a) = a$. Para encontrar el inverso de una suma, comprobamos —reagrupando y permutando términos mediante asociatividad y conmutatividad— que

$$\begin{aligned}(a+b)+\big((-a)+(-b)\big) &= \big(a+(-a)\big)+\big(b+(-b)\big)\\ &= 0+0\\ &= 0.\end{aligned}$$

La escritura intermedia es una abreviación de aplicaciones de asociatividad y conmutatividad. Por unicidad del inverso aditivo, $-(a+b) = (-a)+(-b)$.

Pasemos a las reglas de signos. Por conmutatividad, distributividad y Ejemplo 1.1,

$$\begin{aligned}ab+(-a)b &= ba+b(-a)\\ &= b\big(a+(-a)\big)\\ &= b0\\ &= 0.\end{aligned}$$

Por tanto, $(-a)b$ es un inverso aditivo de ab, y la unicidad da $(-a)b = -(ab)$.

De manera análoga, $a(-b) = -(ab)$.

Aplicando estas identidades dos veces y usando $-(-x) = x$,

$$\begin{aligned}(-a)(-b) &= -\big(a(-b)\big)\\ &= -\big(-(ab)\big)\\ &= ab.\end{aligned}$$

Supongamos ahora $a \neq 0$. Su inverso satisface $aa^{-1} = 1$. No puede ocurrir que $a^{-1} = 0$, porque entonces Ejemplo 1.1 daría $aa^{-1} = a0 = 0$, en contradicción con $aa^{-1} = 1$ y $0 \neq 1$. Luego $a^{-1} \neq 0$.

Además, $a^{-1}a = aa^{-1} = 1$.

Así, a es un inverso multiplicativo de a^{-1}. Por unicidad, $(a^{-1})^{-1} = a$.

Finalmente, sean $a, b \neq 0$. Consideremos $c = a^{-1}b^{-1}$.

Por asociatividad y conmutatividad,

$$\begin{aligned}(ab)c &= (ab)(a^{-1}b^{-1})\\ &= (aa^{-1})(bb^{-1})\\ &= 1.\end{aligned}$$

En particular, ab no puede ser 0: si lo fuera, el miembro izquierdo sería $0c = c0 = 0$, por conmutatividad y Ejemplo 1.1, contradiciendo $1 \neq 0$. Por tanto, $ab \neq 0$ y posee inverso multiplicativo. Como c es un inverso de ab, la unicidad implica $(ab)^{-1} = a^{-1}b^{-1}$.

■

Una técnica que conviene reconocer

Varias partes de la prueba tienen la misma arquitectura:

> para demostrar que dos expresiones son iguales, mostramos que ambas desempeñan el mismo papel caracterizado de manera única.

Así demostramos, por ejemplo, que $(-a)b = -(ab)$: en lugar de manipular directamente el signo menos, probamos que $(-a)b$ es **el inverso aditivo** de ab. Esta forma de razonar será mucho más útil que memorizar una lista independiente de reglas de signos.

Cancelar no significa borrar símbolos. En álgebra escolar se habla de «cancelar» como si ciertos símbolos simplemente desaparecieran. En una demostración conviene saber qué teorema hay detrás.

PROPOSICIÓN 1.3. **Leyes de cancelación y producto nulo.** Sean $a, b, c \in F$.

1. Se tiene
$$a + b = a + c \iff b = c.$$
2. Si $a \neq 0$, entonces
$$ab = ac \iff b = c.$$
3. Se tiene
$$ab = 0 \iff a = 0 \text{ o } b = 0.$$

Demostración.

Para la cancelación aditiva, supongamos primero $a+b = a+c$. Sumando $-a$ a ambos miembros y reagrupando, $(-a) + (a + b) = (-a) + (a + c)$, de donde $b = c$. La implicación recíproca es inmediata por sustitución: si $b = c$, entonces $a + b = a + c$. Para la cancelación multiplicativa, supongamos $a \neq 0$ y $ab = ac$.

Multiplicando ambos miembros por a^{-1}, $a^{-1}(ab) = a^{-1}(ac)$.

Por asociatividad, $(a^{-1}a)b = (a^{-1}a)c$, y por tanto $b = c$.

Nuevamente, la recíproca sigue de sustituir iguales por iguales.

Consideremos ahora el producto nulo. Si $b = 0$, Ejemplo 1.1 da directamente $ab = a0 = 0$. Si $a = 0$, usamos además conmutatividad:

$$ab = 0b = b0 = 0.$$

Así, en cualquiera de los dos casos, $ab = 0$.

Para la otra dirección, supongamos $ab = 0$.

Hay dos casos.

Si $a = 0$, ya tenemos una de las alternativas deseadas. Si $a \neq 0$, existe a^{-1} y podemos multiplicar la igualdad por él: $a^{-1}(ab) = a^{-1}0$.

Entonces $(a^{-1}a)b = 0$, de modo que $b = 0$.

En cualquiera de los casos, $a = 0$ o $b = 0$.

■

Por qué la hipótesis $a \neq 0$ no es decorativa
De $ab = ac$ no podemos concluir siempre $b = c$. Si $a = 0$, entonces $0b = 0c$ para **todos** $b, c \in F$, aunque $b \neq c$. Por eso la cancelación multiplicativa exige $a \neq 0$. Esa hipótesis es exactamente la que permite invocar a^{-1}.

La tercera parte de Proposición 1.3 dice que un cuerpo no tiene divisores de cero no triviales. Más adelante, en álgebra abstracta, esta propiedad aparecerá en contextos más generales. Aquí nos interesa sobre todo porque justifica el método de resolver ecuaciones factorizadas:

$$(x - r)(x - s) = 0 \quad \Longrightarrow \quad x = r \text{ o } x = s.$$

No es una regla independiente sobre polinomios. Es una consecuencia de la estructura de cuerpo.

Despejar una incógnita es un teorema de existencia y unicidad. Ahora podemos formular rigurosamente dos operaciones que hacemos constantemente.

TEOREMA 1.1. **Ecuaciones elementales en un cuerpo.** Sean $a, b \in F$.

1. La ecuación
$$a + x = b$$
tiene una única solución, dada por
$$x = b + (-a).$$
2. Si $a \neq 0$, la ecuación
$$ax = b$$
tiene una única solución, dada por
$$x = a^{-1}b.$$

Idea de la prueba
Cada parte exige distinguir dos trabajos matemáticos:

- **existencia:** exhibir un candidato y comprobar que satisface la ecuación;
- **unicidad:** demostrar que cualquier otra solución debe coincidir con ese candidato.

La fórmula obtenida al «despejar» no sustituye esos dos trabajos: es su conclusión.

Demostración.
Para la primera ecuación proponemos $x = b + (-a)$.
Entonces

$$\begin{aligned} a + x &= a + \big(b + (-a)\big) \\ &= (a + (-a)) + b \\ &= 0 + b \\ &= b, \end{aligned}$$

donde hemos usado asociatividad y conmutatividad para reagrupar. Así, la solución existe. Supongamos ahora que y es cualquier otra solución. Entonces $a + y = b = a + x$.
Por cancelación aditiva, $y = x$.
La solución es única.
Para la segunda ecuación supongamos $a \neq 0$ y propongamos $x = a^{-1}b$.
Entonces

$$\begin{aligned} ax &= a(a^{-1}b) \\ &= (aa^{-1})b \\ &= 1b \\ &= b. \end{aligned}$$

Por tanto, existe una solución. Si y es otra solución, entonces $ay = b = ax$. Como $a \neq 0$, la cancelación multiplicativa de Proposición 1.3 da $y = x$. Así, la solución también es única. ■

Qué significa ahora «hacer la misma operación a ambos lados»
Cuando resolvemos $a + x = b$
«restando a», en realidad estamos sumando el inverso aditivo $-a$ a ambos miembros y usando asociatividad, neutro e inverso.
Cuando resolvemos

$$ax = b, \qquad a \neq 0,$$

«dividiendo por a», en realidad estamos multiplicando por el inverso a^{-1}.
Las reglas escolares son versiones comprimidas de argumentos estructurales.

Resta y división: operaciones derivadas. Ya podemos introducir las dos notaciones que hasta ahora parecían operaciones independientes.

DEFINICIÓN 1.2. **Resta y división.** Sean $a, b \in F$.
La **diferencia** de a y b se define por

$$a - b := a + (-b).$$

Si $b \neq 0$, el **cociente** de a por b se define por

$$\frac{a}{b} := ab^{-1}.$$

Desarrollo optativo: la resta como operación inversa de la suma
La definición anterior también puede obtenerse a partir de la idea elemental de que **restar b debe deshacer la suma de b**. En vez de definir de entrada $a - b$, caractericemos provisionalmente la diferencia como el único elemento $x \in F$ que satisface $b + x = a$.
Sumamos $-b$ a ambos miembros: $(-b) + (b + x) = (-b) + a$.
Por asociatividad, $\big((-b) + b\big) + x = (-b) + a$.
Como $(-b) + b = 0$, obtenemos $0 + x = (-b) + a$, y, por el neutro aditivo y la conmutatividad, $x = (-b) + a = a + (-b)$. Por tanto, la operación caracterizada como «deshacer la suma de b» coincide necesariamente con sumar su inverso aditivo:

$$\boxed{a - b = a + (-b).}$$

En nuestro desarrollo esta igualdad es, desde ahora, **definicional**; el argumento muestra por qué esa definición captura exactamente la operación inversa de sumar b.

Desarrollo optativo: la división como operación inversa del producto
El mismo razonamiento funciona multiplicativamente, con una diferencia esencial: debemos exigir $b \neq 0$ para que exista b^{-1}. En vez de definir de entrada a/b, caractericemos provisionalmente el cociente como el único elemento $x \in F$ que satisface $bx = a$.
Multiplicamos ambos miembros por b^{-1}: $b^{-1}(bx) = b^{-1}a$.
Por asociatividad, $(b^{-1}b)x = b^{-1}a$.
Como $b^{-1}b = 1$, obtenemos $1x = b^{-1}a$, y, por el neutro multiplicativo y la conmutatividad, $x = b^{-1}a = ab^{-1}$. Por tanto, la operación caracterizada como «deshacer la multiplicación por b» coincide necesariamente con multiplicar por su inverso:

$$\boxed{\frac{a}{b} = ab^{-1}}, \qquad b \neq 0.$$

Así como la resta no introduce una nueva operación primitiva, la división tampoco: ambas se construyen a partir de las operaciones del cuerpo y de sus respectivos inversos.

La resta y la división se apoyan así en las dos operaciones iniciales del cuerpo: sumar un inverso aditivo y multiplicar por un inverso multiplicativo. La segunda construcción sólo está disponible cuando existe ese inverso, de donde procede la restricción

$$\boxed{\text{no se divide por 0.}}$$

El número 0 carece de inverso multiplicativo: si existiera c tal que $0c = 1$,

Ejemplo 1.1 daría $0c = 0$, y obtendríamos $0 = 1$, contradiciendo la definición de cuerpo.

De la definición de resta y Proposición 1.2 se obtienen, por ejemplo, $a - (-b) = a + b$ y $a - b = 0 \iff a = b$. La segunda equivalencia también puede leerse mediante Teorema 1.1: la ecuación $a + (-b) = 0$ determina de manera única la relación entre a y b.

Las reglas de fracciones también se demuestran. La notación fraccionaria concentra varias aplicaciones de los axiomas. Conviene establecer una vez las reglas esenciales y dejar de tratarlas como recetas sin fundamento.

Proposición 1.4. **Reglas básicas de cocientes.** Sean $a, b, c, d \in F$.

1. Para todo a,
$$\frac{a}{1} = a.$$
Si $a \neq 0$, entonces
$$\frac{1}{a} = a^{-1}, \qquad \frac{a}{a} = 1.$$
2. Si $b \neq 0$ y $d \neq 0$, entonces
$$\frac{a}{b} = \frac{c}{d} \iff ad = bc.$$
3. Si $b \neq 0$ y $d \neq 0$, entonces
$$\frac{a}{b}\frac{c}{d} = \frac{ac}{bd}.$$
4. Si $b \neq 0$ y $d \neq 0$, entonces
$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd},$$
y
$$\frac{a}{b} - \frac{c}{d} = \frac{ad - bc}{bd}.$$

Demostración.

Como 1 es su propio inverso multiplicativo, $1^{-1} = 1$, y por tanto $\frac{a}{1} = a1^{-1} = a$. Si $a \neq 0$, las otras dos identidades iniciales son simplemente $\frac{1}{a} = 1a^{-1} = a^{-1}$ y $\frac{a}{a} = aa^{-1} = 1$. Para la igualdad de fracciones, supongamos $b, d \neq 0$. Si $\frac{a}{b} = \frac{c}{d}$, entonces, por definición de cociente, $ab^{-1} = cd^{-1}$. Por sustitución de iguales por iguales podemos multiplicar ambos miembros por el mismo elemento bd: $(ab^{-1})(bd) = (cd^{-1})(bd)$. Ahora reducimos cada miembro por separado. En el izquierdo,

$$\begin{aligned}(ab^{-1})(bd) &= ab^{-1}bd\\ &= a(b^{-1}b)d\\ &= a1d\\ &= ad.\end{aligned}$$

En el derecho,

$$\begin{aligned}(cd^{-1})(bd) &= cd^{-1}bd\\ &= cb(d^{-1}d)\\ &= cb\\ &= bc.\end{aligned}$$

Aquí hemos usado asociatividad y conmutatividad del producto, las identidades $b^{-1}b = 1$ y $d^{-1}d = 1$, y el neutro multiplicativo. Por tanto, $ad = bc$.

Recíprocamente, supongamos $ad = bc$.

Como $b, d \neq 0$, existen b^{-1} y d^{-1}. De nuevo por sustitución de iguales por iguales, multiplicamos ambos miembros por $b^{-1}d^{-1}$: $(ad)(b^{-1}d^{-1}) = (bc)(b^{-1}d^{-1})$.

En el miembro izquierdo,

$$\begin{aligned}(ad)(b^{-1}d^{-1}) &= adb^{-1}d^{-1}\\ &= ab^{-1}(dd^{-1})\\ &= ab^{-1}.\end{aligned}$$

En el derecho,

$$\begin{aligned}(bc)(b^{-1}d^{-1}) &= bcb^{-1}d^{-1}\\ &= c(bb^{-1})d^{-1}\\ &= cd^{-1}.\end{aligned}$$

Así, $ab^{-1} = cd^{-1}$, y, por definición de cociente, $\frac{a}{b} = \frac{c}{d}$. Para el producto usamos Proposición 1.2:

$$\begin{aligned}\frac{a}{b}\frac{c}{d} &= (ab^{-1})(cd^{-1})\\ &= ac(b^{-1}d^{-1})\\ &= ac(bd)^{-1}\\ &= \frac{ac}{bd}.\end{aligned}$$

Finalmente,

$$\frac{ad}{bd} = ad(bd)^{-1} = ad\,b^{-1}d^{-1} = ab^{-1} = \frac{a}{b},$$

y análogamente $\frac{bc}{bd} = \frac{c}{d}$.

Por distributividad,

$$\frac{a}{b}+\frac{c}{d}=\frac{ad}{bd}+\frac{bc}{bd}=(ad+bc)(bd)^{-1}=\frac{ad+bc}{bd}.$$

La fórmula para la resta se obtiene sustituyendo c por $-c$ y usando las reglas de signos ya demostradas. ■

Qué es realmente la multiplicación cruzada
La equivalencia

$$\frac{a}{b}=\frac{c}{d}\iff ad=bc\qquad (b,d\neq 0)$$

no introduce una nueva operación llamada «multiplicar en cruz». Es una abreviación de un argumento con inversos. Los denominadores no nulos son hipótesis matemáticas, no detalles de notación.

Auditar una manipulación algebraica completa. Podemos aplicar ahora todo el repertorio a una cadena que en un curso elemental escribiríamos casi automáticamente.

Ejemplo 1.2. **De una cadena escolar a una prueba auditada.** Sean $a,b,c\in F$ con $b\neq 0$. Resolvamos $\frac{x-a}{b}=c$.

La cadena habitual es

$$\frac{x-a}{b}=c\iff x-a=bc\iff x=a+bc.$$

El resultado es correcto, pero ahora podemos explicar cada flecha.

Primer paso. Por definición de cociente, $\frac{x-a}{b}=(x-a)b^{-1}$.

La hipótesis $b\neq 0$ ya es necesaria para que el cociente original esté definido y garantiza además la existencia de b^{-1}. Multiplicar una igualdad por b es legal incluso sin esa hipótesis; la no nulidad será necesaria cuando reduzcamos $b^{-1}b$ a 1.

Multiplicando ambos miembros por b, $\big((x-a)b^{-1}\big)b=cb$.

Reducimos primero el miembro izquierdo:

$$\begin{aligned}\big((x-a)b^{-1}\big)b&=(x-a)(b^{-1}b) && \text{(asociatividad)}\\ &=(x-a)1 && \text{(inverso multiplicativo)}\\ &=x-a && \text{(neutro multiplicativo).}\end{aligned}$$

En el miembro derecho, por conmutatividad, $cb=bc$.

Por tanto, $x-a=bc$.

Además, esta flecha es reversible. Si partimos de $x-a=bc$ y multiplicamos ambos miembros por b^{-1}, obtenemos $(x-a)b^{-1}=(bc)b^{-1}$. Por asociatividad y por $bb^{-1}=1$, $(bc)b^{-1}=c(bb^{-1})=c$, de modo que recuperamos $\frac{x-a}{b}=c$.

Segundo paso. Por definición de resta, $x-a=x+(-a)$, así que la ecuación anterior es $x+(-a)=bc$.

Sumamos a a ambos miembros: $\big(x+(-a)\big)+a=bc+a$.

El miembro izquierdo se reduce paso a paso:

$$\begin{aligned}(x + (-a)) + a &= x + ((-a) + a) && \text{(asociatividad)}\\ &= x + 0 && \text{(inverso aditivo)}\\ &= x && \text{(neutro aditivo).}\end{aligned}$$

Por consiguiente, $x = bc + a$.

Finalmente, por conmutatividad de la suma, $bc + a = a + bc$, y obtenemos

$$\boxed{x = a + bc}.$$

También esta flecha es reversible: sumando $-a$ a ambos miembros de $x = a + bc$ recuperamos $x - a = bc$. Por tanto, no solo hemos encontrado un candidato; hemos mostrado una cadena de equivalencias. En particular, Teorema 1.1 garantiza que la solución es única. Podemos resumir la auditoría así:

Movimiento escolar	Estructura que realmente utiliza
«multiplicar ambos miembros por b»	sustitución en una igualdad; por sí sola no exige $b \neq 0$
«se cancela b»	$b \neq 0$, existencia de b^{-1}, $b^{-1}b = 1$ y neutro multiplicativo
«pasar a sumando»	resta = suma con inverso; asociatividad, inverso y neutro aditivos
«la solución es la única»	reversibilidad de las equivalencias y Teorema 1.1

La última fila es importante. Encontrar un valor que satisface una ecuación prueba **existencia**; demostrar que ningún otro valor puede satisfacerla prueba **unicidad**.

De las operaciones al orden. Las reglas obtenidas dependen unas de otras: la unicidad de los inversos sostiene las reglas de signos y éstas permiten justificar las manipulaciones con cocientes. El esquema recoge esas dependencias:

axiomas de cuerpo
↓
unicidad de neutros e inversos
↓
$a0 = 0$, reglas de signos e inversos
↓
cancelación y producto nulo
↓
existencia y unicidad en ecuaciones
↓
resta y división
↓
reglas de fracciones.

Estas consecuencias quedarán disponibles sin necesidad de reconstruirlas en cada cálculo. Una división o una cancelación exigirá, en cambio, comprobar las condiciones bajo las que fue demostrada. Las pruebas anteriores utilizan únicamente los axiomas de cuerpo, por lo que sus conclusiones valen también en cuerpos distintos del real. Comparar números exige una estructura adicional. Las operaciones del cuerpo, por sí solas, no determinan qué elemento es mayor ni cómo se comporta una desigualdad al sumar o multiplicar sus miembros; esas propiedades corresponden al **orden**.

En lugar de postular de una vez una larga lista de reglas para el símbolo $<$, seguiremos una estrategia más económica: distinguiremos primero qué elementos llamaremos **positivos** y exigiremos tres propiedades básicas. A partir de ellas construiremos el orden y demostraremos sus reglas de manipulación. Esta elección permite prolongar el mismo principio que guió la capa algebraica:

$$\boxed{\text{axiomas mínimos} \longrightarrow \text{consecuencias demostradas} \longrightarrow \text{reglas de uso cotidiano.}}$$

La parte de orden: empezar por la positividad. Sea F un cuerpo. Supongamos que se ha distinguido un subconjunto $F_+ \subseteq F$, cuyos elementos llamaremos **positivos**, y que satisface las propiedades siguientes.

1. **(O1) Clausura de los positivos bajo la suma.** Si $a, b \in F_+$, entonces $a + b \in F_+$.
2. **(O2) Clausura de los positivos bajo el producto.** Si $a, b \in F_+$, entonces $ab \in F_+$.
3. **(O3) Tricotomía respecto de cero.** Para cada $a \in F$ ocurre exactamente una de las tres posibilidades
$$a = 0, \qquad a \in F_+, \qquad -a \in F_+.$$

Diremos entonces que $a > 0 \iff a \in F_+$, y $a < 0 \iff -a \in F_+$.

Para comparar dos elementos cualesquiera definimos

$$\boxed{a < b \iff b - a > 0.}$$

Equivalentemente, $a > b \iff a - b > 0$.

Un cuerpo equipado con una elección de positivos que satisface (O1)–(O3), y con el orden definido de esta manera, se llama **cuerpo ordenado**. En este libro asumiremos que $\mathbb{R}$ posee esta estructura. Obsérvese que todavía no hemos introducido ninguna propiedad de completitud.

Por qué esta formulación es útil
También es posible presentar un cuerpo ordenado tomando la relación $<$ como dato primitivo y postulando directamente tricotomía, transitividad y compatibilidad con las operaciones. Aquí preferimos comenzar por los positivos porque hace visible una dependencia más profunda: varias propiedades que en la escuela suelen aparecer como «reglas de las desigualdades» pueden demostrarse a partir de solo tres exigencias sobre F_+.

De la tricotomía respecto de cero a la tricotomía entre dos números. La condición (O3) habla de un solo número y de su relación con 0. Sin embargo, basta aplicarla a la **diferencia** de dos números para comparar cualquier par.

Dados $a, b \in F$, apliquemos (O3) a $b - a$.

Exactamente una de estas posibilidades ocurre:

$$b - a = 0, \qquad b - a > 0, \qquad -(b - a) > 0.$$

Hagamos explícita la primera equivalencia. Por definición de resta, $b-a=0 \iff b+(-a)=0$. Si sumamos a a ambos miembros, obtenemos $\big(b+(-a)\big)+a=0+a$.

Ahora,

$$\begin{aligned}\big(b+(-a)\big)+a &= b+\big((-a)+a\big) && \text{(asociatividad)}\\ &= b+0 && \text{(inverso aditivo)}\\ &= b && \text{(neutro aditivo),}\end{aligned}$$

mientras que, por conmutatividad y neutro aditivo, $0+a=a+0=a$. Por tanto, $b-a=0$ implica $b=a$, es decir, $a=b$. Recíprocamente, si $a=b$, entonces por sustitución

$$b-a=a-a=a+(-a)=0.$$

Así, $b-a=0 \iff a=b$.

La segunda posibilidad, $b-a>0$, equivale por definición a $a<b$. Para la tercera conviene justificar también la identidad que cambia el orden de la diferencia. Usando la regla ya demostrada para el inverso de una suma,

$$\begin{aligned}-(b-a) &= -\big(b+(-a)\big)\\ &= (-b)+\big(-(-a)\big)\\ &= (-b)+a\\ &= a+(-b)\\ &= a-b.\end{aligned}$$

Por tanto,

$$-(b-a)>0 \iff a-b>0 \iff b<a.$$

En consecuencia, para cualesquiera $a,b\in F$, exactamente una de las afirmaciones

$$\boxed{a<b, \qquad a=b, \qquad b<a}$$

es verdadera.

Esta es la **tricotomía del orden**. En particular, no puede ocurrir simultáneamente

$$a<b \quad \text{y} \quad b<a.$$

Tampoco puede cumplirse $a<a$, porque por definición exigiría $a-a>0$. Pero $a-a=a+(-a)=0$, y (O3), al afirmar que exactamente una de sus tres alternativas ocurre, excluye que 0 sea positivo.

La transitividad también se demuestra. Supongamos

$$a<b \qquad \text{y} \qquad b<c.$$

Por definición,

$$b-a>0 \qquad \text{y} \qquad c-b>0.$$

La clausura de los positivos bajo la suma da $(b-a)+(c-b)>0$. Ahora hacemos explícita la reducción algebraica del miembro izquierdo:

$$\begin{aligned}(b-a)+(c-b) &= \big(b+(-a)\big)+\big(c+(-b)\big)\\ &= \big((-a)+c\big)+\big(b+(-b)\big) && \text{(asociatividad y conmutatividad)}\\ &= \big((-a)+c\big)+0 && \text{(inverso aditivo)}\\ &= (-a)+c && \text{(neutro aditivo)}\\ &= c+(-a) && \text{(conmutatividad)}\\ &= c-a && \text{(definición de resta).}\end{aligned}$$

Por sustitución de iguales por iguales, de $(b-a)+(c-b)>0$ obtenemos entonces $c-a>0$.

Y, por definición del orden,

$$\boxed{a<c.}$$

Así, la transitividad de $<$ no ha sido añadida como un cuarto axioma independiente: sale de (O1), de las reglas algebraicas ya demostradas y de la manera en que definimos la comparación mediante diferencias.

Una idea estructural que conviene retener
Para comparar a y b estudiamos el signo de $b-a$. Este patrón aparecerá constantemente en cálculo. Muchas preguntas acerca de dos cantidades se convierten en preguntas acerca del signo, el tamaño o el valor absoluto de su diferencia.

Del orden estricto al orden débil. Definiremos $a\le b$ como abreviatura de

$$a<b \quad \text{o} \quad a=b.$$

Análogamente, $a\ge b \iff b\le a$.

Las palabras **positivo**, **negativo**, **no negativo** y **no positivo** son entonces comparaciones con 0:

$$a>0, \qquad a<0, \qquad a\ge 0, \qquad a\le 0.$$

La relación $\le$ hereda ahora las propiedades de un orden. Conviene verificarlas una vez.

Reflexividad. Para todo a, sabemos que $a=a$. Por la definición

$$a\le a \iff a<a \quad \text{o} \quad a=a,$$

la segunda alternativa es verdadera. Por tanto, $a\le a$.

Antisimetría. Supongamos

$$a\le b \qquad \text{y} \qquad b\le a.$$

Si $a=b$, no hay nada que demostrar. Supongamos entonces $a\neq b$. La definición de $\le$ obliga en ese caso a que

$$a<b \qquad \text{y} \qquad b<a,$$

pero la tricotomía demuestra que esas dos desigualdades no pueden ser simultáneamente verdaderas. Por tanto, la suposición $a\neq b$ es imposible y necesariamente $a=b$.

Transitividad. Supongamos

$$a \leq b \qquad \text{y} \qquad b \leq c.$$

Si $a = b$, la segunda desigualdad da directamente $a \leq c$ por sustitución. Si $b = c$, la primera da $a \leq c$. En el caso restante tenemos

$$a < b \qquad \text{y} \qquad b < c,$$

y la transitividad del orden estricto ya demostrada produce $a < c$, de donde, por definición, $a \leq c$. Así, $\leq$ es reflexiva, antisimétrica y transitiva. Estas propiedades serán especialmente importantes en §1.3, cuando hablemos de cotas superiores e inferiores.

Las reglas de desigualdad son teoremas. Ya podemos obtener sistemáticamente las reglas que necesitaremos durante todo el libro.

PROPOSICIÓN 1.5. **Leyes básicas de desigualdad en un cuerpo ordenado.** Sean $a, b, c, d \in \mathbb{R}$.

1. **Traslación del orden.** Para todo c,
$$a < b \iff a + c < b + c,$$
y también
$$a \leq b \iff a + c \leq b + c.$$
2. **Suma de desigualdades.** Si $a < b$ y $c < d$, entonces
$$a + c < b + d.$$
La versión correspondiente con $\leq$ también es válida.
3. **Multiplicación por un no negativo.** Si $a \leq b$ y $c \geq 0$, entonces
$$ac \leq bc.$$
Si además $a < b$ y $c > 0$, entonces
$$ac < bc.$$
4. **Multiplicación por un no positivo.** Si $a \leq b$ y $c \leq 0$, entonces
$$ac \geq bc.$$
Si además $a < b$ y $c < 0$, entonces
$$ac > bc.$$
5. **Signo del inverso.** Si $a \neq 0$, entonces a y a^{-1} tienen el mismo signo:
$$a > 0 \iff a^{-1} > 0, \qquad a < 0 \iff a^{-1} < 0.$$
6. **División y orden.** Si $c > 0$, entonces
$$a < b \iff \frac{a}{c} < \frac{b}{c}.$$
Si $c < 0$, entonces
$$a < b \iff \frac{a}{c} > \frac{b}{c}.$$
7. **Orden de los recíprocos positivos.** Si
$$0 < a < b,$$
entonces
$$0 < \frac{1}{b} < \frac{1}{a}.$$

8. **Signo de un producto.** Se tiene
$$ab > 0 \iff (a > 0 \text{ y } b > 0) \text{ o} (a < 0 \text{ y } b < 0),$$
y
$$ab < 0 \iff (a > 0 \text{ y } b < 0) \text{ o} (a < 0 \text{ y } b > 0).$$
9. **Cuadrados.** Para todo $a \in \mathbb{R}$,
$$a^2 \geq 0,$$
y si $a \neq 0$, entonces
$$a^2 > 0.$$
10. **El cuadrado preserva el orden en los no negativos.** Si
$$0 \leq a \leq b,$$
entonces
$$a^2 \leq b^2.$$

Idea de la prueba
Las diez afirmaciones no son reglas independientes. La demostración se apoya repetidamente en cuatro movimientos:

1. traducir $a < b$ a la positividad de $b - a$;
2. usar la clausura de los positivos bajo suma o producto;
3. utilizar las identidades algebraicas ya demostradas en la primera parte de §1.1;
4. traducir nuevamente una positividad en una desigualdad.

El caso de los inversos añade una observación crucial: para dividir una desigualdad necesitamos saber el **signo del divisor**, no solo que sea distinto de cero.

Demostración.

Para la parte 1, comencemos haciendo explícita la identidad algebraica que sostiene la traslación del orden. Por definición de resta y por la regla ya demostrada para el inverso de una suma,

$$\begin{aligned}
(b + c) - (a + c) &= (b + c) + \big(-(a + c)\big) \\
&= (b + c) + \big((-a) + (-c)\big) \\
&= b + \big((-a) + (c + (-c))\big) && \text{(asociatividad y conmutatividad)} \\
&= b + \big((-a) + 0\big) && \text{(inverso aditivo)} \\
&= b + (-a) && \text{(neutro aditivo)} \\
&= b - a && \text{(definición de resta).}
\end{aligned}$$

Por sustitución de iguales por iguales,

$$b - a > 0 \iff (b + c) - (a + c) > 0.$$

Traduciendo ambos extremos mediante la definición del orden, obtenemos $a < b \iff a + c < b + c$. La versión débil también merece hacerse explícita. Supongamos primero $a \leq b$. Si $a = b$, entonces $a + c = b + c$ por sustitución, y por tanto $a + c \leq b + c$. Si $a < b$, la equivalencia estricta recién demostrada da igualmente $a + c < b + c$, luego $a + c \leq b + c$. Así, $a \leq b \Longrightarrow a + c \leq b + c$. Recíprocamente, supongamos $a + c \leq b + c$. Si $a + c = b + c$, la cancelación aditiva ya demostrada da $a = b$. Si $a + c < b + c$, la equivalencia estricta anterior, usada de derecha a izquierda, da $a < b$. En ambos casos, $a \leq b$.

Por consiguiente, $a \leq b \iff a + c \leq b + c$.

Para la parte 2, supongamos primero

$$a < b \qquad \text{y} \qquad c < d.$$

Por la parte 1, podemos trasladar la primera desigualdad sumando c: $a + c < b + c$. Del mismo modo, trasladamos $c < d$ sumando b: $b + c < b + d$. Ahora la transitividad del orden estricto da $a + c < b + c < b + d$, y por tanto

$$\boxed{a + c < b + d.}$$

La versión con orden débil se obtiene sin introducir una regla nueva. Si

$$a \leq b \qquad \text{y} \qquad c \leq d,$$

la parte 1 da $a + c \leq b + c$ y $b + c \leq b + d$.

Como ya demostramos que $\leq$ es transitiva,

$$\boxed{a + c \leq b + d.}$$

Consideremos ahora la parte 3. Supongamos primero

$$a < b \qquad \text{y} \qquad c > 0.$$

Por definición del orden, $b - a > 0$.

Como $c > 0$, la clausura de los positivos bajo el producto, (O2), da $c(b - a) > 0$. Hagamos explícita la expresión que aparece a la izquierda. Por definición de resta, distributividad y las reglas de signos ya demostradas,

$$\begin{aligned} c(b - a) &= c\big(b + (-a)\big) \\ &= cb + c(-a) && \text{(distributividad)} \\ &= cb + (-(ca)) && \text{(regla de signos)} \\ &= bc + (-(ac)) && \text{(conmutatividad)} \\ &= bc - ac && \text{(definición de resta).} \end{aligned}$$

Por sustitución de iguales por iguales, de $c(b - a) > 0$ obtenemos $bc - ac > 0$.

Y como $ac < bc \iff bc - ac > 0$, concluimos

$$\boxed{ac < bc.}$$

Pasemos a la versión no estricta. Supongamos

$$a \leq b \qquad \text{y} \qquad c \geq 0.$$

Por definición,

$$a \leq b \iff a < b \text{ o } a = b,$$

mientras que $c \geq 0 \iff c > 0$ o $c = 0$.

Si $a = b$, la sustitución da $ac = bc$.

Si $c = 0$, entonces Ejemplo 1.1 da

$$ac = a0 = 0 \qquad \text{y} \qquad bc = b0 = 0,$$

de modo que nuevamente $ac = bc$.

En el único caso restante,

$$a < b \qquad \text{y} \qquad c > 0,$$

acabamos de demostrar que $ac < bc$.

Por tanto, en todos los casos,

$$\boxed{ac \leq bc.}$$

Para la parte 4, supongamos primero

$$a < b \qquad \text{y} \qquad c < 0.$$

Por la definición de número negativo, $-c > 0$. Podemos entonces aplicar la parte 3 al factor positivo $-c$: $a(-c) < b(-c)$. Las reglas de signos ya demostradas dan

$$a(-c) = -(ac) \qquad \text{y} \qquad b(-c) = -(bc),$$

de modo que, por sustitución de iguales por iguales, $-ac < -bc$. Ahora trasladamos esta desigualdad sumando $ac + bc$ a ambos miembros. Por la parte 1,

$$(-ac) + (ac + bc) < (-bc) + (ac + bc).$$

Reducimos cada miembro por separado. En el izquierdo,

$$\begin{aligned} (-ac) + (ac + bc) &= \big((-ac) + ac\big) + bc && \text{(asociatividad)} \\ &= 0 + bc && \text{(inverso aditivo)} \\ &= bc && \text{(neutro aditivo).} \end{aligned}$$

En el derecho,

$$\begin{aligned} (-bc) + (ac + bc) &= ac + \big((-bc) + bc\big) && \text{(asociatividad y conmutatividad)} \\ &= ac + 0 && \text{(inverso aditivo)} \\ &= ac && \text{(neutro aditivo).} \end{aligned}$$

Por tanto,

$$\boxed{bc < ac},$$

o, equivalentemente,

$$\boxed{ac > bc}.$$

Pasemos a la versión débil. Supongamos

$$a \leq b \qquad \text{y} \qquad c \leq 0.$$

Por definición,

$$a \leq b \iff a < b \text{ o } a = b,$$

y $c \leq 0 \iff c < 0$ o $c = 0$.

Si $a = b$, la sustitución da $ac = bc$.

Si $c = 0$, entonces Ejemplo 1.1 da

$$ac = a0 = 0 \qquad \text{y} \qquad bc = b0 = 0,$$

y nuevamente $ac = bc$.

En el único caso restante,

$$a < b \qquad \text{y} \qquad c < 0,$$

acabamos de demostrar que $ac > bc$.

Por tanto, en todos los casos,

$$\boxed{ac \geq bc.}$$

Antes de estudiar inversos conviene establecer un hecho pequeño pero decisivo:

$$\boxed{1 > 0.}$$

Como $1 \neq 0$, la tricotomía (O3), aplicada al elemento 1, excluye la alternativa $1 = 0$. Por tanto, exactamente una de las dos afirmaciones

$$1 > 0 \qquad \text{o} \qquad -1 > 0$$

puede ser verdadera. Supongamos, para obtener una contradicción, que $-1 > 0$. Como el producto de dos positivos es positivo, (O2) da $(-1)(-1) > 0$. Por las reglas de signos ya demostradas, $(-1)(-1) = 1$. Sustituyendo iguales por iguales en la desigualdad anterior obtenemos $1 > 0$.

Tendríamos entonces simultáneamente

$$1 > 0 \qquad \text{y} \qquad -1 > 0,$$

lo que contradice la exclusividad de (O3) aplicada a 1. La suposición $-1 > 0$ es, por tanto, imposible. Como una de las dos alternativas debe cumplirse, concluimos

$$\boxed{1 > 0.}$$

Probemos ahora la parte 5. Comencemos con la implicación $a > 0 \Longrightarrow a^{-1} > 0$. Supongamos $a > 0$. En particular, $a \neq 0$, por lo que existe a^{-1}; además, Proposición 1.2 garantiza $a^{-1} \neq 0$. Aplicando (O3) a a^{-1} y excluyendo el caso $a^{-1} = 0$, queda exactamente una de las posibilidades

$$a^{-1} > 0 \qquad \text{o} \qquad -a^{-1} > 0.$$

Supongamos que ocurriera la segunda. Como $a > 0$ y $-a^{-1} > 0$, (O2) implicaría $a(-a^{-1}) > 0$. Ahora reducimos algebraicamente ese producto:

$$\begin{aligned} a(-a^{-1}) &= -(aa^{-1}) && \text{(regla de signos)} \\ &= -1 && \text{(inverso multiplicativo).} \end{aligned}$$

Por sustitución de iguales por iguales obtendríamos $-1 > 0$. Pero ya demostramos $1 > 0$, y (O3) aplicada a 1 prohíbe que 1 y -1 sean positivos simultáneamente. Luego la alternativa $-a^{-1} > 0$ es imposible y necesariamente $a^{-1} > 0$.

Esto prueba $a > 0 \Longrightarrow a^{-1} > 0$.

La dirección recíproca también debe quedar explícita. Supongamos $a^{-1} > 0$. Aplicamos la implicación recién demostrada al elemento a^{-1}. Obtenemos $(a^{-1})^{-1} > 0$. Por Proposición 1.2, $(a^{-1})^{-1} = a$, y por sustitución concluimos $a > 0$.

Por tanto,

$$\boxed{a > 0 \iff a^{-1} > 0.}$$

Pasemos al signo negativo. Supongamos primero $a < 0$. Entonces $a \neq 0$ y, por definición de número negativo, $-a > 0$. La equivalencia positiva que acabamos de demostrar, aplicada a $-a$, da $(-a)^{-1} > 0$. Para identificar este inverso, observemos primero que

$$\begin{aligned}(-a)(-a^{-1}) &= aa^{-1} && \text{(regla de signos)}\\ &= 1 && \text{(inverso multiplicativo).}\end{aligned}$$

Así, $-a^{-1}$ es un inverso multiplicativo de $-a$. Como el inverso multiplicativo es único, $(-a)^{-1} = -a^{-1}$. Sustituyendo esta igualdad en $(-a)^{-1} > 0$, obtenemos $-a^{-1} > 0$. Por definición de número negativo, esto equivale a $a^{-1} < 0$.

Hemos probado entonces $a < 0 \Longrightarrow a^{-1} < 0$.

Para la recíproca, supongamos $a^{-1} < 0$.

Aplicamos la implicación negativa recién demostrada al elemento a^{-1}. Entonces $(a^{-1})^{-1} < 0$.

Usando nuevamente $(a^{-1})^{-1} = a$, concluimos $a < 0$.

Por tanto,

$$\boxed{a < 0 \iff a^{-1} < 0.}$$

Probemos ahora la parte 6. Aquí la división no introduce una regla nueva: por definición, dividir por c significa multiplicar por c^{-1}, y la parte 5 nos permite determinar el signo de ese inverso.

Supongamos primero $c > 0$.

Entonces $c \neq 0$, de modo que los cocientes están definidos, y la parte 5 da $c^{-1} > 0$. Si $a < b$, podemos multiplicar ambos miembros por el número positivo c^{-1}. Por la parte 3, $ac^{-1} < bc^{-1}$.

Por definición de cociente,

$$ac^{-1} = \frac{a}{c} \qquad \text{y} \qquad bc^{-1} = \frac{b}{c},$$

de modo que $\frac{a}{c} < \frac{b}{c}$.

Así hemos probado una dirección: $a < b \Longrightarrow \frac{a}{c} < \frac{b}{c}$.

Para la recíproca, supongamos $\frac{a}{c} < \frac{b}{c}$.

Por definición de cociente, $ac^{-1} < bc^{-1}$.

Como $c > 0$, multiplicar ambos miembros por c conserva el sentido de la desigualdad: $(ac^{-1})c < (bc^{-1})c$. Reducimos ambos miembros usando asociatividad, inverso multiplicativo y neutro:

$$\begin{aligned}(ac^{-1})c &= a(c^{-1}c) && \text{(asociatividad)}\\ &= a1 && \text{(inverso multiplicativo)}\\ &= a && \text{(neutro multiplicativo),}\end{aligned}$$

y, del mismo modo,

$$\begin{aligned}(bc^{-1})c &= b(c^{-1}c)\\ &= b1\\ &= b.\end{aligned}$$

Por sustitución de iguales por iguales obtenemos $a < b$.

Por tanto, cuando $c > 0$,

$$\boxed{a < b \iff \frac{a}{c} < \frac{b}{c}.}$$

Consideremos ahora $c < 0$.

De nuevo $c \neq 0$, y la parte 5 da $c^{-1} < 0$. Si $a < b$, la parte 4 aplicada al factor negativo c^{-1} invierte el orden: $ac^{-1} > bc^{-1}$.

Por definición de cociente, $\frac{a}{c} > \frac{b}{c}$.

Así, $a < b \Longrightarrow \frac{a}{c} > \frac{b}{c}$.

Para demostrar la recíproca, supongamos $\frac{a}{c} > \frac{b}{c}$. Por definición de cociente, esto significa $ac^{-1} > bc^{-1}$, o, escrito con el símbolo $<$, $bc^{-1} < ac^{-1}$. Como $c < 0$, la parte 4 permite multiplicar esta última desigualdad por c e invertir el orden: $(bc^{-1})c > (ac^{-1})c$. Las mismas reducciones algebraicas anteriores dan $b > a$, que equivale a $a < b$.

Por consiguiente, cuando $c < 0$,

$$\boxed{a < b \iff \frac{a}{c} > \frac{b}{c}.}$$

La hipótesis sobre el signo de c cumple, por tanto, dos funciones distintas: garantiza que $c \neq 0$, de modo que la división esté definida, y determina si al multiplicar por c^{-1} el orden se conserva o se invierte.

Para la parte 7, supongamos $0 < a < b$.

En particular,

$$a > 0 \qquad \text{y} \qquad b > 0.$$

Por (O2), el producto de estos dos números positivos también es positivo: $ab > 0$. Así, $ab \neq 0$, existe $(ab)^{-1}$ y, por la parte 5, $(ab)^{-1} > 0$. Podemos entonces multiplicar la desigualdad $a < b$ por la cantidad positiva $(ab)^{-1}$. La parte 3 da $a(ab)^{-1} < b(ab)^{-1}$. Ahora hacemos explícita la simplificación de ambos miembros. Por Proposición 1.2, $(ab)^{-1} = a^{-1}b^{-1}$.

En el miembro izquierdo,

$$\begin{aligned} a(ab)^{-1} &= a(a^{-1}b^{-1}) && \text{(inverso de un producto)} \\ &= (aa^{-1})b^{-1} && \text{(asociatividad)} \\ &= 1b^{-1} && \text{(inverso multiplicativo)} \\ &= b^{-1} && \text{(neutro multiplicativo).} \end{aligned}$$

En el miembro derecho,

$$\begin{aligned} b(ab)^{-1} &= b(a^{-1}b^{-1}) && \text{(inverso de un producto)} \\ &= a^{-1}(bb^{-1}) && \text{(asociatividad y conmutatividad)} \\ &= a^{-1}1 && \text{(inverso multiplicativo)} \\ &= a^{-1} && \text{(neutro multiplicativo).} \end{aligned}$$

Por sustitución de iguales por iguales en la desigualdad anterior obtenemos

$$\boxed{b^{-1} < a^{-1}}.$$

Falta incorporar el extremo izquierdo de la cadena. Como $a > 0$ y $b > 0$, la parte 5 aplicada por separado a ambos números da

$$a^{-1} > 0 \qquad \text{y} \qquad b^{-1} > 0.$$

En particular, $0 < b^{-1} < a^{-1}$.

Finalmente, por definición de cociente,

$$\frac{1}{b} = 1b^{-1} = b^{-1} \qquad \text{y} \qquad \frac{1}{a} = 1a^{-1} = a^{-1}.$$

Por sustitución concluimos

$$\boxed{0 < \frac{1}{b} < \frac{1}{a}}.$$

La inversión del orden no procede de una nueva regla especial para recíprocos: aparece porque hemos multiplicado $a < b$ por el número positivo $(ab)^{-1}$ y después hemos reducido algebraicamente los dos productos resultantes. Probemos la parte 8. Comencemos por la equivalencia para un producto positivo. Supongamos $ab > 0$. En particular, $ab \neq 0$. Si $a = 0$ o $b = 0$, Proposición 1.3 daría $ab = 0$, en contradicción con $ab > 0$. Por tanto,

$$a \neq 0 \qquad \text{y} \qquad b \neq 0.$$

La tricotomía aplicada a a deja entonces exactamente dos posibilidades:

$$a > 0 \qquad \text{o} \qquad a < 0.$$

Como $a \neq 0$, podemos dividir por a. Conviene registrar primero las simplificaciones que utilizaremos. Por definición de cociente,

$$\begin{aligned}\frac{ab}{a} &= (ab)a^{-1} \\ &= b(aa^{-1}) \quad \text{(asociatividad y conmutatividad)} \\ &= b1 \quad \text{(inverso multiplicativo)} \\ &= b \quad \text{(neutro multiplicativo),}\end{aligned}$$

mientras que $\frac{0}{a} = 0a^{-1} = 0$ por definición de cociente y Ejemplo 1.1. Si $a > 0$, de $0 < ab$ y la parte 6 obtenemos $\frac{0}{a} < \frac{ab}{a}$. Las reducciones anteriores dan $0 < b$, es decir, $b > 0$. Si $a < 0$, dividir $0 < ab$ por el número negativo a invierte el sentido: $\frac{0}{a} > \frac{ab}{a}$. Por las mismas reducciones, $0 > b$, es decir, $b < 0$.

Hemos demostrado

$$ab > 0 \Longrightarrow (a > 0 \text{ y } b > 0) \text{ o } (a < 0 \text{ y } b < 0).$$

Probemos la recíproca. Si $a > 0$ y $b > 0$, (O2) da directamente $ab > 0$.

Si $a < 0$ y $b < 0$, entonces

$$-a > 0 \qquad \text{y} \qquad -b > 0.$$

Por (O2), $(-a)(-b) > 0$.

Las reglas de signos ya demostradas dan $(-a)(-b) = ab$. Por sustitución de iguales por iguales, $ab > 0$.

En consecuencia,

$$\boxed{ab > 0 \iff (a > 0 \text{ y } b > 0) \text{ o } (a < 0 \text{ y } b < 0).}$$

Consideremos ahora un producto negativo. Supongamos $ab < 0$. Nuevamente $ab \neq 0$. Si $a = 0$ o $b = 0$, Proposición 1.3 implicaría $ab = 0$, contradicción. Luego

$$a \neq 0 \qquad \text{y} \qquad b \neq 0.$$

La tricotomía aplicada a a deja otra vez los dos casos $a > 0$ y $a < 0$. Si $a > 0$, la parte 6 aplicada a $ab < 0$ permite dividir por a sin cambiar el sentido: $\frac{ab}{a} < \frac{0}{a}$. Usando las reducciones ya establecidas, $b < 0$. Si $a < 0$, dividir por a invierte el sentido: $\frac{ab}{a} > \frac{0}{a}$, de donde $b > 0$.

Por tanto,

$$ab < 0 \Longrightarrow (a > 0 \text{ y } b < 0) \text{ o } (a < 0 \text{ y } b > 0).$$

Falta la recíproca. Supongamos primero

$$a > 0 \qquad \text{y} \qquad b < 0.$$

Entonces $-b > 0$ y, por (O2), $a(-b) > 0$.

Como la regla de signos da $a(-b) = -(ab)$, por sustitución obtenemos $-(ab) > 0$. Por definición de número negativo, esto equivale a $ab < 0$.

Si, en cambio,

$$a < 0 \qquad \text{y} \qquad b > 0,$$

entonces $-a > 0$ y (O2) da $(-a)b > 0$.

La regla de signos $(-a)b = -(ab)$ produce de nuevo $-(ab) > 0$, y por tanto $ab < 0$.

Concluimos

$$\boxed{ab < 0 \iff (a > 0 \text{ y } b < 0) \text{ o } (a < 0 \text{ y } b > 0).}$$

El signo de un producto queda así determinado por una dicotomía estructural: factores con el mismo signo producen un producto positivo y factores con signos opuestos producen un producto negativo. No hemos añadido una nueva regla de signos; la hemos deducido de la tricotomía, de las leyes del orden y de las identidades algebraicas ya demostradas. Para la parte 9 consideremos los tres casos que proporciona la tricotomía. Si $a = 0$, entonces, por Ejemplo 1.1,

$$a^2 = aa = 00 = 0.$$

Supongamos ahora $a > 0$.

Entonces ambos factores de $a^2 = aa$ son positivos, y (O2) da $a^2 > 0$.

Finalmente, supongamos $a < 0$.

Por definición de número negativo, $-a > 0$. Aplicando (O2) a los dos factores positivos $-a$, $(-a)(-a) > 0$. Hagamos explícita la identidad algebraica que permite volver al cuadrado de a. Por definición de cuadrado y por la regla de signos ya demostrada,

$$\begin{aligned} (-a)^2 &= (-a)(-a) \\ &= aa \qquad \text{(producto de dos opuestos)} \\ &= a^2 \qquad \text{(definición de cuadrado).} \end{aligned}$$

Por sustitución de iguales por iguales en $(-a)^2 > 0$, obtenemos $a^2 > 0$. Hemos cubierto las tres posibilidades de la tricotomía. Si $a = 0$, el cuadrado es 0; si $a > 0$ o $a < 0$, el cuadrado es estrictamente positivo. Por tanto, para todo $a \in \mathbb{R}$,

$$\boxed{a^2 \geq 0}.$$

Falta hacer explícito cuándo puede ocurrir la igualdad. Ya vimos que $a = 0 \Longrightarrow a^2 = 0$.

Recíprocamente, supongamos $a^2 = 0$.

Por definición de cuadrado, $aa = 0$.

La ley del producto nulo, Proposición 1.3, afirma que al menos uno de los dos factores debe ser 0. Como ambos factores son el mismo número a, necesariamente $a = 0$.

Así,

$$\boxed{a^2 = 0 \iff a = 0}.$$

En particular, si $a \neq 0$, la tricotomía excluye el primer caso y deja $a > 0$ o $a < 0$; en ambos ya demostramos que

$$\boxed{a^2 > 0}.$$

Finalmente, probemos la parte 10. Supongamos $0 \leq a \leq b$. La hipótesis contiene dos desigualdades:

$$0 \leq a \qquad \text{y} \qquad a \leq b.$$

Por transitividad de $\leq$, $0 \leq b$.

Ahora trasladamos $a \leq b$ sumando $-a$ a ambos miembros. Por la parte 1, $a+(-a) \leq b+(-a)$. Reduciendo ambos miembros mediante inverso aditivo y definición de resta,

$$\boxed{0 \leq b-a}.$$

Para obtener la segunda cantidad no negativa utilizamos la parte 2 en su versión débil. De

$$0 \leq a \qquad \text{y} \qquad 0 \leq b$$

se sigue $0+0 \leq a+b$.

Como $0+0=0$,

$$\boxed{0 \leq a+b}.$$

Tenemos, por tanto, dos factores no negativos:

$$0 \leq b-a \qquad \text{y} \qquad 0 \leq a+b.$$

Aplicamos la parte 3 a la desigualdad $0 \leq b-a$ con el factor no negativo $a+b$. Obtenemos $0(a+b) \leq (b-a)(a+b)$. Por Ejemplo 1.1, $0(a+b)=0$, y así $0 \leq (b-a)(a+b)$. Hagamos explícita ahora la factorización que convierte este producto en una diferencia de cuadrados. Por definición de resta, distributividad, conmutatividad y las reglas de signos,

$$\begin{aligned}
(b-a)(a+b) &= \big(b+(-a)\big)(a+b) \\
&= b(a+b)+(-a)(a+b) && \text{(distributividad)} \\
&= (ba+b^2)+\big((-a)a+(-a)b\big) && \text{(distributividad)} \\
&= (ab+b^2)+\big(-(a^2)+(-(ab))\big) && \text{(conmutatividad y reglas de signos)} \\
&= b^2+\big(ab+(-(ab))\big)+(-(a^2)) && \text{(asociatividad y conmutatividad)} \\
&= b^2+0+(-(a^2)) && \text{(inverso aditivo)} \\
&= b^2-a^2 && \text{(neutro y definición de resta).}
\end{aligned}$$

Por sustitución de iguales por iguales, $0 \leq b^2-a^2$.

Finalmente trasladamos esta desigualdad sumando a^2 a ambos miembros. La parte 1 da $0+a^2 \leq (b^2-a^2)+a^2$.

Reducimos:

$$\begin{aligned}
0+a^2 &= a^2, \\
(b^2-a^2)+a^2 &= \big(b^2+(-(a^2))\big)+a^2 \\
&= b^2+\big(-(a^2)+a^2\big) && \text{(asociatividad)} \\
&= b^2+0 && \text{(inverso aditivo)} \\
&= b^2 && \text{(neutro aditivo).}
\end{aligned}$$

Por tanto,

$$\boxed{a^2 \leq b^2}.$$

Esto demuestra las diez afirmaciones. ■

Después de la prueba
La proposición muestra que hay tres preguntas diferentes antes de «cancelar» un factor en una desigualdad:

1. ¿es el factor distinto de cero?;
2. ¿es positivo?;
3. ¿es negativo?

En una **igualdad**, para cancelar multiplicativamente basta la no nulidad. En una **desigualdad**, la no nulidad no basta: el signo decide si el orden se conserva o se invierte. Si además queremos dividir, la hipótesis de signo cumple simultáneamente dos funciones: garantiza que el divisor no sea 0 y determina qué ocurre con el sentido del orden.

Una pequeña tabla de control. Las reglas anteriores pueden condensarse, una vez demostradas, en la tabla siguiente. Añadimos la columna «Fundamento» para recordar que ninguna fila introduce un axioma nuevo.

Operación aplicada a ambos miembros	Hipótesis	Efecto sobre $<$	Fundamento
sumar c	ninguna	conserva el sentido	parte 1
restar c	ninguna	conserva el sentido	parte 1 aplicada a $-c$
multiplicar por c	$c > 0$	conserva el sentido	parte 3
multiplicar por c	$c < 0$	invierte el sentido	parte 4
dividir por c	$c > 0$	conserva el sentido	partes 5 y 6
dividir por c	$c < 0$	invierte el sentido	partes 5 y 6

La fila de la resta no requiere una regla independiente: restar c significa sumar $-c$, y la parte 1 vale para cualquier elemento del cuerpo, sin hipótesis de signo. También conviene aislar el caso excluido de las filas multiplicativas estrictas. Si $a < b$ y $c = 0$, entonces Ejemplo 1.1 da

$$ac = a0 = 0 \qquad \text{y} \qquad bc = b0 = 0.$$

Por tanto, $ac = bc$, no $ac < bc$ ni $ac > bc$. Una desigualdad estricta colapsa a igualdad al multiplicar ambos miembros por 0. La tabla es, pues, una herramienta de cálculo **derivada**. Su contenido ya está demostrado en Proposición 1.5.

Una regla que necesitaremos al estudiar el hueco racional. Supongamos que a y b son no negativos y que $b \leq a$. Las hipótesis pueden reunirse como $0 \leq b \leq a$. La parte 10 de Proposición 1.5 afirma que, si $0 \leq u \leq v$, entonces $u^2 \leq v^2$. Particularizamos ahora esa afirmación tomando

$$u = b, \qquad v = a.$$

Como sus hipótesis son precisamente $0 \leq b \leq a$, obtenemos

$$\boxed{b^2 \leq a^2}.$$

Así, la regla que utilizaremos después no es una intuición acerca de que «los cuadrados crecen», sino una aplicación directa de la monotonía del cuadrado ya demostrada en los no negativos. Será esencial en §1.4 para comparar los racionales situados a ambos lados de la ecuación $x^2 = 2$.

Esta relectura muestra una ventaja del método axiomático. Podemos auditar una demostración preguntando:

¿qué propiedad estructural autoriza este paso?

En pruebas más largas, esa pregunta ayuda a distinguir una manipulación legítima de una inferencia que solo «parece razonable».

Una desigualdad no es una ecuación: los pasos deben ser reversibles. Consideremos la desigualdad $3x - 7 < 8$.

Antes de transformarla, fijemos una hipótesis que suele quedar implícita. Ya demostramos que $1 > 0$. Por (O1), $2 = 1 + 1 > 0$ y nuevamente $3 = 2 + 1 > 0$. Por tanto, multiplicar o dividir una desigualdad por 3 conserva su sentido.

Primera dirección. Supongamos $3x - 7 < 8$.

Por la parte 1 de Proposición 1.5 podemos sumar 7 a ambos miembros: $(3x - 7) + 7 < 8 + 7$. Reducimos el miembro izquierdo haciendo explícita la resta como suma con inverso:

$$\begin{aligned}(3x - 7) + 7 &= \big(3x + (-7)\big) + 7 \\ &= 3x + \big((-7) + 7\big) && \text{(asociatividad)} \\ &= 3x + 0 && \text{(inverso aditivo)} \\ &= 3x && \text{(neutro aditivo).}\end{aligned}$$

Como $8 + 7 = 15$, obtenemos $3x < 15$.

Ahora usamos la parte 6 de Proposición 1.5. Puesto que $3 > 0$, dividir por 3 conserva el orden: $\frac{3x}{3} < \frac{15}{3}$. Las dos fracciones se reducen mediante la definición de cociente. En el miembro izquierdo,

$$\begin{aligned}\frac{3x}{3} &= (3x)3^{-1} \\ &= x(33^{-1}) && \text{(asociatividad y conmutatividad)} \\ &= x1 && \text{(inverso multiplicativo)} \\ &= x && \text{(neutro multiplicativo).}\end{aligned}$$

Y, como $15 = 3 \cdot 5$,

$$\begin{aligned}\frac{15}{3} &= (3 \cdot 5)3^{-1} \\ &= 5(33^{-1}) && \text{(asociatividad y conmutatividad)} \\ &= 5.\end{aligned}$$

Por sustitución de iguales por iguales concluimos

$$\boxed{x < 5}.$$

Hemos probado hasta aquí solamente $3x - 7 < 8 \Longrightarrow x < 5$.

Dirección recíproca. Supongamos ahora $x < 5$.

Como $3 > 0$, la parte 3 de Proposición 1.5 permite multiplicar ambos miembros por 3 sin invertir el orden: $3x < 3 \cdot 5 = 15$. A continuación aplicamos la parte 1 sumando -7 a ambos miembros: $3x + (-7) < 15 + (-7)$. Por definición de resta y por la aritmética de los enteros,

$$3x + (-7) = 3x - 7 \qquad \text{y} \qquad 15 + (-7) = 8.$$

Por tanto, $3x - 7 < 8$.

Hemos demostrado también $x < 5 \Longrightarrow 3x - 7 < 8$.

Juntando ambas implicaciones,

$$\boxed{3x - 7 < 8 \iff x < 5.}$$

Ahora sí podemos afirmar que el conjunto de soluciones de la desigualdad original es exactamente $(-\infty, 5)$. La diferencia entre una implicación y una equivalencia es esencial: una cadena que solo avanza en un sentido puede producir una **condición necesaria** sin haber caracterizado todavía todas las soluciones. Esta pequeña auditoría anticipa una regla importante para resolver inecuaciones:

una cadena de transformaciones encuentra el conjunto solución solo si controlamos cuáles pasos son reversibles.

Cuando el signo es desconocido, hay que separar casos. El riesgo aumenta cuando multiplicamos por una expresión cuyo signo depende de la incógnita. Consideremos $\frac{2}{x} < 3$. Antes de operar debemos fijar el dominio. Como el denominador no puede ser 0, $x \neq 0$. La tricotomía deja entonces exactamente dos casos admisibles:

$$x > 0 \qquad \text{o} \qquad x < 0.$$

No podemos «multiplicar por x» antes de separar esos casos, porque el signo de x decide si el orden se conserva o se invierte.

Caso 1: $x > 0$. Como el factor es positivo, multiplicar por x es reversible y conserva el sentido de la desigualdad. Por tanto,

$$\frac{2}{x} < 3 \iff 2 < 3x \iff \frac{2}{3} < x,$$

donde en la segunda equivalencia hemos dividido por $3 > 0$. La condición $x > 2/3$ ya implica $x > 0$, de modo que las soluciones de esta rama son exactamente

$$\left(\frac{2}{3}, \infty\right).$$

Caso 2: $x < 0$. Ahora multiplicar por x sigue siendo reversible, pero invierte el orden:

$$\frac{2}{x} < 3 \iff 2 > 3x \iff x < \frac{2}{3}.$$

Dentro de esta rama ya suponemos $x < 0$, y como $2/3 > 0$, todo $x < 0$ satisface automáticamente $x < 2/3$. Por tanto, las soluciones de esta rama son exactamente $(-\infty, 0)$. Las dos ramas agotan el dominio, así que el conjunto solución de la desigualdad original es

$$\boxed{(-\infty, 0) \cup \left(\frac{2}{3}, \infty\right).}$$

El punto $x = 0$ queda excluido desde el comienzo porque la expresión original no está definida allí.

Error frecuente
Antes de multiplicar o dividir una desigualdad por una expresión variable, determine su signo y audite su dominio. Si el signo no está fijado por las hipótesis, separe el dominio en regiones donde sí lo esté. Dentro de cada región use equivalencias reversibles; al final, reúna las soluciones obtenidas mediante una unión.

Intervalos: traducir entre orden y conjuntos. Las desigualdades describen regiones de la recta real. Para evitar repetir expresiones largas, utilizaremos la notación de intervalos.

Definición 1.3. **Intervalos reales.** Sean $a, b \in \mathbb{R}$. Definimos las notaciones siguientes mediante condiciones de pertenencia. No supondremos de entrada que $a \leq b$: cuando los extremos aparezcan en otro orden, las mismas fórmulas determinarán qué conjunto resulta.

El **intervalo abierto** entre a y b es

$$(a, b) = \{x \in \mathbb{R} : a < x < b\}.$$

El **intervalo cerrado** es

$$[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}.$$

Los intervalos **semiabiertos** son

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

y

$$(a, b] = \{x \in \mathbb{R} : a < x \leq b\}.$$

También utilizaremos los intervalos no acotados

$$(a, \infty) = \{x \in \mathbb{R} : x > a\}, \qquad [a, \infty) = \{x \in \mathbb{R} : x \geq a\},$$

$$(-\infty, b) = \{x \in \mathbb{R} : x < b\}, \qquad (-\infty, b] = \{x \in \mathbb{R} : x \leq b\}.$$

Finalmente,

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

Los paréntesis y corchetes codifican pertenencia de los extremos finitos. Así, $2 \in [2, 5)$ pero $5 \notin [2, 5)$. La notación no introduce un nuevo tipo de objeto: es una abreviatura para conjuntos definidos por desigualdades. Por ejemplo, $x \in (-3, 4] \iff -3 < x \leq 4$.

Lectura de la fórmula
La expresión $(-\infty, 5]$ no significa que $-\infty$ sea un número real que actúa como extremo izquierdo. Significa exactamente $\{x \in \mathbb{R} : x \leq 5\}$. Los símbolos ∞ y $-\infty$ funcionan aquí como parte de una notación para describir ausencia de cota en una dirección. **No son números reales.** Por eso nunca «pertenecen» a estos intervalos y la notación usa paréntesis del lado infinito.

Intervalos degenerados y conjuntos vacíos. Como los intervalos se han definido mediante desigualdades, los casos degenerados se deducen de las propias condiciones de pertenencia. Si los extremos coinciden, entonces $[a, a] = \{a\}$, porque $a \leq x \leq a$ obliga a $x = a$. En cambio,

$$(a, a) = [a, a) = (a, a] = \varnothing,$$

pues cada una de esas condiciones exigiría simultáneamente que x estuviera en un lado estricto de a y no más allá del mismo a. Si $a > b$, ninguna de las cuatro condiciones acotadas puede satisfacerse. Por tanto,

$$(a, b) = [a, b] = [a, b) = (a, b] = \varnothing.$$

Esto permite resumir exactamente cuándo estos conjuntos son no vacíos:

$$\begin{aligned}(a, b) \neq \varnothing &\iff a < b,\\ [a, b] \neq \varnothing &\iff a \leq b,\\ [a, b) \neq \varnothing &\iff a < b,\\ (a, b] \neq \varnothing &\iff a < b.\end{aligned}$$

En la primera equivalencia, si $a < b$, basta observar que el punto medio $\frac{a+b}{2}$ satisface $a < (a + b)/2 < b$. En los casos semiabiertos, cuando $a < b$, los propios extremos a o b proporcionan inmediatamente un elemento. Así, los corchetes no «crean» por sí solos un intervalo: las desigualdades que definen la pertenencia deciden si el conjunto contiene un punto, muchos puntos o ninguno.

Antes de seguir.

Antes de seguir
1. Sean $a < b$ y $c < d$. Demuestra, utilizando solo las leyes de orden ya establecidas, que $a - d < b - c$.
Respuesta. De $c < d$, al multiplicar por $-1 < 0$, obtenemos $-d < -c$. Sumando esta desigualdad con $a < b$ mediante la parte 2 de Proposición 1.5, $a + (-d) < b + (-c)$, de donde $a - d < b - c$.

Antes de seguir
2. Supón que $a < b \leq 0$.
¿Qué relación existe entre a^2 y b^2?
Respuesta. Al multiplicar $a < b$ por -1 se invierte el orden: $0 \leq -b < -a$.
La monotonía del cuadrado en los no negativos da $(-b)^2 \leq (-a)^2$, y por las reglas de signos,

$$\boxed{b^2 \leq a^2}.$$

Lo que un cuerpo ordenado todavía no resuelve. Hemos hecho explícitas dos capas distintas de estructura. La primera fue puramente algebraica:

$$\text{axiomas de cuerpo} \longrightarrow \text{reglas de cálculo algebraico.}$$

La segunda añadió positividad y comparación:

axiomas de orden $\longrightarrow$ reglas de desigualdad e intervalos.

Podemos resumir el recorrido de §1.1 así:

$$\boxed{\text{cuerpo} \longrightarrow \text{álgebra derivada} \longrightarrow \text{orden} \longrightarrow \text{desigualdades e intervalos.}}$$

Pero todavía no hemos respondido la pregunta central de este capítulo.

La razón es decisiva:

$$\boxed{\mathbb{Q} \text{ también es un cuerpo ordenado.}}$$

Los racionales satisfacen las mismas leyes algebraicas y de orden que acabamos de imponer a $\mathbb{R}$. Podemos sumar y multiplicar racionales, tomar inversos de racionales no nulos y ordenar racionales de una manera compatible con esas operaciones. En §1.4 demostraremos que, pese a ello, ninguna solución de $x^2 = 2$ pertenece a $\mathbb{Q}$.

Por tanto,

$$\boxed{\text{cuerpo} + \text{orden} \quad \text{no bastan para explicar la propiedad decisiva de } \mathbb{R}.}$$

Para describir esa insuficiencia necesitaremos medir distancias y distinguir entre una cota de un conjunto y su menor cota superior. Ese lenguaje permitirá formular la propiedad adicional de los reales:

$$\boxed{\text{la completitud.}}$$

La secuencia no es accidental. Antes de afirmar que una frontera existe, necesitamos saber con precisión qué significa ser una frontera.

Cambio de régimen: el rigor permanece, la explicación se comprime
En esta primera sección trabajamos deliberadamente **a cámara lenta**. Expandimos asociatividad, conmutatividad, neutros, inversos, sustituciones y reversibilidad para que el lector aprendiera a auditar una prueba y no tratara las reglas escolares como movimientos sin fundamento. A partir de aquí cambia la granularidad. Una vez demostrada una regla, podremos citarla y utilizarla sin reconstruir cada vez toda su genealogía. No volveremos a desarrollar de rutina pasos como $x + 0 = x$, $aa^{-1} = 1$, una reagrupación asociativa o una sustitución de iguales por iguales.
Sí volveremos a abrir una cadena cuando el paso contenga información matemática nueva o una hipótesis fácil de perder: **dominio y no nulidad, signo de un factor, reversibilidad de una equivalencia, separación de casos, orden de cuantificadores, existencia o unicidad, aplicación de completitud, o diseño de una construcción no evidente**.
El rigor no disminuye. Lo que disminuye es el andamiaje visible: el objetivo es que el lector empiece a cargar por sí mismo con las justificaciones que ya aprendió a reconocer.

1.2. Valor absoluto, distancia y desigualdades

Del orden a la distancia. El orden permite decidir cuál de dos números es mayor. Para describir aproximaciones necesitamos además medir cuánto los separa: la distancia de x a 0 y, más generalmente, la distancia de x a un número fijo a. El **valor absoluto** convertirá esa idea de cercanía en desigualdades con las que podamos calcular.

Definición 1.4. **Valor absoluto.** Para $x \in \mathbb{R}$, definimos

$$|x| = \begin{cases} x, & x \geq 0, \\ -x, & x < 0. \end{cases}$$

La definición está bien determinada porque, por tricotomía, todo real satisface exactamente una de las condiciones $x \geq 0$ o $x < 0$. En particular, $x = 0$ pertenece únicamente a la primera rama. El número $|x|$ representa la distancia de x al origen sobre la recta real.

La definición por casos puede parecer puramente algebraica, pero su significado es geométrico. Por ejemplo,

$$|5| = 5, \qquad |-5| = 5,$$

porque los puntos 5 y -5 se encuentran a la misma distancia del origen. Esta observación explica por qué el valor absoluto nunca es negativo: una distancia no puede serlo. Pero, como siempre, la intuición geométrica debe poder traducirse a las reglas del cuerpo ordenado.

Propiedades básicas que no conviene memorizar aisladas.

Proposición 1.6. Para cualesquiera $x, y \in \mathbb{R}$ se cumplen las siguientes propiedades:

1. $|x| \geq 0$, y $|x| = 0$ si y solo si $x = 0$;
2. $|-x| = |x|$;
3. $-|x| \leq x \leq |x|$;
4. $|xy| = |x|\,|y|$;
5. si $a \geq 0$, entonces
$$|x| \leq a \iff -a \leq x \leq a;$$
6. si $a > 0$, entonces
$$|x| < a \iff -a < x < a.$$

Demostración. Las tres primeras afirmaciones se obtienen directamente de la definición por casos. Si $x \geq 0$, entonces $|x| = x$; si $x < 0$, entonces $|x| = -x > 0$. De aquí se sigue $|x| \geq 0$, y la igualdad ocurre exactamente cuando $x = 0$. Aplicando la misma definición a $-x$ se obtiene $|-x| = |x|$, y en ambos casos $-|x| \leq x \leq |x|$.

Para la propiedad del producto controlamos los signos. Si $x, y \geq 0$, entonces $|xy| = xy = |x|\,|y|$.

Si $x, y < 0$, entonces $xy > 0$ y

$$|xy| = xy = (-x)(-y) = |x|\,|y|.$$

Si uno es no negativo y el otro negativo, entonces $xy \leq 0$ y aparece exactamente una negación. Por ejemplo, si $x \geq 0$ e $y < 0$,

$$|xy| = -xy = x(-y) = |x|\,|y|.$$

El caso restante es simétrico. Por tanto, $|xy| = |x|\,|y|$.

Consideremos ahora $a \geq 0$. Supongamos primero $|x| \leq a$.

Como ya sabemos que $-|x| \leq x \leq |x|$, la desigualdad $|x| \leq a$ da inmediatamente $x \leq |x| \leq a$. Al multiplicar $|x| \leq a$ por -1, el orden se invierte: $-a \leq -|x|$.

Junto con $-|x| \leq x$, obtenemos $-a \leq x$.

Por tanto, $-a \leq x \leq a$.

Recíprocamente, supongamos $-a \leq x \leq a$.

Si $x \geq 0$, entonces $|x| = x \leq a$. Si $x < 0$, de $-a \leq x$ se sigue, al multiplicar por -1, que $-x \leq a$; como $|x| = -x$, también $|x| \leq a$. Así,

$$|x| \leq a \iff -a \leq x \leq a.$$

Para la versión estricta supongamos $a > 0$. Si $|x| < a$, entonces $x \leq |x| < a$, de modo que $x < a$. Además, al negar $|x| < a$ obtenemos $-a < -|x|$, y como $-|x| \leq x$, $-a < x$.

Luego $-a < x < a$.

Recíprocamente, si $-a < x < a$, entonces, si $x \geq 0$, tenemos $|x| = x < a$; y si $x < 0$, de $-a < x$ se sigue $-x < a$, por lo que $|x| = -x < a$. En consecuencia,

$$|x| < a \iff -a < x < a.$$

■

Lectura de la fórmula
La equivalencia

$$|x| < a \iff -a < x < a$$

no es un truco de resolución de inequaciones. Dice literalmente:

estar a distancia menor que a del origen equivale a encontrarse entre $-a$ y a.

La desigualdad algebraica y la descripción geométrica son dos lenguajes para la misma región de la recta.

Distancia entre dos puntos. Si $|x|$ mide la distancia de x al origen, para medir la distancia entre dos puntos x e y basta trasladar uno de ellos al origen. La diferencia $x - y$ indica cuánto debemos desplazarnos desde y para llegar a x, y su valor absoluto elimina la orientación del desplazamiento. Definimos entonces, en la recta real, $d(x, y) := |x - y|$. Antes de utilizar esta notación como una verdadera noción de distancia, registremos las propiedades que ya podemos justificar a partir de Proposición 1.6. Para cualesquiera $x, y \in \mathbb{R}$, $d(x, y) \geq 0$.

Además,

$$\begin{aligned} d(x,y) = 0 &\iff |x - y| = 0 \\ &\iff x - y = 0 \\ &\iff x = y. \end{aligned}$$

Por último,

$$d(x, y) = |x - y| = |-(y - x)| = |y - x| = d(y, x).$$

Tenemos así no negatividad, separación de puntos y simetría. Falta una propiedad más profunda: la distancia directa de x a z no debería superar la longitud de un recorrido que pasa por un punto intermedio y. Esa propiedad será consecuencia de la desigualdad triangular para el valor absoluto.

La desigualdad triangular.

Teorema 1.2. **Desigualdad triangular.** Para cualesquiera $x, y \in \mathbb{R}$,

$$|x+y| \leq |x| + |y|.$$

Idea de la prueba
Ya sabemos que cada número queda atrapado entre el negativo y el positivo de su valor absoluto:

$$-|x| \leq x \leq |x|, \qquad -|y| \leq y \leq |y|.$$

La estrategia consiste en obtener por separado una cota inferior y una cota superior para $x+y$, reunirlas en una doble desigualdad y aplicar después la caracterización de $|\cdot|$ demostrada en Proposición 1.6.

Demostración. De

$$-|x| \leq x \qquad \text{y} \qquad -|y| \leq y$$

se obtiene, al sumar, $-(|x| + |y|) \leq x + y$.
Por otra parte, de

$$x \leq |x| \qquad \text{y} \qquad y \leq |y|$$

se obtiene $x + y \leq |x| + |y|$.
Reuniendo ambas cotas, $-(|x| + |y|) \leq x + y \leq |x| + |y|$.
Como $|x| + |y| \geq 0$, la caracterización no estricta de Proposición 1.6 da

$$\boxed{|x+y| \leq |x| + |y|}.$$

■

Después de la prueba
La demostración tiene una arquitectura reutilizable:

1. obtener cotas ordinarias para una expresión;
2. reunirlas en una doble desigualdad simétrica;
3. volver a empaquetarlas como una desigualdad de valor absoluto.

Este movimiento entre **orden**, **valor absoluto** y **distancia** reaparecerá constantemente en cálculo.

Aplicamos ahora el teorema a

$$u = a - c, \qquad v = c - b.$$

Como $u + v = a - b$, $|a - b| \leq |a - c| + |c - b|$.
En términos de la función d,

$$\boxed{d(a,b) \leq d(a,c) + d(c,b).}$$

Por tanto, $d(x, y) = |x - y|$ posee las cuatro propiedades fundamentales de una distancia sobre $\mathbb{R}$: no negatividad, separación de puntos, simetría y desigualdad triangular. No necesitamos

todavía desarrollar la teoría general de espacios métricos; bastará reutilizar estas propiedades cuando aparezca la noción de cercanía.

La desigualdad triangular inversa. La desigualdad triangular también permite controlar cuánto pueden cambiar las distancias al comparar dos puntos.

COROLARIO 1.1. **Desigualdad triangular inversa.** Para cualesquiera $x, y \in \mathbb{R}$,

$$||x| - |y|| \leq |x - y|.$$

Demostración. Escribamos $x = (x - y) + y$.

Por la desigualdad triangular, $|x| \leq |x - y| + |y|$, y por tanto $|x| - |y| \leq |x - y|$.

Intercambiando x e y obtenemos $|y| - |x| \leq |y - x| = |x - y|$.

Para convertir esta segunda desigualdad en una cota inferior de $|x| - |y|$, multiplicamos por -1 e invertimos el orden: $-|x-y| \leq |x| - |y|$. Junto con la primera cota, $-|x-y| \leq |x| - |y| \leq |x-y|$. Como $|x-y| \geq 0$, la caracterización no estricta de Proposición 1.6, aplicada a la cantidad $|x| - |y|$, da

$$\boxed{||x| - |y|| \leq |x - y|.}$$

■

Lectura métrica
Como

$$d(x, 0) = |x|, \qquad d(y, 0) = |y|,$$

la desigualdad triangular inversa puede escribirse como $\big|d(x, 0) - d(y, 0)\big| \leq d(x, y)$. Más generalmente, fijado cualquier $c \in \mathbb{R}$, la misma desigualdad aplicada a $x - c$ e $y - c$ produce

$$\boxed{\big|d(x, c) - d(y, c)\big| \leq d(x, y).}$$

En efecto, $(x - c) - (y - c) = x - y$.
Así, cambiar el punto desde x hasta y no puede modificar su distancia a un mismo punto de referencia c en una cantidad mayor que la propia distancia entre x e y.

Una desigualdad centrada en un punto. En cálculo rara vez nos interesará solamente la distancia al origen. Mucho más frecuente será medir la distancia respecto de un número fijo a. Si $r > 0$, la condición $|x - a| < r$ significa que la distancia de x a a es menor que r. La caracterización de Proposición 1.6 permite traducirla directamente:

$$\begin{aligned} |x - a| < r &\iff -r < x - a < r \\ &\iff a - r < x < a + r. \end{aligned}$$

Por tanto,

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

La misma traducción con orden débil describe la banda cerrada alrededor de a.

PROPOSICIÓN 1.7. Si $a \in \mathbb{R}$ y $r > 0$, entonces

$$|x-a|<r \iff a-r<x<a+r \iff x\in(a-r,a+r),$$

mientras que

$$|x-a|\le r \iff a-r\le x\le a+r \iff x\in[a-r,a+r].$$

Fuera de esas bandas se tiene

$$|x-a|>r \iff x<a-r \text{ o } x>a+r \iff x\in(-\infty,a-r)\cup(a+r,\infty),$$

y

$$|x-a|\ge r \iff x\le a-r \text{ o } x\ge a+r \iff x\in(-\infty,a-r]\cup[a+r,\infty).$$

Las dos formas exteriores merecen una justificación porque introducen una disyunción. Como el orden de $\mathbb{R}$ es total, $|x-a|>r \iff \neg(|x-a|\le r)$. La equivalencia para la banda cerrada transforma esto en $\neg(a-r\le x\le a+r)$. Para que falle esa doble desigualdad debe fallar al menos una de sus dos condiciones. Por tricotomía, $\neg(a-r\le x) \iff x<a-r$, y $\neg(x\le a+r) \iff x>a+r$.

Por tanto,

$$\boxed{|x-a|>r \iff x<a-r \text{ o } x>a+r.}$$

De manera análoga,

$$|x-a|\ge r \iff \neg(|x-a|<r) \iff \neg(a-r<x<a+r),$$

y la negación de la banda abierta da

$$\boxed{|x-a|\ge r \iff x\le a-r \text{ o } x\ge a+r.}$$

La diferencia entre $>$ y $\ge$ aparece exactamente en los dos puntos frontera $a-r$ y $a+r$: pertenecen a la región exterior cerrada, pero no a la exterior estricta.

Error frecuente: olvidar el centro
De $|x-a|<r$ no se sigue $-r<x<r$ salvo que $a=0$. El intervalo está centrado en a, no en el origen: $(a-r,a+r)$.

Dos maneras de leer una misma inequación.

Ejemplo 1.3. Resolvamos $|2x-3|<5$.

Podemos hacerlo algebraicamente o geométricamente.

Lectura algebraica. Por la equivalencia para el valor absoluto, $-5<2x-3<5$.

Sumando 3, $-2<2x<8$, y dividiendo por el número positivo 2, $-1<x<4$.

Por tanto, $x\in(-1,4)$.

Lectura geométrica. Como

$$2x-3=2\left(x-\frac{3}{2}\right),$$

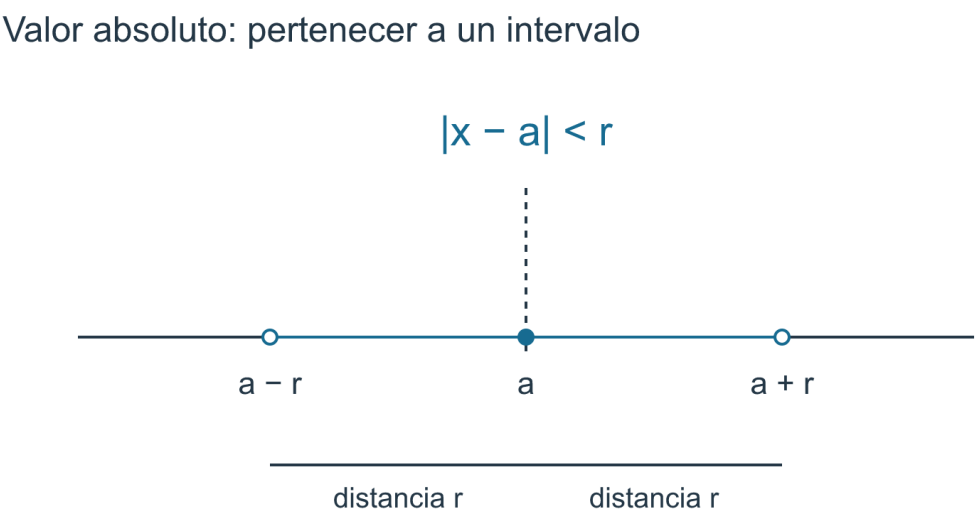


FIGURA 1.1. La condición $|x-a|<r$ describe el intervalo abierto $(a-r, a+r)$: cada punto está a distancia menor que r del centro a.

la multiplicatividad del valor absoluto da

$$|2x-3| = 2\left|x-\frac{3}{2}\right|.$$

Puesto que $2>0$,

$$|2x-3|<5 \iff \left|x-\frac{3}{2}\right|<\frac{5}{2}.$$

Así, x debe encontrarse a distancia menor que $5/2$ del centro $3/2$. El intervalo correspondiente es

$$\left(\frac{3}{2}-\frac{5}{2}, \frac{3}{2}+\frac{5}{2}\right) = (-1,4).$$

Las dos lecturas caracterizan la misma región mediante equivalencias reversibles: una en lenguaje de orden y la otra en lenguaje de distancia.

Estimar no significa calcular exactamente. El valor absoluto será también nuestra herramienta básica para **estimar** cantidades. Supongamos que conocemos una aproximación de x a un punto a, con $r>0$, en el sentido de que $|x-a|<r$. Quizá no sepamos el valor exacto de x, pero podemos controlar su tamaño. Como $x=(x-a)+a$, la desigualdad triangular y la hipótesis dan $|x| \le |x-a|+|a| < r+|a|$.

Así obtenemos

$$\boxed{|x-a|<r \quad \Longrightarrow \quad |x|<|a|+r.}$$

Podemos decir algo más preciso utilizando la desigualdad triangular inversa: $||x| - |a|| \leq |x - a| < r$. Por la caracterización estricta del valor absoluto, $-r < |x| - |a| < r$, y por tanto

$$\boxed{|a| - r < |x| < |a| + r.}$$

La cota superior recupera la estimación anterior; la inferior muestra que, si x está a distancia menor que r de a, su magnitud tampoco puede disminuir respecto de $|a|$ en una cantidad igual o mayor que r. Este pequeño patrón es una de las técnicas fundamentales del cálculo:

$$\boxed{\text{cantidad desconocida} = \text{error respecto de una referencia} + \text{referencia conocida.}}$$

Después aplicamos desigualdades de valor absoluto para transformar esa descomposición en cotas. No estamos hablando todavía de límites. Pero cuando más adelante aparezcan expresiones como $|x - a| < \delta$ o $|u_n - L| < \varepsilon$, su significado geométrico ya no deberá aprenderse de nuevo: serán afirmaciones acerca de **distancias**.

Acumular errores y separar regiones. Si

$$|x - a| < r, \qquad |y - b| < s,$$

con $r, s > 0$, la suma de los errores queda controlada por $|(x + y) - (a + b)| < r + s$. Para obtener la estimación, escribimos

$$(x + y) - (a + b) = (x - a) + (y - b),$$

y aplicamos la desigualdad triangular:

$$|(x + y) - (a + b)| \leq |x - a| + |y - b| < r + s.$$

La misma definición por casos permite resolver la inequación $|x - 1| + |x + 1| \leq 4$. Los puntos -1 y 1 son los lugares donde cambian los signos de los argumentos de los valores absolutos, así que separamos tres regiones.

Si $x \leq -1$, entonces

$$|x - 1| + |x + 1| = (1 - x) + (-x - 1) = -2x.$$

La condición $-2x \leq 4$ equivale a $x \geq -2$, de modo que esta rama aporta $[-2, -1]$.

Si $-1 \leq x \leq 1$, entonces

$$|x - 1| + |x + 1| = (1 - x) + (x + 1) = 2,$$

por lo que todos los puntos de esta región satisfacen la inequación. Esta rama aporta $[-1, 1]$.

Si $x \geq 1$, entonces

$$|x - 1| + |x + 1| = (x - 1) + (x + 1) = 2x.$$

La condición $2x \leq 4$ equivale a $x \leq 2$, así que esta rama aporta $[1, 2]$. Las tres regiones cubren toda la recta real y sus soluciones se reúnen mediante una unión: $[-2, -1] \cup [-1, 1] \cup [1, 2] = [-2, 2]$.

Por tanto,

$$\boxed{x \in [-2, 2]}.$$

Qué hemos ganado. El valor absoluto comenzó como una definición por casos y terminó organizando tres lenguajes que utilizaremos constantemente:

$$\boxed{\text{orden} \longleftrightarrow \text{valor absoluto} \longleftrightarrow \text{distancia.}}$$

Si $r > 0$, la definición de distancia y Proposición 1.7 condensan la traducción fundamental en una sola cadena:

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

Así, una misma afirmación puede leerse como desigualdad de valor absoluto, condición de distancia, doble desigualdad o pertenencia a un intervalo abierto. También disponemos de las dos estimaciones fundamentales $|x+y| \leq |x| + |y|$ y $\big||x| - |y|\big| \leq |x-y|$. La desigualdad triangular controla cómo se acumulan tamaños o errores; la desigualdad triangular inversa controla cuánto puede cambiar una magnitud —o una distancia a un punto fijo— cuando cambia el punto. Las distancias comparan puntos; una frontera de un conjunto exige comparar un mismo candidato con todos sus elementos. Esa exigencia distinguirá al **máximo**, que pertenece al conjunto, del **supremo**, que puede quedar fuera de él. La ecuación $x^2 = 2$ hará necesario examinar esta segunda posibilidad.

1.3. Cotas, máximos, mínimos, supremos e ínfimos

De estar cerca a tener una frontera. En §1.2 aprendimos a medir cercanía. Podemos decir que un punto x está a distancia menor que r de a escribiendo $|x-a| < r$. El cambio de pregunta es también un cambio de escala. La condición anterior compara un punto con otro; hablar de una frontera para un conjunto exigirá comparar un mismo candidato con **todos** los elementos del conjunto a la vez.

En el conjunto de números cuyo cuadrado es menor que 2, por ejemplo, podemos buscar un número que domine a todos sus elementos. Tal número sería una cota superior, aunque quizá otra cota más pequeña cumpliera la misma función. Preguntar por la menor de ellas es distinto de buscar el mayor elemento del conjunto: esa cota mínima puede existir sin pertenecer al conjunto que acota.

Cotas: barreras que no tienen por qué pertenecer al conjunto.

Definición 1.5. **Cotas y acotación.** Sea $A \subseteq \mathbb{R}$.
Un número $M \in \mathbb{R}$ es una **cota superior** de A si

$$\forall a \in A, \qquad a \leq M.$$

Un número $m \in \mathbb{R}$ es una **cota inferior** de A si

$$\forall a \in A, \qquad m \leq a.$$

Decimos que A está **acotado superiormente** si posee al menos una cota superior, y **acotado inferiormente** si posee al menos una cota inferior. Decimos que A está **acotado** si está acotado en ambos sentidos.

La palabra decisiva es **barrera**. Una cota superior M debe encontrarse a la derecha de todo el conjunto, pero no tiene por qué tocarlo ni pertenecer a él. El orden de los cuantificadores importa: primero se fija el candidato M y después debe verificarse la desigualdad $a \leq M$ para **todo** $a \in A$. Por negación,

$$M \text{ no es cota superior de } A \iff \exists a \in A \text{ tal que } a > M.$$

De manera dual,

$$m \text{ no es cota inferior de } A \iff \exists a \in A \text{ tal que } a < m.$$

Por ejemplo, para $A = (0,1)$, los números $1,\ 2,\ 10,\ 10^6$ son todos cotas superiores. También lo es cualquier real $M \geq 1$.

En cambio, 0.9 no es una cota superior. Basta exhibir un elemento de $(0,1)$ que la supere; por ejemplo,

$$\frac{19}{20} \in (0,1) \qquad \text{y} \qquad \frac{19}{20} > \frac{9}{10} = 0.9.$$

Del mismo modo, cualquier real $m \leq 0$ es una cota inferior de $(0,1)$.

Error frecuente
Una cota superior **no tiene que pertenecer al conjunto**. En $A = (0,1)$, el número 1 es una cota superior aunque $1 \notin A$. Confundir «ser una cota» con «ser un elemento» hará imposible comprender correctamente el supremo.

Máximo y mínimo: extremos que sí pertenecen al conjunto. La noción cambia cuando exigimos que la barrera sea además un elemento del propio conjunto.

Definición 1.6. **Máximo y mínimo.** Sea $A \subseteq \mathbb{R}$.

Un elemento $M \in A$ es el **máximo** de A si

$$\forall a \in A, \qquad a \leq M.$$

Un elemento $m \in A$ es el **mínimo** de A si

$$\forall a \in A, \qquad m \leq a.$$

Cuando existen, escribimos

$$M = \max A, \qquad m = \min A.$$

Así, un máximo cumple simultáneamente dos condiciones:

$$\boxed{M \in A \qquad \text{y} \qquad M \text{ es una cota superior de } A.}$$

La primera condición es precisamente la que una cota superior arbitraria no necesita satisfacer. Máximo y mínimo, cuando existen, son únicos. En efecto, si M y N fueran dos máximos de A, entonces $M, N \in A$; como M es máximo, $N \leq M$, y como N es máximo, $M \leq N$. Por antisimetría, $M = N$. El argumento para el mínimo es dual. Por eso las notaciones $\max A$ y $\min A$ designan, cuando existen, números determinados.

Ejemplo 1.4. **Una cota superior que no es máximo.** Consideremos $A = (0,1)$.

El número 1 es una cota superior de A, pero A no tiene máximo.

¿Por qué no hay máximo? Tomemos un elemento cualquiera $a \in (0,1)$. Como $a < 1$, al sumar a y después 1 obtenemos $2a < a + 1 < 2$.

Dividiendo por $2 > 0$, $a < \frac{a+1}{2} < 1$.

Definamos $b = \frac{a+1}{2}$.

Entonces $b \in A$ y $b > a$.

Como el punto $a \in A$ fue arbitrario, hemos probado

$$\forall a \in A \; \exists b \in A \qquad a < b.$$

Por tanto, ningún elemento de A puede ser máximo: desde cualquiera de ellos podemos construir otro elemento permitido situado más a la derecha.

Observa la diferencia:

$$\boxed{1 \text{ está por encima de todo } A, \quad \textit{pero} \quad 1 \notin A.}$$

Por eso 1 puede ser una cota superior sin ser máximo. El argumento muestra una técnica general: para negar que un conjunto tenga máximo, tomamos un elemento arbitrario y construimos otro elemento permitido que lo supera. En §1.4 aplicaremos una técnica semejante, menos inmediata, al conjunto de racionales positivos cuyo cuadrado es menor que 2.

Una cota superior especial: la menor de todas. En $(0,1)$ existen infinitas cotas superiores. Sin embargo, una de ellas ocupa una posición privilegiada: 1. No porque pertenezca al conjunto —no pertenece—, sino porque no existe una cota superior más pequeña. Esta es la idea de **supremo**.

Definición 1.7. **Supremo e ínfimo.** Sea $A \subseteq \mathbb{R}$.

Un número $s \in \mathbb{R}$ es el **supremo** de A si:

1. s es una cota superior de A;
2. si u es cualquier cota superior de A, entonces
$$s \leq u.$$

Es decir, s es la **menor cota superior** de A. Cuando existe, escribimos

$$s = \sup A.$$

De manera dual, un número $i \in \mathbb{R}$ es el **ínfimo** de A si:

1. i es una cota inferior de A;
2. si ℓ es cualquier cota inferior de A, entonces
$$\ell \leq i.$$

Es decir, i es la **mayor cota inferior** de A. Cuando existe, escribimos

$$i = \text{ínf } A.$$

Conviene leer lentamente estas definiciones. En la definición de supremo intervienen dos comparaciones universales distintas. La primera dice

$$\forall a \in A, \qquad a \leq s,$$

y compara el candidato s con los **elementos del conjunto**. La segunda puede escribirse

$$\forall u \in \mathbb{R}, \qquad \left[(\forall a \in A,\ a \le u) \Longrightarrow s \le u\right],$$

y compara s con **todas las cotas superiores** posibles. Por tanto, demostrar que $s = \sup A$ exige dos trabajos diferentes:

$$\boxed{s \text{ domina a todos los elementos de } A \quad + \quad \text{toda cota superior domina a } s.}$$

La lectura del ínfimo es exactamente dual:

$$\forall a \in A, \qquad i \le a,$$

y, para toda cota inferior ℓ, $\ell \le i$.

El supremo no es «el número más grande del conjunto». Esa frase describe, cuando existe, al **máximo**.

El supremo es

$$\boxed{\text{la más pequeña entre todas las barreras superiores.}}$$

El ínfimo es

$$\boxed{\text{la más grande entre todas las barreras inferiores.}}$$

Esta diferencia de una palabra —«elemento» frente a «cota»— cambia toda la teoría.

Máximo frente a supremo. Volvamos a dos intervalos muy parecidos:

$$A = (0, 1), \qquad B = (0, 1].$$

Ya demostramos que A no tiene máximo. En cambio, $\max B = 1$, porque $1 \in B$ y todo $b \in B$ satisface $b \le 1$. La relación entre máximo y supremo puede formularse ahora sin ambigüedad. Si A tiene máximo M, entonces M es una cota superior de A. Además, si u es cualquier otra cota superior, como $M \in A$ debe cumplirse $M \le u$. Por tanto, M es la menor cota superior y

$$\boxed{\sup A = \max A = M.}$$

Recíprocamente, supongamos que $s = \sup A$ existe y que además $s \in A$. Como s es una cota superior,

$$\forall a \in A, \qquad a \le s.$$

La pertenencia $s \in A$ y esta desigualdad universal son precisamente las dos condiciones que definen al máximo. Luego

$$\boxed{\max A = s = \sup A.}$$

En consecuencia, siempre que $\sup A$ exista,

$$\boxed{A \text{ tiene máximo } \iff \sup A \in A,}$$

y, en ese caso, $\max A = \sup A$.

La afirmación dual es igualmente válida: si $\inf A$ existe, entonces A tiene mínimo $\iff \inf A \in A$, y, cuando esto ocurre, $\min A = \inf A$. Aplicado a $B = (0, 1]$, el máximo ya identificado nos

permite concluir inmediatamente $\sup B = \max B = 1$. Para $A = (0,1)$ sabemos, en cambio, que no hay máximo. Todavía falta justificar que 1 es efectivamente su supremo: esa demostración será el objeto del microtramo siguiente.

Por qué 1 es realmente el supremo de $(0,1)$. Probemos ahora las dos cláusulas de la definición de supremo.

1. El número 1 es una cota superior. Sea $x \in (0,1)$. Por definición del intervalo, $0 < x < 1$.

En particular, $x \leq 1$.

Como x fue arbitrario,

$$\forall x \in (0,1), \qquad x \leq 1.$$

Por tanto, 1 es una cota superior de $(0,1)$.

2. Toda cota superior está por encima de 1. Sea u una cota superior cualquiera de $(0,1)$. Como $\frac{1}{2} \in (0,1)$, la definición de cota superior obliga a que $\frac{1}{2} \leq u$.

Supongamos, para obtener una contradicción, que $u < 1$.

Definamos $b = \frac{u+1}{2}$.

De $u < 1$, sumando primero u y después 1, obtenemos $2u < u + 1 < 2$. Como $2 > 0$, al dividir por 2 resulta $u < b < 1$. Además, $u \geq 1/2 > 0$, de modo que $b > u > 0$. Por tanto, $b \in (0,1)$. Pero $b > u$, lo que contradice que u sea una cota superior de $(0,1)$. La suposición $u < 1$ es imposible. Por tricotomía, $1 \leq u$. Como u era una cota superior arbitraria, toda cota superior de $(0,1)$ domina a 1. Junto con la primera parte, esto demuestra

$$\boxed{\sup(0,1) = 1.}$$

El argumento inferior es dual, pero podemos dejarlo también explícito. El número 0 es una cota inferior porque $0 < x$ para todo $x \in (0,1)$. Sea ahora ℓ cualquier cota inferior. Como $1/2 \in (0,1)$, $\ell \leq \frac{1}{2}$. Si supusiéramos $\ell > 0$, el número $c = \frac{\ell}{2}$ satisfaría $0 < c < \ell$ y, como $\ell \leq 1/2 < 1$, $c < 1$. Así $c \in (0,1)$, pero $c < \ell$, contradiciendo que ℓ fuera una cota inferior. Por tanto, toda cota inferior satisface $\ell \leq 0$.

Concluimos

$$\boxed{\text{ínf}(0,1) = 0.}$$

En particular, este intervalo tiene supremo e ínfimo aunque ninguno de los dos pertenezca al conjunto.

El supremo es único. La definición de supremo contiene ya su unicidad. Supongamos que s y t son ambos supremos de un mismo conjunto A. Como t es un supremo, en particular es una cota superior de A. Pero s es la menor cota superior, así que $s \leq t$. Ahora intercambiamos los papeles. Como s es una cota superior de A y t es la menor cota superior, $t \leq s$.

Tenemos, por tanto,

$$s \leq t \qquad \text{y} \qquad t \leq s.$$

Por antisimetría del orden,

$$\boxed{s = t}.$$

Así, un conjunto no puede tener dos supremos distintos: **si el supremo existe, es único**. La demostración para el ínfimo es dual. Si i y j fueran ambos ínfimos de A, entonces j sería una cota inferior y la maximalidad de i daría $j \leq i$; intercambiando los papeles obtendríamos $i \leq j$. Por antisimetría,

$$\boxed{i = j}.$$

En consecuencia,

$$\boxed{\text{si } \sup A \text{ existe, es único;} \qquad \text{si } \operatorname{ínf} A \text{ existe, es único.}}$$

Conviene separar esta afirmación de una cuestión distinta: aquí hemos demostrado **unicidad**, no **existencia**. El argumento dice que puede haber a lo sumo un número con la propiedad de ser supremo o ínfimo; todavía no hemos probado que todo conjunto apropiado posea uno.

Por eso las expresiones

$$\sup A, \qquad \operatorname{ínf} A$$

tienen sentido como números determinados una vez que su existencia ha sido establecida.

Una caracterización operativa: acercarse tanto como queramos. La definición de supremo compara s con **todas las cotas superiores**. Esa formulación es conceptualmente exacta, pero en las demostraciones necesitaremos una versión más operativa. Si s es el supremo, ningún intervalo inmediatamente situado por debajo de s puede quedar completamente vacío de elementos de A. La expresión «acercarnos tanto como queramos» puede formularse sin límites, usando solo desigualdades y cuantificadores.

Proposición 1.8. **Caracterización aproximativa del supremo y del ínfimo.** Sea $A \subseteq \mathbb{R}$ no vacío.

Un número $s \in \mathbb{R}$ satisface

$$s = \sup A$$

si y solo si se cumplen las dos condiciones siguientes:

1. s es una cota superior de A;
2. para todo $\varepsilon > 0$ existe $a \in A$ tal que
$$s - \varepsilon < a \leq s.$$

De manera dual,

$$i = \operatorname{ínf} A$$

si y solo si:

1. i es una cota inferior de A;
2. para todo $\varepsilon > 0$ existe $a \in A$ tal que
$$i \leq a < i + \varepsilon.$$

En la segunda condición para el supremo, la desigualdad $a \leq s$ ya viene garantizada por la primera condición; la conservamos escrita porque muestra geométricamente que a pertenece a la franja $(s - \varepsilon, s]$. Del mismo modo, en la caracterización del ínfimo la condición $i \leq a$ procede de que i es una cota inferior y permite visualizar la franja $[i, i + \varepsilon)$.

Idea de la prueba
Si s fuera el supremo pero existiera una franja $(s-\varepsilon, s]$ sin elementos de A, entonces $s-\varepsilon$ seguiría estando por encima de todo el conjunto. Habríamos encontrado una cota superior menor que s, contradiciendo que s es la menor. En la dirección inversa, si podemos encontrar elementos de A arbitrariamente cerca de s por debajo, ninguna cota superior puede situarse estrictamente por debajo de s: algún elemento del conjunto la sobrepasaría. Para el ínfimo ocurre exactamente la imagen reflejada: una franja $[i, i+\varepsilon)$ no puede quedar vacía, y ninguna cota inferior puede situarse estrictamente por encima de i.

Demostración para el supremo. Supongamos primero que $s = \sup A$.

Por definición, s es una cota superior. Sea ahora $\varepsilon > 0$. Supongamos que no existe ningún $a \in A$ tal que $s - \varepsilon < a$. Negar la existencia significa que, para todo $a \in A$, $\neg(s - \varepsilon < a)$. Como el orden es total, esta negación equivale a $a \leq s - \varepsilon$.

Por tanto,

$$\forall a \in A, \qquad a \leq s - \varepsilon,$$

de modo que $s - \varepsilon$ sería una cota superior de A. Pero, como $\varepsilon > 0$, $s - \varepsilon < s$, lo que contradice que s sea la menor cota superior.

Luego existe $a \in A$ tal que $s - \varepsilon < a$.

Como s es cota superior, además $a \leq s$. Así, $s - \varepsilon < a \leq s$. Esto prueba la propiedad aproximativa. Recíprocamente, supongamos que s es una cota superior y que

$$\forall \varepsilon > 0 \; \exists a \in A \qquad s - \varepsilon < a \leq s.$$

Queremos demostrar que s es la **menor** cota superior. Sea u una cota superior cualquiera de A. Supongamos, para obtener una contradicción, que $u < s$.

Entonces $\varepsilon = s - u > 0$.

Por la propiedad aproximativa existe $a \in A$ tal que $s - \varepsilon < a \leq s$.

Pero $s - \varepsilon = s - (s - u) = u$, así que $u < a$.

Esto contradice que u sea una cota superior, pues toda cota superior debe satisfacer $a \leq u$ para cada $a \in A$. Por tanto, $u < s$ es imposible. Como u era una cota superior arbitraria, $s \leq u$ para toda cota superior u de A. Así, s es la menor cota superior y $s = \sup A$.

Demostración para el ínfimo. Supongamos ahora que $i = \inf A$.

Entonces i es una cota inferior. Sea $\varepsilon > 0$. Si no existiera ningún $a \in A$ con $a < i + \varepsilon$, entonces, para todo $a \in A$, $a \geq i + \varepsilon$. Por tanto, $i + \varepsilon$ sería una cota inferior de A. Pero $i < i + \varepsilon$, lo que contradice que i sea la **mayor** cota inferior.

Luego existe $a \in A$ tal que $a < i + \varepsilon$.

Como i es cota inferior, además $i \leq a$. Por tanto, $i \leq a < i + \varepsilon$. Recíprocamente, supongamos que i es una cota inferior y que

$$\forall \varepsilon > 0 \; \exists a \in A \qquad i \leq a < i + \varepsilon.$$

Sea ℓ una cota inferior cualquiera de A. Queremos demostrar $\ell \leq i$.

Supongamos, por contradicción, que $i < \ell$.

Tomemos $\varepsilon = \ell - i > 0$.

Por la propiedad aproximativa existe $a \in A$ tal que $a < i + \varepsilon = \ell$. Pero esto contradice que ℓ sea una cota inferior, pues debería cumplirse $\ell \leq a$. Por tanto, ninguna cota inferior puede ser mayor que i. Así, $i = \text{ínf } A$.

■

El orden de los cuantificadores no se puede intercambiar
La propiedad del supremo dice

$$\forall \varepsilon > 0 \; \exists a_\varepsilon \in A \qquad s - \varepsilon < a_\varepsilon \leq s.$$

El elemento a_ε **puede depender de** ε. No estamos afirmando

$$\exists a \in A \; \forall \varepsilon > 0 \qquad s - \varepsilon < a \leq s.$$

Esta segunda afirmación sería mucho más fuerte. En efecto, si un mismo $a \in A$ funcionara para todo $\varepsilon > 0$ y tuviéramos $a < s$, podríamos elegir $\varepsilon = \frac{s-a}{2} > 0$. Entonces $s - \varepsilon = \frac{s+a}{2} > a$, contradiciendo $s - \varepsilon < a$. Por tanto tendría que cumplirse $a = s$, y en consecuencia $s \in A$: el supremo sería además un máximo. Así, el patrón $\forall \varepsilon > 0 \; \exists a_\varepsilon$ expresa aproximación a una frontera sin exigir que la frontera pertenezca al conjunto.

Ejemplos de lectura completa.

Ejemplo 1.5. **Máximo, mínimo, supremo e ínfimo en un conjunto elemental.** Sea $C = (-2, 1] \cup \{4\}$.

Entonces

$$\sup C = 4, \qquad \text{máx } C = 4,$$

mientras que $\text{ínf } C = -2$ y C no tiene mínimo.

La parte superior es inmediata porque $4 \in C$ y todo elemento de C es menor o igual que 4. Por tanto, 4 es máximo y, en consecuencia, también supremo. Estudiemos ahora el extremo inferior. Si $x \in C$, entonces o bien $x \in (-2, 1]$ o bien $x = 4$; en ambos casos $-2 < x$. Por tanto, -2 es una cota inferior de C. Sin embargo, $-2 \notin C$, de modo que -2 no puede ser mínimo. Esto todavía no basta para concluir que C carece de mínimo: podría existir otro elemento del conjunto que fuese el menor. Debemos descartarlo. Sea $x \in C$ arbitrario. Si $x = 4$, podemos tomar $y = 1 \in C$, y entonces $y < x$.

Si, en cambio, $x \in (-2, 1]$, definimos $y = \frac{x-2}{2}$.

Como $x > -2$, $y + 2 = \frac{x+2}{2} > 0$, y por tanto $y > -2$. Además, $x - y = \frac{x+2}{2} > 0$, de modo que $y < x$. Puesto que $y < x \leq 1$, concluimos $y \in (-2, 1] \subset C$. Así, en todos los casos,

$$\forall x \in C \; \exists y \in C \qquad y < x.$$

Ningún elemento de C puede ser, por tanto, su mínimo. Falta comprobar que -2 no es solamente una cota inferior, sino la **mayor** cota inferior. Utilicemos la caracterización aproximativa del ínfimo. Sea $\varepsilon > 0$. Necesitamos encontrar $a \in C$ tal que $-2 \leq a < -2 + \varepsilon$.

Si $0 < \varepsilon \leq 2$, tomamos $a = -2 + \frac{\varepsilon}{2}$.

Entonces $-2 < a < -2 + \varepsilon$, y además

$$-2 < a \leq -1 < 1,$$

de modo que $a \in (-2, 1] \subset C$.

Si $\varepsilon > 2$, basta tomar $a = -1 \in C$, pues $-2 < -1 < -2 + \varepsilon$.

Hemos mostrado que para todo $\varepsilon > 0$ existe $a \in C$ con $-2 \leq a < -2 + \varepsilon$. Como -2 es una cota inferior, Proposición 1.8 permite concluir

$$\boxed{\inf C = -2}.$$

Así, este único ejemplo reúne las cuatro nociones:

$$\boxed{\max C = \sup C = 4, \qquad \inf C = -2, \qquad C \text{ no tiene mínimo.}}$$

No todo conjunto tiene máximo, ni todo conjunto tiene una cota. Conviene separar tres fenómenos: estar acotado, alcanzar una frontera y poseer siquiera una barrera en una dirección. El conjunto $(0, 1)$ está acotado superior e inferiormente, pero no tiene máximo ni mínimo. En cambio, $[0, 1]$ también está acotado y sí alcanza ambas fronteras:

$$\min[0, 1] = 0, \qquad \max[0, 1] = 1.$$

Por tanto, estar acotado no obliga a que exista un elemento extremo. En particular,

$$\boxed{A \text{ tiene máximo} \Longrightarrow A \text{ está acotado superiormente,}}$$

pero la implicación recíproca es falsa, como muestra $(0, 1)$.

Consideremos ahora $(0, \infty)$.

Este conjunto está acotado inferiormente —por ejemplo, por 0—, pero no está acotado superiormente. Negar la existencia de una cota superior significa precisamente afirmar

$$\forall M \in \mathbb{R}\ \exists x \in (0, \infty) \qquad x > M.$$

Fijemos, pues, un candidato arbitrario $M \in \mathbb{R}$. Si $M > 0$, tomamos $x = M + 1$; si $M \leq 0$, tomamos $x = 1$. En ambos casos $x \in (0, \infty)$ y $x > M$. Como esto funciona para todo M, ninguna cota superior existe. En consecuencia, $(0, \infty)$ no puede tener máximo: un máximo sería, por definición, una cota superior que además pertenece al conjunto. Tampoco puede tener supremo, porque un supremo debe ser ante todo una cota superior.

Hay todavía otra razón para exigir cuidado con las hipótesis. El conjunto vacío $\varnothing$ está acotado superiormente en el sentido de la definición: dado cualquier $M \in \mathbb{R}$, no existe ningún elemento de $\varnothing$ que pueda violar la condición $a \leq M$. Por tanto, **todo** real es una cota superior de $\varnothing$. Sin embargo, $\varnothing$ no tiene supremo real. En efecto, si s fuera la menor cota superior, entonces $s - 1$ también sería una cota superior y $s - 1 < s$, contradiciendo la minimalidad de s. Así aparecen de manera natural las dos hipótesis que más adelante acompañarán a la propiedad de existencia del supremo:

$$\boxed{A \neq \varnothing \qquad + \qquad A \text{ acotado superiormente.}}$$

Hasta aquí solo hemos comprobado que ambas condiciones son relevantes. **Todavía no hemos demostrado que basten** para garantizar la existencia de un supremo. Afirmar esa suficiencia será precisamente la nueva propiedad de $\mathbb{R}$ que introduciremos en §1.5.

Antes de seguir.

Antes de seguir
1. Sea $A = [-2, 0) \cup (1, 3)$.
Determina $\sup A$, $\inf A$ y decide si existen máximo y mínimo.
Respuesta. Todo elemento de A es menor o igual que 3, de modo que 3 es una cota superior. Para demostrar que es la menor, sea $\varepsilon > 0$ y definamos

$$\delta = \min\left\{\frac{\varepsilon}{2}, \frac{1}{2}\right\} > 0, \qquad a = 3 - \delta.$$

Como $\delta \leq 1/2$, $\frac{5}{2} \leq a < 3$, así que $a \in (1, 3) \subset A$. Además, $\delta < \varepsilon$, y por tanto $3 - \varepsilon < a \leq 3$. La caracterización aproximativa Proposición 1.8 da entonces

$$\boxed{\sup A = 3}.$$

Como $3 \notin A$, el resultado «máximo $\Leftrightarrow$ supremo perteneciente al conjunto» muestra que A no tiene máximo. En el extremo inferior, todo elemento de A satisface $-2 \leq a$, y además $-2 \in A$. Por tanto,

$$\boxed{\min A = -2}.$$

Todo mínimo es también el ínfimo, luego

$$\boxed{\inf A = \min A = -2}.$$

Antes de seguir
2. Sea $A \neq \varnothing$ y $s = \sup A$. Demuestra que, para todo $t < s$, existe $a \in A$ tal que $t < a \leq s$.
Respuesta. Sea $t < s$ arbitrario. Entonces $\varepsilon = s - t > 0$.
La caracterización aproximativa del supremo produce un elemento $a \in A$ tal que $s - \varepsilon < a \leq s$.
Como $s - \varepsilon = s - (s - t) = t$, obtenemos $t < a \leq s$.
Como $t < s$ fue arbitrario, esto vale para todo real situado estrictamente por debajo de s.

El punto al que hemos llegado. La frontera sugerida por $x^2 = 2$ puede describirse ya mediante condiciones precisas. Cada una impone una exigencia distinta al candidato. Una **cota superior** M satisface

$$\forall a \in A, \qquad a \leq M,$$

pero no necesita pertenecer a A.

Un **máximo** M satisface simultáneamente

$$M \in A \qquad \text{y} \qquad \forall a \in A,\ a \leq M.$$

Un **supremo** s es una cota superior con una exigencia adicional: ninguna otra cota superior puede quedar estrictamente por debajo de él. Cuando existe,

$$\boxed{s = \sup A \iff \begin{cases} a \leq s & \text{para todo } a \in A, \\ s \leq u & \text{para toda cota superior } u \text{ de } A. \end{cases}}$$

También hemos demostrado que, cuando existe, el supremo es único. Y si además $\sup A \in A$, entonces esa frontera es alcanzada y coincide con el máximo. La caracterización aproximativa ofrece otra lectura de la misma frontera. Para un conjunto no vacío, una vez que s es cota superior, $s = \sup A$ equivale a exigir

$$\boxed{\forall \varepsilon > 0 \ \exists a_\varepsilon \in A \qquad s - \varepsilon < a_\varepsilon \leq s.}$$

Así, el supremo puede no pertenecer al conjunto, pero ningún descenso positivo desde s sigue quedando por encima de todos sus elementos. Estas caracterizaciones permiten comprobar que un candidato es el supremo y demuestran que, si existe, es único. Su existencia requiere todavía una justificación: la definición, por sí sola, puede describir un objeto que falte en el sistema numérico considerado. La pregunta decisiva queda entonces formulada con precisión:

> si $A \subseteq \mathbb{R}$ es no vacío y está acotado superiormente, ¿existe necesariamente una menor cota superior real?

Los axiomas de cuerpo y orden no bastan para responder afirmativamente. En §1.4 veremos de manera concreta qué puede fallar dentro de $\mathbb{Q}$ mediante la ecuación $x^2 = 2$, y en §1.5 incorporaremos la propiedad adicional de $\mathbb{R}$ que garantiza la existencia de esas fronteras bajo las hipótesis apropiadas:

$$\boxed{\text{la completitud.}}$$

1.4. Por qué los racionales no bastan: el ejemplo de Rudin

Una ecuación demasiado sencilla para causar problemas. Busquemos primero soluciones racionales de $x^2 = 2$. Antes de decidir si tal racional existe, podemos localizar dónde tendría que estar una eventual solución positiva. Como $1^2 < 2 < 2^2$, si $x \geq 0$ y $x^2 = 2$, no puede ocurrir $x \leq 1$: por la monotonía del cuadrado en los no negativos, demostrada en Proposición 1.5, tendríamos $x^2 \leq 1$. Tampoco puede ocurrir $x \geq 2$, pues entonces $x^2 \geq 4$. Por tanto, $1 < x < 2$. Podemos estrechar la localización con racionales intermedios:

$$\left(\frac{4}{3}\right)^2 = \frac{16}{9} < 2, \qquad \left(\frac{3}{2}\right)^2 = \frac{9}{4} > 2.$$

El mismo argumento de monotonía muestra que una solución positiva debería satisfacer $\frac{4}{3} < x < \frac{3}{2}$. Podríamos continuar insertando puntos medios racionales. En cada etapa ocurre una de tres cosas: el punto medio tiene cuadrado menor que 2, mayor que 2 o exactamente igual a 2. En los dos primeros casos conservamos la mitad que todavía encierra una eventual solución; en el tercero habríamos encontrado una solución racional exacta. Este procedimiento puede producir intervalos racionales cada vez más estrechos, pero por sí solo no responde la pregunta decisiva: **¿algún racional satisface exactamente $x^2 = 2$?**

La cuestión precisa es, por tanto:

> ¿podemos demostrar que existe un racional con cuadrado 2, o podemos demostrar que ninguno existe?

La segunda alternativa es la correcta.

Proposición 1.9. **Inexistencia de una solución racional de $x^2 = 2$.** No existe ningún número racional q tal que

$$q^2 = 2.$$

Idea de la prueba
Si un racional q satisficiera $q^2 = 2$, podríamos elegir una representación reducida $q = \frac{m}{n}$ con $m, n \in \mathbb{Z}$, $n \neq 0$, y sin factor común entero mayor que 1. La ecuación obligará primero a que m sea par y después a que n también sea par. Entonces ambos tendrán el factor común 2, contradiciendo que la fracción fuese reducida. El único hecho aritmético adicional que necesitaremos es este: si el cuadrado de un entero es par, entonces el entero es par. Lo justificamos localmente antes de usarlo.

Lema de paridad (demostración local)
Supongamos que $m \in \mathbb{Z}$ y que m^2 es par. Si m no fuera par, sería impar y podríamos escribir $m = 2k + 1$ para algún $k \in \mathbb{Z}$. Entonces

$$m^2 = (2k+1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1,$$

que es impar. Esto contradice que m^2 sea par. Por tanto,

$$\boxed{m^2 \text{ par} \Longrightarrow m \text{ par.}}$$

Demostración. Supongamos, para obtener una contradicción, que existe $q \in \mathbb{Q}$ tal que $q^2 = 2$.

Elegimos una representación reducida $q = \frac{m}{n}$, con $m, n \in \mathbb{Z}$, $n \neq 0$, y sin factor común entero mayor que 1.

Sustituyendo en la ecuación, $\left(\frac{m}{n}\right)^2 = 2$.

Como $n \neq 0$, podemos multiplicar por n^2 y obtener $m^2 = 2n^2$. El miembro derecho es divisible por 2, así que m^2 es par. Por el lema de paridad, m es par. Existe entonces $k \in \mathbb{Z}$ tal que $m = 2k$. Sustituyendo esta expresión en $m^2 = 2n^2$, $(2k)^2 = 2n^2$, de modo que $4k^2 = 2n^2$.

Dividiendo por 2, $n^2 = 2k^2$.

Por la misma razón, n^2 es par y el lema implica que n es par. Existe, pues, $\ell \in \mathbb{Z}$ tal que $n = 2\ell$.

Hemos obtenido simultáneamente

$$m = 2k, \qquad n = 2\ell.$$

Por tanto, m y n tienen el factor común 2, en contradicción con que m/n hubiese sido elegida como una fracción reducida. La suposición inicial es imposible. En consecuencia,

$$\boxed{\forall q \in \mathbb{Q}, \qquad q^2 \neq 2.}$$

■

Después de la prueba
Conviene identificar la arquitectura del argumento.

1. **Objetivo:** demostrar una inexistencia dentro de $\mathbb{Q}$.
2. **Estrategia:** contradicción.
3. **Suposición temporal:** existe $q \in \mathbb{Q}$ con $q^2 = 2$.
4. **Representación útil:** $q = m/n$ en forma reducida.

5. **Mecanismo:** la ecuación fuerza primero $2 \mid m$ y después $2 \mid n$.
6. **Contradicción:** numerador y denominador tienen el factor común 2.

La conclusión debe leerse con precisión. Hemos demostrado

$$\boxed{\text{no existe } q \in \mathbb{Q} \text{ tal que } q^2 = 2.}$$

No hemos demostrado todavía que exista algún número real cuyo cuadrado sea 2. La inexistencia en un dominio y la existencia en otro son afirmaciones lógicamente distintas. Mantendremos esa separación hasta que la completitud nos autorice a construir la frontera real correspondiente.

Qué hemos demostrado y qué no. La proposición anterior establece exactamente

$$\boxed{\neg\exists q \in \mathbb{Q} \qquad q^2 = 2.}$$

Es decir: **dentro de $\mathbb{Q}$ no existe solución** de la ecuación $x^2 = 2$. Es fácil deslizarse, sin advertirlo, desde esta afirmación hasta otra diferente:

$$\exists x \in \mathbb{R} \setminus \mathbb{Q} \qquad x^2 = 2.$$

Pero la segunda proposición contiene una afirmación nueva: **afirma que una solución existe**. La diferencia lógica puede verse separando dos trabajos:

$$\boxed{\begin{array}{lll} \text{(I)} & \neg\exists q \in \mathbb{Q} & q^2 = 2, \\ \text{(II)} & \exists x \in \mathbb{R} & x^2 = 2. \end{array}}$$

La proposición anterior demuestra (I). Todavía no hemos demostrado (II). Solo después de establecer (II) podremos combinar ambas afirmaciones: si existe $x \in \mathbb{R}$ con $x^2 = 2$ y ningún racional satisface esa ecuación, entonces ese x necesariamente cumple $x \notin \mathbb{Q}$, y por tanto $x \in \mathbb{R}\setminus\mathbb{Q}$. Así, demostrar que **ningún racional** posee cierta propiedad no fabrica por sí solo un objeto no racional que la posea.

Una negación de existencia no es una existencia
De

$$\neg\exists q \in \mathbb{Q} \qquad q^2 = 2$$

no podemos concluir, sin una premisa adicional,

$$\exists x \in \mathbb{R} \setminus \mathbb{Q} \qquad x^2 = 2.$$

Para esa conclusión necesitamos primero demostrar que alguna solución existe en $\mathbb{R}$.

Esta distinción explica también una precaución de notación. Todavía no utilizaremos el símbolo $\sqrt{2}$ como nombre de un número real cuya existencia ya hubiese sido establecida. Una notación puede describir qué objeto **queremos** identificar, pero no sustituye una demostración de que tal objeto existe. La existencia real será obtenida más adelante mediante completitud: primero construiremos una frontera real como supremo y después demostraremos que esa frontera tiene cuadrado 2. Solo entonces quedará legitimada, dentro de nuestra cadena lógica, la escritura $\sqrt{2}$.

La diagonal de un cuadrado y el problema aritmético. La geometría hace que la situación sea todavía más provocadora. Consideremos un cuadrado de lado 1. Si llamamos d a la longitud de su diagonal, el teorema de Pitágoras impone la relación $d^2 = 1^2 + 1^2 = 2$. Aquí conviene distinguir dos lenguajes. La geometría nos presenta una **magnitud** —la diagonal— y una relación que esa magnitud debe satisfacer. La aritmética pregunta además:

> ¿puede esa magnitud representarse mediante un número del sistema numérico en el que estamos trabajando?

Si intentáramos representarla por un racional $q > 0$, necesariamente tendría que cumplirse $q^2 = 2$. Pero acabamos de demostrar que ningún racional satisface esa ecuación. Por tanto,

$$\boxed{\begin{array}{c}\text{la longitud de la diagonal de un cuadrado de lado 1}\\ \text{no puede representarse mediante un número racional.}\end{array}}$$

Esta conclusión sigue sin ser una prueba de que, dentro de nuestro desarrollo axiomático, exista ya un número real particular que represente esa longitud. La imagen geométrica funciona aquí como **motivación del problema aritmético**: muestra qué clase de magnitud queremos poder incorporar, mientras que la existencia numérica deberá justificarse después mediante la completitud. El obstáculo tampoco consiste en que los racionales estén «muy separados». Sean $r, s \in \mathbb{Q}$ con $r < s$. Su punto medio $m = \frac{r+s}{2}$ sigue siendo racional. Además, $m - r = \frac{s-r}{2} > 0$ y $s - m = \frac{s-r}{2} > 0$.

Por tanto,

$$\boxed{r < \frac{r+s}{2} < s.}$$

Así, entre dos racionales distintos siempre existe otro racional. Podemos incluso repetir el procedimiento indefinidamente y seguir insertando racionales entre racionales. Pero esta riqueza local del orden no responde una pregunta diferente: si ciertos racionales quedan sistemáticamente a un lado de una frontera y otros al lado opuesto, **¿debe existir un racional que ocupe esa frontera?**

La respuesta será negativa. La densidad entre puntos racionales no impide que falte un punto frontera racional. Para hacer precisa esta diferencia entre «siempre hay puntos intermedios» y «toda frontera existe dentro del sistema», examinaremos ahora el ejemplo introductorio de Rudin.

El ejemplo introductorio de Rudin, paso a paso. En el Ejemplo 1.1 de la tercera edición de *Principles of Mathematical Analysis*, Walter Rudin utiliza la ecuación $p^2 = 2$ para exhibir el hueco de los racionales. Después de demostrar que esa ecuación no tiene solución racional, considera los conjuntos

$$A = \{p \in \mathbb{Q} : p > 0,\ p^2 < 2\}$$

y

$$B = \{p \in \mathbb{Q} : p > 0,\ p^2 > 2\}.$$

Antes de estudiar la fórmula que aparecerá enseguida, conviene describir con precisión la estructura de estos dos conjuntos. En primer lugar, ambos son no vacíos:

$$1 \in A, \qquad 2 \in B,$$

porque $1^2 < 2 < 2^2$.

En segundo lugar, todo racional positivo pertenece exactamente a uno de ellos. En efecto, si $p \in \mathbb{Q}$ y $p > 0$, la tricotomía aplicada a p^2 y 2 da exactamente una de las posibilidades

$$p^2 < 2, \qquad p^2 = 2, \qquad p^2 > 2.$$

La posibilidad intermedia ya fue excluida: ningún racional tiene cuadrado igual a 2. Por tanto,

$$p \in A \qquad \text{o} \qquad p \in B,$$

y las dos alternativas son mutuamente excluyentes. En tercer lugar, todo elemento de A está estrictamente a la izquierda de todo elemento de B. Sean

$$a \in A, \qquad b \in B.$$

Supongamos, para obtener una contradicción, que $b \leq a$. Como $a, b > 0$, tenemos $0 < b \leq a$. La monotonía del cuadrado en los no negativos, Proposición 1.5, da entonces $b^2 \leq a^2$. Pero $a \in A$ implica $a^2 < 2$, así que $b^2 \leq a^2 < 2$, en contradicción con $b \in B$, que exige $b^2 > 2$.

Por tanto,

$$\boxed{\forall a \in A\ \forall b \in B, \qquad a < b.}$$

Tenemos así dos regiones racionales no vacías, disjuntas y ordenadas:

$$\boxed{A \quad \text{queda por debajo del borde,} \qquad B \quad \text{queda por encima del borde.}}$$

Pero todavía no hemos producido ningún punto que ocupe ese borde. El paso siguiente de Rudin es más fino: demostrar que tampoco hay un **último** racional por debajo ni un **primero** por encima. En términos de cuantificadores, debemos probar

$$\boxed{\forall p \in A\ \exists q \in A \qquad p < q,}$$

y

$$\boxed{\forall p \in B\ \exists q \in B \qquad q < p.}$$

Estas dos afirmaciones implicarán, respectivamente, que A no tiene máximo y que B no tiene mínimo. Para lograr ambas cosas con una sola construcción, Rudin introduce, para un racional positivo p,

$$q = p - \frac{p^2 - 2}{p + 2} = \frac{2p + 2}{p + 2} = \frac{2(p + 1)}{p + 2}.$$

La fórmula funciona admirablemente, pero un lector novel puede preguntarse con toda razón:

¿de dónde salió q?

Conviene separar dos tareas. Primero verificaremos que la fórmula satisface exactamente las propiedades cuantificadas que necesitamos. Después reconstruiremos una ruta sistemática para diseñarla.

EJEMPLO 1.6. **El ejemplo de Rudin, sin pasos ocultos.** Para cada $p \in \mathbb{Q}$ con $p > 0$, definamos

$$q = p - \frac{p^2 - 2}{p + 2}.$$

Entonces:

- si $p^2 < 2$, se cumple
$$p < q, \qquad q^2 < 2;$$
- si $p^2 > 2$, se cumple
$$0 < q < p, \qquad q^2 > 2.$$

En consecuencia,

$$\forall p \in A \; \exists q \in A \qquad p < q,$$

y

$$\forall p \in B \; \exists q \in B \qquad q < p.$$

Por tanto, A no tiene máximo y B no tiene mínimo.

Primera pregunta: ¿q sigue siendo racional y positivo? Sí, pero conviene separar tres comprobaciones.

1. La expresión está bien definida. Partimos de $p \in \mathbb{Q}$ con $p > 0$. Entonces $p + 2 > 0$, y, en particular, $p + 2 \neq 0$.

Por tanto, el cociente que aparece en $q = \frac{2(p+1)}{p+2}$ está definido.

2. El nuevo número sigue siendo racional. Como $p \in \mathbb{Q}$ y $\mathbb{Q}$ es un cuerpo,

$$p + 1 \in \mathbb{Q}, \qquad p + 2 \in \mathbb{Q}.$$

Además, $2(p + 1) \in \mathbb{Q}$ y, como $p + 2 \neq 0$, también su cociente pertenece a $\mathbb{Q}$. Por tanto,

$$\boxed{q \in \mathbb{Q}}.$$

3. El nuevo número sigue siendo positivo. De $p > 0$ obtenemos

$$p + 1 > 0, \qquad p + 2 > 0.$$

Como $2 > 0$, $2(p + 1) > 0$.

El cociente de dos números positivos es positivo, luego

$$\boxed{q > 0}.$$

Hemos probado así las dos condiciones de dominio que necesitaremos después:

$$\boxed{q \in \mathbb{Q} \qquad \text{y} \qquad q > 0.}$$

Todavía no sabemos si q pertenece a A o a B. Para decidirlo faltan dos controles distintos: comparar primero q con p y después comparar q^2 con 2.

Segunda pregunta: ¿q se mueve en la dirección correcta? Comparemos ahora q con el punto de partida p. Como $q = \frac{2(p+1)}{p+2}$, podemos escribir

$$\begin{aligned} q - p &= \frac{2(p+1)}{p+2} - p \\ &= \frac{2(p+1) - p(p+2)}{p+2} \\ &= \frac{2 + 2p - p^2 - 2p}{p+2} \\ &= \frac{2 - p^2}{p+2}. \end{aligned}$$

La hipótesis $p > 0$ ya nos dio $p + 2 > 0$. Por tanto, dividir por $p + 2$ **no cambia el signo**: el signo de $q - p$ es exactamente el signo de $2 - p^2$. Si $p \in A$, entonces $p^2 < 2$, de modo que $2 - p^2 > 0$. Así, $q - p > 0$, y por definición del orden

$$\boxed{p < q}.$$

Si, en cambio, $p \in B$, entonces $p^2 > 2$, por lo que $2 - p^2 < 0$.

En consecuencia, $q - p < 0$, y por tanto

$$\boxed{q < p}.$$

Como ningún racional positivo satisface $p^2 = 2$, el numerador $2 - p^2$ nunca es 0 para los valores de p que estamos considerando. Por ello el desplazamiento es siempre **estricto**. Hemos demostrado hasta aquí exactamente

$$\boxed{\begin{aligned} &p \in A \Longrightarrow p < q, \\ &p \in B \Longrightarrow q < p. \end{aligned}}$$

La transformación se mueve, pues, en la dirección correcta: hacia la derecha desde A y hacia la izquierda desde B. Pero esta información todavía no basta para concluir

$$q \in A \qquad \text{o} \qquad q \in B$$

en el caso correspondiente. Moverse en la dirección adecuada no garantiza que no atravesemos la frontera. Por ejemplo, si partiéramos de $p = 1$ y utilizáramos la corrección ingenua $p + (2 - p^2)$, obtendríamos $1 + (2 - 1) = 2$, cuyo cuadrado es $4 > 2$. La dirección del movimiento era correcta, pero el paso fue demasiado grande y terminó al otro lado. Por tanto, queda una tercera obligación: demostrar que la transformación de Rudin **conserva el lado** de la frontera en el que comenzó p.

Tercera pregunta: ¿q permanece en el mismo lado? La comparación entre q y p controla la **dirección** del movimiento, pero todavía debemos verificar que la corrección no atraviese la frontera. Para ello necesitamos comparar q^2 con 2.

Partimos de $q = \frac{2(p+1)}{p+2}$.

Entonces

$$\begin{aligned} q^2 - 2 &= \left(\frac{2(p+1)}{p+2}\right)^2 - 2 \\ &= \frac{4(p+1)^2}{(p+2)^2} - \frac{2(p+2)^2}{(p+2)^2} \\ &= \frac{4(p+1)^2 - 2(p+2)^2}{(p+2)^2}. \end{aligned}$$

Abramos ahora el numerador:

$$\begin{aligned} 4(p+1)^2 - 2(p+2)^2 &= 4(p^2+2p+1) - 2(p^2+4p+4) \\ &= 4p^2 + 8p + 4 - 2p^2 - 8p - 8 \\ &= 2p^2 - 4 \\ &= 2(p^2-2). \end{aligned}$$

Por tanto,

$$\boxed{q^2 - 2 = \frac{2(p^2-2)}{(p+2)^2}.}$$

De la primera comprobación sabemos que $p > 0$, luego $p+2 > 0$ y, en consecuencia, $(p+2)^2 > 0$.

Además, $2 > 0$.

Así, el factor $\frac{2}{(p+2)^2}$ es estrictamente positivo. Multiplicar por él no cambia el signo. Por consiguiente,

$$\boxed{q^2 - 2 \text{ tiene exactamente el mismo signo que } p^2 - 2.}$$

Ahora podemos reunir las tres comprobaciones anteriores.

Si $p \in A$, entonces

$$p \in \mathbb{Q}, \qquad p > 0, \qquad p^2 < 2.$$

La primera pregunta mostró que $q \in \mathbb{Q}$ y $q > 0$; la segunda mostró que $p < q$; y la identidad recién obtenida da $q^2 - 2 < 0$, es decir, $q^2 < 2$.

Por tanto, $q \in A$ y, además, $p < q$.

Como $p \in A$ fue arbitrario,

$$\boxed{\forall p \in A \; \exists q \in A \qquad p < q.}$$

En consecuencia, A no tiene máximo.

Si $p \in B$, entonces

$$p \in \mathbb{Q}, \qquad p > 0, \qquad p^2 > 2.$$

De nuevo, $q \in \mathbb{Q}$ y $q > 0$; la segunda pregunta dio $q < p$; y ahora $q^2 - 2 > 0$, de modo que $q^2 > 2$.

Así, $q \in B$ y $q < p$.

Como $p \in B$ fue arbitrario,

$$\boxed{\forall p \in B\ \exists q \in B \qquad q < p.}$$

Por tanto, B no tiene mínimo. La transformación ha cumplido simultáneamente las dos exigencias que fijamos al comienzo:

$$\boxed{\begin{array}{c} p \in A \Longrightarrow p < q \in A, \\ p \in B \Longrightarrow q \in B \text{ y } q < p. \end{array}}$$

La demostración del ejemplo de Rudin queda así cerrada. La cuestión siguiente ya no será verificar que la fórmula funciona, sino comprender cómo puede diseñarse una transformación con estas propiedades.

Cómo se fabrica una fórmula que parece caída del cielo. Rudin presenta q directamente. Nosotros reconstruiremos una ruta posible para descubrirla. No pretendemos afirmar que este haya sido históricamente el proceso mental exacto mediante el cual se eligió la fórmula. Lo que sí podemos mostrar es que **una transformación con las propiedades necesarias puede diseñarse sistemáticamente**. Partamos de un racional positivo p y midamos su defecto respecto de la ecuación mediante $E(p) = p^2 - 2$. El signo de $E(p)$ indica de qué lado de la frontera se encuentra p: $E(p) < 0 \iff p \in A$, mientras que $E(p) > 0 \iff p \in B$. Por tanto, queremos que la corrección tenga signo opuesto al defecto:

- si $E(p) < 0$, debemos aumentar p;
- si $E(p) > 0$, debemos disminuirlo.

Una familia natural es

$$q_c = p - \frac{p^2 - 2}{p + c},$$

donde elegiremos

$$c \in \mathbb{Q}, \qquad c > 0.$$

La positividad de c cumple una primera función. Como $p > 0$, $p + c > 0$, de modo que el denominador no se anula. Además, $q_c = \frac{cp+2}{p+c}$. Como $c, p > 0$, tanto $cp + 2$ como $p + c$ son positivos. En consecuencia,

$$\boxed{q_c \in \mathbb{Q} \qquad \text{y} \qquad q_c > 0.}$$

Ahora controlemos la **dirección** del movimiento. Restando p,

$$\begin{aligned} q_c - p &= -\frac{p^2 - 2}{p + c} \\ &= \frac{2 - p^2}{p + c}. \end{aligned}$$

Como $p + c > 0$,

$$\boxed{\operatorname{sgn}(q_c - p) = \operatorname{sgn}(2 - p^2).}$$

Así, para cualquier $c > 0$, la transformación se mueve en la dirección deseada. Pero eso no basta: también debemos impedir que atraviese la frontera. Calculemos el nuevo defecto:

$$\begin{aligned} q_c^2 - 2 &= \frac{(cp+2)^2 - 2(p+c)^2}{(p+c)^2} \\ &= \frac{c^2p^2 + 4cp + 4 - 2p^2 - 4cp - 2c^2}{(p+c)^2} \\ &= \frac{(c^2-2)p^2 - 2(c^2-2)}{(p+c)^2} \\ &= \boxed{\frac{(c^2-2)(p^2-2)}{(p+c)^2}}. \end{aligned}$$

Como $(p+c)^2 > 0$, el signo del nuevo defecto queda determinado por los dos factores

$$c^2 - 2 \qquad \text{y} \qquad p^2 - 2.$$

Queremos que $q_c^2 - 2$ tenga el **mismo signo** que $p^2 - 2$. Para ello basta imponer

$$\boxed{c^2 - 2 > 0}.$$

La condición sobre el parámetro ya no es arbitraria. Si eligiéramos un racional positivo con $c^2 < 2$, el factor $c^2 - 2$ sería negativo y el signo del defecto se invertiría: la transformación cruzaría al lado opuesto. La posibilidad $c^2 = 2$ no está disponible dentro de $\mathbb{Q}$, porque ya demostramos que ningún racional tiene cuadrado 2. Por tanto, entre los parámetros racionales positivos, la condición $c^2 > 2$ es exactamente la que nos permite conservar el lado. No necesitamos conocer todavía ningún número real con cuadrado 2 para exhibir un parámetro racional adecuado. Basta escoger $c = 2$, pues $2^2 - 2 = 2 > 0$.

Sustituyendo $c = 2$ obtenemos

$$\boxed{q = p - \frac{p^2 - 2}{p + 2},}$$

que es precisamente la transformación utilizada en el ejemplo.

La técnica escondida

La fórmula puede reconstruirse mediante cuatro decisiones sucesivas:

1. **medir el defecto:** $E(p) = p^2 - 2$;
2. **elegir la dirección:** corregir p con signo opuesto al defecto;
3. **introducir un parámetro racional positivo:** controlar el tamaño de la corrección sin perder racionalidad ni positividad;
4. **imponer un invariante de signo:** exigir que el nuevo defecto permanezca en el mismo lado.

En símbolos, el diseño busca simultáneamente $\operatorname{sgn}(q_c - p) = -\operatorname{sgn}(E(p))$ y $\operatorname{sgn}(E(q_c)) = \operatorname{sgn}(E(p))$. La primera condición mueve el punto **hacia** la frontera; la segunda evita que la atraviese.

La fórmula de Rudin no es, por tanto, la única posible dentro de esta familia. Todo racional positivo c con $c^2 > 2$ produce la misma arquitectura cualitativa. La elección $c = 2$ es especialmente simple porque satisface inmediatamente la condición requerida y deja una expresión elemental.

Ver la máquina funcionando. Las identidades anteriores ya demuestran el comportamiento de la transformación para todo racional positivo. Ahora podemos seguir algunas iteraciones concretas para ver el mecanismo con aritmética exacta. Partamos por debajo del borde con $p = 1$.

Como $T(p) = \frac{2(p+1)}{p+2}$, obtenemos sucesivamente $1 \longmapsto \frac{4}{3} \longmapsto \frac{7}{5} \longmapsto \frac{24}{17}$.

Verifiquemos que cada punto sigue del mismo lado: $1^2 = 1 < 2$,

$$\left(\frac{4}{3}\right)^2 = \frac{16}{9} < 2,$$

$$\left(\frac{7}{5}\right)^2 = \frac{49}{25} < 2,$$

y

$$\left(\frac{24}{17}\right)^2 = \frac{576}{289} < \frac{578}{289} = 2.$$

Además,

$$1 < \frac{4}{3} < \frac{7}{5} < \frac{24}{17}.$$

La cadena se mueve, por tanto, hacia la derecha sin abandonar A. Si comenzamos por encima del borde con $p = 2$, la misma transformación produce $2 \longmapsto \frac{3}{2} \longmapsto \frac{10}{7} \longmapsto \frac{17}{12}$.

Ahora las comprobaciones son $2^2 = 4 > 2$,

$$\left(\frac{3}{2}\right)^2 = \frac{9}{4} > 2,$$

$$\left(\frac{10}{7}\right)^2 = \frac{100}{49} > \frac{98}{49} = 2,$$

y

$$\left(\frac{17}{12}\right)^2 = \frac{289}{144} > \frac{288}{144} = 2.$$

También

$$2 > \frac{3}{2} > \frac{10}{7} > \frac{17}{12}.$$

Esta cadena se mueve hacia la izquierda sin abandonar B. Los cálculos ilustran el comportamiento, pero **una lista de ejemplos no demuestra la afirmación general**. La demostración sigue estando en las dos identidades válidas para todo racional positivo p:

$$\boxed{T(p) - p = \frac{2 - p^2}{p + 2},}$$

y

$$\boxed{T(p)^2 - 2 = \frac{2(p^2 - 2)}{(p + 2)^2}.}$$

La primera controla la dirección del movimiento; la segunda controla el lado de la frontera. Tampoco diremos todavía que estas cadenas «convergen». No hemos definido convergencia de sucesiones ni demostrado ningún teorema que permita extraer de estas iteraciones la existencia de un límite. Por ahora solo estamos observando desigualdades exactas entre racionales.

Antes de seguir
Sea

$$T(p) = p - \frac{p^2 - 2}{p + 2}, \qquad p \geq 1.$$

Usando la identidad ya demostrada para $T(p)^2 - 2$, prueba que $|T(p)^2 - 2| \leq \frac{2}{9}|p^2 - 2|$.
Respuesta. De

$$T(p)^2 - 2 = \frac{2(p^2 - 2)}{(p + 2)^2}$$

obtenemos, usando la multiplicatividad del valor absoluto, $|T(p)^2 - 2| = \frac{2}{(p+2)^2}|p^2 - 2|$, porque $(p + 2)^2 > 0$.
Como $p \geq 1$, $p + 2 \geq 3 > 0$.
La monotonía del cuadrado en los no negativos da $(p + 2)^2 \geq 9$. Como ambas cantidades son positivas, al tomar recíprocos se invierte el orden: $\frac{1}{(p+2)^2} \leq \frac{1}{9}$.
Multiplicando por $2 > 0$, $0 < \frac{2}{(p+2)^2} \leq \frac{2}{9}$.
Finalmente,

$$\boxed{|T(p)^2 - 2| \leq \frac{2}{9}|p^2 - 2|.}$$

Así, para $p \geq 1$, una aplicación de T deja el nuevo defecto absoluto como máximo en $2/9$ del defecto anterior. Esta es una estimación cuantitativa exacta; todavía no la interpretamos como una afirmación de convergencia.

Entonces, ¿dónde está exactamente el hueco? La palabra «hueco» puede inducir una imagen equivocada. No significa que exista alrededor de la frontera un intervalo sin números racionales. Eso sería falso: entre dos racionales distintos siempre podemos insertar otro racional. El fenómeno es más sutil.

Recordemos

$$A = \{p \in \mathbb{Q} : p > 0,\ p^2 < 2\}$$

y

$$B = \{p \in \mathbb{Q} : p > 0,\ p^2 > 2\}.$$

Ya hemos demostrado cuatro hechos:

$$A \neq \varnothing, \qquad B \neq \varnothing,$$

$$\forall a \in A\, \forall b \in B, \qquad a < b,$$

$$\forall a \in A\, \exists a' \in A, \qquad a < a',$$

y

$$\forall b \in B\ \exists b' \in B, \qquad b' < b.$$

Por tanto, A no tiene máximo y B no tiene mínimo. Además, como todo elemento de A está por debajo de todo elemento de B, cada $b \in B$ es una cota superior racional de A. Aquí aparece una distinción importante. Que A no tenga máximo **no basta**, por sí solo, para decir que carece de supremo. Por ejemplo, $(0,1) \cap \mathbb{Q}$ no tiene máximo dentro de $\mathbb{Q}$, pero sí posee la menor cota superior racional 1. Así que el problema no consiste simplemente en que «falta el último elemento de A». La pregunta correcta es más fuerte:

> entre todas las cotas superiores racionales de la región inferior, ¿existe una que sea menor que todas las demás?

Para conectar esta pregunta con la formulación que utilizaremos en §1.5, introducimos $S_{\mathbb{Q}} = \{q \in \mathbb{Q} : q \geq 0,\ q^2 < 2\}$. Este conjunto no es literalmente el mismo que A. Como $0^2 < 2$,

$$\boxed{S_{\mathbb{Q}} = A \cup \{0\}.}$$

Sin embargo, añadir 0 no cambia las cotas superiores. En efecto, como $1 \in A$, toda cota superior u de A satisface $1 \leq u$, y por tanto también $0 \leq u$. Así, toda cota superior de A también domina al nuevo elemento 0. Recíprocamente, toda cota superior de $S_{\mathbb{Q}}$ es automáticamente cota superior de $A \subseteq S_{\mathbb{Q}}$. Luego ambos conjuntos tienen exactamente las mismas cotas superiores racionales. Podemos formular entonces con precisión la cuestión pendiente:

$$\boxed{\text{¿existe } s \in \mathbb{Q} \text{ que sea la menor cota superior de } S_{\mathbb{Q}} \text{ dentro de } \mathbb{Q}?}$$

El trabajo realizado en esta sección ya ha dejado visibles todos los ingredientes que harán fracasar esa posibilidad: por debajo siempre podemos subir, por encima siempre podemos bajar y ningún racional satisface $q^2 = 2$. Pero conviene mantener separadas **intuición estructural** y **demostración formal**. En §1.5 volveremos a $S_{\mathbb{Q}}$ y probaremos explícitamente, en el lenguaje de supremos, que $S_{\mathbb{Q}}$ es no vacío y está acotado superiormente en $\mathbb{Q}$, pero no posee una menor cota superior racional. Ese será el sentido preciso del hueco:

> no falta un intervalo de racionales;
> falta en $\mathbb{Q}$ el punto frontera exigido por la propiedad del supremo.

El paso a $\mathbb{R}$ no consistirá, por tanto, en «rellenar espacios visibles» entre racionales, sino en imponer una propiedad estructural que garantice la existencia de fronteras bajo hipótesis precisas.

Una segunda lectura del ejemplo de Rudin. El ejemplo clásico comprime mucho más que una manipulación ingeniosa. Conviene releerlo ahora como un pequeño mapa de habilidades matemáticas.

1. Traducir una afirmación verbal a cuantificadores. Decir que A no tiene máximo no significa comprobar algunos elementos. Exige demostrar

$$\forall p \in A\ \exists q \in A \qquad p < q.$$

Del mismo modo, decir que B no tiene mínimo exige

$$\forall p \in B\ \exists q \in B \qquad q < p.$$

El elemento nuevo puede depender del punto de partida, pero la construcción debe funcionar para **todo** punto admisible.

2. Diseñar un objeto que satisfaga varias condiciones a la vez. No basta producir un racional mayor que p cuando $p \in A$, ni uno menor cuando $p \in B$. El nuevo número debe conservar simultáneamente:

$$q \in \mathbb{Q}, \qquad q > 0,$$

y el lado de la desigualdad respecto de 2: $p^2 < 2 \Longrightarrow q^2 < 2$,

$$p^2 > 2 \Longrightarrow q^2 > 2.$$

Por eso el problema no era simplemente «mover p», sino moverlo **sin abandonar la región correcta**.

3. Separar las funciones de cada cálculo. Las dos identidades centrales responden a preguntas distintas:

$$q - p = \frac{2 - p^2}{p + 2}$$

controla la **dirección** del movimiento, mientras que

$$q^2 - 2 = \frac{2(p^2 - 2)}{(p + 2)^2}$$

controla el **invariante de signo** que impide cruzar de A a B o de B a A. Ninguna de las dos comprobaciones reemplaza a la otra.

4. Interpretar la construcción en el nivel estructural correcto. De las identidades anteriores obtenemos

$$A \text{ no tiene máximo} \qquad \text{y} \qquad B \text{ no tiene mínimo.}$$

Pero, como acabamos de advertir, de aquí todavía no se sigue que A —o $S_{\mathbb{Q}}$— carezca de supremo racional. La ausencia de un elemento extremo y la ausencia de una menor cota superior son afirmaciones diferentes. El paso conceptual decisivo consiste precisamente en cambiar de pregunta:

$$\boxed{\text{¿existe dentro de } \mathbb{Q} \text{ una menor cota superior para } S_{\mathbb{Q}}\,?}$$

La transformación de Rudin ya contiene la maquinaria que necesitaremos para responder, pero la respuesta formal pertenece a §1.5.

Mapa de dependencias
El argumento puede leerse también como una cadena de dependencias:

cuerpo ordenado
↓
signos, desigualdades y monotonía del cuadrado
↓
transformación racional controlada
↓
A sin máximo y B sin mínimo
↓
pregunta por la existencia de una frontera como supremo.

La demostración de que ningún racional satisface $q^2 = 2$ elimina además la única posibilidad de que la frontera buscada estuviera representada por un racional con cuadrado exactamente igual a 2.

El verdadero tema que ha aparecido no es, por tanto, la raíz cuadrada de 2 como objeto aislado. La pregunta general es:

> si un conjunto no vacío está acotado superiormente, ¿qué propiedad del sistema numérico garantiza que exista dentro de ese mismo sistema una menor cota superior?

Ya disponemos del vocabulario necesario para formular la respuesta. La propiedad adicional será la **completitud**.

La frontera que falta en los racionales. Las operaciones racionales permiten mejorar las aproximaciones sin alcanzar una solución racional de la ecuación. La prueba aritmética establece que

$$q^2 \neq 2 \qquad \text{para todo } q \in \mathbb{Q}.$$

Poder insertar otro racional entre dos racionales tampoco resuelve el problema. Esa posibilidad da nuevas aproximaciones, pero deja abierta la existencia de la frontera que todas ellas sugieren. La transformación del ejemplo organiza las aproximaciones mediante un defecto cuyo signo podemos controlar:

defecto → corrección dirigida → control del invariante → nuevo objeto

La transformación tampoco prueba por sí sola que exista una solución real de $x^2 = 2$. Para pasar de las aproximaciones a una frontera real necesitamos una propiedad que asegure la existencia de los supremos pertinentes. En §1.5 formularemos el **axioma de completitud**, demostraremos que $\mathbb{Q}$ falla en la propiedad del supremo y obtendremos una frontera real. En §1.6 probaremos que esa frontera da efectivamente un número cuyo cuadrado es 2, sin invocar continuidad ni límites.

1.5. Completitud: la propiedad que falta en $\mathbb{Q}$

Una definición no garantiza una existencia. Para demostrar que $s = \sup A$ debemos comprobar que s es cota superior y que ninguna cota superior queda por debajo de s. Ambas comprobaciones parten de un candidato; cuando aún no disponemos de él, aparece otra pregunta:

¿qué ocurre si A es no vacío y está acotado superiormente, pero no sabemos de antemano cuál debería ser su menor cota superior?

Nada de lo demostrado hasta ahora garantiza que esa menor cota exista dentro del sistema numérico en el que estamos trabajando. Esta distinción es fundamental. Una definición responde a la pregunta

$$\boxed{\text{¿qué propiedades tendría el objeto si existiera?}}$$

mientras que un teorema o un axioma de existencia responde a otra:

$$\boxed{\text{¿tenemos derecho a afirmar que tal objeto existe?}}$$

En §1.4 ya vimos una advertencia de este tipo: conocer la ecuación $x^2 = 2$ no nos autorizaba todavía a suponer que existía una solución positiva en nuestro dominio. Ahora aparece el mismo problema en un nivel estructural.

El axioma de completitud. La pregunta abierta al final del microtramo anterior era:

$$A \neq \varnothing, \qquad A \text{ acotado superiormente} \quad \stackrel{?}{\Longrightarrow} \quad \exists s \in \mathbb{R}\ P_A(s).$$

La propiedad adicional que adoptaremos para $\mathbb{R}$ responde afirmativamente a esa pregunta.

Axioma de completitud
Todo subconjunto no vacío $A \subseteq \mathbb{R}$ que esté acotado superiormente posee un supremo en $\mathbb{R}$. En símbolos,

$$\boxed{A \neq \varnothing \quad \text{y} \quad \exists M \in \mathbb{R}\ \forall a \in A,\ a \leq M \quad \Longrightarrow \quad \exists s \in \mathbb{R}\ P_A(s).}$$

Equivalentemente, existe $s \in \mathbb{R}$ tal que $s = \sup A$.

La fuerza nueva del axioma está en el cuantificador

$$\boxed{\exists s \in \mathbb{R}.}$$

La definición de supremo ya nos decía qué debía satisfacer un candidato; la completitud afirma que, bajo las hipótesis $A \neq \varnothing$ y

$$A \text{ acotado superiormente,}$$

ese candidato **existe dentro de** $\mathbb{R}$. Además, en §1.3 ya demostramos que el supremo, si existe, es único. Por tanto, combinando aquella unicidad con el axioma de completitud obtenemos inmediatamente

$$\boxed{A \neq \varnothing, \quad A \subseteq \mathbb{R}, \quad A \text{ acotado superiormente} \quad \Longrightarrow \quad \exists!\, s \in \mathbb{R} \text{ tal que } s = \sup A.}$$

Así, la completitud no redefine la noción de supremo. Añade exactamente la garantía de existencia que faltaba:

$$\boxed{\text{definición} + \text{unicidad} + \text{completitud} \quad \Longrightarrow \quad \text{existencia y unicidad del supremo.}}$$

¿Por qué lo llamamos axioma? La palabra **axioma** describe el papel que una afirmación desempeña dentro de un desarrollo determinado. No significa que la afirmación carezca de justificación en cualquier contexto ni que deba adoptarse siempre como punto de partida. En este libro hemos elegido una vía axiomática. No construiremos $\mathbb{R}$ a partir de objetos más elementales; asumiremos que los números reales forman un cuerpo ordenado y añadiremos como propiedad fundamental la completitud recién enunciada. Así, dentro de nuestra cadena lógica,

$$\boxed{\text{cuerpo} + \text{orden} + \text{completitud}}$$

son datos estructurales de partida para $\mathbb{R}$, y las consecuencias posteriores deberán deducirse de ellos. Hay otra vía posible. Si se construye $\mathbb{R}$ —por ejemplo, mediante cortes de Dedekind o mediante una construcción apropiada a partir de sucesiones racionales—, entonces las propiedades de cuerpo, orden y completitud deben demostrarse para el objeto construido. En ese contexto, la propiedad del supremo aparece como **teorema de la construcción**, no como axioma inicial. Por tanto, la diferencia no está en el contenido matemático de la propiedad, sino en su **posición dentro de la cadena de dependencias**:

$$\boxed{\begin{array}{c}\text{presentación axiomática:} \quad \text{completitud} \longrightarrow \text{consecuencias},\\ \text{presentación constructiva:} \quad \text{construcción de } \mathbb{R} \longrightarrow \text{demostración de completitud.}\end{array}}$$

En nuestra presentación utilizaremos la primera ruta. Conviene añadir una precisión lógica. En este punto acabamos de **adoptar** la completitud como hipótesis adicional; todavía no hemos demostrado formalmente dentro del capítulo que no pueda deducirse de los axiomas de cuerpo y orden. Esa separación quedará certificada más adelante en esta misma sección. Demostraremos que $\mathbb{Q}$ es un cuerpo ordenado pero contiene un conjunto no vacío y acotado superiormente que no posee supremo racional. Entonces tendremos un ejemplo de estructura que satisface cuerpo y orden pero falla la propiedad del supremo. Por ello, la completitud contiene información que cuerpo y orden, por sí solos, no fuerzan.

Terminología
En este capítulo, **completitud** significará específicamente la **propiedad del supremo** que acabamos de adoptar. Más adelante aparecerán otras nociones de completitud, por ejemplo formuladas mediante sucesiones de Cauchy. No las utilizaremos aquí para justificar este axioma: la teoría rigurosa de sucesiones todavía no ha sido desarrollada y las relaciones entre esas formulaciones deberán demostrarse en el lugar correspondiente.

El objetivo de esta presentación no es, por tanto, construir los reales, sino fijar con precisión qué estructura de $\mathbb{R}$ utilizaremos para desarrollar el cálculo y qué resultados dependen de cada parte de esa estructura.

Qué autoriza el axioma y qué no. El axioma de completitud es una herramienta de **existencia** con hipótesis precisas. Antes de invocarlo conviene auditar tres datos.

1. Sistema ambiente. Debemos estar considerando $A \subseteq \mathbb{R}$.

La acotación y el supremo se entienden entonces dentro de $\mathbb{R}$. Si un conjunto de racionales se considera como subconjunto de $\mathbb{R}$, la completitud puede producir un supremo real, pero no afirma que ese supremo pertenezca a $\mathbb{Q}$.

2. No vacuidad. Debemos verificar $A \neq \varnothing$.

En una demostración suele bastar exhibir un testigo concreto $a_0 \in A$.

3. Acotación superior en $\mathbb{R}$. Debemos producir al menos un número $M \in \mathbb{R}$ tal que

$$\forall a \in A, \qquad a \leq M.$$

Solo después de estas comprobaciones el axioma autoriza exactamente la afirmación

$$\boxed{\exists s \in \mathbb{R} \qquad s = \sup A.}$$

La palabra **exactamente** importa. La existencia procede del axioma de completitud; la unicidad no. La unicidad del supremo ya fue demostrada en §1.3 a partir de su definición y de la antisimetría del orden. Por tanto, al combinar ambos resultados obtenemos

$$\boxed{\begin{array}{rcl} A \subseteq \mathbb{R}, & & \\ A \neq \varnothing, & \Longrightarrow & \exists!\, s \in \mathbb{R} \text{ tal que } s = \sup A. \\ \exists M \in \mathbb{R}\ \forall a \in A,\ a \leq M & & \end{array}}$$

El cuantificador $\exists!$ es, pues, una **conclusión combinada**: completitud aporta existencia y el resultado previo de §1.3 aporta unicidad. Conviene observar también qué **no** forma parte de las hipótesis. No necesitamos que A esté acotado inferiormente. Por ejemplo, $A = (-\infty, 0]$ es no vacío y está acotado superiormente en $\mathbb{R}$ por 0, aunque no está acotado inferiormente. La completitud se aplica igualmente. Veamos ahora dos fallos de hipótesis. Para $A = \varnothing$, la condición de no vacuidad falla. El conjunto vacío está acotado superiormente de manera vacua, pero eso no basta para aplicar el axioma. Para $A = (0, \infty)$, la no vacuidad sí se cumple, pero no existe ninguna cota superior real. Por tanto, tampoco podemos invocar completitud. En ambos casos la conclusión del axioma queda fuera de alcance. Lógicamente,

$$\boxed{\text{no poder aplicar un teorema} \not\Longrightarrow \text{haber demostrado la negación de su conclusión.}}$$

Cualquier afirmación adicional deberá justificarse por otro argumento. Hay otras tres cosas que el axioma tampoco afirma.

Primero, no dice que el supremo pertenezca al conjunto. Para $A = (0, 1)$, la completitud garantiza la existencia de un supremo real, y ya demostramos que $\sup A = 1$, pero $1 \notin A$.

Por tanto,

$$\boxed{\text{completitud garantiza un supremo, no necesariamente un máximo.}}$$

Segundo, el axioma no calcula el supremo. Puede garantizar que existe un número $s = \sup A$ sin proporcionar una fórmula explícita para s.

Tercero, tampoco entrega automáticamente propiedades adicionales del objeto construido. Si más adelante definimos un conjunto mediante una ecuación o una desigualdad y obtenemos su supremo por completitud, todavía tendremos que demostrar por separado que esa frontera posee la propiedad concreta que buscamos.

Protocolo de uso de completitud
Antes de escribir «por completitud», audite:

1. **ambiente:** ¿el conjunto está siendo considerado dentro de $\mathbb{R}$?;
2. **testigo:** ¿hemos exhibido al menos un elemento del conjunto?;
3. **cota:** ¿hemos exhibido una cota superior en $\mathbb{R}$?;

4. **conclusión exacta:** ¿estamos usando el axioma solamente para afirmar existencia?

Si después escribimos «existe un único supremo», la unicidad proviene del resultado ya demostrado en §1.3. Cualquier identificación o propiedad adicional del supremo requiere trabajo posterior.

El axioma está formulado de manera unilateral, para cotas superiores. En el microtramo siguiente veremos que no hace falta postular por separado la existencia de ínfimos: la versión inferior puede deducirse de esta misma propiedad.

El ínfimo no necesita un segundo axioma. El axioma de completitud fue formulado únicamente para conjuntos no vacíos y acotados **superiormente**. Podríamos postular una segunda propiedad, dual, para conjuntos acotados inferiormente. No hace falta: la existencia de ínfimos puede deducirse del mismo axioma.

Proposición 1.10. **Existencia de ínfimos a partir del axioma de completitud.** Todo subconjunto no vacío $A \subseteq \mathbb{R}$ que esté acotado inferiormente posee un ínfimo en $\mathbb{R}$.

Idea de la prueba
Reuniremos en un solo conjunto **todas las cotas inferiores de** A:

$$L = \{\ell \in \mathbb{R} : \ell \leq a \text{ para todo } a \in A\}.$$

La hipótesis de acotación inferior hará que L sea no vacío. La hipótesis $A \neq \varnothing$ nos permitirá escoger un elemento de A que funcione como cota superior de L. Entonces podremos aplicar completitud a L. Su supremo resultará ser exactamente la mayor de todas las cotas inferiores de A, es decir, $\inf A$.

Demostración. Sea $A \subseteq \mathbb{R}$ no vacío y acotado inferiormente. Definamos $L = \{\ell \in \mathbb{R} : \forall a \in A,\ \ell \leq a\}$.

Por definición, L es el conjunto de todas las cotas inferiores de A.

1. L **es no vacío.** Como A está acotado inferiormente, existe algún $\ell_0 \in \mathbb{R}$ tal que

$$\forall a \in A, \qquad \ell_0 \leq a.$$

Por la definición de L, $\ell_0 \in L$.

Por tanto, $L \neq \varnothing$.

2. L **está acotado superiormente.** Como $A \neq \varnothing$, podemos escoger $a_0 \in A$.

Sea $\ell \in L$ arbitrario. Puesto que ℓ es una cota inferior de A, debe cumplirse en particular $\ell \leq a_0$.

Así,

$$\forall \ell \in L, \qquad \ell \leq a_0.$$

Por tanto, a_0 es una cota superior de L. Ya hemos verificado exactamente las dos hipótesis que necesita completitud:

$$L \neq \varnothing \qquad \text{y} \qquad L \text{ está acotado superiormente.}$$

Existe entonces un número real $s = \sup L$. Debemos demostrar que este mismo número es el ínfimo de A. Para ello verificaremos las dos cláusulas de la definición.

3. s es una cota inferior de A. Sea $a \in A$ arbitrario. Para todo $\ell \in L$ tenemos $\ell \leq a$, porque cada elemento de L es una cota inferior de A. Esto significa que a es una cota superior de L.

Como $s = \sup L$ es la **menor** cota superior de L, $s \leq a$. El elemento $a \in A$ era arbitrario. Por tanto,

$$\forall a \in A, \qquad s \leq a,$$

y así s es una cota inferior de A. Equivalentemente, $s \in L$.

4. Ninguna cota inferior de A es mayor que s. Sea ℓ una cota inferior cualquiera de A. Entonces, por definición, $\ell \in L$.

Pero $s = \sup L$ es, en particular, una cota superior de L. Por tanto, $\ell \leq s$.

Hemos demostrado simultáneamente que

$$s \leq a \qquad \text{para todo } a \in A,$$

y que

$$\ell \leq s \qquad \text{para toda cota inferior } \ell \text{ de } A.$$

Estas son exactamente las dos cláusulas que caracterizan al ínfimo. En consecuencia,

$$\boxed{s = \text{ínf } A.}$$

■

La demostración proporciona además una identidad conceptual útil. Si

$$L = \{\ell \in \mathbb{R} : \ell \text{ es cota inferior de } A\},$$

entonces

$$\boxed{\text{ínf } A = \sup L.}$$

Más aún, como acabamos de demostrar que $s = \sup L$ pertenece al propio L,

$$\boxed{\text{ínf } A = \sup L = \text{máx } L.}$$

Es decir: el ínfimo de A es literalmente **la mayor de todas sus cotas inferiores**.

Dónde se usó cada hipótesis
Las dos hipótesis sobre A desempeñan funciones distintas:

- A **acotado inferiormente** $\Longrightarrow L \neq \varnothing$;
- $A \neq \varnothing \Longrightarrow$ podemos escoger $a_0 \in A$, que sirve como cota superior de L.

Solo después de esas dos verificaciones entra la completitud, una única vez:

$$L \neq \varnothing \quad + \quad L \text{ acotado superiormente} \quad \Longrightarrow \quad \sup L \text{ existe.}$$

El resto de la prueba utiliza únicamente las definiciones de cota, supremo e ínfimo y las propiedades del orden.

Por tanto, no necesitamos añadir una segunda «propiedad de completitud para ínfimos». La versión inferior ya está contenida en la propiedad del supremo.

Ahora volvamos a $\mathbb{Q}$. La importancia del axioma de completitud sería difícil de apreciar si la misma propiedad fuese válida también en $\mathbb{Q}$. Veamos que no lo es.

Consideremos nuevamente $S_{\mathbb{Q}} = \{q \in \mathbb{Q} : q \geq 0,\ q^2 < 2\}$.

Antes de preguntar por su supremo debemos auditar las dos hipótesis de existencia.

1. $S_{\mathbb{Q}}$ es no vacío. El número 1 es racional y satisface

$$1 \geq 0 \qquad \text{y} \qquad 1^2 = 1 < 2.$$

Por tanto, $1 \in S_{\mathbb{Q}}$, y así $S_{\mathbb{Q}} \neq \varnothing$.

2. $S_{\mathbb{Q}}$ está acotado superiormente en $\mathbb{Q}$. Mostremos que 2 es una cota superior racional. Sea $q \in S_{\mathbb{Q}}$.

Entonces

$$q \geq 0 \qquad \text{y} \qquad q^2 < 2.$$

Si ocurriera $q \geq 2$, tendríamos $0 \leq 2 \leq q$, y la monotonía del cuadrado en los no negativos, Proposición 1.5, daría $4 = 2^2 \leq q^2$. Esto contradice $q^2 < 2$. Por tanto, $q < 2$, y en particular $q \leq 2$.

Como $q \in S_{\mathbb{Q}}$ fue arbitrario,

$$\forall q \in S_{\mathbb{Q}}, \qquad q \leq 2.$$

Así, $2 \in \mathbb{Q}$ es una cota superior de $S_{\mathbb{Q}}$ dentro del sistema racional.

Hemos verificado, pues,

$$\boxed{S_{\mathbb{Q}} \neq \varnothing \qquad \text{y} \qquad S_{\mathbb{Q}} \text{ está acotado superiormente en } \mathbb{Q}.}$$

Si $\mathbb{Q}$ satisficiera la propiedad del supremo que hemos adoptado para $\mathbb{R}$, estas dos condiciones obligarían a que existiera una menor cota superior **racional**. Conviene fijar la notación. Escribiremos $s = \sup_{\mathbb{Q}} S_{\mathbb{Q}}$ para decir que $s \in \mathbb{Q}$ satisface simultáneamente

$$\forall q \in S_{\mathbb{Q}}, \qquad q \leq s,$$

y

$$\forall u \in \mathbb{Q}, \qquad \left[(\forall q \in S_{\mathbb{Q}},\ q \leq u) \Longrightarrow s \leq u\right].$$

Es decir, s tendría que ser una cota superior racional y, además, la menor entre **todas las cotas superiores racionales**. La siguiente proposición demuestra que tal s no existe.

Proposición 1.11. **Fallo de la propiedad del supremo en $\mathbb{Q}$.** El conjunto

$$S_{\mathbb{Q}} = \{q \in \mathbb{Q} : q \geq 0,\ q^2 < 2\}$$

es no vacío y está acotado superiormente en $\mathbb{Q}$, pero no tiene supremo en $\mathbb{Q}$.

Idea de la prueba
Supongamos que existiera $s = \sup_{\mathbb{Q}} S_{\mathbb{Q}}$.
Como $1 \in S_{\mathbb{Q}}$ y s sería una cota superior, tendríamos $s \geq 1$. En particular, podremos aplicar a s la transformación racional estudiada en §1.4. La tricotomía deja exactamente tres posibilidades:

$$s^2 < 2, \qquad s^2 = 2, \qquad s^2 > 2.$$

Cada caso ataca una parte distinta de la definición de supremo.

- Si $s^2 < 2$, la transformación producirá un elemento de $S_{\mathbb{Q}}$ estrictamente mayor que s. Así, s dejará de ser siquiera una **cota superior**.
- Si $s^2 > 2$, producirá una **cota superior racional** estrictamente menor que s. Así, s dejará de ser la **menor** cota superior.
- Si $s^2 = 2$, contradiremos el resultado de §1.4: ningún racional tiene cuadrado 2.

Los tres casos serán imposibles.

Demostración. Supongamos, para obtener una contradicción, que existe un racional $s = \sup_{\mathbb{Q}} S_{\mathbb{Q}}$.

Por la definición fijada arriba, s es en particular una cota superior racional de $S_{\mathbb{Q}}$:

$$\forall q \in S_{\mathbb{Q}}, \qquad q \leq s.$$

Como $1 \in S_{\mathbb{Q}}$, obtenemos $1 \leq s$.

En particular,

$$s > 0 \qquad \text{y} \qquad s + 2 > 0,$$

de modo que $s + 2 \neq 0$.

Aplicaremos al candidato s la misma transformación estudiada en §1.4: $T(s) = s - \frac{s^2-2}{s+2} = \frac{2(s+1)}{s+2}$.

Como $s \in \mathbb{Q}$, también

$$s + 1 \in \mathbb{Q}, \qquad s + 2 \in \mathbb{Q}.$$

Además, $s + 2 \neq 0$. Puesto que $\mathbb{Q}$ es un cuerpo, concluimos

$$\boxed{T(s) \in \mathbb{Q}}.$$

Por otra parte,

$$s + 1 > 0, \qquad s + 2 > 0, \qquad 2 > 0,$$

así que

$$\boxed{T(s) > 0}.$$

Ya demostramos en §1.4 las dos identidades decisivas:

$$\boxed{T(s) - s = \frac{2 - s^2}{s + 2}}$$

y

$$\boxed{T(s)^2 - 2 = \frac{2(s^2-2)}{(s+2)^2}.}$$

La primera controla la **dirección** del desplazamiento; la segunda controla el **lado** respecto de 2. Consideremos los tres casos que proporciona la tricotomía para s^2 y 2.

Caso 1: $s^2 < 2$**.** Entonces $2 - s^2 > 0$.

Como $s + 2 > 0$, la primera identidad da $T(s) - s > 0$.

Por definición del orden,

$$\boxed{T(s) > s}.$$

Ahora debemos comprobar que este nuevo racional sigue perteneciendo al conjunto. De $s^2 < 2$ obtenemos $s^2 - 2 < 0$. Además, $\frac{2}{(s+2)^2} > 0$, pues $2 > 0$ y $(s+2)^2 > 0$. Multiplicar $s^2 - 2$ por este factor positivo conserva su signo. Por la segunda identidad, $T(s)^2 - 2 < 0$, y por tanto

$$\boxed{T(s)^2 < 2}.$$

Reunamos las condiciones de pertenencia:

$$T(s) \in \mathbb{Q}, \qquad T(s) > 0, \qquad T(s)^2 < 2.$$

Luego

$$\boxed{T(s) \in S_{\mathbb{Q}}}.$$

Pero s es una cota superior de $S_{\mathbb{Q}}$. Por consiguiente, todo elemento del conjunto debe satisfacer $q \leq s$.

Particularizando en el elemento $q = T(s) \in S_{\mathbb{Q}}$, $T(s) \leq s$.

Esto contradice $T(s) > s$.

Por tanto,

$$\boxed{s^2 < 2 \text{ es imposible.}}$$

Caso 2: $s^2 > 2$**.** Entonces $2 - s^2 < 0$.

Como ya sabemos que $s + 2 > 0$, la identidad $T(s) - s = \frac{2-s^2}{s+2}$ da $T(s) - s < 0$.

Por definición del orden,

$$\boxed{T(s) < s}.$$

La segunda identidad controla ahora el lado de la frontera. De $s^2 > 2$ se sigue $s^2 - 2 > 0$.

Además, $\frac{2}{(s+2)^2} > 0$.

Por tanto,

$$T(s)^2 - 2 = \frac{2(s^2-2)}{(s+2)^2} > 0,$$

y así

$$\boxed{T(s)^2 > 2}.$$

Hasta aquí solo sabemos que $T(s)$ es un racional positivo situado por encima del nivel 2 y que además $T(s) < s$. Para contradecir la **minimalidad** de s todavía debemos demostrar que $T(s)$ es una cota superior de $S_{\mathbb{Q}}$.

Sea, pues, $q \in S_{\mathbb{Q}}$ arbitrario. Entonces

$$q \geq 0 \qquad \text{y} \qquad q^2 < 2.$$

Supongamos, para obtener una contradicción, que $q \geq T(s)$. Ya demostramos antes de separar los casos que $T(s) > 0$.

Por tanto, $0 < T(s) \leq q$.

La monotonía del cuadrado en los no negativos, Proposición 1.5, implica $T(s)^2 \leq q^2$. Pero acabamos de demostrar $T(s)^2 > 2$, mientras que la pertenencia $q \in S_{\mathbb{Q}}$ exige $q^2 < 2$.

Obtendríamos entonces

$$2 < T(s)^2 \leq q^2 < 2,$$

una contradicción.

Así, la suposición $q \geq T(s)$ es imposible y necesariamente $q < T(s)$.

Como $q \in S_{\mathbb{Q}}$ era arbitrario,

$$\boxed{\forall q \in S_{\mathbb{Q}}, \qquad q < T(s).}$$

En particular,

$$\forall q \in S_{\mathbb{Q}}, \qquad q \leq T(s).$$

Por tanto, $T(s)$ es una cota superior de $S_{\mathbb{Q}}$. Y, como ya establecimos que $T(s) \in \mathbb{Q}$, se trata de una **cota superior racional**. Ahora entra la segunda cláusula de $s = \sup_{\mathbb{Q}} S_{\mathbb{Q}}$. Como $T(s) \in \mathbb{Q}$ es una cota superior racional, la minimalidad de s obliga a que $s \leq T(s)$.

Pero antes habíamos obtenido $T(s) < s$.

Las dos desigualdades son incompatibles. Por tanto,

$$\boxed{s^2 > 2 \text{ es imposible.}}$$

Caso 3: $s^2 = 2$. Recordemos que, desde el comienzo de la contradicción, hemos supuesto $s \in \mathbb{Q}$.

Pero en §1.4 demostramos

$$\forall q \in \mathbb{Q}, \qquad q^2 \neq 2.$$

Particularizando esa afirmación en el racional s, obtenemos $s^2 \neq 2$. Esto contradice directamente la hipótesis del caso. Por tanto,

$$\boxed{s^2 = 2 \text{ es imposible.}}$$

La tricotomía aplicada a s^2 y 2 afirma que ocurre exactamente una de las tres posibilidades

$$s^2 < 2, \qquad s^2 = 2, \qquad s^2 > 2.$$

Hemos demostrado que las tres son imposibles bajo la suposición $s = \sup_{\mathbb{Q}} S_{\mathbb{Q}}$. Por tanto, esa suposición inicial debe ser falsa:

$$\boxed{\neg\exists s \in \mathbb{Q} \text{ tal que } s = \sup_{\mathbb{Q}} S_{\mathbb{Q}}.}$$

Equivalentemente,

$$\boxed{S_{\mathbb{Q}} \text{ no posee supremo en } \mathbb{Q}.}$$

Junto con la no vacuidad y la acotación superior ya demostradas, concluimos que $\mathbb{Q}$ falla la propiedad del supremo.

■

Qué hizo realmente la transformación de Rudin. En §1.4 introdujimos

$$T(p) = p - \frac{p^2 - 2}{p + 2}$$

para demostrar que no existe un último racional positivo con cuadrado menor que 2 ni un primero con cuadrado mayor que 2. En la prueba que acabamos de cerrar, las mismas identidades desempeñan una función lógica más precisa. La transformación no intenta **encontrar** un supremo racional. Hace exactamente lo contrario: permite **refutar cualquier candidato racional** según el lado del nivel 2 en que se encuentre. Supongamos que $s \in \mathbb{Q}$ pretende satisfacer las dos cláusulas de $s = \sup_{\mathbb{Q}} S_{\mathbb{Q}}$.

Hay tres posibilidades.

1. Si $s^2 < 2$, falla la condición de cota superior. La transformación produce

$$T(s) \in S_{\mathbb{Q}} \qquad \text{y} \qquad T(s) > s.$$

Así existe un elemento permitido situado por encima de s. En símbolos,

$$\boxed{s^2 < 2 \quad \Longrightarrow \quad \exists q \in S_{\mathbb{Q}} \text{ tal que } q > s.}$$

Por tanto, s no puede dominar a todos los elementos de $S_{\mathbb{Q}}$.

2. Si $s^2 > 2$, falla la minimalidad. En este caso la transformación produce un racional $T(s)$ tal que $T(s) < s$ y, además,

$$\forall q \in S_{\mathbb{Q}}, \qquad q \leq T(s).$$

Es decir, $T(s)$ es una cota superior racional estrictamente menor que s. Por tanto,

$$\boxed{s^2 > 2 \quad \Longrightarrow \quad \exists u \in \mathbb{Q} \text{ cota superior de } S_{\mathbb{Q}} \text{ con } u < s.}$$

Así, aun cuando s fuese una cota superior, no podría ser la menor.

3. Si $s^2 = 2$, falla la racionalidad del candidato. Este caso no lo descarta la transformación, sino Proposición 1.9:

$$\boxed{\forall q \in \mathbb{Q}, \qquad q^2 \neq 2.}$$

Por eso los tres casos atacan exactamente las tres posibilidades abiertas por la tricotomía, pero no todos con la misma herramienta: T elimina las dos desigualdades estrictas y la inexistencia racional de una raíz elimina la igualdad. Podemos condensar la arquitectura así:

$$\boxed{\begin{aligned} s^2 < 2 &\implies s \text{ no es cota superior,} \\ s^2 > 2 &\implies s \text{ no es la menor cota superior,} \\ s^2 = 2 &\implies \text{contradicción con } s \in \mathbb{Q}. \end{aligned}}$$

En consecuencia, **todo candidato racional falla** antes de poder satisfacer simultáneamente las dos cláusulas de supremo. Esta es la función estructural de la transformación de Rudin: no aproxima por sí sola un objeto cuya existencia ya conozcamos, ni produce una frontera racional. Es un **mecanismo de exclusión** que muestra cómo fracasa cada candidato posible dentro de $\mathbb{Q}$. El resultado tiene además una consecuencia lógica que habíamos dejado pendiente. Ya sabemos que $\mathbb{Q}$ es un cuerpo ordenado, pero acabamos de exhibir en él un conjunto no vacío y acotado superiormente que no posee supremo racional. Por tanto,

$$\boxed{\text{cuerpo} + \text{orden} \not\Longrightarrow \text{completitud.}}$$

Así, la completitud contiene información estructural genuinamente adicional. El «hueco» no consiste simplemente en que «faltan irracionales», sino en que $\mathbb{Q}$ no garantiza la existencia interna de ciertas fronteras que su propio orden permite describir.

El mismo problema dentro de $\mathbb{R}$. Cambiemos ahora el sistema ambiente. Consideremos $S_{\mathbb{R}} = \{x \in \mathbb{R} : x \geq 0,\ x^2 < 2\}$. La desigualdad que define el conjunto es la misma que antes, pero ahora los candidatos y las cotas pertenecen a $\mathbb{R}$. Por eso podemos intentar aplicar el axioma de completitud. Antes debemos verificar sus dos hipótesis.

1. $S_{\mathbb{R}}$ es no vacío. El número 1 es real y satisface

$$1 \geq 0 \qquad \text{y} \qquad 1^2 = 1 < 2.$$

Por tanto,

$$\boxed{1 \in S_{\mathbb{R}}},$$

y así $S_{\mathbb{R}} \neq \varnothing$.

2. $S_{\mathbb{R}}$ está acotado superiormente. Mostremos que 2 es una cota superior. Sea $x \in S_{\mathbb{R}}$. Entonces

$$x \geq 0 \qquad \text{y} \qquad x^2 < 2.$$

Si ocurriera $x \geq 2$, tendríamos $0 \leq 2 \leq x$.

La monotonía del cuadrado en los no negativos, Proposición 1.5, daría $4 = 2^2 \leq x^2$, en contradicción con $x^2 < 2$. Por tanto, $x < 2$, y en particular $x \leq 2$.

Como $x \in S_{\mathbb{R}}$ fue arbitrario,

$$\boxed{\forall x \in S_{\mathbb{R}}, \qquad x \leq 2.}$$

Así, 2 es una cota superior real de $S_{\mathbb{R}}$. Hemos verificado exactamente las hipótesis del axioma:

$$S_{\mathbb{R}} \neq \varnothing \qquad \text{y} \qquad S_{\mathbb{R}} \text{ está acotado superiormente en } \mathbb{R}.$$

Por completitud existe un número real

$$\boxed{\alpha = \sup S_{\mathbb{R}}.}$$

EJEMPLO 1.7. **La completitud produce la frontera antes de que sepamos identificarla.** La afirmación $\alpha = \sup S_{\mathbb{R}}$ significa que $\alpha \in \mathbb{R}$ satisface simultáneamente

$$\forall x \in S_{\mathbb{R}}, \qquad x \leq \alpha,$$

y

$$\forall u \in \mathbb{R}, \qquad \left[(\forall x \in S_{\mathbb{R}},\ x \leq u) \Longrightarrow \alpha \leq u\right].$$

La completitud garantiza la **existencia** de ese número. La unicidad procede de la unicidad general del supremo demostrada en §1.3.

Podemos extraer inmediatamente dos cotas para α. Como $1 \in S_{\mathbb{R}}$ y α es una cota superior, $1 \leq \alpha$.

Por otra parte, 2 es una cota superior de $S_{\mathbb{R}}$ y α es la menor de todas ellas. Por tanto, $\alpha \leq 2$.

Así,

$$\boxed{1 \leq \alpha \leq 2.}$$

En este punto sabemos rigurosamente que existe una frontera real y sabemos dónde está localizada. Pero completitud no nos ha entregado todavía la propiedad que motivó toda la construcción.

No hemos demostrado aún $\alpha^2 = 2$.

Tampoco tenemos derecho todavía a escribir $\alpha = \sqrt{2}$, porque la notación $\sqrt{2}$ como nombre de una raíz real positiva solo quedará legitimada después de probar esa existencia e identificación. El trabajo se divide, por tanto, en dos etapas distintas:

$$\boxed{\begin{array}{c}\text{completitud}\\ \downarrow \\ \text{existe } \alpha = \sup S_{\mathbb{R}} \\ \downarrow \\ \text{demostrar que } \alpha^2 = 2.\end{array}}$$

La primera etapa acaba de cerrarse. La segunda será el objetivo de §1.6, y requerirá un argumento adicional de álgebra y orden: la completitud fabrica el candidato, pero no lo identifica automáticamente.

El sistema ambiente importa. Conviene separar dos cambios que podrían confundirse. En el apartado anterior pasamos de $S_{\mathbb{Q}}$ a $S_{\mathbb{R}}$ y, al mismo tiempo, pasamos de trabajar en $\mathbb{Q}$ a trabajar en $\mathbb{R}$. Para aislar el papel del sistema ambiente, mantengamos ahora **fijo el conjunto** $S_{\mathbb{Q}} = \{q \in \mathbb{Q} : q \geq 0,\ q^2 < 2\}$ y cambiemos únicamente el universo en el que buscamos sus cotas. Si trabajamos dentro de $\mathbb{Q}$, las cotas superiores admisibles deben ser racionales. Podemos reunirlas en $U_{\mathbb{Q}} = \{u \in \mathbb{Q} : \forall q \in S_{\mathbb{Q}},\ q \leq u\}$. Preguntar por $\sup_{\mathbb{Q}} S_{\mathbb{Q}}$ significa preguntar si $U_{\mathbb{Q}}$ posee un menor elemento. Acabamos de demostrar que no lo posee.

Pero como $\mathbb{Q} \subset \mathbb{R}$, podemos considerar exactamente los mismos elementos de $S_{\mathbb{Q}}$ como un subconjunto de $\mathbb{R}$. Entonces cambia el universo de cotas posibles: $U_{\mathbb{R}} = \{u \in \mathbb{R} : \forall q \in S_{\mathbb{Q}},\ q \leq u\}$.

El conjunto $S_{\mathbb{Q}}$ sigue siendo no vacío y 2 sigue siendo una cota superior. Por completitud de $\mathbb{R}$ existe, por tanto,

$$\boxed{\beta = \sup_{\mathbb{R}} S_{\mathbb{Q}}.}$$

Este nuevo supremo no puede ser racional. En efecto, supongamos que $\beta \in \mathbb{Q}$. Como β es una cota superior real de $S_{\mathbb{Q}}$, sería también una cota superior racional, es decir, $\beta \in U_{\mathbb{Q}}$. Además, si $u \in U_{\mathbb{Q}}$, entonces u es también una cota superior real; por la minimalidad de β en $\mathbb{R}$, $\beta \leq u$. Así, β sería el menor elemento de $U_{\mathbb{Q}}$, contradiciendo que $S_{\mathbb{Q}}$ no posee supremo en $\mathbb{Q}$. Luego

$$\boxed{\beta \notin \mathbb{Q}.}$$

No hay contradicción entre

$$\sup_{\mathbb{Q}} S_{\mathbb{Q}} \quad \text{no existe}$$

y

$$\sup_{\mathbb{R}} S_{\mathbb{Q}} = \beta \quad \text{sí existe.}$$

Son dos preguntas diferentes porque las cotas admisibles pertenecen a sistemas distintos. Un candidato real no racional sencillamente **no estaba disponible** en el primer problema. Tampoco necesitamos identificar todavía β con el número $\alpha = \sup S_{\mathbb{R}}$ obtenido en el apartado anterior. Esa sería una afirmación adicional acerca de la relación entre los racionales y los reales. Para el argumento presente basta algo más elemental y más importante: **el conjunto se ha mantenido fijo y solo ha cambiado el sistema ambiente**. Por eso, cuando exista ambigüedad, escribiremos si hace falta

$$\sup_{\mathbb{Q}} A \qquad \text{o} \qquad \sup_{\mathbb{R}} A.$$

En nuestro desarrollo habitual, una vez fijado $A \subseteq \mathbb{R}$, la notación abreviada $\sup A$ significará siempre el supremo calculado en $\mathbb{R}$.

Antes de seguir.

Antes de seguir
1. Cambio de ambiente y retorno. Sea $A \subseteq \mathbb{Q}$ no vacío. Supón que, considerado como subconjunto de $\mathbb{R}$, posee supremo $s = \sup_{\mathbb{R}} A$ y que además $s \in \mathbb{Q}$.
Demuestra, sin utilizar densidad, que A posee supremo en $\mathbb{Q}$ y que

$$\boxed{\sup_{\mathbb{Q}} A = s.}$$

Respuesta. Debemos verificar las dos cláusulas del supremo, ahora dentro de $\mathbb{Q}$.
Como $s = \sup_{\mathbb{R}} A$, en particular

$$\forall a \in A, \qquad a \leq s.$$

Además, $s \in \mathbb{Q}$. Por tanto, s es una **cota superior racional** de A. Sea ahora $u \in \mathbb{Q}$ cualquier otra cota superior racional de A. Entonces

$$\forall a \in A, \qquad a \leq u.$$

Pero $u \in \mathbb{R}$, de modo que u es también una cota superior real de A. Como s es la menor cota superior en $\mathbb{R}$, $s \leq u$. Así, s es menor o igual que toda cota superior racional. Reuniendo ambas condiciones,

$$\boxed{s = \sup_{\mathbb{Q}} A.}$$

El argumento utiliza únicamente la definición de supremo y la inclusión $\mathbb{Q} \subseteq \mathbb{R}$; no necesita ningún resultado de densidad.

Antes de seguir

2. Reflejar un conjunto para aplicar completitud. Sea $A \subseteq \mathbb{R}$ no vacío y acotado inferiormente, y define $-A = \{-a : a \in A\}$.

Sin citar Proposición 1.10, verifica que $-A$ satisface las hipótesis del axioma de completitud y determina qué número aparece naturalmente como candidato para ínf A.

Respuesta. Como $A \neq \varnothing$, podemos escoger $a_0 \in A$.

Entonces $-a_0 \in -A$, así que $-A \neq \varnothing$.

Como A está acotado inferiormente, existe $m \in \mathbb{R}$ tal que

$$\forall a \in A, \qquad m \leq a.$$

Al multiplicar por -1 se invierte el orden: $-a \leq -m$.

Por tanto,

$$\forall (-a) \in -A, \qquad -a \leq -m,$$

y $-m$ es una cota superior de $-A$.

Ya podemos aplicar completitud:

$$\boxed{s = \sup(-A) \text{ existe en } \mathbb{R}.}$$

Como la reflexión $x \mapsto -x$ invierte el orden, el candidato natural para la mayor cota inferior de A es

$$\boxed{-s = -\sup(-A).}$$

Aquí solo hemos fabricado y motivado el candidato. La demostración completa de que realmente satisface las dos cláusulas del ínfimo reaparecerá más adelante como ejercicio de transferencia.

La diferencia decisiva. A esta altura ya podemos aislar con precisión qué aporta la completitud. Los axiomas de cuerpo y de orden permiten operar, comparar, definir cotas y formular qué significaría ser una menor cota superior. Pero no obligan a que esa frontera pertenezca al sistema. El ejemplo racional lo demuestra. El conjunto $S_{\mathbb{Q}} = \{q \in \mathbb{Q} : q \geq 0,\ q^2 < 2\}$ es no vacío y está acotado superiormente en $\mathbb{Q}$, pero no posee supremo racional. Por tanto,

$$\boxed{\text{cuerpo ordenado} \not\Longrightarrow \text{completitud.}}$$

La completitud de $\mathbb{R}$ añade precisamente una afirmación de **existencia**:

$$\boxed{A \subseteq \mathbb{R}, \quad A \neq \varnothing, \quad A \text{ acotado superiormente} \quad \Longrightarrow \quad \sup A \in \mathbb{R}.}$$

La diferencia estructural no consiste solo en decir que $\mathbb{R}$ contiene números que $\mathbb{Q}$ no contiene. El punto decisivo es que, en $\mathbb{R}$, determinados problemas de frontera que pueden formularse mediante el orden tienen garantizada una solución interna. Esto permite separar dos trabajos que conviene no confundir:

completitud: existe la frontera
álgebra y orden: identificamos qué propiedad satisface.

En §1.6 aplicaremos exactamente esta arquitectura. Ya sabemos que existe $\alpha = \sup\{x \in \mathbb{R} : x \geq 0,\ x^2 < 2\}$. Lo que falta no es producir otro candidato, sino demostrar que el candidato garantizado por completitud satisface $\alpha^2 = 2$. Esa identificación se obtendrá sin continuidad ni límites, utilizando únicamente las herramientas algebraicas y de orden ya desarrolladas.

1.6. Completitud en acción: existencia de raíces

La frontera ya existe; ahora debemos identificarla. En §1.5 la completitud hizo exactamente el trabajo que le correspondía. Para $S_{\mathbb{R}} = \{x \in \mathbb{R} : x \geq 0,\ x^2 < 2\}$ quedó garantizada la existencia de un número real $\alpha = \sup S_{\mathbb{R}}$. A partir de aquí el problema cambia de naturaleza. Ya no necesitamos producir una frontera: necesitamos **identificarla**. En concreto, debemos demostrar

$$\boxed{\alpha = \sup S_{\mathbb{R}} \quad \Longrightarrow \quad \alpha^2 = 2.}$$

La completitud garantiza que α existe, y deja por determinar qué ecuación satisface. Para identificarlo usaremos el álgebra y el orden junto con las dos condiciones del supremo: es cota superior y ninguna cota superior puede ser menor. No utilizaremos continuidad de $x \mapsto x^2$, ni límites, ni sucesiones convergentes. Esos conceptos todavía no forman parte de nuestra teoría. La estrategia será perturbativa. Si α^2 fuese menor o mayor que 2, intentaremos desplazar ligeramente a α y hacer fracasar una de las dos cláusulas que definen al supremo. La igualdad $\alpha^2 = 2$ aparecerá entonces como la única posibilidad compatible con ambas.

Qué significaría que el supremo estuviese en el lugar equivocado. Para aislar la lógica del argumento, reemplacemos temporalmente 2 por un real positivo arbitrario a y consideremos $S_a = \{x \in \mathbb{R} : x \geq 0,\ x^2 < a\}$. Supongamos que ya hemos verificado las hipótesis necesarias para que exista $\alpha = \sup S_a$. La tricotomía deja exactamente tres posibilidades:

$$\alpha^2 < a, \qquad \alpha^2 = a, \qquad \alpha^2 > a.$$

La igualdad es la conclusión buscada. Las otras dos alternativas deben fracasar por razones distintas, y cada una entra en conflicto con una cláusula diferente de la definición de supremo.

Primer modo de fracaso: α dejaría de ser una cota superior. Si $\alpha^2 < a$, existe un margen positivo $a - \alpha^2 > 0$.

Si podemos elegir $h > 0$ suficientemente pequeño para que $(\alpha + h)^2 < a$, entonces $\alpha + h \in S_a$ y, al mismo tiempo, $\alpha + h > \alpha$. Eso contradice la primera obligación del supremo:

$$\forall x \in S_a, \qquad x \leq \alpha.$$

Segundo modo de fracaso: α dejaría de ser la menor cota superior. Si $\alpha^2 > a$, existe un exceso positivo $\alpha^2 - a > 0$.

Si podemos encontrar $c < \alpha$ tal que $c^2 > a$, entonces ningún elemento de S_a puede alcanzar a c. En efecto, si $x \in S_a$ y $x \geq c$, la monotonía del cuadrado en los no negativos daría $x^2 \geq c^2 > a$, en contradicción con $x^2 < a$. Por tanto, c sería una cota superior de S_a con $c < \alpha$, lo que contradice la segunda obligación del supremo: ser la menor de todas las cotas superiores. La arquitectura queda así completamente separada:

$$\boxed{\begin{array}{lcl} \alpha^2 < a & \Longrightarrow & \text{fracasa la condición de cota superior,} \\ \alpha^2 > a & \Longrightarrow & \text{fracasa la condición de minimalidad.} \end{array}}$$

Una vez construidas las dos perturbaciones, la tricotomía dejará únicamente $\alpha^2 = a$. El trabajo que queda es, por tanto, cuantitativo: elegir las perturbaciones de manera que las desigualdades anteriores queden garantizadas por el álgebra ya disponible.

Cómo se diseña una perturbación controlada. En la prueba que sigue habremos establecido que

$$a > 0 \qquad \text{y} \qquad \alpha > 0.$$

Esto nos permite convertir las dos contradicciones anteriores en elecciones explícitas.

Perturbación hacia la derecha. Supongamos primero $\alpha^2 < a$ y definamos el defecto $\delta = a - \alpha^2 > 0$.

Queremos elegir $h > 0$ de modo que $(\alpha + h)^2 < a$. Como $(\alpha + h)^2 = \alpha^2 + h(2\alpha + h)$, basta controlar el término añadido. Si imponemos $h < \alpha$, entonces $2\alpha + h < 3\alpha$. Por ello es suficiente exigir además $3\alpha h < \delta$. Una sola elección garantiza simultáneamente ambas condiciones:

$$\boxed{h = \frac{1}{2}\,\text{mín}\left\{\alpha, \frac{\delta}{3\alpha}\right\}.}$$

En efecto,

$$0 < h < \alpha \qquad \text{y} \qquad 3\alpha h < \delta,$$

de modo que $(\alpha + h)^2 < \alpha^2 + \delta = a$.

La perturbación produce así un elemento de S_a estrictamente mayor que α.

Perturbación hacia la izquierda. Supongamos ahora $\alpha^2 > a$ y midamos el exceso mediante $E = \alpha^2 - a > 0$.

Buscamos un número $c < \alpha$ cuyo cuadrado siga por encima de a. Definimos

$$\boxed{c = \alpha - \frac{E}{2\alpha} = \frac{1}{2}\left(\alpha + \frac{a}{\alpha}\right).}$$

Como $E > 0$, $c < \alpha$, y como $a, \alpha > 0$, $c > 0$.

Además, el nuevo exceso puede calcularse exactamente:

$$\begin{aligned} c^2 - a &= \left(\alpha - \frac{\alpha^2 - a}{2\alpha}\right)^2 - a \\ &= \frac{(\alpha^2 - a)^2}{4\alpha^2} \\ &> 0. \end{aligned}$$

Por tanto, $c^2 > a$.

Las dos construcciones responden al mismo patrón:

medir el defecto $\longrightarrow$ elegir una corrección $\longrightarrow$ controlar el nuevo defecto.

Ya no queda ninguna elección heurística pendiente. Podemos utilizar estas perturbaciones dentro de la demostración formal.

TEOREMA 1.3. **Existencia y unicidad de la raíz cuadrada no negativa.** Para todo número real $a \geq 0$ existe un único número real $\alpha \geq 0$ tal que

$$\alpha^2 = a.$$

Cuando $a > 0$, este número es positivo.

Idea de la prueba
Para $a > 0$ consideraremos el conjunto de los números no negativos cuyo cuadrado aún queda por debajo de a. La completitud se utilizará una sola vez, para garantizar que ese conjunto tiene supremo. Las dos perturbaciones construidas inmediatamente antes excluirán después las posibilidades $\alpha^2 < a$ y $\alpha^2 > a$. La unicidad se obtendrá al final mediante la monotonía estricta del cuadrado en los números no negativos.

Demostración. Si $a = 0$, el número 0 satisface $0^2 = 0$, de modo que la existencia es inmediata. La unicidad se demostrará al final junto con el caso general.

Supongamos ahora que $a > 0$ y definamos $S_a = \{x \in \mathbb{R} : x \geq 0,\ x^2 < a\}$.

Antes de invocar completitud debemos verificar que S_a es no vacío y está acotado superiormente.

1. S_a es no vacío. Sea $r = \frac{a}{1+a}$.

Como $a > 0$, se tiene

$$0 < r < 1 \qquad \text{y} \qquad r < a.$$

Por tanto, $r^2 < r < a$, y así $r \in S_a$.

2. S_a está acotado superiormente. El número $a + 1$ es una cota superior. En efecto, si $x \in S_a$ y fuese $x \geq a + 1$, como ambos números son no negativos, la monotonía del cuadrado daría $x^2 \geq (a + 1)^2 > a$, en contradicción con $x^2 < a$.

Ahora sí interviene la completitud: como S_a es no vacío y está acotado superiormente, el axioma del supremo garantiza la existencia de $\alpha = \sup S_a$. Este es el único paso de la demostración en el que se usa completitud.

Además, como $r \in S_a$ y $r > 0$, $\alpha \geq r > 0$.

Queda identificar este supremo. Por tricotomía, exactamente una de las relaciones

$$\alpha^2 < a, \qquad \alpha^2 = a, \qquad \alpha^2 > a$$

debe cumplirse.

Caso 1: $\alpha^2 < a$**.** Pongamos $\delta = a - \alpha^2 > 0$ y elijamos, como en la perturbación hacia la derecha construida arriba,

$$h = \frac{1}{2}\,\text{mín}\left\{\alpha, \frac{\delta}{3\alpha}\right\}.$$

La construcción precedente garantiza

$$h > 0 \qquad \text{y} \qquad (\alpha + h)^2 < a.$$

Por tanto, $\alpha + h \in S_a$, pero $\alpha + h > \alpha$. Esto contradice que α sea una cota superior de S_a. Luego $\alpha^2 < a$ es imposible.

Caso 2: $\alpha^2 > a$**.** Definamos, como en la perturbación hacia la izquierda, $c = \alpha - \frac{\alpha^2 - a}{2\alpha} = \frac{1}{2}\left(\alpha + \frac{a}{\alpha}\right)$.

La construcción precedente garantiza

$$0 < c < \alpha \qquad \text{y} \qquad c^2 > a.$$

Veamos que c es una cota superior de S_a. Si $x \in S_a$ y fuese $x \geq c$, entonces, como $x, c \geq 0$, la monotonía del cuadrado implicaría $x^2 \geq c^2 > a$, en contradicción con $x^2 < a$. Por consiguiente, todo $x \in S_a$ satisface $x < c$, de modo que c es una cota superior de S_a. Pero $c < \alpha$, lo cual contradice que α sea la menor cota superior de S_a. Luego $\alpha^2 > a$ es imposible. La tricotomía deja entonces una única posibilidad:

$$\boxed{\alpha^2 = a.}$$

Esto demuestra la existencia. Como además $\alpha > 0$ cuando $a > 0$, queda probada también la afirmación de positividad. Falta la unicidad. Sean $u, v \geq 0$ tales que

$$u^2 = a, \qquad v^2 = a.$$

Si $u < v$, la monotonía estricta del cuadrado en los números no negativos daría $u^2 < v^2$, contradicción. Del mismo modo, $v < u$ es imposible. Por tricotomía, $u = v$. Por consiguiente, para todo $a \geq 0$ existe exactamente una raíz cuadrada no negativa. ■

Dónde entró realmente la completitud. La prueba anterior permite localizar con precisión la dependencia lógica. Primero verificamos, usando solo álgebra y orden, que

$$S_a \neq \varnothing \qquad \text{y} \qquad S_a \text{ está acotado superiormente.}$$

Solo entonces interviene la completitud, y lo hace una única vez:

$$\boxed{\alpha = \sup S_a \text{ existe.}}$$

El axioma no afirma que $\alpha^2 = a$ ni proporciona la unicidad de la raíz. A partir de $\alpha = \sup S_a$, esas tareas se resuelven con las propiedades del supremo, las perturbaciones ya construidas y el orden:

- si $\alpha^2 < a$, una perturbación hacia la derecha produce un elemento de S_a mayor que α, contradiciendo que α sea cota superior;
- si $\alpha^2 > a$, una perturbación hacia la izquierda produce una cota superior menor que α, contradiciendo su minimalidad;
- descartadas ambas desigualdades, la tricotomía obliga a que $\alpha^2 = a$;
- una vez obtenida la existencia, la monotonía estricta del cuadrado en los no negativos da la unicidad.

La arquitectura completa queda así:

$$\boxed{\text{hipótesis del supremo} \xrightarrow{\text{completitud}} \text{existencia de } \alpha \xrightarrow{\text{álgebra + orden}} \alpha^2 = a \xrightarrow{\text{monotonía}} \text{unicidad.}}$$

Este patrón será recurrente en cálculo: una propiedad de existencia fabrica un candidato; después, argumentos adicionales lo identifican y determinan si es único.

Ahora sí podemos definir $\sqrt{a}$. Hasta este punto habíamos evitado utilizar la notación de raíz cuadrada como si la existencia del número designado fuese automática. El teorema anterior elimina esa dificultad: para cada $a \geq 0$ existe un único real no negativo cuyo cuadrado es a.

Definimos, por tanto,

$$\boxed{\sqrt{a}}$$

como **el único número real no negativo** que satisface $(\sqrt{a})^2 = a$.

En particular, $\sqrt{a} \geq 0$.

El símbolo $\sqrt{a}$ designa un solo número; el signo $\pm$ no forma parte de esta definición. Si $a < 0$, ningún real tiene cuadrado igual a a, porque todo cuadrado real es no negativo. Por ello, dentro de $\mathbb{R}$, la notación $\sqrt{a}$ queda reservada para $a \geq 0$. Conviene distinguir ahora la raíz no negativa de las soluciones de la ecuación $x^2 = a$.

Si $a = 0$, la única solución es $x = 0 = \sqrt{0}$. Si $a > 0$, entonces $\sqrt{a} > 0$, y tanto $\sqrt{a}$ como $-\sqrt{a}$ tienen cuadrado a. Recíprocamente, sea $x \in \mathbb{R}$ tal que $x^2 = a$. Como $a > 0$, tenemos $x \neq 0$. Si $x > 0$, la unicidad de la raíz no negativa obliga a que $x = \sqrt{a}$. Si $x < 0$, entonces $-x > 0$ y $(-x)^2 = x^2 = a$, de modo que, por la misma unicidad, $-x = \sqrt{a}$ y por tanto $x = -\sqrt{a}$.

Así, para $a > 0$,

$$\boxed{x^2 = a \iff x = \sqrt{a} \text{ o } x = -\sqrt{a}.}$$

La unicidad del teorema se refiere, pues, a la **raíz cuadrada no negativa**; la ecuación $x^2 = a$ tiene dos soluciones reales distintas cuando $a > 0$.

La raíz que faltaba desde §1.4. Podemos regresar finalmente a la ecuación $x^2 = 2$. Como $2 > 0$, el teorema garantiza que existe un único número real positivo $\sqrt{2}$ tal que $(\sqrt{2})^2 = 2$. La existencia y la notación están ahora justificadas. Además, podemos localizar este número sin recurrir a aproximaciones decimales. Como $1^2 < 2 = (\sqrt{2})^2$, no puede ocurrir $\sqrt{2} \leq 1$: en tal caso, de $0 \leq \sqrt{2} \leq 1$ y la monotonía del cuadrado en los no negativos se seguiría $2 = (\sqrt{2})^2 \leq 1$, una contradicción. Del mismo modo, como

$$(\sqrt{2})^2 = 2 < 4 = 2^2,$$

no puede ocurrir $\sqrt{2} \geq 2$, pues $0 \leq 2 \leq \sqrt{2}$ implicaría $4 \leq (\sqrt{2})^2 = 2$.

Por tanto,

$$\boxed{1 < \sqrt{2} < 2.}$$

En §1.4 demostramos que ningún racional tiene cuadrado igual a 2. Como $\sqrt{2} \in \mathbb{R}$ y $(\sqrt{2})^2 = 2$, necesariamente

$$\boxed{\sqrt{2} \in \mathbb{R} \setminus \mathbb{Q}.}$$

Ahora sí podemos reunir las dos afirmaciones que al comienzo debían mantenerse separadas:

$$\boxed{\begin{array}{c}\text{existe un número real positivo cuyo cuadrado es 2,}\\ \text{y ningún número racional posee esa propiedad.}\end{array}}$$

El «hueco» racional queda así expresado con precisión: no faltaba un intervalo sin números racionales, sino una frontera que $\mathbb{Q}$ no podía realizar internamente. En $\mathbb{R}$, la completitud garantiza la existencia de esa frontera y el argumento anterior la identifica como $\sqrt{2}$.

Una prueba sin continuidad. Más adelante, una vez desarrollada la continuidad, podremos obtener también la existencia de raíces mediante resultados como el teorema del valor intermedio. Aquí esa ruta sería lógicamente prematura: ni la continuidad ni ese teorema forman todavía parte de nuestra teoría. La dependencia construida en este capítulo va en la dirección opuesta:

$$\boxed{\text{completitud} \longrightarrow \text{existencia del supremo} \xrightarrow{\text{álgebra + orden}} \text{existencia de raíces} \longrightarrow \text{cálculo posterior.}}$$

En particular, la demostración no ha utilizado:

- límites;
- convergencia de sucesiones;
- continuidad;
- teorema del valor intermedio;
- Bolzano–Weierstrass.

La raíz se obtuvo únicamente con álgebra del cuerpo, propiedades del orden y completitud. Más adelante podremos demostrar el mismo hecho por otras vías, pero sin invertir retroactivamente esta cadena de dependencias.

Antes de seguir.

Antes de seguir
1. Demuestra que, para todo $x \in \mathbb{R}$,

$$\boxed{\sqrt{x^2} = |x|}.$$

Respuesta. Como $x^2 \geq 0$, la raíz $\sqrt{x^2}$ está definida. Además, $|x| \geq 0$ y $|x|^2 = x^2$.
Así, $|x|$ es un número no negativo cuyo cuadrado es x^2. Por la unicidad de la raíz cuadrada no negativa,

$$\boxed{\sqrt{x^2} = |x|}.$$

En particular, no puede reemplazarse $|x|$ por x sin conocer previamente el signo de x.

Antes de seguir
2. Si $0 \le a < b$, demuestra que $\sqrt{a} < \sqrt{b}$.
Respuesta. Supongamos, por contradicción, que $\sqrt{a} \ge \sqrt{b}$.
Como ambas raíces son no negativas, tendríamos $0 \le \sqrt{b} \le \sqrt{a}$. La monotonía del cuadrado en los no negativos daría entonces

$$b = (\sqrt{b})^2 \le (\sqrt{a})^2 = a,$$

lo cual contradice $a < b$. Por tanto,

$$\boxed{\sqrt{a} < \sqrt{b}}.$$

Del primer hueco a una herramienta permanente. La ecuación $x^2 = 2$ fue el problema que hizo visible la insuficiencia de $\mathbb{Q}$, pero el resultado obtenido no pertenece solo a ese ejemplo. Hemos demostrado que

$$\boxed{\forall a \ge 0 \quad \exists!\, \alpha \ge 0 \quad \alpha^2 = a.}$$

Así, para cada real no negativo existe una raíz cuadrada no negativa bien determinada. La completitud aporta el paso de existencia decisivo: cuando $a > 0$, garantiza el supremo del conjunto adecuado. Después, el álgebra y el orden identifican ese supremo mediante $\alpha^2 = a$, y la monotonía del cuadrado proporciona la unicidad. La sección deja, por tanto, algo más que la existencia de $\sqrt{2}$: deja una arquitectura reutilizable,

$$\boxed{\text{conjunto adecuado} \xrightarrow{\text{completitud}} \text{supremo} \xrightarrow{\text{álgebra + orden}} \text{propiedad buscada.}}$$

En §1.7 aplicaremos la completitud de otra manera. Ya no construiremos la solución de una ecuación: supondremos que $\mathbb{N}$ está acotado superiormente y mostraremos que la existencia de un supremo conduce a una contradicción. Obtendremos así la **propiedad arquimediana**: ningún número real puede ser una cota superior de todos los naturales.

1.7. La propiedad arquimediana

¿Pueden los naturales quedar atrapados bajo un número real? El conjunto de los naturales ofrece otra ocasión de usar la completitud. Si tuviera una cota superior real, tendría también un supremo; veremos que la posibilidad de pasar de cada natural a su sucesor impide que ese supremo sea una cota. La contradicción probará que ninguna cota real puede dominar a todos los naturales. La afirmación que queremos obtener es familiar desde la aritmética:

por grande que sea un número real, siempre existe un número natural mayor.

Lo que debemos demostrar no es una colección de casos particulares, sino la afirmación universal

$$\forall x \in \mathbb{R}\ \exists n \in \mathbb{N} \quad n > x.$$

En nuestro desarrollo esta propiedad no se dará por supuesta: la deduciremos de la completitud.

Teorema 1.4. **Propiedad arquimediana.** El conjunto $\mathbb{N}$ no está acotado superiormente en $\mathbb{R}$. Equivalentemente, para todo $x \in \mathbb{R}$ existe $n \in \mathbb{N}$ tal que

$$n > x.$$

Idea de la prueba
Supongamos, por contradicción, que $\mathbb{N}$ está acotado superiormente. Como es no vacío, la completitud produciría $\alpha = \sup \mathbb{N}$. La caracterización aproximativa del supremo, aplicada con $\varepsilon = 1$, daría un natural n situado a menos de una unidad por debajo de α: $\alpha - 1 < n \leq \alpha$. Pero entonces $n + 1$ seguiría siendo natural y satisfaría $n + 1 > \alpha$, contradiciendo que α fuese una cota superior.

Demostración. Supongamos, para obtener una contradicción, que $\mathbb{N}$ está acotado superiormente en $\mathbb{R}$. Como $\mathbb{N}$ es no vacío, el axioma de completitud garantiza la existencia de $\alpha = \sup \mathbb{N}$.

Puesto que $1 > 0$, podemos aplicar la caracterización aproximativa del supremo, Proposición 1.8, con $\varepsilon = 1$. Existe entonces $n \in \mathbb{N}$ tal que $\alpha - 1 < n \leq \alpha$. De la primera desigualdad, sumando 1, obtenemos $\alpha < n + 1$. Pero $n + 1 \in \mathbb{N}$. Como α es una cota superior de $\mathbb{N}$, también debería cumplirse $n + 1 \leq \alpha$, lo cual es imposible. Por tanto, $\mathbb{N}$ no está acotado superiormente en $\mathbb{R}$. ■

Después de la prueba
La dependencia es especialmente limpia:

- **completitud:** bajo la hipótesis de acotación, produce $\alpha = \sup \mathbb{N}$;
- **propiedad aproximativa del supremo:** con $\varepsilon = 1$, produce un natural n con $\alpha - 1 < n$;
- **aritmética de $\mathbb{N}$:** de $n \in \mathbb{N}$ se obtiene $n + 1 \in \mathbb{N}$;
- **contradicción:** ese sucesor satisface $n + 1 > \alpha$, aunque α debía dominar a todos los naturales.

Aquí la completitud no identifica un objeto que queramos conservar: fabrica, bajo una hipótesis falsa, un supremo cuya existencia hace visible la contradicción.

Dos maneras de leer la misma propiedad. Decir que $\mathbb{N}$ no está acotado superiormente equivale exactamente a afirmar

$$\boxed{\forall x \in \mathbb{R}\ \exists n \in \mathbb{N} \qquad n > x.}$$

Veamos ambas direcciones. Si $\mathbb{N}$ no está acotado superiormente y fijamos $x \in \mathbb{R}$, entonces x no puede ser una cota superior de $\mathbb{N}$. Por definición de cota, debe existir algún $n \in \mathbb{N}$ tal que $n > x$. Recíprocamente, supongamos que para todo $x \in \mathbb{R}$ existe $n \in \mathbb{N}$ con $n > x$. Si $\mathbb{N}$ tuviera una cota superior $M \in \mathbb{R}$, al particularizar la propiedad en $x = M$ obtendríamos un natural $n > M$, contradiciendo que M domine a todos los naturales. La forma cuantificada es la que utilizaremos cuando necesitemos elegir un natural **suficientemente grande**. De ella se obtiene inmediatamente una versión escalada. Sean

$$u > 0, \qquad v \in \mathbb{R}.$$

Definamos $R = \max\left\{\frac{v}{u}, 0\right\}$.

Por la propiedad arquimediana existe $n \in \mathbb{N}$ tal que $n > R$. Como $R \geq 0$, se tiene $n > 0$, de modo que $n \in \mathbb{N}_{>0}$. Además, $n > R \geq \frac{v}{u}$. Multiplicando por $u > 0$ conservamos el orden y obtenemos $nu > v$.

Por tanto,

$$\boxed{\forall u > 0\ \forall v \in \mathbb{R}\ \exists n \in \mathbb{N}_{>0} \qquad nu > v.}$$

Así, cualquier escala positiva fija puede hacerse superar cualquier umbral real mediante un múltiplo natural suficientemente grande.

Naturales grandes, recíprocos pequeños. La propiedad arquimediana también permite fabricar escalas positivas tan pequeñas como se desee.

Corolario 1.2. **Recíprocos arbitrariamente pequeños.** Para todo $\varepsilon > 0$ existe $n \in \mathbb{N}_{>0}$ tal que

$$\frac{1}{n} < \varepsilon.$$

Demostración. Sea $\varepsilon > 0$. Entonces $\frac{1}{\varepsilon} > 0$.

Por la propiedad arquimediana existe $n \in \mathbb{N}$ tal que $n > \frac{1}{\varepsilon}$. Como el miembro derecho es positivo, también $n > 0$, de modo que $n \in \mathbb{N}_{>0}$. Además, ambos miembros de la desigualdad anterior son positivos; al tomar recíprocos se invierte el orden y obtenemos $\frac{1}{n} < \varepsilon$.

Esto prueba el resultado. ■

La afirmación completa es

$$\boxed{\forall \varepsilon > 0\ \exists n \in \mathbb{N}_{>0} \qquad \frac{1}{n} < \varepsilon.}$$

El natural se elige **después** de fijar ε y, por tanto, puede depender de esa tolerancia. Al disminuir ε, puede ser necesario escoger un natural mayor. Este patrón de cuantificadores aparecerá repetidamente en cálculo.

Lectura de la fórmula
Para $n \in \mathbb{N}_{>0}$ y $\varepsilon > 0$,

$$\boxed{\frac{1}{n} < \varepsilon \iff n > \frac{1}{\varepsilon}.}$$

Así, si queremos garantizar $\frac{1}{n} < 10^{-6}$, basta elegir $n > 10^6$. Por ejemplo, $n = 1\,000\,001$ satisface la condición.

En $\mathbb{R}$ no hay infinitésimos positivos. El corolario anterior puede leerse también como una afirmación de exclusión.

Corolario 1.3. **Ausencia de infinitésimos reales positivos.** No existe $\eta \in \mathbb{R}$ tal que

$$\eta > 0$$

y

$$\eta < \frac{1}{n} \qquad \text{para todo } n \in \mathbb{N}_{>0}.$$

Demostración. Supongamos que existiera tal $\eta > 0$. Aplicando Corolario 1.2 con $\varepsilon = \eta$, existiría $n \in \mathbb{N}_{>0}$ tal que $\frac{1}{n} < \eta$.

La propiedad atribuida a η exigiría para ese mismo natural $\eta < \frac{1}{n}$, lo cual es imposible. Por tanto, no existe tal η. ■

Entre los reales no negativos, la conclusión puede escribirse de manera equivalente como

$$\boxed{\eta \geq 0 \quad \Longrightarrow \quad \left[\left(\forall n \in \mathbb{N}_{>0},\ \eta < \frac{1}{n}\right) \iff \eta = 0\right].}$$

En este pasaje, **infinitésimo** designa únicamente la propiedad hipotética de ser un real positivo menor que todos los recíprocos $1/n$. No introducimos aquí una clase adicional de números ni una teoría de infinitesimales. Este resultado no requiere una nueva aplicación de completitud: es una consecuencia inmediata de la propiedad arquimediana, a través de Corolario 1.2.

La propiedad arquimediana no es lo mismo que la completitud. Hemos deducido la propiedad arquimediana a partir de la completitud. Esta implicación no debe confundirse con una equivalencia entre ambas propiedades de un cuerpo ordenado. El contraejemplo ya está disponible: $\mathbb{Q}$. Sea $q \in \mathbb{Q}$. Como $\mathbb{Q} \subseteq \mathbb{R}$, el teorema arquimediano Teorema 1.4, aplicado al real q, proporciona un natural n tal que $n > q$. Por tanto, $\mathbb{N}$ no está acotado superiormente dentro de $\mathbb{Q}$: el cuerpo ordenado $\mathbb{Q}$ es arquimediano.

Por otra parte, Proposición 1.11 mostró que $S_{\mathbb{Q}} = \{q \in \mathbb{Q} : q \geq 0,\ q^2 < 2\}$ es no vacío y está acotado superiormente en $\mathbb{Q}$, pero no posee supremo en $\mathbb{Q}$. Así, $\mathbb{Q}$ no es completo.

En consecuencia,

$$\boxed{\text{completitud} \Longrightarrow \text{propiedad arquimediana}, \qquad \text{propiedad arquimediana} \not\Longrightarrow \text{completitud}.}$$

Las dos propiedades tienen alcances distintos. La arquimedianidad controla la escala del orden: los naturales superan cualquier elemento y los recíprocos $1/n$ pueden hacerse menores que cualquier cantidad positiva. La completitud añade un principio de existencia para fronteras: garantiza supremos bajo las hipótesis correspondientes. El cuerpo $\mathbb{Q}$ satisface la primera propiedad y falla la segunda.

Encajonar un real entre dos enteros consecutivos. Para la siguiente sección necesitaremos convertir la propiedad arquimediana en una herramienta de localización respecto de los enteros.

Lema 1.1. **Encajonamiento entero.** Para todo $x \in \mathbb{R}$ existe $m \in \mathbb{Z}$ tal que

$$m \leq x < m + 1.$$

Usaremos una consecuencia elemental del **principio de inducción**: todo subconjunto no vacío de $\mathbb{N}$ posee un elemento mínimo. Para justificarla localmente, recordemos que la inducción afirma que, si una propiedad $P(n)$ vale para $n = 0$ y $P(k)$ implica $P(k+1)$ para todo $k \in \mathbb{N}$, entonces vale para todo $n \in \mathbb{N}$.

Supongamos que un conjunto no vacío $A \subseteq \mathbb{N}$ no tuviera mínimo. Entonces $0 \notin A$. Si ninguno de $0, 1, \ldots, k$ pertenece a A, tampoco puede pertenecer $k+1$: en tal caso, $k+1$ sería el mínimo de A. Por inducción, ningún natural pertenecería a A, contradicción. Queda así justificado el principio de buena ordenación que utilizaremos a continuación.

Demostración. Sea $x \in \mathbb{R}$. Por la propiedad arquimediana podemos elegir $k \in \mathbb{N}$ tal que $k > |x|$.

Entonces $k > 0$ y, como $x \geq -|x|$, $x + k \geq k - |x| > 0$.

Consideremos $A = \{n \in \mathbb{N} : n > x + k\}$.

La propiedad arquimediana, aplicada a $x+k$, garantiza que A es no vacío. Por buena ordenación, existe $n_0 = \min A$. Como $x + k > 0$, tenemos $0 \notin A$ y por tanto $n_0 \geq 1$. En particular, $n_0 - 1 \in \mathbb{N}$. La minimalidad de n_0 implica que $n_0 - 1 \notin A$, de modo que $n_0 - 1 \leq x + k$.

Por otra parte, $n_0 \in A$, así que $x + k < n_0$.

Reuniendo ambas desigualdades, $n_0 - 1 \leq x + k < n_0$.

Restando k, $n_0 - k - 1 \leq x < n_0 - k$.

Definamos $m = n_0 - k - 1 \in \mathbb{Z}$.

Entonces

$$\boxed{m \leq x < m + 1},$$

como queríamos demostrar. ■

Lectura del lema
El argumento combina dos hechos distintos. La propiedad arquimediana garantiza que siempre hay naturales a la derecha de un real; la buena ordenación permite escoger el **primero** de esos naturales después de trasladar el problema a una posición positiva. Más adelante el entero m quedará asociado a la función piso, $\lfloor x \rfloor$, pero aquí solo necesitamos su existencia. Esa existencia será el engranaje que permitirá construir un racional entre dos reales cualesquiera.

Preparación para la densidad. Supongamos que $a < b$.

Entonces $b - a > 0$.

Por Corolario 1.2 podemos elegir $n \in \mathbb{N}_{>0}$ de modo que $\frac{1}{n} < b - a$. Como $n > 0$, esta desigualdad equivale a $1 < n(b-a) = nb - na$, y por tanto $na + 1 < nb$. Apliquemos ahora Lema 1.1 al real na. Existe $k \in \mathbb{Z}$ tal que $k \leq na < k + 1$.

Si definimos $m = k + 1$, entonces $m \in \mathbb{Z}$ y

$$na < m = k + 1 \leq na + 1 < nb.$$

Así obtenemos $na < m < nb$.

Al dividir por $n > 0$ resulta $a < \frac{m}{n} < b$, y $m/n \in \mathbb{Q}$.

Esta es la construcción que formalizaremos como teorema en §1.8. Conviene retener la arquitectura antes que las letras concretas:

propiedad arquimediana
$\longrightarrow$ escala $1/n$ menor que $b - a$
$\longrightarrow$ intervalo escalado de longitud > 1
$\longrightarrow$ entero intermedio
$\longrightarrow$ racional intermedio.

Antes de seguir.

Antes de seguir
1. Sean $k > 0$ y $\varepsilon > 0$. Demuestra que existe $n \in \mathbb{N}_{>0}$ tal que $\frac{k}{n} < \varepsilon$.
Respuesta. Como $k/\varepsilon > 0$, la propiedad arquimediana permite elegir $n \in \mathbb{N}$ tal que $n > \frac{k}{\varepsilon}$. En particular, $n > 0$, así que $n \in \mathbb{N}_{>0}$. Multiplicando por $\varepsilon > 0$, $n\varepsilon > k$, y dividiendo por $n > 0$,

$$\boxed{\frac{k}{n} < \varepsilon}.$$

Antes de seguir
2. Supón que $x \geq 0$ y que existe $k > 0$ tal que

$$x \leq \frac{k}{n} \qquad \text{para todo } n \in \mathbb{N}_{>0}.$$

Demuestra que $x = 0$.
Respuesta. Supongamos, por contradicción, que $x > 0$. Como $k/x > 0$, la propiedad arquimediana permite elegir $n \in \mathbb{N}$ tal que $n > \frac{k}{x}$.
Entonces $n > 0$. Multiplicando por $x > 0$, $nx > k$, y dividiendo por $n > 0$, $x > \frac{k}{n}$. Esto contradice la hipótesis, que exige $x \leq k/n$ para todo $n \in \mathbb{N}_{>0}$. Por tanto, $x > 0$ es imposible. Como $x \geq 0$, la tricotomía da

$$\boxed{x = 0}.$$

Lo que exporta esta sección. La sección ha convertido la propiedad arquimediana en tres herramientas concretas:

$$\boxed{\begin{array}{c} \forall x \in \mathbb{R}\ \exists n \in \mathbb{N} :\ n > x, \\ \forall \varepsilon > 0\ \exists n \in \mathbb{N}_{>0} :\ \dfrac{1}{n} < \varepsilon, \\ \forall x \in \mathbb{R}\ \exists m \in \mathbb{Z} :\ m \leq x < m + 1. \end{array}}$$

Su dependencia interna es precisa. La primera es la propiedad arquimediana. La segunda se deduce directamente de ella aplicándola a $1/\varepsilon$. La tercera combina arquimedianidad con la buena ordenación de $\mathbb{N}$ para seleccionar el primer natural adecuado después de una traslación. Podemos resumir esa arquitectura como

$$\boxed{\text{completitud} \longrightarrow \text{propiedad arquimediana} \longrightarrow \begin{cases} \text{naturales suficientemente grandes,} \\ \text{recíprocos } 1/n \text{ arbitrariamente pequeños,} \end{cases}}$$

mientras que

$$\boxed{\text{propiedad arquimediana} + \text{ buena ordenación de } \mathbb{N} \longrightarrow \text{encajonamiento entero.}}$$

En §1.8 combinaremos las dos últimas herramientas. Si $a < b$, elegiremos n con $1/n < b - a$; después, el encajonamiento entero producirá un entero m con $na < m < nb$, y por tanto un racional m/n dentro de (a, b). Una vez establecida la densidad racional, la existencia del

irracional $\sqrt{2}$ permitirá obtener también densidad irracional mediante traslación. La elección del denominador controla la anchura de la escala; la elección del entero sitúa el punto en el intervalo. Éstas serán las dos decisiones que sostengan la prueba de densidad.

1.8. Entre dos reales siempre hay más números

Qué significa que un conjunto sea denso. Un intervalo puede ser muy estrecho y seguir conteniendo puntos de un conjunto dado. La propiedad de densidad exige que esto ocurra en todo intervalo abierto no vacío, cualquiera que sea su posición o su anchura.

Definición 1.8. **Densidad en la recta real.** Sea $D \subseteq \mathbb{R}$. Diremos que D es **denso en** $\mathbb{R}$ si, para cualesquiera $a, b \in \mathbb{R}$ con $a < b$, existe $d \in D$ tal que

$$a < d < b.$$

Equivalentemente, todo intervalo abierto no vacío contiene algún elemento de D:

$$\boxed{\forall a, b \in \mathbb{R}, \quad a < b \Longrightarrow D \cap (a, b) \neq \varnothing.}$$

Las dos formulaciones expresan la misma condición: encontrar $d \in D$ con $a < d < b$ equivale a encontrar un elemento de la intersección $D \cap (a, b)$. La palabra *denso* describe un comportamiento local respecto de los intervalos. Un conjunto denso puede omitir muchos puntos; la definición exige que ninguno de esos puntos ausentes abra un intervalo no vacío completamente separado del conjunto. En símbolos, si $a < b$, no puede ocurrir $D \cap (a, b) = \varnothing$. En esta sección demostraremos dos hechos que conviene mantener simultáneamente: $\mathbb{Q}$ es denso en $\mathbb{R}$ y $\mathbb{R} \setminus \mathbb{Q}$ es denso en $\mathbb{R}$. Así, todo intervalo abierto no vacío contiene racionales e irracionales. La anchura del intervalo puede ser tan pequeña como se quiera mientras sus extremos sigan siendo distintos.

Densidad de los racionales.

Teorema 1.5. **Densidad de $\mathbb{Q}$ en $\mathbb{R}$.** Si $a, b \in \mathbb{R}$ y $a < b$, entonces existe $r \in \mathbb{Q}$ tal que

$$a < r < b.$$

Idea de la prueba
Buscamos un racional de la forma m/n, con $m \in \mathbb{Z}$ y $n \in \mathbb{N}_{>0}$. Como $n > 0$,

$$a < \frac{m}{n} < b \iff na < m < nb.$$

La estrategia consiste en elegir n de modo que el intervalo escalado (na, nb) tenga anchura mayor que 1. Por Corolario 1.2 podemos imponer $\frac{1}{n} < b - a$, lo que equivale a $nb - na > 1$. Después, Lema 1.1 aplicado a na produce un entero inmediatamente a su derecha; la desigualdad anterior garantiza que ese entero todavía queda a la izquierda de nb.

Demostración. Sean $a, b \in \mathbb{R}$ con $a < b$. Entonces $b - a > 0$. Por Corolario 1.2 existe $n \in \mathbb{N}_{>0}$ tal que $\frac{1}{n} < b - a$.

Como $n > 0$, multiplicando por n obtenemos $1 < n(b - a) = nb - na$, y por tanto $na + 1 < nb$. Apliquemos Lema 1.1 al real na. Existe $k \in \mathbb{Z}$ tal que $k \leq na < k + 1$.

Definamos $m = k + 1$.

Entonces $m \in \mathbb{Z}$ y

$$na < m = k + 1 \leq na + 1 < nb.$$

En particular, $na < m < nb$.

Dividiendo por $n > 0$ conservamos el orden: $a < \frac{m}{n} < b$.

Como $m \in \mathbb{Z}$ y $n \in \mathbb{N}_{>0}$, $r := \frac{m}{n} \in \mathbb{Q}$.

Por consiguiente,

$$\boxed{a < r < b}.$$

Esto demuestra la densidad de $\mathbb{Q}$ en $\mathbb{R}$. ■

Después de la prueba

Los extremos a y b son reales arbitrarios; no se ha supuesto que sean racionales. La construcción introduce racionalidad únicamente al elegir $r = \frac{m}{n}$. La prueba puede reconstruirse mediante cuatro movimientos:

1. **medir el intervalo:** $b - a > 0$;
2. **elegir una escala:** $1/n < b - a$;
3. **insertar un entero:** $na < m < nb$;
4. **desescalar:** $a < m/n < b$.

La dependencia estructural también queda localizada. Corolario 1.2 procede de la propiedad arquimediana; Lema 1.1 combina arquimedianidad con la buena ordenación de $\mathbb{N}$. En la arquitectura adoptada por este capítulo, la completitud interviene por tanto de manera indirecta, a través de §1.7.

Un ejemplo construido, no adivinado.

Ejemplo 1.8. **Un racional entre $\sqrt{2}$ y $3/2$.** Construyamos un número racional r que satisfaga $\sqrt{2} < r < \frac{3}{2}$.

Primero necesitamos una escala $1/n$ menor que el hueco $\frac{3}{2} - \sqrt{2}$. La elección $n = 12$ funciona porque $\frac{1}{12} < \frac{3}{2} - \sqrt{2}$ es equivalente a $\sqrt{2} < \frac{17}{12}$, y esta última desigualdad puede verificarse elevando al cuadrado números positivos: $2 < \frac{289}{144}$ porque $288 < 289$. Ahora escalamos el extremo izquierdo: $12\sqrt{2}$.

Sabemos que $16 < 12\sqrt{2} < 17$, pues

$$16^2 = 256 < 288 = (12\sqrt{2})^2 < 289 = 17^2.$$

Así, el entero inmediatamente superior a $12\sqrt{2}$ es 17, y al dividir por 12 obtenemos

$$\boxed{\sqrt{2} < \frac{17}{12} < \frac{3}{2}.}$$

El racional no apareció por ensayo decimal: fue producido por la arquitectura de la demostración.

Densidad de los irracionales. Ahora queremos demostrar algo aparentemente más difícil:

entre dos reales cualesquiera existe un irracional.

Podríamos intentar construirlo directamente dentro del intervalo. Hay una ruta más limpia: ya conocemos un irracional concreto, $\sqrt{2}$, y acabamos de demostrar que podemos insertar racionales en cualquier intervalo. La idea será **trasladar** el intervalo.

TEOREMA 1.6. **Densidad de los irracionales en $\mathbb{R}$.** Si $a, b \in \mathbb{R}$ y $a < b$, entonces existe

$$\xi \in \mathbb{R} \setminus \mathbb{Q}$$

tal que

$$a < \xi < b.$$

Idea de la prueba
Restemos $\sqrt{2}$ a ambos extremos. Como $a - \sqrt{2} < b - \sqrt{2}$, la densidad racional permite escoger $r \in \mathbb{Q}$ con $a - \sqrt{2} < r < b - \sqrt{2}$. Al volver a sumar $\sqrt{2}$, el número $\xi = r + \sqrt{2}$ queda dentro de (a, b). Y debe ser irracional: si fuese racional, al restarle el racional r concluiríamos que $\sqrt{2}$ es racional.

Demostración. Sean $a, b \in \mathbb{R}$ con $a < b$.
Restando $\sqrt{2}$ a ambos lados, $a - \sqrt{2} < b - \sqrt{2}$.
Por la densidad de $\mathbb{Q}$ en $\mathbb{R}$, existe $r \in \mathbb{Q}$ tal que $a - \sqrt{2} < r < b - \sqrt{2}$.
Sumando $\sqrt{2}$ obtenemos $a < r + \sqrt{2} < b$.
Definamos $\xi = r + \sqrt{2}$.
Solo falta verificar que ξ es irracional. Supongamos, para obtener una contradicción, que $\xi \in \mathbb{Q}$. Como $r \in \mathbb{Q}$ y los racionales son cerrados bajo la resta, tendríamos $\sqrt{2} = \xi - r \in \mathbb{Q}$, contradiciendo el resultado de §1.4.
Por tanto, $\xi \in \mathbb{R} \setminus \mathbb{Q}$, y además $a < \xi < b$. ■

Lectura de la demostración
La densidad irracional no necesitó una nueva aplicación de completitud ni otra versión de la propiedad arquimediana. Una vez obtenida la densidad racional y conocido un irracional, bastó combinar:

$$\boxed{\text{traslación del intervalo} + \text{densidad racional} + \text{racional} + \text{irracional} = \text{irracional}.}$$

La última afirmación también ha sido demostrada dentro del argumento: si la suma fuera racional, restar el sumando racional volvería racional al irracional original.

Racionales e irracionales dentro del mismo intervalo. Volvamos al intervalo

$$\left(\sqrt{2}, \frac{3}{2}\right).$$

Ya hemos construido el racional $\frac{17}{12}$. También podemos construir inmediatamente un irracional dentro del mismo intervalo. Como ya verificamos que $\frac{1}{12} < \frac{3}{2} - \sqrt{2}$, tenemos $\sqrt{2} < \sqrt{2} + \frac{1}{12} < \frac{3}{2}$. Y $\sqrt{2} + \frac{1}{12}$ es irracional, porque la suma de $\sqrt{2}$ con un racional no puede ser racional. Por tanto, dentro del mismo intervalo encontramos simultáneamente números de ambos tipos. Pero pertenecer al mismo intervalo no determina el orden relativo entre los puntos construidos; debemos compararlos.

En efecto,

$$\frac{17}{12} - \left(\sqrt{2} + \frac{1}{12}\right) = \frac{4}{3} - \sqrt{2} < 0,$$

porque

$$\left(\frac{4}{3}\right)^2 = \frac{16}{9} < 2.$$

Así obtenemos la cadena correcta

$$\boxed{\sqrt{2} < \frac{17}{12} < \sqrt{2} + \frac{1}{12} < \frac{3}{2}.}$$

Error frecuente
Una construcción demuestra pertenencia a un intervalo, pero no autoriza a ordenar entre sí los objetos construidos sin una comparación adicional. Este tipo de auditoría será especialmente importante cuando trabajemos con varias aproximaciones simultáneas.

Arbitrariamente cerca, sin hablar todavía de límites. La densidad puede reescribirse en un lenguaje que conecta directamente con §1.2. Sea $x \in \mathbb{R}$ y sea $\varepsilon > 0$. El intervalo $(x-\varepsilon, x+\varepsilon)$ es abierto y no vacío. Por densidad contiene un racional q y un irracional ξ. Así,

$$|q - x| < \varepsilon, \qquad |\xi - x| < \varepsilon.$$

Si queremos además que los puntos sean distintos de x, podemos aplicar los teoremas al intervalo $(x, x+\varepsilon)$ y obtener

$$0 < q - x < \varepsilon, \qquad 0 < \xi - x < \varepsilon.$$

Por tanto:

$$\boxed{\forall x \in \mathbb{R}\ \forall \varepsilon > 0, \quad \begin{array}{l} \exists q \in \mathbb{Q} :\ 0 < |q - x| < \varepsilon, \\ \exists \xi \in \mathbb{R} \setminus \mathbb{Q} :\ 0 < |\xi - x| < \varepsilon. \end{array}}$$

Esto justifica rigurosamente la expresión «hay racionales e irracionales arbitrariamente cerca de todo real». No estamos afirmando todavía que exista una **sucesión** de racionales o irracionales que converja a x. Esa reformulación pertenece al capítulo de sucesiones. Aquí solo utilizamos cuantificadores, intervalos y distancia.

Todo intervalo abierto contiene infinitos de ambos tipos. Los teoremas de densidad garantizan al menos un punto de cada tipo. En realidad garantizan mucho más. Tomemos un intervalo abierto no vacío (a, b). Por densidad racional existe $r_1 \in \mathbb{Q} \cap (a, b)$. Aplicando otra vez densidad al intervalo (a, r_1) obtenemos un racional distinto r_2. Repitiendo el argumento tantas veces como queramos, para cada $N \in \mathbb{N}_{>0}$ podemos producir N racionales distintos dentro de (a, b). Por tanto, el intervalo contiene infinitos racionales. El mismo argumento, usando densidad de los irracionales en cada subintervalo, muestra que contiene infinitos irracionales.

Así:

todo intervalo abierto no vacío contiene infinitos racionales e infinitos irracionales.

Conviene notar lo que **no** hemos usado para llegar aquí: ni cardinalidades, ni teoría de conjuntos infinitos más avanzada, ni convergencia. Para descartar que haya solo una cantidad finita, basta poder producir tantos puntos distintos como cualquier número natural prefijado.

Antes de seguir.

Antes de seguir
1. Sea $D \subseteq \mathbb{R}$ denso y sean $\lambda \neq 0$, $c \in \mathbb{R}$. Demuestra que $\lambda D + c = \{\lambda d + c : d \in D\}$ también es denso en $\mathbb{R}$.
Respuesta. Dados $a < b$, si $\lambda > 0$ aplicamos densidad de D al intervalo

$$\left(\frac{a-c}{\lambda}, \frac{b-c}{\lambda}\right).$$

Si $\lambda < 0$, invertimos los extremos. En ambos casos obtenemos $d \in D$ con $a < \lambda d + c < b$.

Antes de seguir
2. Demuestra que $\sqrt{2} + \mathbb{Q} = \{\sqrt{2} + q : q \in \mathbb{Q}\}$ es un conjunto denso formado exclusivamente por irracionales.
Respuesta. Es denso por la pregunta anterior, tomando $D = \mathbb{Q}$, $\lambda = 1$ y $c = \sqrt{2}$. Si $\sqrt{2} + q$ fuese racional para algún $q \in \mathbb{Q}$, al restar q concluiríamos que $\sqrt{2}$ es racional, contradicción.

Lo que exporta esta sección. Ahora conocemos dos hechos simultáneos:

$$\boxed{\mathbb{Q} \text{ es denso en } \mathbb{R}, \qquad \mathbb{R} \setminus \mathbb{Q} \text{ es denso en } \mathbb{R}.}$$

Por tanto, la recta no se divide en regiones racionales e irracionales. Ambas clases están entrelazadas a toda escala de intervalo. La cadena estructural construida en las últimas secciones es ya considerable:

$$\boxed{\text{completitud} \to \text{propiedad arquimediana} \to \text{mallas } 1/n \to \text{densidad racional} \to \text{densidad irracional}.}$$

En §1.9 volveremos a la completitud desde otra dirección. En lugar de insertar puntos dentro de un intervalo, estudiaremos una familia de intervalos cerrados que se van encajando unos dentro de otros y preguntaremos:

> si continuamos estrechando indefinidamente, ¿queda necesariamente algún punto común?

La respuesta será el **principio de intervalos encajados**, y su prueba volverá a utilizar el supremo.

1.9. Intervalos encajados y bisección

Una pregunta distinta sobre los «huecos» de la recta. Las aproximaciones por intervalos permiten conservar información de una etapa a la siguiente. Consideremos una familia de intervalos cerrados $I_0 \supseteq I_1 \supseteq I_2 \supseteq \cdots$ en la que cada intervalo queda contenido en el anterior. Podemos pensar que cada etapa conserva solamente una región más pequeña de la recta. Cada intervalo puede ser no vacío sin que ello resuelva todavía la existencia de un punto común a todos. Al exigir además que sean cerrados, acotados y estén encajados, la completitud permitirá construir ese punto.

Qué significa que los intervalos estén encajados. Escribamos

$$I_n = [a_n, b_n], \qquad n \in \mathbb{N},$$

y supongamos $I_{n+1} \subseteq I_n$ para todo n.

La inclusión impone dos movimientos simultáneos:

$$a_n \leq a_{n+1}, \qquad b_{n+1} \leq b_n.$$

Los extremos izquierdos pueden desplazarse hacia la derecha, mientras que los extremos derechos pueden desplazarse hacia la izquierda. Pero hay una condición adicional que nunca se pierde: $a_n \leq b_n$. Cada intervalo sigue siendo no vacío. Podemos representar la situación esquemáticamente como

$$\underbrace{[a_0,\ a_1,\ \underbrace{a_2, \ldots, b_2}_{I_2}\ , b_1\ }_{I_1}, b_0].$$

El punto decisivo será considerar **todos los extremos izquierdos a la vez**.

Teorema 1.7. **Principio de intervalos encajados.** Sea

$$I_n = [a_n, b_n], \qquad n \in \mathbb{N},$$

una familia de intervalos cerrados y acotados no vacíos tal que

$$I_{n+1} \subseteq I_n$$

para todo $n \in \mathbb{N}$. Entonces

$$\boxed{\bigcap_{n\in\mathbb{N}} I_n \neq \varnothing.}$$

Idea de la prueba

Queremos fabricar un único número que quede simultáneamente a la derecha de todos los extremos izquierdos y a la izquierda de todos los extremos derechos. La completitud sugiere considerar $A = \{a_n : n \in \mathbb{N}\}$ y tomar su supremo. Para que esto funcione debemos verificar dos cosas:

1. A es no vacío y está acotado superiormente;
2. si $x = \sup A$, entonces $x \leq b_n$ para **cada** n.

La segunda parte es donde el encajamiento hace el trabajo decisivo.

Demostración. Consideremos $A = \{a_n : n \in \mathbb{N}\}$.

El conjunto A es no vacío porque, por ejemplo, $a_0 \in A$. Veamos que está acotado superiormente. De hecho, afirmamos algo más fuerte: **cada** extremo derecho b_n es una cota superior de A.

Fijemos $n \in \mathbb{N}$ y tomemos cualquier $a_m \in A$.

Si $m \geq n$, entonces $I_m \subseteq I_n$.

Como $a_m \in I_m$, también $a_m \in I_n = [a_n, b_n]$, y por tanto $a_m \leq b_n$.

Si $m < n$, entonces el encajamiento da $I_n \subseteq I_m$, y en particular $a_m \leq a_n \leq b_n$. Así, en todos los casos, $a_m \leq b_n$. Por consiguiente, b_n es una cota superior de A para todo n. Como A es no vacío y está acotado superiormente, el axioma de completitud garantiza que existe $x = \sup A$.

Ahora fijemos nuevamente n.

Como $a_n \in A$ y x es cota superior de A, $a_n \leq x$.

Por otra parte, acabamos de demostrar que b_n es una cota superior de A. Como x es la **menor** cota superior, $x \leq b_n$. Por tanto, $a_n \leq x \leq b_n$, de modo que $x \in I_n$. Como n era arbitrario, $x \in I_n$ para todo $n \in \mathbb{N}$. Así, $x \in \bigcap_{n\in\mathbb{N}} I_n$, y la intersección es no vacía. ■

Después de la prueba
La completitud se utilizó una sola vez:

$$A \text{ no vacío y acotado} \Longrightarrow \sup A \text{ existe.}$$

Todo el resto de la demostración consiste en mostrar que esa frontera pertenece a cada intervalo. También conviene observar que el teorema garantiza **existencia**, pero no **unicidad**. Si $I_n = [0, 1]$ para todo n, entonces $\bigcap_{n\in\mathbb{N}} I_n = [0, 1]$, que contiene infinitos puntos.

Por qué importa que los intervalos sean cerrados. El teorema anterior no dice simplemente que «los intervalos encajados tienen un punto común». La palabra **cerrados** es esencial.

Ejemplo 1.9. **Intervalos abiertos encajados con intersección vacía.** Para $n \in \mathbb{N}$, definamos

$$J_n = \left(0, \frac{1}{n+1}\right).$$

Entonces $J_{n+1} \subseteq J_n$ para todo n, pero

$$\boxed{\bigcap_{n\in\mathbb{N}} J_n = \varnothing.}$$

Demostración. Supongamos que existe $x \in \bigcap_{n\in\mathbb{N}} J_n$.

Como $x \in J_0 = (0, 1)$, tenemos $x > 0$.

Por la propiedad arquimediana, existe $m \in \mathbb{N}_{>0}$ tal que $\frac{1}{m} < x$.

Tomando $n = m$, tenemos

$$\frac{1}{n+1} < \frac{1}{n} = \frac{1}{m} < x.$$

Pero pertenecer a J_n exigiría $x < \frac{1}{n+1}$, una contradicción. Luego la intersección es vacía. ■

¿Qué ocurrió geométricamente? Los intervalos se estrechan hacia 0, pero 0 fue excluido de todos ellos. En los intervalos cerrados

$$\left[0, \frac{1}{n+1}\right]$$

el punto 0 sí permanece disponible y pertenece a toda la familia.

Error frecuente
Tampoco basta decir simplemente «cerrados» si permitimos intervalos no acotados. Por ejemplo, $K_n = [n, \infty)$ son cerrados y encajados, pero $\bigcap_{n\in\mathbb{N}} K_n = \varnothing$ por la propiedad arquimediana. En `#thm-t1-0006` trabajamos con intervalos de la forma $[a_n, b_n]$: son cerrados **y acotados**. Esa acotación permite que los extremos izquierdos tengan cotas superiores y hace posible aplicar completitud.

Bisección: conservar siempre una mitad. El principio de intervalos encajados se vuelve especialmente concreto en el procedimiento de **bisección**. Comencemos con un intervalo cerrado $I_0 = [a_0, b_0]$.

Su punto medio es $c_0 = \frac{a_0+b_0}{2}$.

El intervalo queda dividido en dos mitades cerradas:

$$[a_0, c_0] \qquad \text{y} \qquad [c_0, b_0].$$

Elegimos una de ellas y la llamamos I_1. Después repetimos el mismo procedimiento con I_1, obteniendo I_2, y así sucesivamente. No importa todavía **qué regla** decide cuál de las dos mitades conservamos. Lo único necesario para la estructura es que en cada etapa escojamos una mitad cerrada del intervalo anterior. Entonces $I_{n+1} \subseteq I_n$ y, si escribimos $I_n = [a_n, b_n]$, su longitud satisface

$$\boxed{b_n - a_n = \frac{b_0 - a_0}{2^n}.}$$

Esta fórmula se demuestra inmediatamente por inducción: cada paso divide la longitud anterior por 2.

Existencia y unicidad son dos preguntas distintas. Por `#thm-t1-0006`, toda cadena de bisecciones cerradas posee al menos un punto común. Pero ahora tenemos información adicional: sus longitudes pueden hacerse tan pequeñas como queramos. Esa pequeñez forzará que no puedan sobrevivir dos puntos distintos.

Corolario 1.4. **Unicidad en una cadena de bisecciones.** Sea

$$I_0 \supseteq I_1 \supseteq I_2 \supseteq \cdots$$

una familia obtenida por bisecciones sucesivas de un intervalo cerrado y acotado $I_0 = [a_0, b_0]$. Entonces existe un único número real x tal que

$$x \in I_n$$

para todo $n \in \mathbb{N}$.

Idea de la prueba de unicidad
La existencia ya está resuelta por el principio de intervalos encajados. Para la unicidad, supongamos que dos puntos distintos $x < y$ sobrevivieran a todas las etapas. Como la distancia $y - x > 0$ es fija, basta encontrar una etapa cuya longitud sea menor que esa distancia. Entonces x e y no podrían caber simultáneamente dentro del mismo intervalo. La propiedad arquimediana y la desigualdad $2^n \geq n + 1$ que demostraremos enseguida mediante inducción nos permitirán encontrar esa etapa sin usar límites.

Lema auxiliar. Para todo $n \in \mathbb{N}$, se tiene $2^n \geq n+1$. La demostración es por inducción. En $n=0$, $2^0 = 1 = 0+1$. Si $2^n \geq n+1$, entonces

$$2^{n+1} = 2 \cdot 2^n \geq 2(n+1) \geq n+2,$$

pues $2(n+1) - (n+2) = n \geq 0$. Así queda establecido el paso inductivo.

Demostración. La existencia de al menos un punto común se sigue de `#thm-t1-0006`.

Para demostrar unicidad, supongamos que existen dos puntos distintos $x < y$ que pertenecen a todos los I_n.

Sea $L = b_0 - a_0$.

Si $L = 0$, entonces I_0 ya contiene un solo punto y la unicidad es inmediata. Supongamos, pues, $L > 0$. Como $y - x > 0$, la propiedad arquimediana permite elegir $N \in \mathbb{N}$ tan grande que $N+1 > \frac{L}{y-x}$. Por el lema auxiliar que acabamos de probar, $2^N \geq N+1$. Por tanto, $2^N > \frac{L}{y-x}$, y, como todas las cantidades son positivas, $\frac{L}{2^N} < y - x$.

Pero $x, y \in I_N = [a_N, b_N]$, así que $y - x \leq b_N - a_N$.

Por la construcción por bisección, $b_N - a_N = \frac{L}{2^N}$.

Llegamos entonces a $y - x \leq \frac{L}{2^N} < y - x$, una contradicción.

No pueden existir dos puntos comunes distintos. Como ya sabemos que existe al menos uno, el punto común es único. ■

El criterio general detrás de la bisección. La prueba anterior revela que la potencia 2^n no es lo esencial. Lo que realmente utilizamos fue esta propiedad:

> para toda distancia positiva ε, existe alguna etapa N cuya longitud satisface $b_N - a_N < \varepsilon$.

Si una familia de intervalos cerrados y encajados tiene longitudes arbitrariamente pequeñas, el principio de intervalos encajados garantiza existencia y el argumento de distancia garantiza unicidad. Todavía no expresamos esta condición diciendo que $b_n - a_n \to 0$, porque la noción de convergencia de sucesiones pertenece a `T1-C04`. La formulación cuantificada anterior contiene exactamente la información que necesitamos sin adelantar esa teoría.

Un ejemplo: bisección alrededor de $\sqrt{2}$.

Ejemplo 1.10. **Encajonando $\sqrt{2}$ por bisección.** Partimos de $I_0 = [1, 2]$.

Como $1^2 < 2 < 2^2$, sabemos que $\sqrt{2} \in I_0$.

En cada etapa tomamos el punto medio c_n de I_n:

- si $c_n^2 < 2$, conservamos la mitad derecha $[c_n, b_n]$;
- si $c_n^2 > 2$, conservamos la mitad izquierda $[a_n, c_n]$;
- si $c_n^2 = 2$, hemos encontrado exactamente $\sqrt{2}$: podemos detener el procedimiento o, si queremos mantener la cadena de bisecciones, conservar cualquiera de las dos mitades cerradas, que contienen c_n.

En todos los casos el intervalo conservado sigue conteniendo $\sqrt{2}$.

Las primeras etapas son: $I_0 = [1, 2]$.

El punto medio es $\frac{3}{2}$, y

$$\left(\frac{3}{2}\right)^2 = \frac{9}{4} > 2,$$

por lo que

$$I_1 = \left[1, \frac{3}{2}\right].$$

El nuevo punto medio es $\frac{5}{4}$, y

$$\left(\frac{5}{4}\right)^2 = \frac{25}{16} < 2,$$

así que

$$I_2 = \left[\frac{5}{4}, \frac{3}{2}\right].$$

El siguiente punto medio es $\frac{11}{8}$, y

$$\left(\frac{11}{8}\right)^2 = \frac{121}{64} < 2,$$

por lo que

$$I_3 = \left[\frac{11}{8}, \frac{3}{2}\right].$$

Después aparece $\frac{23}{16}$, y

$$\left(\frac{23}{16}\right)^2 = \frac{529}{256} > 2,$$

de modo que

$$I_4 = \left[\frac{11}{8}, \frac{23}{16}\right].$$

Cada intervalo de esta cadena contiene a $\sqrt{2}$ y tiene la mitad de la longitud del anterior. En caso de encontrar la raíz exactamente y optar por continuar, conservar una de las dos mitades mantiene esta regla de longitud. El corolario de unicidad garantiza que **ningún otro real** puede permanecer en todos los intervalos. Así la bisección produce una caracterización progresivamente más precisa de un número cuya existencia ya conocemos:

$$\boxed{\{\sqrt{2}\} = \bigcap_{n\in\mathbb{N}} I_n.}$$

No estamos diciendo todavía que la sucesión de extremos o la sucesión de puntos medios «converja». Esa lectura será demostrada más adelante. Aquí hemos obtenido la afirmación de intersección directamente mediante completitud y arquimedianidad.

Existencia, unicidad y algoritmo. La bisección proporciona una regla para construir intervalos cada vez más estrechos. Los teoremas anteriores justifican qué determina ese procedimiento: la completitud garantiza un punto común a los intervalos cerrados, no vacíos, acotados y encajados, y la reducción arbitraria de sus longitudes garantiza que ese punto es único. Al separar la regla de construcción de estos dos argumentos podemos saber qué parte del procedimiento asegura la existencia y cuál descarta otras posibilidades.

Antes de seguir.

Antes de seguir
1. Determina exactamente la intersección de

$$I_n = \left[-\frac{1}{n+1}, \frac{2}{n+1}\right], \qquad n \in \mathbb{N}.$$

Respuesta. El número 0 pertenece a todos los intervalos. Si $x > 0$, la propiedad arquimediana permite elegir n con $2/(n+1) < x$, por lo que $x \notin I_n$. Si $x < 0$, elegimos n con $1/(n+1) < -x$, y entonces $x < -1/(n+1)$, de modo que tampoco pertenece a todos los intervalos. Por tanto,

$$\boxed{\bigcap_{n\in\mathbb{N}} I_n = \{0\}}.$$

Antes de seguir
2. Da una familia de intervalos cerrados y encajados cuya intersección no sea un único punto y explica qué hipótesis adicional de unicidad falla.
Respuesta. La familia constante $I_n = [0, 1]$ es cerrada y encajada, pero $\bigcap_n I_n = [0, 1]$. Falla la condición de que las longitudes puedan hacerse menores que cualquier $\varepsilon > 0$.

Lo que exporta esta sección. Hemos obtenido una nueva manifestación de la completitud:

$$\boxed{I_0 \supseteq I_1 \supseteq I_2 \supseteq \cdots, \quad I_n \text{ cerrado y acotado} \Longrightarrow \bigcap_{n\in\mathbb{N}} I_n \neq \varnothing.}$$

Y, para una cadena de bisecciones,

$$\boxed{\text{existencia por completitud} + \text{longitudes arbitrariamente pequeñas} \Longrightarrow \text{un único punto común.}}$$

El supremo, la propiedad arquimediana, la densidad y los intervalos encajados resuelven problemas de existencia de formas diferentes. Al combinarlos habrá que reconocer cuál corresponde a la situación y qué hipótesis permiten utilizarlo.

1.10. Laboratorio de completitud

Buscar la frontera de un conjunto, insertar un número en un intervalo y controlar una aproximación exigen elecciones diferentes. Los problemas siguientes combinan esas tareas: antes de calcular habrá que reconocer qué se busca y comprobar qué resultado permite construirlo. Una forma de organizar esa decisión es

leer el objetivo
→ identificar la forma de existencia o estimación
→ localizar la herramienta
→ probar
→ auditar dónde entró la completitud

La última pregunta es importante. No todo lo que hacemos en este capítulo utiliza directamente el axioma del supremo. Algunas herramientas dependen de él de manera indirecta y otras, como las propiedades elementales del valor absoluto, proceden solamente de la estructura de cuerpo ordenado.

Laboratorio 1 — Una frontera que existe pero no se alcanza. Consideremos $A = \{x \in \mathbb{R} : x \geq 0,\ x^2 < 5\}$.

Problema. Determinar $\sup A$ y decidir si A tiene máximo.

Diagnóstico. El enunciado contiene dos preguntas diferentes.

1. **Frontera:** necesitamos identificar la menor cota superior.
2. **Extremo alcanzado:** debemos decidir si esa frontera pertenece a A.

El teorema de raíces de §1.6 nos proporciona ya el número real $\sqrt{5}$. La pregunta es si ese número desempeña exactamente el papel de supremo.

Solución. Primero probemos que $\sqrt{5}$ es cota superior de A.

Si $x \in A$, entonces $0 \leq x$ y $x^2 < 5 = (\sqrt{5})^2$.

Como el cuadrado es creciente sobre los reales no negativos, $x < \sqrt{5}$. Por tanto, $\sqrt{5}$ es cota superior. Falta demostrar que es la **menor**. Utilicemos la caracterización aproximativa del supremo. Sea $\varepsilon > 0$ y definamos

$$\delta = \min\left\{\frac{\varepsilon}{2}, \frac{\sqrt{5}}{2}\right\} > 0, \qquad u = \sqrt{5} - \delta.$$

Como $\delta \leq \sqrt{5}/2$, tenemos $u > 0$. Además, $u < \sqrt{5}$, por lo que $u^2 < 5$.

Así, $u \in A$. Y como $\delta \leq \varepsilon/2 < \varepsilon$, $\sqrt{5} - \varepsilon < u < \sqrt{5}$.

Hemos encontrado un elemento de A dentro de toda franja positiva situada inmediatamente debajo de $\sqrt{5}$. En consecuencia,

$$\boxed{\sup A = \sqrt{5}.}$$

¿Hay máximo? No. En efecto, si $x \in A$, entonces $x < \sqrt{5}$. El punto medio $y = \frac{x+\sqrt{5}}{2}$ satisface $x < y < \sqrt{5}$. Como $y \geq 0$, se sigue que $y^2 < 5$, de modo que $y \in A$ y $y > x$. Por tanto, ningún elemento de A puede ser el mayor:

$$\boxed{A \text{ no tiene máximo.}}$$

Lectura de la solución. El problema obliga a separar nuevamente

$$\boxed{\text{supremo} \neq \text{máximo.}}$$

La frontera existe y está perfectamente determinada, pero queda fuera del conjunto porque la condición que define A es estricta: $x^2 < 5$. La completitud no aparece escrita en la última línea de la prueba. Está **aguas arriba**: fue la propiedad que permitió demostrar en §1.6 la existencia de $\sqrt{5}$.

Laboratorio 2 — Una sola elección para dos exigencias. Sean $M > 0$ y $\varepsilon > 0$.

Problema. Demostrar que existe $n \in \mathbb{N}_{>0}$ tal que simultáneamente $n > M$ y $\frac{1}{n} < \varepsilon$.

Diagnóstico. Tenemos dos condiciones sobre el mismo natural n. No conviene elegir un natural para cada una y esperar que coincidan. Debemos convertir ambas exigencias en una única condición suficiente. La segunda desigualdad queda garantizada si $n > \frac{1}{\varepsilon}$. Así que basta fabricar un natural mayor que **dos números reales a la vez**.

Solución. Sea

$$R = \text{máx}\left\{M, \frac{1}{\varepsilon}\right\}.$$

Por la propiedad arquimediana existe $n \in \mathbb{N}$ con $n > R$.

Entonces, en particular, $n > M$ y $n > \frac{1}{\varepsilon}$.

Como $n > 0$ y $\varepsilon > 0$, la segunda desigualdad implica $\frac{1}{n} < \varepsilon$.

Por tanto,

$$\boxed{\exists n \in \mathbb{N}_{>0} \quad n > M \quad \text{y} \quad \frac{1}{n} < \varepsilon.}$$

Por qué importa. Este pequeño argumento contiene un patrón que aparecerá constantemente en cálculo:

varias restricciones $\to$ una cota común mediante máx $\to$ una sola elección que satisface todas

La propiedad arquimediana, a su vez, fue deducida de completitud en §1.7.

Laboratorio 3 — Aproximar desde lados distintos y con naturalezas distintas. Sea $x \in \mathbb{R}$ y sea $\varepsilon > 0$.

Problema. Encontrar un racional q y un irracional ξ tales que

$$x - \varepsilon < q < x < \xi < x + \varepsilon.$$

Diagnóstico. La desigualdad separa el problema en dos intervalos abiertos no vacíos:

$$(x - \varepsilon, x) \qquad \text{y} \qquad (x, x + \varepsilon).$$

La palabra clave es **insertar**. Eso apunta directamente a densidad.

Solución. Por la densidad de $\mathbb{Q}$ existe $q \in \mathbb{Q}$ tal que $x - \varepsilon < q < x$. Por la densidad de los irracionales existe $\xi \in \mathbb{R} \setminus \mathbb{Q}$ tal que $x < \xi < x + \varepsilon$.

Por tanto,

$$\boxed{x - \varepsilon < q < x < \xi < x + \varepsilon.}$$

En lenguaje de distancia,

$$0 < |q - x| < \varepsilon, \qquad 0 < |\xi - x| < \varepsilon.$$

El mismo punto x, por tanto, puede aproximarse arbitrariamente por números de **dos naturalezas aritméticas distintas** y además podemos prescribir de qué lado deben encontrarse.

Laboratorio 4 — Perturbar un punto sin abandonar un intervalo. Supongamos

$$|x-a|<r, \qquad r>0,$$

y sea $\eta>0$.

Problema. Demostrar que existen un racional q y un irracional ξ distintos de x tales que

$$|q-x|<\eta, \qquad |\xi-x|<\eta,$$

y, al mismo tiempo,

$$|q-a|<r, \qquad |\xi-a|<r.$$

Diagnóstico. Sabemos que x está dentro del intervalo centrado en a y radio r. Antes de usar densidad debemos averiguar **cuánto margen queda hasta la frontera**.

Ese margen es $r-|x-a|>0$.

Queremos movernos menos que ese margen y menos que la tolerancia η.

Solución. Definamos $\delta=\frac{1}{2}\min\{\eta,\ r-|x-a|\}>0$.

Por densidad racional podemos elegir $q\in\mathbb{Q}\cap(x,x+\delta)$.

Así, $q\neq x$ y $|q-x|<\delta<\eta$.

Además, por desigualdad triangular,

$$|q-a|\le|q-x|+|x-a|<\delta+|x-a|.$$

Como $\delta<r-|x-a|$, obtenemos $|q-a|<r$.

De manera análoga, por densidad de los irracionales podemos elegir $\xi\in(\mathbb{R}\setminus\mathbb{Q})\cap(x-\delta,x)$.

Entonces $\xi\neq x$, $|\xi-x|<\eta$ y $|\xi-a|<r$.

Por tanto, podemos perturbarnos alrededor de x por cantidades arbitrariamente pequeñas, usando racionales o irracionales, **sin salir del intervalo inicial**.

Qué técnicas se combinaron. Aquí no bastaba con decir «los racionales son densos». La estructura completa fue

valor absoluto → margen hasta la frontera → densidad → desigualdad triangular

Este tipo de control local será esencial cuando estudiemos límites y continuidad.

Laboratorio 5 — Bisección con certificado exacto de error. Sabemos por §1.6 que $\sqrt{5}$ existe y que $2<\sqrt{5}<3$.

Problema. Aplicar cuatro pasos de bisección a $[2,3]$ para producir un intervalo racional de longitud $1/16$ que contenga a $\sqrt{5}$.

Diagnóstico. No necesitamos aproximaciones decimales. En cada punto medio basta comparar su cuadrado con 5.

Solución. Partimos de $I_0=[2,3]$.

El primer punto medio es $\frac{5}{2}$, y

$$\left(\frac{5}{2}\right)^2=\frac{25}{4}>5.$$

Por tanto,

$$I_1 = \left[2, \frac{5}{2}\right].$$

El segundo punto medio es $\frac{9}{4}$, y

$$\left(\frac{9}{4}\right)^2 = \frac{81}{16} > 5,$$

de modo que

$$I_2 = \left[2, \frac{9}{4}\right].$$

El tercer punto medio es $\frac{17}{8}$, y

$$\left(\frac{17}{8}\right)^2 = \frac{289}{64} < 5.$$

Así,

$$I_3 = \left[\frac{17}{8}, \frac{9}{4}\right].$$

El cuarto punto medio es $\frac{35}{16}$, y

$$\left(\frac{35}{16}\right)^2 = \frac{1225}{256} < 5.$$

En consecuencia,

$$I_4 = \left[\frac{35}{16}, \frac{9}{4}\right].$$

Su longitud es exactamente $\frac{9}{4} - \frac{35}{16} = \frac{36-35}{16} = \frac{1}{16}$.
Hemos obtenido el certificado

$$\boxed{\frac{35}{16} < \sqrt{5} < \frac{9}{4}}$$

con un intervalo de incertidumbre de longitud $1/16$. Si continuáramos el procedimiento indefinidamente, §1.9 garantiza que la cadena de intervalos cerrados tendría un único punto común, precisamente $\sqrt{5}$. No hemos necesitado todavía hablar de una sucesión que converge.

Laboratorio 6 — Auditoría de hipótesis. Decidamos cuáles de las siguientes afirmaciones son verdaderas y, sobre todo, **qué resultado del capítulo las justifica o qué contraejemplo las destruye**.

Afirmación A.

Todo subconjunto no vacío de $\mathbb{R}$ acotado superiormente tiene supremo real.

Verdadera. Es exactamente el axioma de completitud adoptado en §1.5.

Afirmación B.

Todo subconjunto no vacío de $\mathbb{Q}$ acotado superiormente tiene supremo racional.

Falsa. El conjunto $S_{\mathbb{Q}} = \{q \in \mathbb{Q} : q \geq 0,\ q^2 < 2\}$ es no vacío y está acotado superiormente en $\mathbb{Q}$, pero §1.5 demostró que no posee supremo racional.

Afirmación C.

Toda familia encajada de intervalos abiertos y acotados tiene intersección no vacía.

Falsa. Ya conocemos

$$\left(0, \frac{1}{n+1}\right),$$

cuya intersección es vacía.

Afirmación D.

Toda familia encajada de intervalos cerrados y acotados de $\mathbb{R}$ tiene intersección no vacía.

Verdadera. Es el principio de intervalos encajados de §1.9.

Afirmación E.

Si $|x - a| < r$ y $|y - x| < r - |x - a|$, entonces $|y - a| < r$.

Verdadera. Por desigualdad triangular, $|y - a| \leq |y - x| + |x - a| < r$.

Aquí no necesitamos invocar completitud: basta la estructura de valor absoluto obtenida del cuerpo ordenado.

Afirmación F.

Todo intervalo abierto no vacío de $\mathbb{R}$ contiene un racional y un irracional.

Verdadera. Es precisamente la densidad demostrada en §1.8.

Dónde está realmente la completitud. El laboratorio permite ahora ordenar las dependencias sin mezclar niveles. No todo parte directamente del axioma del supremo. La arquitectura es más precisa:

$$\boxed{\text{cuerpo + orden} \Longrightarrow \text{intervalos, valor absoluto y desigualdades}}$$

mientras que

$$\boxed{\text{completitud} \Longrightarrow \text{supremos generales, raíces, arquimedianidad e intervalos encajados}}$$

y después

$$\boxed{\text{arquimedianidad} \Longrightarrow \text{densidad racional} \Longrightarrow \text{densidad irracional}}$$

usando en el último paso la existencia de un irracional fijo, como $\sqrt{2}$. Esta distinción es importante. Una demostración rigurosa no solo debe saber **qué teorema** utiliza; debe saber también **de qué hipótesis estructurales depende ese teorema**.

Un mapa de decisiones antes del banco final. Cuando aparezca un problema de este capítulo, estas preguntas ayudan a elegir estrategia.

- ¿Se pide una frontera de un conjunto? Busquemos cotas y pensemos en sup o ínf.
- ¿Hay que garantizar que esa frontera existe? Revisemos las hipótesis de completitud.
- ¿Aparece una tolerancia positiva y necesitamos algo muy grande o muy pequeño? Pensemos en la propiedad arquimediana.
- ¿Hay que insertar un número en un intervalo? Pensemos en densidad.
- ¿Tenemos una familia de aproximaciones por intervalos cerrados cada vez más estrechos? Pensemos en intervalos encajados y bisección.
- ¿Aparece $|x-a|$? Traduzcamos entre distancia, intervalo y desigualdad antes de manipular símbolos.
- ¿Se elimina una hipótesis de un teorema? Antes de intentar demostrar la nueva afirmación, busquemos un contraejemplo.

La elección de herramienta nace de la condición que debemos satisfacer. Identificar esa condición permite comprobar al terminar si el argumento obtuvo la frontera, el punto o la cota que se pedía.

1.11. Ejercicios y soluciones

Llegamos al bloque de práctica sistemática de este capítulo. Hasta aquí las herramientas aparecieron acompañadas por motivación, estrategias modelo y laboratorios guiados. En esta sección disminuye deliberadamente la ayuda: el lector debe decidir qué parte de la estructura de $\mathbb{R}$ es pertinente y justificar cada dependencia. El banco contiene exactamente cuarenta ejercicios, organizados en siete niveles.

Nivel	Función principal	Cantidad
A	reconocimiento estructural, definiciones e intervalos	7
B	aplicación directa: álgebra, desigualdades y estimaciones	7
C	combinación estructural: cuerpo ordenado, cotas y completitud	7
D	hipótesis esenciales, reversibilidad y diagnóstico	6
E	contraejemplos y fronteras de los teoremas	5
F	descubrimiento guiado	5
G	síntesis y desafío	3
Total		**40**

El banco ha sido depurado con una regla adicional: **no pedir como ejercicio la reproducción literal de una demostración que el desarrollo ya resolvió línea por línea**. Cuando reaparece una idea conocida, la tarea exige transferencia: combinarla con otra herramienta, generalizarla, diagnosticar una hipótesis, construir un contraejemplo o resolver una variante cuya estrategia no esté escrita de antemano. Todos los ejercicios pueden resolverse usando los

resultados demostrados en este capítulo, el principio de inducción aquí enunciado y el álgebra escolar. No es necesario —ni está permitido en las soluciones canónicas— invocar convergencia de sucesiones, límites, continuidad, teorema del valor intermedio, Bolzano–Weierstrass o resultados posteriores.

Las soluciones siguen también el cambio de régimen pedagógico del capítulo: citan por nombre los resultados ya establecidos y no expanden de nuevo asociatividad, neutros o sustituciones rutinarias, salvo cuando el propio ejercicio sea una auditoría axiomática.

Cómo trabajar esta sección
Antes de consultar una solución, deja por escrito cuatro cosas:

1. qué se supone y cuál es el dominio;
2. qué debe demostrarse, construirse o calcularse;
3. qué resultado previo parece pertinente;
4. si ese resultado depende directamente de completitud o solo de estructura de cuerpo ordenado.

En una inecuación racional añade una quinta pregunta: **¿qué puntos están excluidos o pueden cambiar la forma algebraica del problema?**

Nivel A — Reconocimiento estructural, definiciones e intervalos. Ejercicio A1. Intersección de dos bandas de distancia. Describe el conjunto

$$S = \{x \in \mathbb{R} : |x - 2| < 3 \text{ y } |x + 1| \leq 2\}$$

como una desigualdad compuesta y como un intervalo. Indica con cuidado qué extremos pertenecen al conjunto.

Ejercicio A2. Tres transformaciones del mismo orden. Supón

$$0 < a < b, \qquad c < 0.$$

Ordena correctamente, justificando cada caso:

1. ac, bc y 0;
2. a/c, b/c y 0;
3. $1/a$, $1/b$ y 0.

Ejercicio A3. Dos componentes y cuatro extremos. Sea $A = (-2, 1] \cup [3, 5)$.

Determina $\sup A$, $\inf A$ y decide si A tiene máximo y mínimo. Justifica la diferencia entre frontera y pertenencia.

Ejercicio A4. Axioma, definición o resultado demostrado. Clasifica cada afirmación en una de las categorías **axioma**, **definición** o **resultado demostrado**. Cuando corresponda, identifica el axioma `C1--C9` o el resultado de §1.1 que la respalda.

1. $a(b + c) = ab + ac$.
2. $a - b := a + (-b)$.
3. $a0 = 0$.
4. Si $a \neq 0$ y $ab = ac$, entonces $b = c$.
5. $a < b$ significa que $b - a$ es positivo.

Explica por qué confundir estas categorías puede ocultar una dependencia lógica en una demostración.

Ejercicio A5. Auditar completitud sin calcular el supremo. Para cada conjunto decide si el axioma de completitud garantiza **directamente** la existencia de un supremo real. En cada caso identifica la hipótesis que se verifica o falla.

1. $\{x \in \mathbb{R} : |x-2| < 1\}$;
2. $[0, \infty)$;
3. $\varnothing$;
4. $\{1/n : n \in \mathbb{N}_{>0}\}$;
5. $(-\infty, 0]$.

No calcules el supremo salvo que sea necesario para justificar una cota.

Ejercicio A6. Dependencia estructural. Clasifica cada hecho según dependa de **cuerpo ordenado**, **completitud directamente** o **una consecuencia previa de completitud**:

1. $|x+y| \le |x| + |y|$;
2. todo conjunto no vacío acotado superiormente tiene supremo;
3. para todo $\varepsilon > 0$ existe n con $1/n < \varepsilon$;
4. entre dos reales distintos hay un racional.

Ejercicio A7. Intervalos encajados. Considera

$$I_n = \left[1 - \frac{1}{n+1},\, 1 + \frac{1}{n+1}\right].$$

1. Verifica que $I_{n+1} \subseteq I_n$.
2. Comprueba que $1 \in I_n$ para todo n.
3. Explica por qué el principio de intervalos encajados garantiza una intersección no vacía, sin afirmar todavía que 1 sea el único punto común.

Nivel B — Aplicación directa: álgebra, desigualdades y estimaciones. Ejercicio B1. Valor absoluto anidado. Resuelve completamente $\big||x-1|-2\big| \le 1$.

Entrega el conjunto solución como unión de intervalos cerrados.

Ejercicio B2. Una estimación alrededor de 2. Supón que $|x-2| < \frac{1}{10}$.

Demuestra, sin usar aproximaciones decimales, que $|x^2 - 4| < \frac{41}{100}$.

Ejercicio B3. La diferencia entre dos recíprocos. Sean $a, b \neq 0$.

1. Demuestra que $a^{-1} - b^{-1} = \frac{b-a}{ab}$.
2. Usa esta identidad, junto con las leyes de signo ya demostradas, para recuperar el hecho de que
$$0 < a < b \quad \Longrightarrow \quad \frac{1}{b} < \frac{1}{a},$$
sin citar directamente la parte 7 de Proposición 1.5.

Ejercicio B4. Una elección arquimediana explícita. Encuentra un natural n que satisfaga simultáneamente

$$n > 200, \qquad \frac{1}{n} < \frac{1}{137}.$$

Explica por qué tu elección funciona.

Ejercicio B5. Tres pasos de bisección para $\sqrt{10}$. Comienza con $[3, 4]$ y conserva en cada paso la mitad cerrada que contiene $\sqrt{10}$. Realiza tres bisecciones y entrega un intervalo racional final $[a, b]$ con $a < \sqrt{10} < b$.

Justifica cada decisión comparando cuadrados.

Ejercicio B6. Regla de signos para un cociente. Sean $a, b \neq 0$. Demuestra, sin hacer una tabla memorizada, que

$$\frac{a}{b} > 0 \iff (a > 0 \text{ y } b > 0) \text{ o } (a < 0 \text{ y } b < 0),$$

y que

$$\frac{a}{b} < 0 \iff (a > 0 \text{ y } b < 0) \text{ o } (a < 0 \text{ y } b > 0).$$

Debes reducir el problema al signo de b^{-1} y al teorema del signo de un producto.

Ejercicio B7. Inecuación racional con valor absoluto I. Resuelve completamente

$$\left|\frac{x-1}{x+2}\right| \leq \frac{|x-3|}{2}.$$

Tu solución debe registrar dominio, puntos críticos, transformación algebraica equivalente, análisis de signos y verificación de extremos.

Nivel C — Combinación estructural: cuerpo ordenado, cotas y completitud.
Ejercicio C1. Escalar un supremo por un número positivo. Sea $A \subseteq \mathbb{R}$ no vacío y acotado superiormente, y sea $c > 0$. Define $cA = \{ca : a \in A\}$.

Demuestra que $\sup(cA) = c \sup A$.

Ejercicio C2. El ínfimo como supremo reflejado. Sea $A \subseteq \mathbb{R}$ no vacío y acotado inferiormente, y define $-A = \{-a : a \in A\}$.

Demuestra que $\inf A = -\sup(-A)$.

Ejercicio C3. Alternar racionales e irracionales. Sean $a < b$ y $N \in \mathbb{N}_{>0}$. Demuestra que existen racionales $q_1, \ldots, q_N$ e irracionales $\xi_1, \ldots, \xi_N$ tales que

$$a < q_1 < \xi_1 < q_2 < \xi_2 < \cdots < q_N < \xi_N < b.$$

No basta afirmar que ambos conjuntos son densos: organiza una construcción que garantice simultáneamente todo el orden indicado.

Ejercicio C4. Una ecuación afín completa. Sean $a, b, c \in F$, donde F es un cuerpo y $a \neq 0$. Considera $ax + b = c$.

1. Construye un candidato explícito para x.
2. Verifica que satisface la ecuación.
3. Demuestra que ninguna otra solución es posible.

La solución debe combinar las dos ecuaciones elementales del capítulo; no basta escribir una cadena escolar de «pasar términos».

Ejercicio C5. Producto de raíces cuadradas. Sean $a, b \geq 0$. Utiliza el teorema de existencia y unicidad de raíces cuadradas para demostrar $\sqrt{ab} = \sqrt{a}\,\sqrt{b}$.

Ejercicio C6. Inecuación racional con valor absoluto II. Resuelve completamente

$$\frac{|x-2|}{|x+1|} \geq \frac{|x|}{2}.$$

Debes identificar el dominio, todos los puntos críticos relevantes y justificar cualquier operación de elevar al cuadrado.

Ejercicio C7. Dos cadenas encajadas que no pueden terminar en puntos distintos. Sean (I_n) y (J_n) dos familias de intervalos cerrados, no vacíos y encajados. Supón que las longitudes de ambas familias pueden hacerse menores que cualquier $\varepsilon > 0$ y que

$$I_n \cap J_n \neq \varnothing \qquad \text{para todo } n.$$

Demuestra que las dos familias tienen el **mismo** único punto común.

Nivel D — Hipótesis esenciales, reversibilidad y diagnóstico. Ejercicio D1. Supremo de una unión: diagnosticar una fórmula falsa. Sean $A, B \subseteq \mathbb{R}$ no vacíos y acotados superiormente. Un estudiante afirma $\sup(A \cup B) = \sup A + \sup B$.

1. Da un contraejemplo.
2. Formula la identidad correcta.
3. Demuéstrala a partir de la definición de supremo.

Ejercicio D2. Cancelar puede borrar una solución. Resuelve en un cuerpo ordenado $(x-1)(x+2) = (x-1)(3x-4)$.

Un estudiante cancela inmediatamente el factor $x-1$ y obtiene una sola solución. Explica por qué ese procedimiento es incompleto y determina **todas** las soluciones.

Ejercicio D3. ¿Eliminar finitos puntos cambia el supremo? Sea $A \subseteq \mathbb{R}$ no vacío y acotado superiormente, sea $F \subseteq A$ finito y supón $A \setminus F \neq \varnothing$. Un estudiante afirma:

> «Eliminar finitos puntos nunca cambia el supremo».

1. Da un contraejemplo.
2. Demuestra que la afirmación sí es correcta bajo la hipótesis adicional $\sup A \notin F$.

Ejercicio D4. Intervalos racionales encajados alrededor de un punto que no es racional. Parte de $I_0 = [1, 2]$ y aplica bisección conservando siempre la mitad cerrada que contiene a $\sqrt{2}$.

1. Explica por qué todos los extremos de I_n son racionales.
2. Demuestra que las longitudes son 2^{-n}.
3. Explica por qué $\bigcap_n I_n = \{\sqrt{2}\}$ como subconjuntos de $\mathbb{R}$.
4. Concluye que, vistos como intervalos de $\mathbb{Q}$, tienen intersección vacía.

Explica por qué este fenómeno muestra que densidad y completitud son propiedades diferentes.

Ejercicio D5. Inecuación racional con valor absoluto III. Resuelve completamente

$$\frac{|x-2| + |x+1|}{|x-1|} \leq 3.$$

La solución debe separar todos los casos determinados por los argumentos de los valores absolutos y por el punto excluido del denominador.

Ejercicio D6. Una prueba correcta en otro momento, pero circular aquí. Un estudiante propone demostrar la existencia de $\sqrt{3}$ así:

> «La función x^2 es continua; como $1^2 < 3 < 2^2$, el teorema del valor intermedio produce un $c \in (1, 2)$ con $c^2 = 3$».

Explica por qué este argumento no es admisible dentro de este capítulo, aunque llegará a ser matemáticamente válido más adelante. Indica cuál es la dependencia correcta en este capítulo.

Nivel E — Contraejemplos y fronteras de los teoremas. Ejercicio E1. Cuatro comportamientos con la misma frontera superior. Construye cuatro subconjuntos no vacíos y acotados de $\mathbb{R}$, todos con supremo 1, que presenten respectivamente estos comportamientos:

1. tienen máximo y mínimo;
2. no tienen máximo, pero sí mínimo;
3. tienen máximo, pero no mínimo;
4. no tienen ni máximo ni mínimo.

Justifica cada elección.

Ejercicio E2. El mismo supremo con comportamientos inferiores opuestos. Encuentra dos conjuntos $A, B \subseteq \mathbb{R}$ tales que $\sup A = \sup B = 0$, pero A esté acotado inferiormente y B no. Determina además si existe $\inf A$ y explica por qué B no posee ínfimo real.

Ejercicio E3. La división por cero no es una simplificación pendiente. Un estudiante escribe

$$0x = 0 \quad \Longrightarrow \quad x = \frac{0}{0} \quad \Longrightarrow \quad x = 1,$$

invocando informalmente la regla $a/a = 1$.

Refuta el argumento de dos maneras complementarias:

1. demuestra que la ecuación $0x = 0$ tiene **todos** los elementos del cuerpo como soluciones;

2. explica desde la definición de división por qué $0/0$ no está definido y por qué la identidad $a/a = 1$ exige $a \neq 0$.

Identifica exactamente el primer paso ilegítimo de la cadena.

Ejercicio E4. Cerrados y cada vez más pequeños, pero no encajados. Construye una familia de intervalos cerrados y no vacíos I_n tal que

1. sus longitudes puedan hacerse menores que cualquier $\varepsilon > 0$;
2. $\bigcap_n I_n = \varnothing$;
3. la razón del fracaso sea exactamente que la familia **no** es encajada.

Explica por qué esto no contradice el principio de intervalos encajados.

Ejercicio E5. Transportar el hueco racional. Sea $S_{\mathbb{Q}} = \{q \in \mathbb{Q} : q \geq 0,\ q^2 < 2\}$ y define $T = \{2q + 1 : q \in S_{\mathbb{Q}}\} \subseteq \mathbb{Q}$.

1. Demuestra que T es no vacío y está acotado superiormente en $\mathbb{Q}$.
2. Demuestra que T no tiene supremo racional.

La segunda parte debe reducir un supuesto $\sup_{\mathbb{Q}} T$ a un supuesto supremo racional de $S_{\mathbb{Q}}$.

Nivel F — Descubrimiento guiado. Ejercicio F1. Redescubrir el supremo de una traslación. Sea $A \subseteq \mathbb{R}$ no vacío y acotado superiormente, sea $s = \sup A$, y fija $c \in \mathbb{R}$. Define $A + c = \{a + c : a \in A\}$.

Utiliza la **caracterización aproximativa** del supremo —no una fórmula previamente memorizada— para demostrar

$$\boxed{\sup(A + c) = s + c}.$$

Ejercicio F2. Una sola elección para muchas exigencias. Sean $M_1, \ldots, M_r \in \mathbb{R}$ y $\varepsilon_1, \ldots, \varepsilon_s > 0$, con $r, s \geq 1$. Demuestra que existe un único natural elegido **una sola vez**, $n \in \mathbb{N}_{>0}$, que satisface simultáneamente

$$n > M_i \quad (i = 1, \ldots, r)$$

y

$$\frac{1}{n} < \varepsilon_j \quad (j = 1, \ldots, s).$$

Ejercicio F3. Densidad racional con control del denominador. Sean $a < b$ y $N \in \mathbb{N}_{>0}$. Demuestra que existen $m \in \mathbb{Z}$ y $n \in \mathbb{N}_{>0}$ tales que $n > N$ y $a < \frac{m}{n} < b$.

No basta citar densidad de $\mathbb{Q}$: debes adaptar su construcción para imponer además la cota inferior sobre el denominador.

Ejercicio F4. Presupuesto de bisección sin calcular todos los puntos medios. Parte de un intervalo de longitud 1 que contiene a $\sqrt{7}$ y aplica bisección conservando siempre una mitad que contenga la raíz.

1. Demuestra por inducción que después de n bisecciones la longitud es 2^{-n}.
2. Determina el menor n que garantiza una longitud estrictamente menor que $1/100$.

3. Explica qué certificado de localización de $\sqrt{7}$ proporciona esa etapa, aunque no calcules sus extremos concretos.

Ejercicio F5. Reconstruir $(-1)a = -a$ sin usar las reglas de signos. Sea F un cuerpo y $a \in F$.

Puedes utilizar la unicidad del inverso aditivo y el resultado ya demostrado $a0 = 0$, pero **no** las identidades $(-a)b = -(ab)$ ni $a(-b) = -(ab)$.

1. Demuestra que $(-1)a$ es un inverso aditivo de a.
2. Concluye que $(-1)a = -a$.
3. Particulariza el resultado para deducir $(-1)(-1) = 1$.

Identifica dónde intervienen distributividad, conmutatividad y unicidad.

Nivel G — Síntesis y desafío. Ejercicio G1. Supremo de una suma de conjuntos. Sean $A, B \subseteq \mathbb{R}$ no vacíos y acotados superiormente. Define $A + B = \{a + b : a \in A,\ b \in B\}$.

Demuestra que

$$\boxed{\sup(A + B) = \sup A + \sup B.}$$

Ejercicio G2. Una propiedad de corte obtenida de completitud. Sean $L, U \subseteq \mathbb{R}$ no vacíos, disjuntos, con $L \cup U = \mathbb{R}$, y supón que

$$\ell < u \qquad \text{para todo } \ell \in L,\ u \in U.$$

1. Demuestra que L está acotado superiormente.
2. Sea $c = \sup L$. Demuestra que $\ell \le c \le u$ para todo $\ell \in L$ y $u \in U$.
3. Como $c \in L \cup U$, demuestra que ocurre exactamente una de estas dos posibilidades: $c = \max L$ o $c = \min U$.

Esta es una forma elemental de propiedad de corte de la recta real.

Ejercicio G3. Construir una raíz cúbica sin continuidad. Sea $a > 0$ y define $S = \{x \in \mathbb{R} : x \ge 0,\ x^3 < a\}$.

1. Demuestra que S es no vacío y está acotado superiormente.
2. Define $\alpha = \sup S$.
3. Demuestra, mediante perturbaciones algebraicas explícitas, que ni $\alpha^3 < a$ ni $\alpha^3 > a$ son posibles.
4. Concluye que existe un único $\alpha > 0$ tal que $\alpha^3 = a$.

No utilices continuidad ni el teorema del valor intermedio.

Soluciones. Las soluciones siguen el mismo orden. En los ejercicios técnicos se incluyen los controles de dominio y de signos; en los problemas de síntesis se explicita la estrategia y la dependencia estructural decisiva.

Soluciones del nivel A. **Solución A1.**

Las dos condiciones equivalen a $-1 < x < 5$ y $-3 \leq x \leq 1$.

Debemos tomar la intersección. Por tanto, $-1 < x \leq 1$, y

$$\boxed{S = (-1, 1]}.$$

El extremo -1 queda excluido por la primera desigualdad estricta; el extremo 1 satisface ambas condiciones.

Solución A2.

Como $c < 0$, multiplicar $a < b$ por c invierte el orden; además, los dos productos son negativos:

$$\boxed{bc < ac < 0}.$$

Dividir por $c < 0$ también invierte el orden, y ambos cocientes son negativos:

$$\boxed{\frac{b}{c} < \frac{a}{c} < 0}.$$

Finalmente, para recíprocos positivos se invierte el orden:

$$\boxed{0 < \frac{1}{b} < \frac{1}{a}}.$$

Las tres cadenas usan el mismo dato $a < b$, pero cada transformación exige controlar el signo del factor o divisor.

Solución A3.

Todo elemento de A es menor que 5, y en el componente $[3, 5)$ hay elementos arbitrariamente próximos a 5 por la izquierda. Por tanto, $\sup A = 5$, pero $5 \notin A$, así que no hay máximo. Del mismo modo, -2 es cota inferior y el componente $(-2, 1]$ contiene puntos arbitrariamente próximos a -2 por la derecha. Luego $\inf A = -2$, y como $-2 \notin A$, tampoco hay mínimo.

Solución A4.

La clasificación es la siguiente.

1. **Axioma.**
$$a(b + c) = ab + ac$$
es exactamente `C9`, la distributividad.

2. **Definición.**
$$a - b := a + (-b)$$
es la definición de resta como operación derivada de la suma y el inverso aditivo.

3. **Resultado demostrado.**
$$a0 = 0$$
no figura entre `C1--C9`. Se demuestra a partir de los axiomas; en §1.1 aparece como la prueba auditada Ejemplo 1.1.

4. **Resultado demostrado.** La implicación
$$a \neq 0, \quad ab = ac \Longrightarrow b = c$$
es la cancelación multiplicativa de Proposición 1.3. La hipótesis $a \neq 0$ no es decorativa: permite usar el inverso multiplicativo de a.
5. **Definición.** El orden se introduce mediante la positividad: $a < b \iff b - a > 0$. Por tanto esta equivalencia fija el significado de $<$ dentro de la estructura ordenada.

La distinción importa porque un axioma puede usarse como punto de partida, una definición fija significado y un resultado demostrado solo puede reutilizarse después de haber sido establecido. Si tratamos una consecuencia como si fuese axioma, podemos ocultar precisamente la dependencia que una auditoría de prueba debe hacer visible.

Solución A5.

1. Sí. El conjunto es no vacío —por ejemplo, contiene a 2— y está acotado superiormente, por ejemplo por 3.
2. No. Es no vacío, pero no está acotado superiormente.
3. No. Falla la hipótesis de no vacuidad.
4. Sí. Contiene a 1 y está acotado superiormente por 1.
5. Sí. Es no vacío y 0 es una cota superior.

La quinta parte subraya que el axioma solo exige acotación **superior**; el conjunto puede ser ilimitado hacia abajo.

Solución A6.

1. La desigualdad triangular se obtiene de **cuerpo ordenado** y las propiedades del valor absoluto; no necesita completitud.
2. La existencia general de supremos para conjuntos no vacíos y acotados superiormente es **completitud directamente**.
3. $1/n < \varepsilon$ se deduce de la **propiedad arquimediana**, que en nuestro desarrollo es una **consecuencia previa de completitud**.
4. La densidad racional se deduce de arquimedianidad y encajonamiento entero; por tanto depende **indirectamente de completitud** en la cadena adoptada por este capítulo.

La clasificación depende de nuestra arquitectura de pruebas, no solo de que un resultado sea verdadero en $\mathbb{R}$.

Solución A7.

Tenemos $\frac{1}{n+2} < \frac{1}{n+1}$.

Por tanto, $1 - \frac{1}{n+1} < 1 - \frac{1}{n+2}$ y $1 + \frac{1}{n+2} < 1 + \frac{1}{n+1}$.

Así, los extremos del intervalo siguiente quedan dentro del anterior y $I_{n+1} \subseteq I_n$. Además, $1 - \frac{1}{n+1} \leq 1 \leq 1 + \frac{1}{n+1}$, de modo que $1 \in I_n$ para todo n. Como todos los I_n son cerrados, no vacíos, acotados y encajados, el principio de intervalos encajados garantiza $\bigcap_n I_n \neq \varnothing$. Ese teorema por sí solo garantiza existencia, no unicidad. La unicidad requeriría además explotar que las longitudes se hacen arbitrariamente pequeñas.

Soluciones del nivel B. **Solución B1.**

La desigualdad equivale a $1 \leq |x-1| \leq 3$.

La cota superior da $-2 \leq x \leq 4$, mientras que $|x-1| \geq 1$ equivale a

$$x \leq 0 \quad \text{o} \quad x \geq 2.$$

Intersectando ambas condiciones,

$$\boxed{x \in [-2, 0] \cup [2, 4]}.$$

Solución B2.

Factorizamos: $|x^2 - 4| = |x-2|\,|x+2|$.

Para controlar el segundo factor escribimos $x + 2 = (x-2) + 4$.

Por desigualdad triangular,

$$|x+2| \leq |x-2| + 4 < \frac{1}{10} + 4 = \frac{41}{10}.$$

Luego $|x^2 - 4| < \frac{1}{10} \cdot \frac{41}{10} = \frac{41}{100}$.

La idea es típica de las estimaciones futuras: el dato controla $|x-2|$ y fabricamos a partir de él un control del factor restante.

Solución B3.

Por las reglas de cocientes, $a^{-1} - b^{-1} = \frac{1}{a} - \frac{1}{b} = \frac{b-a}{ab}$. Si $0 < a < b$, entonces $b - a > 0$ y $ab > 0$. Dividir un positivo por un positivo produce un número positivo, de modo que $\frac{1}{a} - \frac{1}{b} > 0$.

Por definición del orden,

$$\boxed{\frac{1}{b} < \frac{1}{a}}.$$

Solución B4.

La elección más sencilla es

$$\boxed{n = 201}.$$

Claramente $201 > 200$. Además, como $201 > 137 > 0$, al tomar recíprocos se invierte el orden: $\frac{1}{201} < \frac{1}{137}$. Por tanto satisface simultáneamente ambas exigencias.

Solución B5.

Partimos de $I_0 = [3, 4]$, porque $3^2 = 9 < 10 < 16 = 4^2$.

Primer punto medio:

$$m_1 = \frac{7}{2}, \qquad \left(\frac{7}{2}\right)^2 = \frac{49}{4} > 10.$$

Conservamos

$$I_1 = \left[3, \frac{7}{2}\right].$$

Segundo punto medio:

$$m_2 = \frac{13}{4}, \qquad \left(\frac{13}{4}\right)^2 = \frac{169}{16} > 10$$

porque $169 > 160$. Luego

$$I_2 = \left[3, \frac{13}{4}\right].$$

Tercer punto medio:

$$m_3 = \frac{25}{8}, \qquad \left(\frac{25}{8}\right)^2 = \frac{625}{64} < 10$$

porque $625 < 640$. Por tanto,

$$I_3 = \left[\frac{25}{8}, \frac{13}{4}\right].$$

Así,

$$\boxed{\frac{25}{8} < \sqrt{10} < \frac{13}{4}.}$$

Solución B6.

Por definición, $\frac{a}{b} = ab^{-1}$.

La parte 5 de Proposición 1.5 dice que b^{-1} tiene el mismo signo que b. La parte 8 caracteriza el signo de un producto: es positivo cuando los factores tienen el mismo signo y negativo cuando tienen signos opuestos. Sustituyendo el signo de b^{-1} por el de b obtenemos exactamente

$$\boxed{\frac{a}{b} > 0 \iff (a > 0, b > 0) \text{ o } (a < 0, b < 0)}$$

y

$$\boxed{\frac{a}{b} < 0 \iff (a > 0, b < 0) \text{ o } (a < 0, b > 0)}.$$

Solución B7.

Queremos resolver

$$\left|\frac{x-1}{x+2}\right| \leq \frac{|x-3|}{2}.$$

1. Dominio y puntos críticos. El denominador exige $x \neq -2$.

Los argumentos de los valores absolutos se anulan en $x = 1$ y $x = 3$. Estos puntos, junto con el polo -2, deben registrarse aunque una transformación posterior produzca fronteras adicionales.

2. Eliminación segura de denominadores y valores absolutos. Para $x \neq -2$, $2|x+2| > 0$. Multiplicamos sin cambiar el sentido: $2|x-1| \leq |x-3|\,|x+2|$.

Ambos lados son no negativos, así que elevar al cuadrado es una equivalencia: $4(x-1)^2 \leq (x-3)^2(x+2)^2$. Llevando todo a un lado y factorizando,

$$(x-3)^2(x+2)^2 - 4(x-1)^2 = (x-4)(x+1)(x^2+x-8).$$

Las raíces del factor cuadrático son

$$r_- = \frac{-1-\sqrt{33}}{2}, \qquad r_+ = \frac{-1+\sqrt{33}}{2}.$$

El orden relevante es

$$r_- < -2 < -1 < r_+ < 3 < 4.$$

3. Análisis de signos. Debemos resolver

$$(x-4)(x+1)(x-r_-)(x-r_+) \geq 0.$$

Todos los ceros son simples, por lo que el signo alterna al atravesarlos. Como el polinomio es positivo para x grande y positivo, resulta no negativo en $(-\infty, r_-] \cup [-1, r_+] \cup [4, \infty)$. El punto excluido $x = -2$ se encuentra en una región que ya no pertenece a la solución, pero debe mantenerse fuera del dominio en todo momento.

4. Extremos. En r_-, -1, r_+ y 4 se obtiene igualdad, y ninguno es un polo; por eso se incluyen.

Por consiguiente,

$$\boxed{(-\infty, \tfrac{-1-\sqrt{33}}{2}] \cup [-1, \tfrac{-1+\sqrt{33}}{2}] \cup [4, \infty)}.$$

La comprobación final respeta el dominio original y todos los extremos proceden de equivalencias algebraicas reversibles.

Soluciones del nivel C. **Solución C1.**

Sea $s = \sup A$.

Como $a \leq s$ para todo $a \in A$ y $c > 0$, $ca \leq cs$.

Así, cs es cota superior de cA. Para demostrar que es la menor, sea $\varepsilon > 0$. Como $s = \sup A$, existe $a \in A$ tal que $s - \frac{\varepsilon}{c} < a \leq s$.

Multiplicando por $c > 0$, $cs - \varepsilon < ca \leq cs$.

Por la caracterización aproximativa del supremo,

$$\boxed{\sup(cA) = c \sup A.}$$

Solución C2.

Sea $t = \sup(-A)$.

Para todo $a \in A$, el número $-a$ pertenece a $-A$, luego $-a \leq t$.

Multiplicando por -1, $a \geq -t$.

Así, $-t$ es cota inferior de A. Si ℓ es cualquier cota inferior de A, entonces $\ell \leq a$ para todo $a \in A$, por lo que $-a \leq -\ell$ para todo $-a \in -A$. Así, $-\ell$ es cota superior de $-A$. Como t es la menor de ellas, $t \leq -\ell$, y, al multiplicar por -1, $-t \geq \ell$. Por tanto, $-t$ es la mayor cota inferior de A:

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

Solución C3.
Sea $h = \frac{b-a}{2N+1} > 0$.
Para cada $j = 1, \dots, N$, por densidad de $\mathbb{Q}$ elegimos

$$q_j \in \mathbb{Q} \cap \big(a + (2j-2)h,\, a + (2j-1)h\big),$$

y por densidad de los irracionales elegimos

$$\xi_j \in (\mathbb{R} \setminus \mathbb{Q}) \cap \big(a + (2j-1)h,\, a + 2jh\big).$$

Los subintervalos fueron escogidos en orden y son disjuntos. Por tanto,

$$a < q_1 < \xi_1 < q_2 < \xi_2 < \cdots < q_N < \xi_N < a + 2Nh < b.$$

Esto demuestra la afirmación.

Solución C4.
La ecuación equivale primero a $ax = c - b$.
Como $a \neq 0$, el candidato natural es $x_0 = a^{-1}(c-b)$.
En efecto, $ax_0 + b = (aa^{-1})(c-b) + b = c$.
Si y es otra solución, entonces $ay + b = c = ax_0 + b$. La cancelación aditiva da $ay = ax_0$, y la cancelación multiplicativa —válida porque $a \neq 0$— da $y = x_0$. Por tanto,

$$\boxed{x = a^{-1}(c-b)}$$

es la solución única.

Solución C5.
Como $a, b \geq 0$, también

$$\sqrt{a} \geq 0, \qquad \sqrt{b} \geq 0,$$

de modo que $\sqrt{a}\,\sqrt{b} \geq 0$.
Además, $(\sqrt{a}\,\sqrt{b})^2 = (\sqrt{a})^2(\sqrt{b})^2 = ab$.
El teorema de raíces cuadradas afirma que existe un **único** número no negativo cuyo cuadrado es ab. Como $\sqrt{a}\,\sqrt{b}$ tiene esas dos propiedades, debe coincidir con él:

$$\boxed{\sqrt{ab} = \sqrt{a}\,\sqrt{b}.}$$

La unicidad es el paso que legitima la identificación final.

Solución C6.
Resolvemos

$$\frac{|x-2|}{|x+1|} \geq \frac{|x|}{2}.$$

1. Dominio. Debe cumplirse $x \neq -1$.
Los argumentos de los valores absolutos se anulan en $x = -1, 0, 2$; el primero es además un polo.

2. Transformación equivalente. Para $x \neq -1$, multiplicamos por $2|x+1| > 0$: $2|x-2| \geq |x|\,|x+1|$.

Ambos miembros son no negativos, por lo que podemos elevar al cuadrado sin introducir soluciones espurias: $4(x-2)^2 \geq x^2(x+1)^2$.

La diferencia factoriza como

$$4(x-2)^2 - x^2(x+1)^2 = -(x-1)(x+4)(x^2-x+4).$$

El cuadrático satisface

$$\Delta = (-1)^2 - 16 = -15 < 0$$

y tiene coeficiente principal positivo, así que $x^2 - x + 4 > 0$ para todo real x. Por tanto, la inecuación equivale a $-(x-1)(x+4) \geq 0$, o $(x-1)(x+4) \leq 0$.

Esto ocurre para $-4 \leq x \leq 1$.

3. Reincorporación del dominio. Debemos eliminar $x = -1$.

Los extremos -4 y 1 producen igualdad y son admisibles. En consecuencia,

$$\boxed{[-4,-1) \cup (-1,1].}$$

Los puntos 0 y 2 eran críticos para la forma original de los valores absolutos; la transformación mediante cuadrados los absorbió en una equivalencia global válida, pero registrarlos evita perder de vista la estructura original.

Solución C7.

Por el principio de intervalos encajados y el criterio de longitudes arbitrariamente pequeñas, cada familia tiene un único punto común. Llamémoslos x para (I_n) e y para (J_n).

Supongamos $x \neq y$ y sea $\varepsilon = \frac{|x-y|}{3} > 0$.

Elige una etapa N en la que las longitudes de I_N y J_N sean menores que ε; podemos tomar el máximo de dos etapas si fuera necesario. Por hipótesis existe $z \in I_N \cap J_N$.

Como $x, z \in I_N$ y $y, z \in J_N$,

$$|x-z| < \varepsilon, \qquad |y-z| < \varepsilon.$$

La desigualdad triangular produce

$$|x-y| \leq |x-z| + |z-y| < 2\varepsilon = \frac{2}{3}|x-y|,$$

contradicción. Luego $x = y$.

Soluciones del nivel D. **Solución D1.**

Tomemos $A = B = (0,1)$. Entonces

$$\sup(A \cup B) = 1, \qquad \sup A + \sup B = 2,$$

así que la fórmula propuesta es falsa.

La identidad correcta es

$$\boxed{\sup(A \cup B) = \max\{\sup A, \sup B\}}.$$

Sea s el máximo del miembro derecho. Todo elemento de $A \cup B$ pertenece a uno de los dos conjuntos y, por tanto, es menor o igual que s; luego s es cota superior de la unión. Si u es cualquier cota superior de $A \cup B$, entonces también es cota superior de A y de B. Por ello

$$\sup A \leq u, \qquad \sup B \leq u,$$

y en consecuencia $s \leq u$. Así s es la menor cota superior.

Solución D2.

La cancelación de $x - 1$ solo es válida bajo la hipótesis $x - 1 \neq 0$. Si $x = 1$, ambos miembros de la ecuación original son 0, de modo que $x = 1$ es una solución y se perdería al cancelar. En el caso $x \neq 1$ sí podemos cancelar y obtenemos $x + 2 = 3x - 4$, de donde $x = 3$.

Por tanto,

$$\boxed{x \in \{1, 3\}}.$$

Equivalentemente, restando los miembros y factorizando se llega a $(x - 1)(6 - 2x) = 0$, y la ley del producto nulo produce los mismos dos casos.

Solución D3.

La afirmación sin hipótesis adicional es falsa. Si

$$A = \{0, 1\}, \qquad F = \{1\},$$

entonces $\sup A = 1$, pero $\sup(A \setminus F) = 0$.

Ahora sea $s = \sup A$ y supongamos $s \notin F$. Como $A \setminus F \subseteq A$, el número s es cota superior de $A \setminus F$. Si $F = \varnothing$, no hay nada que probar. Supongamos $F \neq \varnothing$ y sea $m = \max F$. Como todos los elementos de F son menores o iguales que s y $s \notin F$, tenemos $m < s$.

Dado $\varepsilon > 0$, toma $\delta = \frac{1}{2} \min\{\varepsilon, s - m\} > 0$.

Por la caracterización aproximativa del supremo existe $a \in A$ tal que $s - \delta < a \leq s$. Además $s - \delta > m$, por lo que $a \notin F$. Así $a \in A \setminus F$ y $s - \varepsilon < a \leq s$.

La caracterización aproximativa concluye

$$\boxed{\sup(A \setminus F) = s = \sup A}.$$

Solución D4.

El intervalo inicial tiene extremos racionales. El punto medio de dos racionales es racional, de modo que cada bisección vuelve a producir extremos racionales. Cada paso divide la longitud por 2; como $|I_0| = 1$, $|I_n| = 2^{-n}$. Los intervalos son cerrados, no vacíos y encajados, así que tienen un punto común en $\mathbb{R}$. Como sus longitudes pueden hacerse menores que cualquier número positivo, ese punto es único. La regla de elección mantiene a $\sqrt{2}$ dentro de todos ellos, luego $\bigcap_n I_n = \{\sqrt{2}\}$.

Pero $\sqrt{2} \notin \mathbb{Q}$. Por tanto, $\bigcap_n (I_n \cap \mathbb{Q}) = \varnothing$.

Los racionales siguen siendo densos —cada intervalo contiene racionales—, pero el punto común exigido por completitud puede faltar dentro de $\mathbb{Q}$.

Solución D5.
Resolvemos

$$\frac{|x-2|+|x+1|}{|x-1|} \leq 3.$$

1. Dominio y puntos críticos. El denominador exige $x \neq 1$.
Los argumentos cambian de signo en

$$-1, \quad 1, \quad 2.$$

Estos tres puntos dividen la recta en cuatro regiones.
Caso 1: $x < -1$**.**

$$|x-2| = 2-x, \quad |x+1| = -x-1, \quad |x-1| = 1-x.$$

Entonces $\frac{1-2x}{1-x} \leq 3$.
Como $1-x > 0$, $1-2x \leq 3-3x \iff x \leq 2$, lo cual es verdadero en todo el caso. Obtenemos $(-\infty, -1)$.
Caso 2: $-1 \leq x < 1$**.**
El numerador vale $(2-x)+(x+1) = 3$, y el denominador $1-x > 0$. Así,

$$\frac{3}{1-x} \leq 3 \iff 1 \leq 1-x \iff x \leq 0.$$

Dentro del caso obtenemos $[-1, 0]$.
Caso 3: $1 < x < 2$**.**
El numerador sigue siendo 3, pero ahora $|x-1| = x-1$:

$$\frac{3}{x-1} \leq 3 \iff 1 \leq x-1 \iff x \geq 2.$$

No hay soluciones en $(1, 2)$.
Caso 4: $x \geq 2$**.**

$$|x-2| = x-2, \quad |x+1| = x+1, \quad |x-1| = x-1.$$

Por tanto, $\frac{2x-1}{x-1} \leq 3$.
Como $x-1 > 0$, $2x-1 \leq 3x-3 \iff x \geq 2$, válido en todo el caso. Obtenemos $[2, \infty)$.
Recomposición.

$$(-\infty, -1) \cup [-1, 0] = (-\infty, 0].$$

El punto $x = 1$ permanece excluido y no aparece en la solución. Finalmente,

$$\boxed{(-\infty, 0] \cup [2, \infty).}$$

Los extremos 0 y 2 satisfacen la igualdad original, por lo que se incluyen.

Solución D6.

El argumento será válido una vez que dispongamos de continuidad y del teorema del valor intermedio. El problema es su **posición lógica** dentro de este libro. En este capítulo todavía no hemos demostrado continuidad de x^2 ni el teorema del valor intermedio. Más aún, esos resultados posteriores se apoyarán en propiedades estructurales de $\mathbb{R}$ cuya raíz está precisamente en la completitud. Usarlos ahora para justificar una consecuencia que estamos extrayendo de completitud invertiría la dirección de dependencia comprometida por el proyecto. La vía autorizada en este capítulo es:

$$\boxed{\text{completitud} \to \alpha = \sup\{x \geq 0 : x^2 < 3\} \to \alpha^2 = 3.}$$

No afirmamos que la prueba con IVT sea matemáticamente falsa; afirmamos que es **prematura y circular respecto de nuestra arquitectura**.

Soluciones del nivel E. **Solución E1.**

Una elección elemental es

$$A_1 = [0,1], \qquad A_2 = [0,1), \qquad A_3 = (0,1], \qquad A_4 = (0,1).$$

Todos tienen supremo 1. Además:

- A_1 tiene máximo 1 y mínimo 0;
- A_2 no tiene máximo y tiene mínimo 0;
- A_3 tiene máximo 1 y no tiene mínimo;
- A_4 no tiene máximo ni mínimo.

La pertenencia de los puntos frontera, no el valor del supremo o del ínfimo, decide la existencia de extremos alcanzados.

Solución E2.

Podemos tomar

$$A = (-1,0), \qquad B = (-\infty,0).$$

En ambos casos 0 es la menor cota superior, así que $\sup A = \sup B = 0$. El conjunto A está acotado inferiormente y $\inf A = -1$.

En cambio, B no está acotado inferiormente: dado cualquier $m \in \mathbb{R}$, existe un elemento de B menor que m. Por ello no puede existir un ínfimo real de B. El ejemplo muestra que la existencia de un supremo es una propiedad unilateral: no implica acotación por abajo.

Solución E3.

La ecuación $0x = 0$ no determina un valor particular de x. Por el resultado $0x = 0$, válido para todo $x \in F$, **cada elemento del cuerpo es una solución**. Así, si el cuerpo contiene al menos dos elementos —y por definición $0 \neq 1$—, la ecuación tiene al menos las dos soluciones distintas

$$x = 0 \qquad \text{y} \qquad x = 1.$$

Esto ya refuta la conclusión $x = 1$ como solución única. Ahora examinemos la notación $0/0$. La división fue definida por $\frac{a}{b} := ab^{-1}$ solo cuando $b \neq 0$. Para formar $0/0$ necesitaríamos un inverso multiplicativo 0^{-1}, es decir, un elemento y tal que $0y = 1$. Pero $0y = 0$ para todo y, mientras que $0 \neq 1$. Tal inverso no existe. Por eso

$$\boxed{\frac{0}{0} \text{ no está definido.}}$$

La identidad $\frac{a}{a} = 1$ solo fue demostrada bajo la hipótesis $a \neq 0$. Sustituir $a = 0$ borra precisamente la condición que hace legal la división. El **primer paso ilegítimo** de la cadena es, por tanto,

$$0x = 0 \quad \Longrightarrow \quad x = \frac{0}{0},$$

porque intenta dividir ambos miembros por un elemento que no posee inverso multiplicativo.

Solución E4.

Para $k \in \mathbb{N}_{>0}$ definamos

$$I_{2k} = \left[0, \frac{1}{k}\right], \qquad I_{2k-1} = \left[1, 1 + \frac{1}{k}\right].$$

Todos son intervalos cerrados y no vacíos, y sus longitudes son $1/k$, que pueden hacerse menores que cualquier $\varepsilon > 0$. Sin embargo, un punto que perteneciera a todos los intervalos debería pertenecer simultáneamente a $I_2 = [0, 1]$ y a todos los intervalos impares que se concentran junto a 1, y también a $I_4 = [0, 1/2]$; de hecho, para $k \geq 2$, los intervalos pares están contenidos en $[0, 1/2]$ mientras los impares están contenidos en $[1, 2]$. Por tanto, $\bigcap_n I_n = \varnothing$. La familia no es encajada: por ejemplo, $I_1 = [1, 2]$ e $I_2 = [0, 1]$ no satisfacen $I_2 \subseteq I_1$. No se viola ninguna hipótesis del teorema porque precisamente falta el encajamiento.

Solución E5.

Como $1 \in S_{\mathbb{Q}}$, tenemos $3 \in T$, así que T es no vacío. Además 2 es cota superior de $S_{\mathbb{Q}}$, por lo que $2q + 1 \leq 5$ para todo $q \in S_{\mathbb{Q}}$; así, 5 es una cota superior racional de T.

Supongamos ahora que $t \in \mathbb{Q}$ fuese $t = \sup_{\mathbb{Q}} T$.

Definamos $s = \frac{t-1}{2} \in \mathbb{Q}$.

Para todo $q \in S_{\mathbb{Q}}$, de $2q + 1 \leq t$ se sigue $q \leq s$, así que s es cota superior racional de $S_{\mathbb{Q}}$. Si u es cualquier cota superior racional de $S_{\mathbb{Q}}$, entonces $2u + 1$ es cota superior racional de T. Como t es la menor de estas cotas, $t \leq 2u + 1$, y por tanto $s \leq u$. Así s sería el supremo racional de $S_{\mathbb{Q}}$, contradiciendo el resultado del capítulo. Luego T no tiene supremo en $\mathbb{Q}$.

Soluciones del nivel F. **Solución F1.**

Como $a \leq s$ para todo $a \in A$, $a + c \leq s + c$, de modo que $s + c$ es cota superior de $A + c$. Sea ahora $\varepsilon > 0$. Como $s = \sup A$, existe $a \in A$ tal que $s - \varepsilon < a \leq s$.

Sumando c, $(s + c) - \varepsilon < a + c \leq s + c$.

El punto $a + c$ pertenece a $A + c$, así que la caracterización aproximativa del supremo da

$$\boxed{\sup(A + c) = s + c}.$$

Solución F2.
Define

$$R = \max\left\{M_1, \dots, M_r, \frac{1}{\varepsilon_1}, \dots, \frac{1}{\varepsilon_s}\right\}.$$

Por la propiedad arquimediana existe $n \in \mathbb{N}_{>0}$ con $n > R$. Entonces $n > M_i$ para todo i y $n > \frac{1}{\varepsilon_j}$ para todo j. Como $n, \varepsilon_j > 0$, esto último equivale a $1/n < \varepsilon_j$. Una sola elección satisface todas las restricciones.

Solución F3.
Como $b - a > 0$, por la propiedad arquimediana podemos elegir $n \in \mathbb{N}_{>0}$ tal que

$$n > \max\left\{N, \frac{1}{b-a}\right\}.$$

Entonces $n > N$ y $na + 1 < nb$.
Por el lema de encajonamiento entero existe $k \in \mathbb{Z}$ con $k \le na < k + 1$.
Tomando $m = k + 1$ obtenemos

$$na < m \le na + 1 < nb.$$

Dividiendo por $n > 0$,

$$\boxed{a < \frac{m}{n} < b},$$

con el denominador además sujeto a $n > N$.

Solución F4.
Cada bisección divide la longitud por 2. Si $L_0 = 1$, entonces $L_n = 2^{-n}$; esto puede formalizarse por inducción.
Queremos $2^{-n} < \frac{1}{100}$, equivalentemente $2^n > 100$. Como

$$2^6 = 64 \le 100 < 128 = 2^7,$$

el menor valor es

$$\boxed{n = 7}.$$

En esa etapa obtenemos un intervalo racional cerrado $[a_7, b_7]$ tal que $\sqrt{7} \in [a_7, b_7]$ y $b_7 - a_7 = \frac{1}{128} < \frac{1}{100}$. Ese es un certificado exacto de incertidumbre, independientemente de que hayamos escrito los extremos.

Solución F5.
Por conmutatividad del producto, $1a = a$. Entonces

$$\begin{aligned} a + (-1)a &= 1a + (-1)a \\ &= (1 + (-1))a \quad \text{(distributividad)} \\ &= 0a \quad\quad\quad\quad\quad \text{(inverso aditivo de 1)} \\ &= 0. \end{aligned}$$

Así, $(-1)a$ es un inverso aditivo de a. Por unicidad,

$$\boxed{(-1)a = -a}.$$

Tomando $a = -1$ obtenemos $(-1)(-1) = -(-1) = 1$, donde la última igualdad usa que el inverso aditivo del inverso de 1 vuelve a ser 1.

Soluciones del nivel G. **Solución G1.**

Sea

$$\alpha = \sup A, \qquad \beta = \sup B.$$

Primero comprobamos que $\alpha + \beta$ es cota superior de $A + B$. Si $a \in A$ y $b \in B$, entonces

$$a \leq \alpha, \qquad b \leq \beta,$$

y sumando, $a + b \leq \alpha + \beta$.

Ahora debemos demostrar que ninguna cota menor funciona. Sea $\varepsilon > 0$. Por la caracterización aproximativa del supremo existen $a \in A$ y $b \in B$ tales que $\alpha - \frac{\varepsilon}{2} < a \leq \alpha$,

$$\beta - \frac{\varepsilon}{2} < b \leq \beta.$$

Sumando, $\alpha + \beta - \varepsilon < a + b \leq \alpha + \beta$.

Como $a + b \in A + B$, la caracterización aproximativa vuelve a decir que

$$\boxed{\sup(A + B) = \alpha + \beta = \sup A + \sup B.}$$

La completitud interviene para garantizar la existencia de α y β; el resto es orden y la caracterización del supremo.

Solución G2.

La condición de separación es

$$\ell < u \qquad \text{para todo } \ell \in L \text{ y todo } u \in U.$$

Como $U \neq \varnothing$, elegimos $u_0 \in U$. Para todo $\ell \in L$, $\ell < u_0$, de modo que u_0 es cota superior de L. Como $L \neq \varnothing$, completitud garantiza $c = \sup L$. Por definición de supremo, $\ell \leq c$ para todo $\ell \in L$. Por otro lado, cada $u \in U$ es cota superior de L, porque todo $\ell \in L$ satisface $\ell < u$. Como c es la menor cota superior, $c \leq u$ para todo $u \in U$.

Así,

$$\boxed{\ell \leq c \leq u}$$

para todos $\ell \in L$, $u \in U$.

Como $L \cup U = \mathbb{R}$, el real c pertenece a uno de los dos conjuntos. Son disjuntos, así que pertenece exactamente a uno.

- Si $c \in L$, como c domina a todo elemento de L, tenemos $c = \max L$.
- Si $c \in U$, como $c \leq u$ para todo $u \in U$, tenemos $c = \min U$.

Por tanto ocurre exactamente una de las dos posibilidades:

$$\boxed{c = \max L \quad \text{o} \quad c = \min U.}$$

Esta propiedad traduce completitud en lenguaje de cortes: una separación ordenada de la recta posee un punto frontera real.

Solución G3.

Sea $a > 0$ y $S = \{x \geq 0 : x^3 < a\}$.

La estrategia replica, con un grado algebraico mayor, la construcción de raíces cuadradas: completitud produce un candidato frontera y dos perturbaciones descartan que su cubo quede por debajo o por encima de a.

1. No vacuidad y acotación.

El número $t = \frac{a}{1+a}$ es positivo y pertenece a S. En efecto, si $a \geq 1$, entonces $0 < t < 1$ y $t^3 < 1 \leq a$; si $0 < a < 1$, entonces $0 < t < a < 1$, por lo que $t^3 < a^3 < a$. Además, $a + 1$ es cota superior de S. Si $x \geq a + 1$, entonces $x > 1$ y $x > a$ cuando $a < 1$, o $x > a \geq 1$ cuando $a \geq 1$; en ambos casos $x^3 > a$, de modo que tal x no pertenece a S.

Por completitud existe $\alpha = \sup S$.

Como $t \in S$ y $t > 0$, tenemos $\alpha > 0$.

2. No puede ocurrir $\alpha^3 < a$.

Supongamos $\alpha^3 < a$ y definamos $\delta = a - \alpha^3 > 0$.

Elegimos

$$h = \frac{1}{2} \min \left\{ 1, \frac{\delta}{3\alpha^2 + 3\alpha + 1} \right\} > 0.$$

Entonces $h < 1$ y $h(3\alpha^2 + 3\alpha + 1) < \delta$.

Como $h < 1$,

$$3\alpha h \leq 3\alpha, \qquad h^2 < 1,$$

y por tanto

$$\begin{aligned} (\alpha + h)^3 &= \alpha^3 + h(3\alpha^2 + 3\alpha h + h^2) \\ &< \alpha^3 + h(3\alpha^2 + 3\alpha + 1) \\ &< \alpha^3 + \delta \\ &= a. \end{aligned}$$

Así $\alpha + h \in S$, pero $\alpha + h > \alpha$, contradiciendo que α sea cota superior.

3. No puede ocurrir $\alpha^3 > a$.

Supongamos ahora $\alpha^3 > a$ y pongamos $\delta = \alpha^3 - a > 0$.

Elegimos

$$h = \frac{1}{2} \min \left\{ \alpha, \frac{\delta}{3\alpha^2} \right\} > 0$$

y definimos $c = \alpha - h$.

Como $h < \alpha$, $c > 0$. Además, $\alpha^3 - c^3 = 3\alpha^2 h - 3\alpha h^2 + h^3$.

Como $0 < h < \alpha$, el término $-3\alpha h^2 + h^3 = h^2(h - 3\alpha) < 0$, de modo que $\alpha^3 - c^3 < 3\alpha^2 h < \delta$.

Por tanto, $c^3 > \alpha^3 - \delta = a$.

Si $x \in S$ y $x \geq c$, como $x, c \geq 0$, la identidad

$$x^3 - c^3 = (x - c)(x^2 + xc + c^2) \geq 0$$

justifica directamente que $x^3 \geq c^3 > a$. Esto contradice la condición $x^3 < a$ que define a S. Así todo $x \in S$ satisface $x < c$. Por tanto c es una cota superior de S, pero $c < \alpha$, contradiciendo que α sea la menor cota superior. Los dos casos estrictos son imposibles. Por tricotomía,

$$\boxed{\alpha^3 = a.}$$

4. Unicidad.

Sean $0 < u < v$. Entonces

$$v^3 - u^3 = (v - u)(v^2 + uv + u^2) > 0.$$

Así el cubo es estrictamente creciente sobre los reales no negativos. Por consiguiente, no pueden existir dos números no negativos distintos con cubo a. Hemos demostrado, sin continuidad, que para todo $a > 0$ existe un único $\alpha > 0$ tal que

$$\boxed{\alpha^3 = a.}$$

La prueba muestra de nuevo la arquitectura central del capítulo:

$$\boxed{\text{conjunto de aproximantes} \to \text{supremo} \to \text{perturbaciones} \to \text{identificación de la frontera.}}$$

Capítulo 2

Funciones reales: estructura, composición, inversas y gráficas

Durante buena parte de la matemática escolar una función aparece casi siempre acompañada por una fórmula:

$$f(x) = x^2, \qquad f(x) = \frac{1}{x}, \qquad f(x) = \sigma(x).$$

Estas expresiones permiten calcular valores, pero dejan sin responder preguntas que el cálculo necesita formular con precisión: qué entradas están admitidas, en qué conjunto deben quedar las salidas y qué regla asigna una salida a cada entrada. Con la misma fórmula podemos definir funciones que tienen propiedades distintas, porque el dominio y el codominio forman parte del objeto que estudiamos. Para decidir qué podemos afirmar sobre una función necesitamos responder

¿qué información necesitamos para saber realmente qué función tenemos delante?

Un cambio en las entradas admitidas puede eliminar repeticiones de valores; un cambio en el conjunto de llegada puede dejar salidas sin alcanzar. Esas diferencias determinarán cuándo podemos recuperar la entrada a partir de la salida y cuándo tiene sentido encadenar dos funciones.

Un modelo periódico que podemos definir sin trigonometría. Para disponer de ejemplos de oscilación sin presuponer seno ni coseno, construiremos dos funciones por tramos. Todo real t pertenece a un único intervalo $[4m-1, 4m+3)$, con $m \in \mathbb{Z}$; la existencia y unicidad de ese entero se deducen de la propiedad arquimediana. Escribamos $r = t - 4m \in [-1, 3)$ y definamos

$$\sigma(t) = \begin{cases} r, & -1 \le r \le 1, \\ 2 - r, & 1 < r < 3, \end{cases} \qquad \kappa(t) = \sigma(t+1), \qquad \omega = 2.$$

Ambas funciones están definidas en toda la recta y sus valores pertenecen a $[-1, 1]$. Tienen período $4 = 2\omega$, y la definición da directamente $\sigma(1+2n) = (-1)^n$, $\sigma(1+4n) = 1$ y $\sigma(3+4n) = -1$ para todo entero n. Las fórmulas afines coinciden en las uniones: en $r = 1$ dan uno, y al pasar de $r = 3$ a $r = -1$ dan menos uno. Cuando dispongamos de la definición de continuidad, estos empalmes probarán continuidad en toda la recta. No afirmamos diferenciabilidad en los vértices. Este modelo triangular servirá en los capítulos iniciales para separar acotación, convergencia, continuidad y derivabilidad. Las funciones circulares se construirán desde el cálculo cuando contemos con la teoría de la integral; no necesitamos anticiparlas para probar ninguno de esos fenómenos.

Convenciones conjuntistas indispensables

En lo que sigue trabajaremos con conjuntos entendidos en su sentido matemático habitual, sin detenernos en su construcción fundacional. La notación $x \in A$ significa que x pertenece a A;

$A \subseteq B$ significa que todo elemento de A pertenece también a B. Dos conjuntos son iguales si tienen exactamente los mismos elementos. El conjunto vacío se denota por $\varnothing$.

Para conjuntos E, F usamos $E \cup F = \{x : x \in E \text{ o } x \in F\}$ (unión), $E \cap F = \{x : x \in E \text{ y } x \in F\}$ (intersección) y $E \setminus F = \{x : x \in E \text{ y } x \notin F\}$ (diferencia). Estas igualdades funcionan también como instrucciones de prueba: demostrar una pertenencia equivale a verificar la condición situada a la derecha. Si $E \subseteq U$, su **complemento relativo a** U es $U \setminus E$: sin especificar el conjunto ambiente, la palabra «complemento» resulta ambigua.

Un **par ordenado** (x, y) registra dos coordenadas en posiciones diferentes; $(x, y) = (u, v)$ equivale a $x = u$ e $y = v$. El **producto cartesiano** de A y B es $A \times B := \{(x, y) : x \in A,\ y \in B\}$. Esta convención permitirá definir una gráfica como conjunto de pares, sin exigir al lector un capítulo previo de teoría de conjuntos. Por último, $\forall x \in A$ se lee «para todo x de A»; $\exists x \in A$, «existe al menos un x de A»; y $\exists! x \in A$, «existe exactamente uno». Estas expresiones se interpretarán y utilizarán siempre dentro de las definiciones y pruebas concretas del cálculo.

2.1. Una fórmula no es todavía una función

Cuatro veces la misma fórmula. Comencemos con una expresión que no presenta ninguna dificultad algebraica: x^2.

Consideremos ahora las cuatro asignaciones

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = x^2,$$

$$g\colon [0, \infty) \to [0, \infty), \qquad g(x) = x^2,$$

$$h\colon \mathbb{R} \to [0, \infty), \qquad h(x) = x^2,$$

y

$$k\colon [0, \infty) \to \mathbb{R}, \qquad k(x) = x^2.$$

Las cuatro reglas elevan al cuadrado cada entrada admitida. Para compararlas debemos mirar también qué números podemos introducir y qué conjunto hemos declarado como destino de sus valores. En f podemos introducir números negativos y positivos, mientras que en g y k solo podemos introducir números no negativos. Por tanto, $f(-2)$ tiene sentido, pero $g(-2)$ y $k(-2)$ no están definidos.

También cambia el conjunto en el que declaramos que viven las salidas. En f y k el conjunto de llegada es todo $\mathbb{R}$. Sin embargo, ningún cuadrado real es negativo. En g y h, en cambio, el conjunto de llegada se ha ajustado precisamente a los números no negativos. Esta diferencia afectará pronto a preguntas como las siguientes:

- ¿pueden dos entradas distintas producir la misma salida?;
- ¿se obtiene realmente cada elemento del conjunto de llegada?;
- ¿podemos recuperar una entrada a partir de una salida?

Por ejemplo, $f(2) = f(-2) = 4$, mientras que en g no tenemos la pareja de entradas 2 y -2, porque -2 ni siquiera pertenece a su dominio. Y todo número no negativo aparece como cuadrado de algún número no negativo, gracias al teorema de existencia de raíces demostrado en `T1-C02`; por tanto, g alcanza todo su conjunto de llegada. Con la terminología que formalizaremos en §2.3, el contraste completo será este:

- f no es inyectiva ni sobreyectiva: 2 y -2 producen la misma salida, y ningún real negativo es alcanzado;
- g es biyectiva: entre los números no negativos no aparecen dos entradas distintas con el mismo cuadrado, y todo número no negativo es el cuadrado de su raíz no negativa;
- h es sobreyectiva pero no inyectiva: alcanza todo su codominio $[0, \infty)$, pero, por ejemplo, $h(2) = h(-2)$;
- k es inyectiva pero no sobreyectiva: al restringir las entradas a $[0, \infty)$ desaparece la duplicación producida por x y $-x$, pero siguen sin alcanzarse los números negativos de su codominio $\mathbb{R}$.

Las definiciones de §2.3 precisarán estas propiedades. Los ejemplos permiten reconocer desde ahora lo que deberán distinguir: una misma fórmula puede admitir entradas diferentes o dejar sin alcanzar parte del conjunto de llegada.

la misma fórmula puede determinar funciones estructuralmente distintas.

Ejemplo 2.1. **Una misma regla, cuatro funciones distintas.** Las funciones f, g, h, k anteriores utilizan la regla $x \mapsto x^2$, pero difieren en dominio o codominio. En particular:

- f admite entradas negativas y declara como codominio todo $\mathbb{R}$;
- g solo admite entradas no negativas y declara como codominio $[0, \infty)$;
- h admite todas las entradas reales, pero su codominio es $[0, \infty)$;
- k solo admite entradas no negativas, aunque su codominio declarado sea todo $\mathbb{R}$.

La fórmula aislada x^2 no contiene por sí sola toda esta información.

Antes de seguir
Si alguien escribe únicamente $p(x) = \frac{1}{x}$,
¿conoces ya el dominio de p?
No necesariamente. La expresión exige al menos que $x \neq 0$, pero podría estar describiendo, por ejemplo, $p\colon \mathbb{R} \setminus \{0\} \to \mathbb{R}$, o una restricción como $p\colon (0, \infty) \to (0, \infty)$. La fórmula ayuda a determinar qué entradas **son posibles**, pero el dominio matemático de una función es un dato que debe quedar fijado.

Qué significa realmente $f\colon A \to B$. La notación de conjuntos y los cuantificadores permiten reunir esas exigencias en una definición.

Definición 2.1. **Función, dominio, codominio e imagen.** Sean A y B conjuntos. Una **función**

$$f\colon A \to B$$

es una regla de asignación que a **cada** elemento $x \in A$ hace corresponder **exactamente un** elemento $y \in B$. Ese elemento se denota por $f(x)$. El conjunto A se llama **dominio** de f y B se llama **codominio**. El conjunto de valores que la función efectivamente toma,

$$\operatorname{Im}(f) = \{f(x) : x \in A\},$$

se llama **imagen** o **recorrido** de f. Siempre se cumple

$$\operatorname{Im}(f) \subseteq B.$$

La definición exige existencia y unicidad: para todo $x \in A$ debe haber una salida $f(x) \in B$, y una misma entrada no puede recibir dos salidas diferentes. La expresión «cada entrada tiene exactamente una salida» reúne ambas condiciones. Expresada mediante cuantificadores, la condición puede leerse esquemáticamente como

$$\forall x \in A \; \exists! y \in B \quad \text{tal que } y \text{ es la salida asignada a } x.$$

La notación $f(x)$ presupone que esa existencia y esa unicidad ya están garantizadas por la regla que define f. Esta reformulación permite diagnosticar dos maneras distintas en que una supuesta definición puede fracasar.

Primer fallo posible: alguna entrada no tiene salida. Supongamos que declaramos

$$r\colon \mathbb{R} \to \mathbb{R}, \qquad r(x) = \frac{1}{x}.$$

La fórmula funciona para todo real distinto de cero, pero en $x = 0$ exige dividir por cero. No existe ningún número real que pueda ser $r(0)$ de acuerdo con esa regla. Por tanto, la declaración anterior **no define una función sobre todo** $\mathbb{R}$. Podemos repararla cambiando el dominio:

$$r\colon \mathbb{R} \setminus \{0\} \to \mathbb{R}, \qquad r(x) = \frac{1}{x}.$$

Ahora sí existe una salida real para cada entrada permitida. Este ejemplo enseña una regla de lectura que será permanente:

antes de calcular con una fórmula, audita dónde está definida.

Segundo fallo posible: una entrada tiene más de una salida. Consideremos ahora la instrucción

> a cada $x \geq 0$, asígnele un número real y que satisfaga $y^2 = x$.

Para $x = 4$, tanto $y = 2$ como $y = -2$ satisfacen la condición. La existencia está garantizada, pero la unicidad falla. Así formulada, la instrucción **no define una función** de $[0, \infty)$ en $\mathbb{R}$. En `T1-C02` demostramos, sin embargo, que para cada $x \geq 0$ existe un único número **no negativo** cuyo cuadrado es x. Al añadir esa condición obtenemos una función perfectamente definida:

$$s\colon [0, \infty) \to [0, \infty), \qquad s(x) = \sqrt{x}.$$

La diferencia entre ambas instrucciones es exactamente la diferencia entre $\exists y$ y $\exists! y$. La elección de la raíz no negativa resuelve así la exigencia de unicidad que faltaba en la primera instrucción.

Lo que no exige la definición. La unicidad se refiere a la salida de una entrada fijada. Permite que entradas diferentes compartan un mismo valor, como ocurre con la regla

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = x^2,$$

es una función aunque $f(2) = f(-2) = 4$.

La definición exige **una sola salida para cada entrada**, no **una sola entrada para cada salida**. La segunda condición es más fuerte y recibirá el nombre de **inyectividad** en §2.3. Conviene mantener separados ambos niveles desde ahora:

ser función: una entrada no tiene dos salidas;
ser inyectiva: dos entradas no comparten una salida.

Dominio, codominio e imagen no son la misma cosa. Volvamos a

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = x^2.$$

Su dominio es $\mathbb{R}$.

Su codominio declarado también es $\mathbb{R}$. Pero su imagen es $[0, \infty)$, porque todo cuadrado es no negativo y, recíprocamente, todo real no negativo posee una raíz cuadrada real.

Así, $\operatorname{Im}(f) = [0, \infty) \subsetneq \mathbb{R}$.

El **codominio** forma parte de la declaración; la **imagen** recoge los valores que la regla alcanza al recorrer el dominio. Por eso calcular la imagen exige examinar la asignación, aunque el codominio ya esté fijado.

En cambio, para

$$h\colon \mathbb{R} \to [0, \infty), \qquad h(x) = x^2,$$

tenemos $\operatorname{Im}(h) = [0, \infty)$, de modo que imagen y codominio coinciden. Al cambiar el codominio cambia el conjunto de valores que preguntamos si la función alcanza. La asignación punto a punto puede permanecer intacta y, aun así, cambiar la respuesta a esa pregunta.

Error frecuente
No leas automáticamente $f\colon A \to B$ como si significara que f toma **todos** los valores de B. La flecha solo declara que las salidas de f pertenecen a B. La afirmación adicional $\operatorname{Im}(f) = B$ tendrá un nombre propio —sobreyectividad— y deberá demostrarse cuando sea necesaria.

Una convención sobre la identidad de las funciones. En distintos textos pueden encontrarse convenciones ligeramente diferentes acerca de qué datos se consideran parte de una función. En este libro adoptaremos una convención especialmente útil para el cálculo y para controlar composiciones e inversas:

> en una declaración $f\colon A \to B$, **dominio, codominio y regla de asignación forman parte de los datos de la función**.

Por ello, dos funciones serán iguales cuando coincidan esos datos.

Definición 2.2. **Igualdad, restricción y extensión.** Sean

$$f\colon A \to B, \qquad g\colon C \to D.$$

Diremos que $f = g$ si

$$A = C, \qquad B = D,$$

y

$$f(x) = g(x) \qquad \text{para todo } x \in A.$$

Si $E \subseteq A$, la **restricción** de f a E es la función

$$f|_E \colon E \to B$$

definida por

$$f|_E(x) = f(x) \qquad (x \in E).$$

Si $A \subseteq C$ y $g \colon C \to B$ satisface

$$g|_A = f,$$

diremos que g es una **extensión** de f.

La igualdad de funciones es, por tanto, más exigente que la igualdad de dos fórmulas escritas sobre una hoja.

Consideremos

$$f \colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = x^2,$$

y

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

En el dominio de g se cumple $f(x) = g(x)$. Pero $f \neq g$, porque sus dominios son diferentes. En realidad, $g = f|_{[0,\infty)}$. Así, **restricción** es el nombre preciso para lo que hemos hecho: conservar la regla y permitir menos entradas.

Restricción: perder entradas puede ganar estructura. Al restringir una función conservamos sus valores en las entradas admitidas, pero podemos eliminar pares de entradas que compartían una salida.

La función

$$f \colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = x^2,$$

produce el mismo valor en x y $-x$. Pero su restricción $f|_{[0,\infty)} \colon [0, \infty) \to \mathbb{R}$ ya no contiene simultáneamente esas dos entradas salvo en 0. Más adelante veremos que esta restricción es precisamente lo que permite invertir la operación de elevar al cuadrado sobre los números no negativos. Este ejemplo anticipa un principio general:

cambiar el dominio puede cambiar las propiedades de una función.

No cambia los valores que ya estaban definidos en los puntos conservados. Cambia el **objeto funcional** sobre el que formulamos preguntas.

Extensión: conservar lo conocido y definir más. La operación inversa a restringir es extender.

Por ejemplo, consideremos

$$u \colon [0, \infty) \to \mathbb{R}, \qquad u(x) = x.$$

La función identidad sobre $\mathbb{R}$,

$$U \colon \mathbb{R} \to \mathbb{R}, \qquad U(x) = x,$$

es una extensión de u porque coincide con ella en cada $x \geq 0$. Pero una extensión no tiene por qué ser única. También podemos definir

$$V(x) = \begin{cases} x, & x \geq 0, \\ 0, & x < 0, \end{cases}$$

como función $V\colon \mathbb{R} \to \mathbb{R}$. Entonces V también extiende a u.

Por tanto,

conocer una función en un subconjunto no determina, en general, una extensión única.

Esta observación será importante mucho más adelante, pero aquí solo necesitamos aprender a distinguir los objetos.

Una función no necesita una fórmula cerrada. Las expresiones algebraicas son una manera de comunicar una asignación. También una tabla, una descripción verbal o un algoritmo pueden determinar una salida única para cada entrada del dominio. Podemos describir una función mediante una tabla. Por ejemplo, si $A = \{a, b, c, d\}$, la tabla

x	a	b	c	d
$T(x)$	2	-1	2	5

define una función $T\colon A \to \mathbb{R}$.

Obsérvese nuevamente que $T(a) = T(c) = 2$ no causa ningún problema: cada entrada sigue teniendo una única salida. También podemos usar una descripción verbal. Por ejemplo,

$$\ell\colon \{\text{palabras finitas del alfabeto latino}\} \to \mathbb{N}$$

puede asignar a cada palabra su número de letras.

O podemos usar un algoritmo.

Ejemplo 2.2. **Una función descrita por un procedimiento.** Para cada $n \in \mathbb{N}_{>0}$, sea

$$d(n) = \text{número de divisores positivos de } n.$$

Esto define una función

$$d\colon \mathbb{N}_{>0} \to \mathbb{N}_{>0}.$$

Por ejemplo,

$$d(1) = 1, \qquad d(6) = 4, \qquad d(12) = 6,$$

porque

$$\{m \in \mathbb{N}_{>0} : m \mid 12\} = \{1, 2, 3, 4, 6, 12\}.$$

Para calcular $d(n)$ podemos emplear un procedimiento: examinar los enteros positivos hasta n y contar cuáles dividen exactamente a n. Esa descripción puede ser menos eficiente que una fórmula especializada, pero determina una salida única para cada entrada. Por tanto, determina una función.

La existencia de una función no depende de que poseamos una expresión algebraica breve para ella.

Fórmula, función y gráfica: tres niveles distintos. Podemos resumir ya la primera distinción del capítulo. Una **fórmula** es una manera posible de especificar una regla. Una **función** es el objeto que incluye entradas admitidas, salidas admitidas y una asignación única para cada entrada. Una **gráfica**, que estudiaremos formalmente en §2.7, será un conjunto de pares $(x, f(x))$ que representa geométricamente esa asignación. Los tres niveles se relacionan, pero no deben confundirse:

$$\boxed{\text{fórmula o regla} \longrightarrow \text{función completa} \longrightarrow \text{representaciones de la función.}}$$

Una misma función puede describirse por fórmulas distintas en regiones diferentes. Una misma fórmula puede utilizarse para funciones con dominios distintos. Y una gráfica puede ayudarnos a comprender una función sin sustituir la información precisa de dominio y codominio.

Cómo leer desde ahora una declaración funcional. Ante una escritura como $f\colon A \to B$, conviene hacer cuatro preguntas antes de manipularla:

1. ¿Cuál es el dominio A? Es decir, ¿qué entradas están permitidas?
2. ¿Cuál es el codominio B? Es decir, ¿dónde deben vivir las salidas?
3. ¿Cuál es la regla que determina $f(x)$ para cada $x \in A$?
4. ¿Qué conjunto constituye realmente la imagen $\operatorname{Im}(f)$?

La cuarta pregunta puede requerir una demostración. Las tres primeras forman parte de la especificación de la función.

Antes de seguir
Considera

$$p\colon [-2,2] \to \mathbb{R}, \qquad p(x) = x^2.$$

- El dominio es $[-2,2]$.
- El codominio es $\mathbb{R}$.
- La imagen es $[0,4]$.

¿Por qué la imagen no es simplemente $\mathbb{R}$? Porque el codominio declara **salidas permitidas**, mientras que la imagen registra **salidas realmente alcanzadas**. En este caso $p(x)$ nunca es negativo ni puede superar 4 cuando $|x| \leq 2$.

De valores individuales a conjuntos de valores. Una función también permite estudiar regiones enteras de su dominio y su codominio. Podemos reunir las salidas producidas por un subconjunto de entradas o buscar todas las entradas que llegan a una región fijada del conjunto de llegada. La imagen y la preimagen de conjuntos darán una forma precisa a esas dos operaciones.

2.2. Imágenes y preimágenes

En la sección anterior distinguimos tres conjuntos asociados a una función $f\colon A \to B$: el dominio A, el codominio B y la imagen total $\operatorname{Im}(f)$. Ahora vamos a refinar esa mirada. Al elegir una región $E \subseteq A$, la **imagen** reúne los valores producidos por sus elementos. Al fijar una región $F \subseteq B$, la **preimagen** reúne las entradas que terminan dentro de ella. La diferencia entre ambas

operaciones está en qué conjunto fijamos y en qué dirección seguimos la asignación. La diferencia de dirección conviene verla desde el principio:

$$\boxed{\begin{array}{ccc} E \subseteq A & \xrightarrow{\ f\ } & f(E) \subseteq B, \\ f^{-1}(F) \subseteq A & \xleftarrow{\ f\ } & F \subseteq B. \end{array}}$$

La imagen avanza en el sentido de la función. La preimagen retrocede desde un conjunto de posibles salidas hacia las entradas que llegan a él.

De la imagen de un punto a la imagen de un conjunto. Cuando $x \in A$, ya sabemos qué significa $f(x)$. Es un **elemento** de B: la salida asignada a la entrada x. Ahora sea $E \subseteq A$. En lugar de introducir una sola entrada, dejamos que x recorra todo E. Reunimos entonces todas las salidas obtenidas.

Definición 2.3. **Imagen y preimagen de subconjuntos.** Sea

$$f\colon A \to B$$

una función.
Si $E \subseteq A$, la **imagen de E bajo f** es el subconjunto de B

$$f(E) = \{f(x) : x \in E\}.$$

Equivalentemente,

$$y \in f(E) \iff \exists x \in E \text{ tal que } f(x) = y.$$

Si $F \subseteq B$, la **preimagen** o **imagen inversa de F bajo f** es el subconjunto de A

$$f^{-1}(F) = \{x \in A : f(x) \in F\}.$$

Equivalentemente,

$$x \in f^{-1}(F) \iff f(x) \in F.$$

La definición contiene una asimetría lógica importante. Para demostrar que un elemento y pertenece a $f(E)$ debemos encontrar **alguna entrada** $x \in E$ que produzca y. La pertenencia a una imagen contiene, por tanto, una afirmación existencial:

$$y \in f(E) \iff \exists x\,(x \in E \wedge f(x) = y).$$

En cambio, para decidir si x pertenece a $f^{-1}(F)$ no tenemos que buscar ningún testigo adicional. Basta comprobar dónde cae su imagen: $x \in f^{-1}(F) \iff f(x) \in F$. Las convenciones conjuntistas anteriores se vuelven operativas: abrir una definición transforma una afirmación sobre conjuntos en una condición sobre elementos y cuantificadores.

Un punto no es un conjunto. La notación puede parecer casi idéntica, pero debemos distinguir $f(x)$ de $f(E)$. Si $x \in A$, entonces $f(x)$ es un elemento de B. Si $E \subseteq A$, entonces $f(E)$ es un subconjunto de B.

Por ejemplo, si

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = x^2,$$

entonces $f(2) = 4$, mientras que $f(\{2\}) = \{4\}$.

La primera salida es un número. La segunda es un conjunto que contiene ese número.

En general, $f(\{x\}) = \{f(x)\}$.

Mantener esta diferencia de tipos permite combinar las notaciones sin confundir un valor con el conjunto que lo contiene.

El recorrido es una imagen particular. En §2.1 definimos $\mathrm{Im}(f) = \{f(x) : x \in A\}$.

Con la notación recién introducida podemos escribir simplemente

$$\boxed{\mathrm{Im}(f) = f(A).}$$

El recorrido de una función es, por tanto, la imagen de **todo su dominio**. Esto permite separar dos afirmaciones que se confunden con frecuencia: $f(A) \subseteq B$ siempre, porque todas las salidas deben pertenecer al codominio, mientras que $f(A) = B$ es una afirmación adicional: dice que no queda ningún elemento del codominio sin alcanzar. En §2.3 daremos nombre y estructura lógica a esa propiedad.

Seguir un conjunto hacia delante. Volvamos a la función

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = x^2.$$

Tomemos $E = [-1, 2]$.

Queremos calcular $f(E)$.

Evaluar únicamente los extremos sugeriría $[f(-1), f(2)] = [1, 4]$.

Eso sería incorrecto, porque $0 \in E$ y $f(0) = 0$.

El conjunto imagen contiene, por tanto, valores menores que 1. Para determinar la imagen mediante un conjunto candidato necesitamos las dos inclusiones: toda salida producida desde E debe pertenecer al candidato, y cada valor del candidato debe proceder de alguna entrada de E. La segunda inclusión es la que impide añadir valores que la función nunca alcanza. Primero, si $x \in [-1, 2]$, entonces $|x| \le 2$, de modo que $0 \le x^2 \le 4$.

Por tanto, $f([-1, 2]) \subseteq [0, 4]$.

Ahora tomemos un $y \in [0, 4]$. Por la existencia de raíces establecida en `T1-C02`, existe $\sqrt{y} \in [0, 2]$. En particular, $\sqrt{y} \in [-1, 2]$, y $f(\sqrt{y}) = (\sqrt{y})^2 = y$.

Así, cada $y \in [0, 4]$ pertenece a $f([-1, 2])$. Luego $[0, 4] \subseteq f([-1, 2])$.

Concluimos que

$$\boxed{f([-1, 2]) = [0, 4].}$$

Este cálculo revela una diferencia importante entre **evaluar una función en puntos** y **determinar la imagen de un conjunto**. La segunda tarea suele ser un problema de conjuntos: debemos caracterizar todas las salidas posibles, no solamente calcular algunos valores.

Error frecuente
Para una función cualquiera no existe una regla general que permita obtener la imagen de un intervalo evaluando solamente sus extremos. En este capítulo aprenderemos más adelante qué información adicional —por ejemplo, ciertas propiedades de orden— permite controlar mejor las imágenes de intervalos. Aquí no debemos suponerla de antemano.

Retroceder desde las salidas: la preimagen. Consideremos de nuevo

$$f(x) = x^2, \qquad f\colon \mathbb{R} \to \mathbb{R},$$

pero ahora partamos del conjunto de salidas $F = [1, 4]$. Buscamos ahora los números reales cuyo cuadrado queda entre 1 y 4; el conjunto de salidas está fijado y debemos recuperar las entradas que llegan a él.

Por definición, $f^{-1}([1, 4]) = \{x \in \mathbb{R} : 1 \leq x^2 \leq 4\}$.

La condición $1 \leq x^2 \leq 4$ es equivalente a $1 \leq |x| \leq 2$.

Por tanto,

$$\boxed{f^{-1}([1, 4]) = [-2, -1] \cup [1, 2].}$$

Observa la geometría algebraica del resultado: un solo intervalo situado en el lado no negativo del codominio tiene dos ramas de antecedentes, porque x y $-x$ poseen el mismo cuadrado. No hemos definido todavía la inyectividad formalmente, pero ya estamos viendo qué aspecto tiene su fracaso: una misma salida puede provenir de más de una entrada.

Una misma función, imágenes y preimágenes.

Ejemplo 2.3. **Imágenes y preimágenes bajo $x \mapsto x^2$.** Sea

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = x^2.$$

Entonces: $f([-1, 2]) = [0, 4]$,

$$f(\{-2, 0, 3\}) = \{0, 4, 9\},$$

$$f^{-1}([1, 4]) = [-2, -1] \cup [1, 2],$$

$$f^{-1}(\{4\}) = \{-2, 2\},$$

$$f^{-1}(\{0\}) = \{0\},$$

y $f^{-1}(\{-1\}) = \varnothing$.

La última igualdad no significa que la preimagen «no exista». La preimagen existe y es un conjunto perfectamente definido; ocurre simplemente que ningún número real tiene cuadrado -1.

Este ejemplo reúne tres posibilidades que conviene distinguir:

- una salida puede tener varios antecedentes;
- una salida puede tener exactamente un antecedente;
- una salida puede no tener antecedentes.

En todos los casos, la **preimagen del conjunto** está bien definida.

Resolver una ecuación como calcular una preimagen. La notación de preimagen permite reinterpretar una tarea algebraica familiar. Resolver $f(x) = y$ para una salida fija $y \in B$ equivale a calcular $f^{-1}(\{y\})$.

En efecto,

$$x \in f^{-1}(\{y\}) \iff f(x) \in \{y\} \iff f(x) = y.$$

Por ejemplo, para $f(x) = x^2$,

$$f^{-1}(\{4\}) = \{x \in \mathbb{R} : x^2 = 4\} = \{-2, 2\}.$$

Así, la solución de la ecuación $x^2 = 4$ puede verse como la preimagen del conjunto unitario $\{4\}$. Pero una preimagen general es más amplia que resolver una sola ecuación. Calcular $f^{-1}([1, 4])$ significa resolver una **condición de pertenencia**: $f(x) \in [1, 4]$, es decir, $1 \leq f(x) \leq 4$. Para $f(x) = x^2$, eso conduce a una inecuación, no a una ecuación puntual.

Podemos resumir:

$$\boxed{\begin{aligned} f^{-1}(\{y\}) &= \text{conjunto de soluciones de } f(x) = y, \\ f^{-1}(F) &= \text{conjunto de soluciones de } f(x) \in F. \end{aligned}}$$

Un ejemplo lineal: la dirección importa. Sea

$$g\colon \mathbb{R} \to \mathbb{R}, \qquad g(x) = 2x - 3.$$

Calculemos primero la imagen de $E = [0, 2]$.

Si $x \in [0, 2]$, entonces $0 \leq x \leq 2$.

Multiplicando por 2 y restando 3 obtenemos $-3 \leq 2x - 3 \leq 1$.

Por tanto, $g([0, 2]) \subseteq [-3, 1]$.

Recíprocamente, si $y \in [-3, 1]$, definimos $x = \frac{y+3}{2}$.

De $-3 \leq y \leq 1$ se sigue que $0 \leq x \leq 2$, y además

$$g(x) = 2\left(\frac{y+3}{2}\right) - 3 = y.$$

Luego

$$\boxed{g([0, 2]) = [-3, 1].}$$

Ahora calculemos la preimagen de $F = [-1, 5]$.

Por definición,

$$g^{-1}([-1, 5]) = \{x \in \mathbb{R} : -1 \leq 2x - 3 \leq 5\}.$$

Sumando 3 y dividiendo por 2 obtenemos $1 \leq x \leq 4$.

Así,

$$\boxed{g^{-1}([-1, 5]) = [1, 4].}$$

Los dos cálculos pueden parecer similares porque la función es algebraicamente sencilla, pero responden a preguntas opuestas:

- para hallar $g(E)$, partimos de las **entradas** y caracterizamos sus salidas;
- para hallar $g^{-1}(F)$, partimos de las **salidas permitidas** y caracterizamos las entradas que llegan a ellas.

La preimagen no exige una función inversa. La notación de preimagen requiere distinguir la operación sobre conjuntos de una posible función inversa.

Para

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = x^2,$$

acabamos de escribir $f^{-1}([1,4])$.

Sin embargo, ya vimos en §2.1 que esta función toma el mismo valor en 2 y -2, y tampoco alcanza los reales negativos. Más adelante demostraremos que una función con estas características no posee una función inversa $f^{-1}\colon \mathbb{R} \to \mathbb{R}$. El significado depende de qué objeto se está construyendo. La expresión $f^{-1}(F)$ que hemos definido en esta sección significa **preimagen del conjunto** F **bajo** f. Está definida para cualquier función $f\colon A \to B$ y cualquier subconjunto $F \subseteq B$. La expresión $f^{-1}(y)$ cuando f^{-1} es una **función inversa** pertenece a otra construcción, que estudiaremos en §2.5 y que solo existe bajo hipótesis adicionales.

La notación f^{-1} tiene dos usos
No concluyas que existe una función inversa solo porque aparece una expresión como $f^{-1}(F)$. La preimagen de conjuntos existe para toda función. Cuando una verdadera función inversa exista, el contexto indicará qué significado tiene f^{-1}. Más adelante veremos además que ambos usos son compatibles en el caso biyectivo.

Tampoco significa tomar el recíproco. Existe una segunda confusión frecuente. Si f es una función real, la expresión $\frac{1}{f}$ puede utilizarse, donde tenga sentido, para la función definida por

$$\left(\frac{1}{f}\right)(x) = \frac{1}{f(x)}.$$

Esto no tiene relación con la preimagen y tampoco es, en general, una función inversa. Por ejemplo, si $f(x) = x^2$, entonces $\frac{1}{f(x)} = \frac{1}{x^2}$ para $x \neq 0$.

En cambio, $f^{-1}([1,4]) = [-2,-1] \cup [1,2]$ es un **conjunto de entradas**. Son objetos de tipos completamente diferentes.

Comprobación de tipos. Una manera eficaz de evitar errores es preguntar siempre qué clase de objeto espera cada notación.

Si $f\colon A \to B$, entonces:

- para formar $f(x)$ necesitamos $x \in A$;
- para formar $f(E)$ necesitamos $E \subseteq A$;
- para formar $f^{-1}(F)$ como preimagen necesitamos $F \subseteq B$.

Y los resultados son: $f(x) \in B$,

$$f(E) \subseteq B,$$

$$f^{-1}(F) \subseteq A.$$

Podemos organizarlo así:

entrada de la notación	operación	resultado
$x \in A$	$f(x)$	elemento de B
$E \subseteq A$	$f(E)$	subconjunto de B
$F \subseteq B$	$f^{-1}(F)$	subconjunto de A

Esta «comprobación de tipos» es especialmente útil cuando dominio y codominio son conjuntos de naturaleza distinta.

Cuatro pruebas de consistencia. Antes de pasar a ejemplos más elaborados, las definiciones nos proporcionan cuatro controles inmediatos.

Primero, $f(\varnothing) = \varnothing$.

No hay entradas en el conjunto vacío, por lo que no puede producirse ninguna salida desde él.

Segundo, $f^{-1}(\varnothing) = \varnothing$.

Ninguna entrada puede tener su imagen dentro de un conjunto que no contiene elementos.

Tercero, $f(A) = \mathrm{Im}(f)$, como ya observamos.

Cuarto,

$$\boxed{f^{-1}(B) = A.}$$

Esta última igualdad merece atención. Toda entrada $x \in A$ satisface $f(x) \in B$ por la propia definición de función $f\colon A \to B$. Por tanto, todas las entradas pertenecen a la preimagen del codominio. No necesitamos que f alcance todos los elementos de B. Así, puede ocurrir simultáneamente que $f(A) \neq B$ pero $f^{-1}(B) = A$. La primera igualdad fallida habla de qué **salidas se alcanzan**. La segunda, siempre verdadera, habla de dónde están autorizadas a caer las salidas.

Preimagen de un valor que no se alcanza. Sea otra vez

$$f(x) = x^2, \qquad f\colon \mathbb{R} \to \mathbb{R}.$$

Como -1 pertenece al codominio, tiene sentido preguntar por $f^{-1}(\{-1\})$. Pero ningún $x \in \mathbb{R}$ satisface $x^2 = -1$. Por tanto, $f^{-1}(\{-1\}) = \varnothing$. Esto muestra por qué debemos distinguir «pertenecer al codominio» de «ser alcanzado por la función». El codominio determina qué preguntas de preimagen están tipadas correctamente. La imagen total determina cuáles de esas preguntas tienen antecedentes efectivos.

Un método práctico para calcular imágenes. Cuando una función está dada por una fórmula y queremos calcular $f(E)$, conviene seguir este esquema:

1. **Propón un conjunto candidato** para las salidas posibles.
2. **Primera inclusión:** toma $x \in E$ y demuestra que $f(x)$ pertenece al candidato.
3. **Segunda inclusión:** toma una salida candidata y y construye, si es posible, un $x \in E$ tal que $f(x) = y$.
4. Concluye la igualdad por doble inclusión.

El paso 3 suele ser el más informativo. No basta saber que ningún valor se sale del conjunto candidato; debemos saber también que no hemos incluido valores que la función nunca alcanza desde E. Este método reaparecerá en §2.3 al estudiar sobreyectividad: allí la construcción de un antecedente será precisamente el núcleo de la demostración.

Un método práctico para calcular preimágenes. Para calcular $f^{-1}(F)$ el procedimiento suele ser más directo:

1. escribe la definición $f^{-1}(F) = \{x \in A : f(x) \in F\}$;
2. traduce la pertenencia $f(x) \in F$ a ecuaciones, inecuaciones o condiciones conocidas;
3. resuelve esas condiciones dentro del dominio A;
4. expresa el conjunto resultante de forma conveniente.

Por ejemplo, si $f(x) = x^2$ y $F = (1, 4]$, entonces

$$x \in f^{-1}((1,4]) \iff 1 < x^2 \leq 4 \iff 1 < |x| \leq 2.$$

Luego

$$\boxed{f^{-1}((1,4]) = [-2,-1) \cup (1,2].}$$

Nótese que no hemos «invertido la fórmula x^2». Hemos traducido una condición de pertenencia y la hemos resuelto.

Cuántos antecedentes tiene una salida. La imagen y la preimagen ya nos permiten formular con precisión las dos preguntas que motivaron la apertura del capítulo.

Para $f\colon A \to B$, podemos preguntar:

1. ¿cada $y \in B$ pertenece a $f(A)$?
2. para cada $y \in B$, ¿el conjunto $f^{-1}(\{y\})$ contiene como máximo un elemento?

La primera pregunta trata de si se alcanza todo el codominio. La segunda trata de si una salida puede provenir de dos entradas distintas. En §2.3 convertiremos esas preguntas en las definiciones formales de **sobreyectividad** e **inyectividad**. Por ahora basta conservar el mapa conceptual:

imagen: qué salidas produce un conjunto de entradas;
preimagen: qué entradas terminan en un conjunto de salidas.

Y queda pendiente una cuestión más sutil. Si combinamos subconjuntos mediante uniones, intersecciones o complementos, ¿cómo se comportan sus imágenes y preimágenes? Algunas igualdades funcionan siempre; otras requieren hipótesis adicionales. En §2.6, la comparación entre las dos inclusiones de cada igualdad permitirá reconocer qué leyes valen para toda función y cuáles necesitan hipótesis adicionales.

2.3. Inyectividad, sobreyectividad y biyectividad

La cantidad de antecedentes de cada salida permite distinguir dos propiedades de una función $f\colon A \to B$. La primera miraba hacia el codominio:

> ¿todo $y \in B$ es realmente alcanzado por alguna entrada?

La segunda miraba hacia las fibras de un valor:

> si dos entradas producen la misma salida, ¿tienen necesariamente que ser la misma entrada?

Alcanzar todo el codominio exige un antecedente para cada salida; impedir que dos entradas compartan una salida exige que ese antecedente, cuando exista, sea único. Podemos estudiar ambas condiciones por separado.

Tres propiedades y tres preguntas distintas.

DEFINICIÓN 2.4. **Inyectividad, sobreyectividad y biyectividad.** Sea

$$f\colon A \to B.$$

Diremos que f es **inyectiva** si, para cualesquiera $x_1, x_2 \in A$,

$$f(x_1) = f(x_2) \quad \Longrightarrow \quad x_1 = x_2.$$

Diremos que f es **sobreyectiva** si cada elemento del codominio es imagen de alguna entrada; es decir, si

$$\forall y \in B\ \exists x \in A \quad f(x) = y.$$

Equivalentemente,

$$f(A) = B.$$

Diremos que f es **biyectiva** si es simultáneamente inyectiva y sobreyectiva.

Las tres palabras responden, por tanto, a preguntas muy concretas:

- **inyectiva:** ¿pueden dos entradas distintas compartir una salida?;
- **sobreyectiva:** ¿queda algún elemento del codominio sin ser alcanzado?;
- **biyectiva:** ¿cada salida del codominio corresponde a una entrada y solo a una?

La última frase debe leerse con precisión. Una función biyectiva establece una correspondencia uno a uno entre el dominio y el codominio. La existencia de una función inversa formalizará esta idea en §2.5; no necesitamos todavía ese resultado para trabajar con la definición.

Inyectividad: empezar suponiendo igualdad de imágenes. La forma lógica de la inyectividad ya nos indica cómo suele comenzar una demostración.

Queremos probar

$$\forall x_1, x_2 \in A, \qquad f(x_1) = f(x_2) \Longrightarrow x_1 = x_2.$$

Por tanto, el esqueleto estándar es:

1. toma $x_1, x_2 \in A$ arbitrarios;
2. supone que $f(x_1) = f(x_2)$;
3. usa la regla de la función y las propiedades del dominio para deducir $x_1 = x_2$.

Consideremos, por ejemplo,

$$p\colon \mathbb{R} \to \mathbb{R}, \qquad p(x) = 3x - 7.$$

Para probar que p es inyectiva, sean $x_1, x_2 \in \mathbb{R}$ y supongamos $p(x_1) = p(x_2)$.

Entonces $3x_1 - 7 = 3x_2 - 7$.

Sumando 7 y dividiendo por 3 obtenemos $x_1 = x_2$.

Luego p es inyectiva.

Obsérvese que no hemos comenzado escogiendo una salida arbitraria. Para la inyectividad, la información inicial útil es la **igualdad de dos imágenes**.

Error frecuente
Demostrar que para cada $x \in A$ existe un único valor $f(x)$ **no demuestra inyectividad**. Eso demuestra que la regla está bien definida como función: una entrada no recibe dos salidas. La inyectividad exige una afirmación diferente: $f(x_1) = f(x_2) \Longrightarrow x_1 = x_2$. Es decir, dos entradas no pueden compartir una misma salida salvo que en realidad sean la misma entrada.

Cómo demostrar que una función no es inyectiva. Negar la definición produce un criterio muy útil. Una función $f\colon A \to B$ **no** es inyectiva si existen $x_1, x_2 \in A$ tales que $x_1 \neq x_2$ pero $f(x_1) = f(x_2)$. Por tanto, para refutar la inyectividad basta un solo par de entradas distintas con la misma imagen.

Para

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = x^2,$$

tenemos $2 \neq -2$ pero $f(2) = 4 = f(-2)$.

Ese par constituye un contraejemplo completo a la inyectividad. En el lenguaje de §2.2, la misma observación puede expresarse diciendo que $f^{-1}(\{4\}) = \{-2, 2\}$ contiene más de un elemento.

Sobreyectividad: comenzar con una salida arbitraria. La forma lógica de la sobreyectividad es diferente:

$$\forall y \in B \; \exists x \in A \quad f(x) = y.$$

Ahora el procedimiento natural es:

1. toma un $y \in B$ arbitrario;
2. busca o construye una entrada $x \in A$ que pueda producirlo;
3. verifica que esa entrada pertenece realmente al dominio;
4. comprueba que $f(x) = y$.

Volvamos a

$$p\colon \mathbb{R} \to \mathbb{R}, \qquad p(x) = 3x - 7.$$

Sea $y \in \mathbb{R}$ arbitrario. Queremos resolver $3x - 7 = y$ para x. La ecuación sugiere $x = \frac{y+7}{3}$. Como $y \in \mathbb{R}$, también $\frac{y+7}{3} \in \mathbb{R}$, de modo que la entrada propuesta pertenece al dominio. Finalmente,

$$p\left(\frac{y+7}{3}\right) = 3\left(\frac{y+7}{3}\right) - 7 = y.$$

Como el argumento funciona para todo $y \in \mathbb{R}$, la función p es sobreyectiva. Este ejemplo ilustra un principio general: una prueba de sobreyectividad es, con frecuencia, una **prueba de existencia parametrizada por una salida arbitraria**.

Cómo demostrar que una función no es sobreyectiva. Para refutar la sobreyectividad debemos exhibir al menos un elemento del codominio que no sea alcanzado.

La negación de

$$\forall y \in B\ \exists x \in A, \qquad f(x) = y$$

es

$$\exists y \in B\ \forall x \in A, \qquad f(x) \neq y.$$

Así, basta encontrar un $y \in B$ tal que $f^{-1}(\{y\}) = \varnothing$.

Por ejemplo, para

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = x^2,$$

el número -1 pertenece al codominio, pero $f^{-1}(\{-1\}) = \varnothing$.

Luego f no es sobreyectiva.

La misma conclusión se obtiene a partir del recorrido: $f(\mathbb{R}) = [0, \infty) \neq \mathbb{R}$.

Una misma fórmula, las cuatro posibilidades. Ahora podemos volver con plena precisión al contraste que abrió el capítulo.

Ejemplo 2.4. **Cuatro clasificaciones con la regla $x \mapsto x^2$.** Consideremos

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = x^2,$$

$$g\colon [0, \infty) \to [0, \infty), \qquad g(x) = x^2,$$

$$h\colon \mathbb{R} \to [0, \infty), \qquad h(x) = x^2,$$

y

$$k\colon [0, \infty) \to \mathbb{R}, \qquad k(x) = x^2.$$

Entonces:

1. f no es inyectiva ni sobreyectiva;
2. g es biyectiva;
3. h es sobreyectiva pero no inyectiva;
4. k es inyectiva pero no sobreyectiva.

La raíz no negativa proporciona antecedentes, y la factorización de una diferencia de cuadrados permite comprobar cuándo son únicos.

1. La función f no es inyectiva. Ya vimos que $f(2) = f(-2) = 4$ con $2 \neq -2$.

2. La función f no es sobreyectiva. Ningún real negativo pertenece a su imagen; por ejemplo, $f^{-1}(\{-1\}) = \varnothing$.

3. La función g es inyectiva. Sean $x_1, x_2 \in [0, \infty)$ y supongamos $x_1^2 = x_2^2$.

Entonces $(x_1 - x_2)(x_1 + x_2) = 0$.

Si $x_1 - x_2 = 0$, ya tenemos $x_1 = x_2$. Si $x_1 + x_2 = 0$, como ambos números son no negativos, necesariamente $x_1 = x_2 = 0$. En ambos casos $x_1 = x_2$. Por tanto, g es inyectiva.

4. La función g es sobreyectiva. Sea $y \in [0,\infty)$. Por el resultado de existencia de raíces de `T1-C02`, existe $\sqrt{y} \in [0,\infty)$ y $g(\sqrt{y}) = (\sqrt{y})^2 = y$.

Así, cada elemento del codominio tiene un antecedente. Luego g es sobreyectiva y, por tanto, biyectiva.

5. La función h es sobreyectiva. El mismo argumento con $\sqrt{y}$ funciona para todo $y \in [0,\infty)$.

6. La función h no es inyectiva. De nuevo, $h(2) = h(-2)$.

7. La función k es inyectiva. Su dominio coincide con el de g, y el argumento anterior para entradas no negativas demuestra la inyectividad.

8. La función k no es sobreyectiva. Su codominio es $\mathbb{R}$, pero ningún número negativo es un cuadrado real.

La tabla completa es entonces

función	inyectiva	sobreyectiva
$f\colon \mathbb{R} \to \mathbb{R}$	no	no
$g\colon [0,\infty) \to [0,\infty)$	sí	sí
$h\colon \mathbb{R} \to [0,\infty)$	no	sí
$k\colon [0,\infty) \to \mathbb{R}$	sí	no

Una sola expresión algebraica ha producido las cuatro combinaciones posibles.

El dominio controla colisiones. Comparemos

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = x^2,$$

con su restricción $k = f|_{[0,\infty)}$.

La primera no es inyectiva y la segunda sí. Al eliminar los números negativos desaparecen los pares de entradas distintas x y $-x$ que producían la misma salida. La fórmula conserva sus valores en las entradas que siguen admitidas. Esto sugiere una operación frecuente:

> cuando una función falla en ser inyectiva porque distintas regiones del dominio repiten valores, puede ser útil restringir el dominio a una región donde esas colisiones desaparezcan.

Pero la palabra «puede» importa. Restringir el dominio no garantiza automáticamente la inyectividad. Si restringimos x^2 al dominio $[-2,-1] \cup [1,2]$, siguen apareciendo pares x y $-x$ con la misma imagen. Por tanto, una restricción debe elegirse por una razón matemática concreta.

El codominio controla la sobreyectividad. Comparemos ahora

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = x^2,$$

con

$$h\colon \mathbb{R} \to [0,\infty), \qquad h(x) = x^2.$$

Dominio y regla son iguales. Sin embargo, f no es sobreyectiva y h sí. La diferencia está únicamente en el codominio. Para cualquier función $u\colon A \to B$, si sabemos que su imagen es $u(A) = C \subseteq B$, podemos considerar la misma asignación con codominio ajustado a su imagen:

$$\widetilde{u}\colon A \to C, \qquad \widetilde{u}(x) = u(x).$$

Por construcción, $\widetilde{u}(A) = C$, así que $\widetilde{u}$ es sobreyectiva. Según la convención de §2.1, este cambio de codominio define una función distinta de u. La asignación de valores permanece igual; lo que cambia es el conjunto de llegada que forma parte de sus datos. Esta observación evita una frase imprecisa que aparece a veces en cálculos informales: «toda función es sobreyectiva sobre su imagen». La idea es correcta si se entiende que estamos **cambiando el codominio y, por tanto, formando una nueva función**.

Qué cambia y qué no al modificar dominio o codominio. Conviene reunir las observaciones anteriores.

- **Restringir el dominio** puede eliminar colisiones y convertir una función no inyectiva en inyectiva.
- Restringir el dominio también puede hacer que se pierdan valores antes alcanzados y, por tanto, puede destruir sobreyectividad.
- **Cambiar el codominio**, manteniendo la misma regla y el mismo dominio, no altera si dos entradas producen o no la misma salida; por tanto, no altera la cuestión de la inyectividad mientras el nuevo codominio siga conteniendo todas las imágenes.
- En cambio, el codominio es parte esencial de la sobreyectividad: una misma asignación puntual puede ser sobreyectiva hacia un codominio y no serlo hacia otro mayor.

Esta sensibilidad explica por qué una declaración funcional completa $f\colon A \to B$ no puede sustituirse impunemente por una fórmula aislada.

Dos pruebas, dos arquitecturas lógicas. Supongamos que queremos demostrar que

$$p\colon \mathbb{R} \to \mathbb{R}, \qquad p(x) = 5x + 1$$

es biyectiva.

No conviene escribir una sola cadena de cálculos sin indicar qué propiedad estamos probando. Una prueba rigurosa tiene dos partes conceptualmente distintas.

Inyectividad. Sean $x_1, x_2 \in \mathbb{R}$ y supongamos $p(x_1) = p(x_2)$.

Entonces $5x_1 + 1 = 5x_2 + 1$, de donde $x_1 = x_2$.

Sobreyectividad. Sea $y \in \mathbb{R}$ arbitrario. Tomemos $x = \frac{y-1}{5}$.

Este x pertenece al dominio y satisface

$$p(x) = 5\left(\frac{y-1}{5}\right) + 1 = y.$$

Por tanto, p es inyectiva y sobreyectiva; luego es biyectiva. La diferencia entre las dos mitades se ve incluso antes de calcular:

inyectividad	$\forall x_1\, \forall x_2$
sobreyectividad	$\forall y\, \exists x$

En una, comparamos dos entradas arbitrarias. En la otra, debemos construir una entrada que puede depender de una salida arbitraria.

Negaciones útiles para buscar contraejemplos. Las negaciones también tienen formas diferentes. Para refutar la inyectividad buscamos

$$\boxed{\exists x_1, x_2 \in A : \quad x_1 \neq x_2 \quad \text{y} \quad f(x_1) = f(x_2).}$$

Para refutar la sobreyectividad buscamos

$$\boxed{\exists y \in B: \quad \forall x \in A, \quad f(x) \neq y.}$$

En lenguaje de preimágenes:

- no inyectiva: alguna preimagen puntual contiene al menos dos elementos;
- no sobreyectiva: alguna preimagen puntual es vacía.

Esto hace visible una conexión muy útil entre §§2.2 y 2.3. Para cada $y \in B$, observemos el conjunto $f^{-1}(\{y\})$.

Entonces:

- la sobreyectividad exige que **ninguno** de esos conjuntos sea vacío;
- la inyectividad exige que **ninguno** tenga más de un elemento;
- la biyectividad exige que cada uno contenga **exactamente un** elemento.

Podemos condensarlo así:

propiedad	tamaño de $f^{-1}(\{y\})$
inyectiva	≤ 1 para todo $y \in B$
sobreyectiva	≥ 1 para todo $y \in B$
biyectiva	$= 1$ para todo $y \in B$

No estamos usando cardinalidad infinita ni una teoría nueva de tamaños de conjuntos. Solo distinguimos aquí entre vacío, unitario y «más de un elemento».

Antes de seguir. Cuando leas una afirmación de la forma

«f es biyectiva»,

no intentes demostrarla de una sola vez. Pregunta primero:

1. ¿cómo probaré la inyectividad?
2. ¿cómo probaré la sobreyectividad?

Y cuando leas

«f no es biyectiva»,

recuerda que basta con que falle **una** de las dos propiedades. No necesitas demostrar que fallan ambas. Las dos obligaciones se conservan al estudiar una cadena de funciones: habrá que seguir qué ocurre con una igualdad de salidas y cómo se construyen antecedentes de un valor final. Las nociones disponibles se relacionan así:

$$\boxed{\text{función completa} \to \text{imagen/preimagen} \to \text{inyectividad/sobreyectividad/biyectividad.}}$$

2.4. Composición e identidad

Muchas expresiones del cálculo encadenan operaciones: el resultado de una transformación pasa a ser la entrada de otra. Para describir la cadena debemos precisar tanto las reglas que se aplican como los valores que cada etapa admite. Por ejemplo, para calcular $(3x-1)^2$ podemos pensar en dos etapas: $x \longmapsto 3x-1 \longmapsto (3x-1)^2$. La primera etapa produce un valor intermedio; la segunda toma precisamente ese valor como nueva entrada. La **composición de funciones** formaliza esta idea de encadenamiento. Pero hay una cuestión que debemos resolver antes de

escribir fórmulas: **¿puede la segunda función aceptar todas las salidas que realmente produce la primera?**

Compatibilidad entre las dos etapas. Sean $f\colon A \to B$ y $g\colon D \to C$.

Queremos aplicar primero f y después g. Si partimos de $x \in A$, la primera etapa produce $f(x) \in B$. Para que tenga sentido escribir $g(f(x))$, necesitamos además que ese valor $f(x)$ pertenezca al dominio D de g. Por tanto, la condición exacta es

$$\boxed{f(A) \subseteq D.}$$

Cuando se cumple, cada salida efectiva de f es una entrada permitida para g. El caso más frecuente se presenta como una cadena tipada $A \xrightarrow{f} B \xrightarrow{g} C$, es decir,

$$f\colon A \to B, \qquad g\colon B \to C.$$

Aquí la compatibilidad es automática, porque $f(A) \subseteq B$ y B es precisamente el dominio de g. La condición general $f(A) \subseteq D$ es un poco más flexible: el codominio declarado de f no tiene por qué coincidir literalmente con el dominio de g si sabemos que **las salidas que f toma de hecho** sí pertenecen a ese dominio.

Error frecuente
No basta con que las fórmulas puedan escribirse una junto a la otra de manera informal. Considera

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = x^2,$$

y

$$g\colon (0,\infty) \to \mathbb{R}, \qquad g(y) = \frac{1}{y}.$$

Para $x \neq 0$, la expresión $g(f(x)) = 1/x^2$ tiene sentido. Pero $f(0) = 0$ y $0 \notin (0,\infty)$, de modo que $g \circ f$ **no está definida como función sobre todo** $\mathbb{R}$. La fórmula $1/x^2$ no corrige silenciosamente el problema de dominio. Habría que restringir la primera etapa a $\mathbb{R} \setminus \{0\}$ o modificar legítimamente los datos de las funciones.

Definición 2.5. **Composición e identidad.** Sean

$$f\colon A \to B, \qquad g\colon D \to C,$$

funciones tales que

$$f(A) \subseteq D.$$

La **composición de g con f** es la función

$$g \circ f\colon A \to C$$

definida por

$$(g \circ f)(x) = g(f(x)) \qquad (x \in A).$$

Para cualquier conjunto A, la **función identidad sobre** A es

$$\mathrm{id}_A \colon A \to A$$

dada por

$$\mathrm{id}_A(x) = x \qquad (x \in A).$$

La notación $g \circ f$ debe leerse de derecha a izquierda en el orden de aplicación:

$$\boxed{(g \circ f)(x) = g(f(x)) = \text{primero } f\text{, después } g.}$$

El símbolo $\circ$ no significa multiplicación. Tampoco autoriza a intercambiar el orden de las funciones.

Componer es construir una función nueva. Tomemos

$$f \colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = 3x - 1,$$

y

$$g \colon \mathbb{R} \to [0, \infty), \qquad g(y) = y^2.$$

Entonces $g \circ f \colon \mathbb{R} \to [0, \infty)$ y

$$(g \circ f)(x) = g(f(x)) = g(3x - 1) = (3x - 1)^2.$$

La expresión compuesta puede parecer una sola fórmula, pero estructuralmente conserva las dos etapas:

$$\mathbb{R} \xrightarrow{x \mapsto 3x-1} \mathbb{R} \xrightarrow{y \mapsto y^2} [0, \infty).$$

Esta perspectiva será importante muchas veces: una función complicada puede entenderse descomponiéndola en funciones más simples, y una función nueva puede construirse encadenando transformaciones conocidas.

El orden importa: la composición no es conmutativa. El hecho de que $g \circ f$ exista no implica que $f \circ g$ exista. Para la segunda composición hay que realizar **otra** auditoría de dominios. Y aun cuando ambas existan, no tienen por qué ser iguales.

Ejemplo 2.5. **Dos composiciones existentes y distintas.** Sean

$$f \colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = x + 1,$$

y

$$g \colon \mathbb{R} \to \mathbb{R}, \qquad g(x) = x^2.$$

Ambas composiciones están definidas sobre todo $\mathbb{R}$. Sin embargo, $(g \circ f)(x) = g(x+1) = (x+1)^2$, mientras que $(f \circ g)(x) = f(x^2) = x^2 + 1$.

Por ejemplo, en $x = 1$,

$$(g \circ f)(1) = 4, \qquad (f \circ g)(1) = 2.$$

Por tanto,

$$\boxed{g \circ f \neq f \circ g.}$$

Este contraejemplo basta para descartar una supuesta ley de conmutatividad universal. La composición tiene otra propiedad, en cambio, que sí es universal cuando las funciones están tipadas de manera compatible: **la asociatividad**.

Asociatividad: tres etapas, una sola cadena. Supongamos

$$f\colon A \to B, \qquad g\colon B \to C, \qquad h\colon C \to D.$$

Podemos componer primero f con g y después aplicar h: $h \circ (g \circ f)$. O podemos componer primero g con h y después anteponer f: $(h \circ g) \circ f$. La intuición sugiere que ambas rutas ejecutan las mismas tres etapas

$$x \longmapsto f(x) \longmapsto g(f(x)) \longmapsto h(g(f(x))).$$

Pero, según nuestra convención de igualdad de funciones, para demostrar que las dos composiciones son iguales no basta con observar que la fórmula final «parece la misma». Debemos comprobar primero que poseen el mismo dominio y codominio, y después que coinciden en cada punto.

Proposición 2.1. **Asociatividad de la composición.** Sean

$$f\colon A \to B, \qquad g\colon B \to C, \qquad h\colon C \to D.$$

Entonces

$$\boxed{h \circ (g \circ f) = (h \circ g) \circ f.}$$

Idea de la prueba. Las dos funciones empiezan en A y terminan en D. Por tanto, una vez auditados los tipos, basta comprobar que asignan la misma salida a cada $x \in A$.

Demostración. Como $g \circ f\colon A \to C$, la composición $h \circ (g \circ f)\colon A \to D$ está bien definida. Del mismo modo, $h \circ g\colon B \to D$, por lo que $(h \circ g) \circ f\colon A \to D$ también está bien definida.

Así, ambas funciones tienen dominio A y codominio D.

Sea ahora $x \in A$. Entonces

$$\begin{aligned} \big(h \circ (g \circ f)\big)(x) &= h\big((g \circ f)(x)\big) \\ &= h(g(f(x))) \\ &= (h \circ g)(f(x)) \\ &= \big((h \circ g) \circ f\big)(x). \end{aligned}$$

Como las dos funciones tienen el mismo dominio, el mismo codominio y coinciden en cada $x \in A$, son iguales. □

Después de la prueba
La asociatividad permite escribir sin ambigüedad $h \circ g \circ f$ cuando los tipos son compatibles. Los paréntesis pueden omitirse porque cualquier agrupación admisible produce la misma función. Pero **no** podemos alterar el orden: $h \circ g \circ f$ significa aplicar primero f, después g y finalmente h.

La identidad no transforma nada. La función identidad

$$\mathrm{id}_A \colon A \to A, \qquad \mathrm{id}_A(x) = x,$$

conserva cada entrada. Por eso, al incorporarla en el lugar compatible de una composición, obtendremos los mismos valores que antes: será el elemento neutro de esa operación. Si $f \colon A \to B$, entonces podemos componer f por la derecha con id_A y por la izquierda con id_B.

Proposición 2.2. **Leyes de identidad.** Para toda función

$$f \colon A \to B,$$

se cumple

$$\boxed{\mathrm{id}_B \circ f = f \qquad \text{y} \qquad f \circ \mathrm{id}_A = f.}$$

Demostración. Las composiciones están bien definidas porque $f(A) \subseteq B$ y $\mathrm{id}_A(A) = A$.

Además, $\mathrm{id}_B \circ f \colon A \to B$, igual que f. Para cada $x \in A$, $(\mathrm{id}_B \circ f)(x) = \mathrm{id}_B(f(x)) = f(x)$. Por la convención de igualdad de funciones, $\mathrm{id}_B \circ f = f$.

Análogamente, $f \circ \mathrm{id}_A \colon A \to B$, y para cada $x \in A$, $(f \circ \mathrm{id}_A)(x) = f(\mathrm{id}_A(x)) = f(x)$.

Por tanto, $f \circ \mathrm{id}_A = f$. □

La posición de la identidad importa para que los tipos coincidan. A la izquierda de f aparece la identidad del **codominio** B; a la derecha aparece la identidad del **dominio** A.

Qué sucede con inyectividad y sobreyectividad al componer. La sección anterior separó dos arquitecturas de prueba:

- para inyectividad, partimos de una igualdad de salidas;
- para sobreyectividad, partimos de una salida arbitraria y construimos un antecedente.

La composición permite ver ambas arquitecturas en cadena.

Proposición 2.3. **Preservación por composición.** Sean

$$f \colon A \to B, \qquad g \colon B \to C.$$

Entonces:

1. si f y g son inyectivas, $g \circ f$ es inyectiva;
2. si f y g son sobreyectivas, $g \circ f$ es sobreyectiva;
3. si f y g son biyectivas, $g \circ f$ es biyectiva.

Idea de la prueba. Para la inyectividad, una igualdad al final de la cadena debe retroceder primero a través de g y después a través de f. Para la sobreyectividad, comenzamos en el codominio final y construimos antecedentes en sentido contrario al orden de aplicación: primero uno bajo g y luego uno bajo f.

Demostración.

1. Inyectividad. Supongamos que f y g son inyectivas. Sean $x_1, x_2 \in A$ y supongamos $(g \circ f)(x_1) = (g \circ f)(x_2)$.

Por definición de composición, $g(f(x_1)) = g(f(x_2))$.

Como g es inyectiva, $f(x_1) = f(x_2)$.

Y como f es inyectiva, $x_1 = x_2$.

Por tanto, $g \circ f$ es inyectiva.

2. Sobreyectividad. Supongamos que f y g son sobreyectivas. Sea $z \in C$ arbitrario.

Como g es sobreyectiva, existe $y \in B$ tal que $g(y) = z$. Como f es sobreyectiva, existe $x \in A$ tal que $f(x) = y$.

Entonces

$$(g \circ f)(x) = g(f(x)) = g(y) = z.$$

Hemos encontrado, para un $z \in C$ arbitrario, un $x \in A$ cuya imagen por $g \circ f$ es z. Por tanto, $g \circ f$ es sobreyectiva.

3. Biyectividad. Si f y g son biyectivas, ambas son inyectivas y sobreyectivas. Por las partes 1 y 2, $g \circ f$ es inyectiva y sobreyectiva; por tanto, es biyectiva. □

Leer la composición como una cadena de dependencias. Hay una diferencia instructiva entre calcular y demostrar. Para **evaluar** $(g \circ f)(x)$, avanzamos en el orden natural de la cadena: $x \xrightarrow{f} f(x) \xrightarrow{g} g(f(x))$. Para demostrar **sobreyectividad**, en cambio, la construcción de testigos recorre la cadena hacia atrás: $z \in C \xleftarrow{g} y \in B \xleftarrow{f} x \in A$. Y para demostrar **inyectividad**, una igualdad de salidas también retrocede:

$$g(f(x_1)) = g(f(x_2)) \Longrightarrow f(x_1) = f(x_2) \Longrightarrow x_1 = x_2.$$

En una biyección, cada salida tiene exactamente un antecedente. Podemos entonces intentar usar ese antecedente como el valor de una nueva función que recorra la asignación en sentido contrario. La composición y la identidad permitirán comprobar que deshace la función original:

función completa → imagen/preimagen → clasificación → composición e identidad.

2.5. Funciones inversas: cuándo se puede deshacer una función

Para invertir una función $f\colon A \to B$ debemos asignar a cada elemento de B un único antecedente en A. La biyectividad garantiza las dos condiciones que requiere esa asignación: cada salida tiene antecedente y ese antecedente es único. La idea de «deshacer» una función es familiar en ejemplos elementales. Si una regla suma 5, esperamos deshacerla restando 5; si multiplica por 3, esperamos deshacerla dividiendo por 3. Pero una función general no tiene por qué admitir una operación inversa. Antes de buscar fórmulas necesitamos identificar la condición estructural que hace posible la recuperación de la entrada.

Supongamos, por ejemplo,

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

La función es sobreyectiva, pero no inyectiva. Si la salida es 4, hay dos entradas posibles:

$$f(2) = 4, \qquad f(-2) = 4.$$

Por tanto, intentar «deshacer» f enviando 4 a su antecedente no produce una salida única.

En cambio, para

$$k\colon [0, \infty) \to \mathbb{R}, \qquad k(x) = x^2,$$

la función es inyectiva, pero no sobreyectiva. Cada valor alcanzado tiene un único antecedente, pero un número como -1 no tiene antecedente alguno. Una supuesta inversa definida sobre todo el codominio $\mathbb{R}$ fallaría ahora por existencia. Los dos obstáculos son exactamente los que distinguimos al definir una función:

sobreyectividad $\Rightarrow$ existencia del antecedente,	inyectividad $\Rightarrow$ unicidad del antecedente.

La biyectividad reúne ambas condiciones.

Qué significa realmente deshacer una función. La composición y la identidad nos permiten formular la idea sin depender de una fórmula algebraica.

Definición 2.6. **Función inversa.** Sea

$$f\colon A \to B.$$

Una función

$$g\colon B \to A$$

se llama **inversa** de f si satisface simultáneamente

$$g \circ f = \mathrm{id}_A$$

y

$$f \circ g = \mathrm{id}_B\,.$$

El teorema siguiente demostrará que, cuando tal función existe, es única. En ese caso se denota por

$$f^{-1}\colon B \to A.$$

Las dos identidades expresan dos maneras de recuperar el punto de partida. Si comenzamos con $x \in A$, aplicamos primero f y después la inversa, debemos volver a x: $f^{-1}(f(x)) = x$. Si comenzamos con $y \in B$, aplicamos primero la inversa y después f, debemos volver a y: $f(f^{-1}(y)) = y$. Una inversa debe, por tanto, estar definida en la dirección contraria y deshacer a f **en ambos sentidos**. Cada identidad comprueba la recuperación de un punto de partida distinto.

No confundir tres notaciones
Las expresiones siguientes son conceptualmente distintas:

1. **Función inversa:** si $f\colon A \to B$ es biyectiva, $f^{-1}\colon B \to A$.
2. **Preimagen de un conjunto:** para cualquier función $f\colon A \to B$ y cualquier $E \subseteq B$, $f^{-1}(E) = \{x \in A : f(x) \in E\}$.
3. **Recíproco de una función real:** cuando $f(x) \neq 0$, $\frac{1}{f}(x) = \frac{1}{f(x)}$.

La preimagen existe como operación sobre subconjuntos aunque f no sea biyectiva. La función inversa, en cambio, solo existirá bajo una condición que demostraremos enseguida.

La condición exacta: biyectividad. La sobreyectividad dará un antecedente para cada entrada de la inversa, y la inyectividad asegurará que su valor quede determinado. En la dirección recíproca, las dos identidades de composición permitirán recuperar esas propiedades.

TEOREMA 2.1. **Existencia de la función inversa.** Sea

$$f\colon A \to B.$$

Entonces f admite una función inversa

$$f^{-1}\colon B \to A$$

si y solo si f es biyectiva. Cuando existe, la inversa es única.

Idea de la prueba. En un sentido, si una función g deshace a f, las identidades de composición obligan a que f no identifique dos entradas distintas y a que todo elemento del codominio tenga antecedente. En el otro sentido, si f es biyectiva, para cada $y \in B$ existe exactamente un $x \in A$ con $f(x) = y$; definimos la inversa asignando a y ese único antecedente.

Demostración. Probaremos las dos implicaciones por separado.

($\Rightarrow$) **Supongamos que f admite inversa.** Sea $g\colon B \to A$ una inversa de f. Entonces

$$g \circ f = \mathrm{id}_A, \qquad f \circ g = \mathrm{id}_B\,.$$

Primero demostraremos que f es inyectiva. Sean $x_1, x_2 \in A$ y supongamos $f(x_1) = f(x_2)$.

Aplicamos g a ambos miembros: $g(f(x_1)) = g(f(x_2))$.

Como $g \circ f = \mathrm{id}_A$, $x_1 = x_2$.

Por tanto, f es inyectiva.

Ahora demostraremos que f es sobreyectiva. Sea $y \in B$ arbitrario. Como $g(y) \in A$, podemos elegir $x = g(y)$.

Entonces

$$f(x) = f(g(y)) = (f \circ g)(y) = \mathrm{id}_B(y) = y.$$

Así, todo $y \in B$ posee un antecedente bajo f. Por tanto, f es sobreyectiva.

Concluimos que f es biyectiva.

($\Leftarrow$) **Supongamos ahora que f es biyectiva.** Queremos construir una función de B en A que recupere antecedentes.

Sea $y \in B$. Como f es sobreyectiva, existe al menos un $x \in A$ tal que $f(x) = y$. Como f es inyectiva, ese x es único: si $x_1, x_2 \in A$ satisfacen $f(x_1) = y = f(x_2)$, entonces $f(x_1) = f(x_2)$, y la inyectividad implica $x_1 = x_2$. Por tanto, para cada $y \in B$ existe un único $x \in A$ con $f(x) = y$. Podemos definir $g\colon B \to A$ por la regla

$$g(y) = \text{el único } x \in A \text{ tal que } f(x) = y.$$

Verifiquemos ahora que g deshace a f en ambos sentidos. Sea $x \in A$. El elemento x es un antecedente de $f(x)$. Como $g(f(x))$ está definido precisamente como el único antecedente de $f(x)$, $g(f(x)) = x$.

Por tanto, $g \circ f = \mathrm{id}_A$.

Sea ahora $y \in B$. Por definición de $g(y)$, $f(g(y)) = y$.

Luego $f \circ g = \mathrm{id}_B$.

Así, g es una inversa de f. Nos falta justificar la unicidad. Supongamos que $g, h\colon B \to A$ son ambas inversas de f. Entonces

$$g \circ f = \mathrm{id}_A, \qquad f \circ h = \mathrm{id}_B\,.$$

Usando las leyes de identidad y la asociatividad demostradas en §2.4,

$$\begin{aligned} g &= g \circ \mathrm{id}_B \\ &= g \circ (f \circ h) \\ &= (g \circ f) \circ h \\ &= \mathrm{id}_A \circ h \\ &= h. \end{aligned}$$

Por tanto, la inversa es única. Podemos denotarla sin ambigüedad por f^{-1}. □

Después de la prueba
La equivalencia anterior puede recordarse por su estructura lógica, no por una frase aislada:

sobreyectiva	$\Longleftrightarrow$	cada salida tiene al menos un antecedente,
inyectiva	$\Longleftrightarrow$	cada salida tiene a lo sumo un antecedente,
biyectiva	$\Longleftrightarrow$	cada salida tiene exactamente un antecedente.

La última línea es precisamente lo que permite convertir «buscar antecedentes» en una función $B \to A$.

Las identidades de inversión. Una vez que f es biyectiva, el teorema nos permite escribir $f^{-1}\colon B \to A$ y las dos identidades fundamentales quedan

$$\boxed{f^{-1} \circ f = \mathrm{id}_A, \qquad f \circ f^{-1} = \mathrm{id}_B\,.}$$

Punto a punto,

$$\boxed{f^{-1}(f(x)) = x \quad (x \in A), \qquad f(f^{-1}(y)) = y \quad (y \in B).}$$

Además, f^{-1} también es biyectiva. En efecto, acabamos de ver que f es una inversa de f^{-1}, porque las mismas dos identidades se leen ahora en el sentido contrario. Por tanto,

$$\boxed{(f^{-1})^{-1} = f.}$$

La inversa de una composición invierte el orden. La composición de biyecciones vuelve a ser biyectiva por §2.4. Así que, si

$$f\colon A \to B, \qquad g\colon B \to C$$

son biyectivas, la función $g \circ f\colon A \to C$ admite inversa. Para deshacer dos etapas, la intuición correcta es recorrerlas en orden contrario. Si primero aplicamos f y después g, debemos deshacer primero g y después f.

COROLARIO 2.1. **Inversa de una composición.** Si

$$f\colon A \to B, \qquad g\colon B \to C$$

son biyectivas, entonces

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

Idea de la prueba. Como la inversa de una función es única, basta demostrar que $f^{-1} \circ g^{-1}$ deshace a $g \circ f$ por ambos lados.

Demostración. Por el teorema anterior existen

$$f^{-1}\colon B \to A, \qquad g^{-1}\colon C \to B.$$

Por tanto, $f^{-1} \circ g^{-1}\colon C \to A$ está bien definida.

Usando asociatividad e identidades,

$$\begin{aligned}(f^{-1} \circ g^{-1}) \circ (g \circ f) &= f^{-1} \circ (g^{-1} \circ g) \circ f \\ &= f^{-1} \circ \mathrm{id}_B \circ f \\ &= f^{-1} \circ f \\ &= \mathrm{id}_A\,.\end{aligned}$$

De manera análoga,

$$\begin{aligned}(g \circ f) \circ (f^{-1} \circ g^{-1}) &= g \circ (f \circ f^{-1}) \circ g^{-1} \\ &= g \circ \mathrm{id}_B \circ g^{-1} \\ &= g \circ g^{-1} \\ &= \mathrm{id}_C\,.\end{aligned}$$

Así, $f^{-1} \circ g^{-1}$ es una inversa de $g \circ f$. Por unicidad, $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$. □

El orden se invierte porque debemos deshacer primero la última operación aplicada. Si la cadena hace $A \xrightarrow{f} B \xrightarrow{g} C$, la cadena inversa debe regresar como $C \xrightarrow{g^{-1}} B \xrightarrow{f^{-1}} A$.

La raíz cuadrada aparece como una inversa. Volvamos a la familia de funciones con la que comenzó el capítulo.

EJEMPLO 2.6. **El cuadrado restringido y la raíz cuadrada.** Consideremos

$$q\colon [0,\infty) \to [0,\infty), \qquad q(x) = x^2.$$

En §2.3 demostramos que q es biyectiva: es inyectiva porque en el dominio no negativo el cuadrado conserva el orden estricto, y es sobreyectiva porque cada $y \geq 0$ posee una raíz cuadrada no negativa. Por el Teorema Teorema 2.1, q admite inversa. Para $y \geq 0$, el único $x \geq 0$ que satisface $x^2 = y$ es $x = \sqrt{y}$.

Por tanto,

$$q^{-1}\colon [0,\infty) \to [0,\infty), \qquad q^{-1}(y) = \sqrt{y}.$$

Las identidades de inversión dicen

$$\sqrt{x^2} = x \qquad (x \geq 0)$$

y

$$(\sqrt{y})^2 = y \qquad (y \geq 0).$$

La restricción $x \geq 0$ en la primera identidad es esencial. Sobre todo $\mathbb{R}$, $\sqrt{x^2} = |x|$, no x en general. Precisamente por eso la función $x \mapsto x^2$ sobre $\mathbb{R}$ no puede tener como inversa a la raíz cuadrada: antes hay que eliminar la duplicación x y $-x$ restringiendo legítimamente el dominio.

Este ejemplo muestra una idea que reaparecerá muchas veces: una función que no es invertible globalmente puede volverse invertible después de restringir su dominio de modo que sea inyectiva y tomando como codominio exactamente la imagen pertinente.

Inversa no significa recíproco. Consideremos la función biyectiva

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = 2x + 1.$$

Para encontrar su inversa resolvemos $y = 2x + 1$ para x: $x = \frac{y-1}{2}$.

Por tanto, $f^{-1}(y) = \frac{y-1}{2}$.

En cambio, el recíproco de los valores de f es $\frac{1}{f(x)} = \frac{1}{2x+1}$, que ni siquiera está definido en $x = -\frac{1}{2}$.

Así,

$$\boxed{f^{-1} \neq \frac{1}{f}}$$

en general. El exponente -1 en f^{-1} indica **inversión respecto de la composición**, no inversión multiplicativa de valores.

Función inversa y preimagen: misma notación, objetos distintos. Terminemos resolviendo la ambigüedad que apareció por primera vez en §2.2. Si $f\colon A \to B$ es una función cualquiera y $E \subseteq B$, la escritura $f^{-1}(E)$ designa la **preimagen del conjunto** E: $f^{-1}(E) = \{x \in A : f(x) \in E\}$. Esto tiene sentido aunque f no sea biyectiva. Si f es biyectiva, en cambio, también existe la función inversa $f^{-1}\colon B \to A$. Entonces, para un punto $y \in B$, $f^{-1}(y)$ es un elemento de A: el único antecedente de y. El tipo de argumento permite distinguir ambos usos:

$$\boxed{\begin{array}{lcll} y \in B & \Longrightarrow & f^{-1}(y) \in A & \text{si existe la función inversa,} \\ E \subseteq B & \Longrightarrow & f^{-1}(E) \subseteq A & \text{para cualquier función.} \end{array}}$$

Cuando f es biyectiva, ambos significados son compatibles: la preimagen de un conjunto E está formada exactamente por las imágenes de sus elementos bajo la función inversa. Pero no debemos usar esa compatibilidad para olvidar que la operación de preimagen existe en un contexto mucho más general. La idea central de la sección puede resumirse así:

$$\boxed{\begin{array}{c} \text{cada salida tiene exactamente un antecedente} \\ \Longleftrightarrow \text{ la función es biyectiva} \\ \Longleftrightarrow \text{ existe una función inversa.} \end{array}}$$

En §2.6 volveremos a imágenes y preimágenes, ahora con una pregunta distinta: cómo interactúan con uniones, intersecciones, diferencias y complementos, y en qué punto la inyectividad se vuelve necesaria para transformar una inclusión en igualdad.

2.6. El álgebra de imágenes y preimágenes

Al combinar conjuntos mediante uniones, intersecciones o diferencias podemos comparar dos procedimientos: combinarlos antes de tomar imágenes o preimágenes, o efectuar esas operaciones primero y combinar los resultados. La igualdad entre ambos procedimientos deberá demostrarse en cada caso.

Sea $f\colon A \to B$.

Si $E, F \subseteq A$ y $S, T \subseteq B$, podemos preguntar, por ejemplo, si es legítimo escribir $f(E \cup F) = f(E) \cup f(F)$,

$$f(E \cap F) = f(E) \cap f(F),$$

o $f^{-1}(S \cap T) = f^{-1}(S) \cap f^{-1}(T)$.

Las tres fórmulas parecen naturales. Dos de ellas son siempre correctas. La restante puede fallar. Demostrar ambas inclusiones permitirá decidir cuáles de estas identidades valen siempre. Si una inclusión falla, el punto donde se detenga el argumento puede revelar qué hipótesis hace falta o cómo construir un contraejemplo.

Una primera ley: la inclusión se conserva hacia delante y hacia atrás. Empecemos por una propiedad que no presenta ninguna sorpresa. Si $E \subseteq F \subseteq A$, entonces todo elemento de E pertenece también a F. Por tanto, cualquier valor obtenido usando una entrada de E también se obtiene usando una entrada de F:

$$\boxed{E \subseteq F \Longrightarrow f(E) \subseteq f(F).}$$

Del mismo modo, si $S \subseteq T \subseteq B$, y $x \in f^{-1}(S)$, entonces $f(x) \in S \subseteq T$, de modo que $x \in f^{-1}(T)$. Por tanto,

$$\boxed{S \subseteq T \Longrightarrow f^{-1}(S) \subseteq f^{-1}(T).}$$

Tanto la imagen como la preimagen son, pues, **monótonas respecto de la inclusión de conjuntos**. Aquí la palabra «monótona» se refiere solamente al orden dado por $\subseteq$ entre subconjuntos; no estamos introduciendo todavía la monotonía de funciones reales que estudiaremos en §2.9.

Una igualdad muy plausible. Probemos ahora algo menos automático. Supongamos que $E, F \subseteq A$. Es tentador afirmar

$$\boxed{f(E \cap F) = f(E) \cap f(F).}$$

La igualdad parece razonable: si una entrada pertenece a ambos conjuntos, su imagen pertenece a ambas imágenes. Pero una igualdad de conjuntos exige **dos inclusiones**. Auditémoslas por separado.

Primera inclusión. Queremos demostrar $f(E \cap F) \subseteq f(E) \cap f(F)$.

Sea $y \in f(E \cap F)$.

Por definición de imagen, existe $x \in E \cap F$ tal que $f(x) = y$.

Como $x \in E \cap F$, tenemos simultáneamente

$$x \in E \qquad \text{y} \qquad x \in F.$$

Por tanto, $y = f(x) \in f(E)$ y también $y = f(x) \in f(F)$.

Así, $y \in f(E) \cap f(F)$.

Hemos demostrado, sin ninguna hipótesis adicional sobre f, que

$$\boxed{f(E \cap F) \subseteq f(E) \cap f(F).}$$

La primera inclusión vale para cualquier función; queda examinar la recíproca.

Intentemos ahora la inclusión inversa. Queremos probar $f(E) \cap f(F) \subseteq f(E \cap F)$.

Sea $y \in f(E) \cap f(F)$.

Entonces

$$y \in f(E) \qquad \text{y} \qquad y \in f(F).$$

La primera pertenencia nos da **algún** $u \in E$ tal que $f(u) = y$. La segunda nos da **algún** $v \in F$ tal que $f(v) = y$.

Por tanto, $f(u) = f(v)$.

La igualdad de las dos imágenes todavía no identifica sus antecedentes. Para concluir que $y \in f(E \cap F)$ necesitaríamos encontrar **una sola entrada** que perteneciera simultáneamente a E y a F. Pero lo único que hemos obtenido son dos posibles testigos:

$$u \in E, \qquad v \in F, \qquad f(u) = f(v).$$

Nada de lo demostrado hasta ahora permite afirmar que $u = v$.

La prueba fallida ha descubierto una hipótesis
El punto donde la demostración se detiene no es ruido algebraico. Es información matemática precisa.
Necesitamos convertir $f(u) = f(v)$ en $u = v$.
Pero esa es exactamente la definición de **inyectividad**. La prueba fallida nos ha dicho qué hipótesis faltaba.

Si f es inyectiva, entonces de $f(u) = f(v)$ obtenemos $u = v$. Llamemos x a ese elemento común. Como $x = u \in E$ y $x = v \in F$, tenemos $x \in E \cap F$.

Además, $f(x) = y$.

Por tanto, $y \in f(E \cap F)$.

Así, bajo inyectividad, la inclusión inversa sí queda demostrada. Antes de formular el resultado correcto, debemos comprobar que la hipótesis añadida es realmente necesaria y no solo un artefacto de nuestra estrategia de prueba.

Ejemplo 2.7. **La imagen no conserva siempre las intersecciones.** Sea

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = x^2,$$

y consideremos

$$E = \{-1\}, \qquad F = \{1\}.$$

Como $E \cap F = \varnothing$, tenemos $f(E \cap F) = f(\varnothing) = \varnothing$.
Sin embargo, $f(E) = \{1\}$ y $f(F) = \{1\}$.
Por tanto, $f(E) \cap f(F) = \{1\}$.
En consecuencia,

$$\boxed{f(E \cap F) = \varnothing \subsetneq \{1\} = f(E) \cap f(F).}$$

El mecanismo del fallo es exactamente el diagnosticado en la prueba: los dos testigos son distintos, $-1 \neq 1$, pero tienen la misma imagen, $f(-1) = f(1) = 1$.

Ya podemos registrar la versión correcta de las leyes básicas de imagen.

Proposición 2.4. **Leyes básicas de la imagen.** Sea $f\colon A \to B$ y sean $E, F \subseteq A$.

1. Si $E \subseteq F$, entonces
$$f(E) \subseteq f(F).$$
2. La imagen conserva uniones:
$$f(E \cup F) = f(E) \cup f(F).$$
3. Para intersecciones siempre se tiene
$$f(E \cap F) \subseteq f(E) \cap f(F).$$
4. Si además f es inyectiva, entonces
$$f(E \cap F) = f(E) \cap f(F).$$

Ya hemos demostrado los apartados 1, 3 y 4. Probemos el apartado 2 para completar la proposición.

Idea de la prueba de la unión. Pertenecer a $E \cup F$ significa pertenecer a E **o** a F. Esa disyunción pasa sin pérdida a las imágenes porque un mismo testigo x basta para justificar una de las dos alternativas.

Demostración. Sea $y \in f(E \cup F)$. Entonces existe $x \in E \cup F$ tal que $f(x) = y$. Como $x \in E$ o $x \in F$, tenemos $y \in f(E)$ o $y \in f(F)$. Por tanto, $y \in f(E) \cup f(F)$.

Esto demuestra $f(E \cup F) \subseteq f(E) \cup f(F)$.

Para la inclusión inversa, sea $y \in f(E) \cup f(F)$. Entonces $y \in f(E)$ o $y \in f(F)$. En el primer caso existe $x \in E$ con $f(x) = y$; como $E \subseteq E \cup F$, ese mismo x pertenece a $E \cup F$, y por tanto $y \in f(E \cup F)$. El segundo caso es análogo.

Luego $f(E) \cup f(F) \subseteq f(E \cup F)$. y, por doble inclusión,

$$\boxed{f(E \cup F) = f(E) \cup f(F).}$$

□

Qué cambió lógicamente entre unión e intersección. La diferencia entre las dos leyes aparece en los testigos que exige cada pertenencia. Para la unión, $y \in f(E \cup F)$ significa que existe un x tal que

$$x \in E \cup F \qquad \text{y} \qquad f(x) = y.$$

La condición $x \in E \cup F$ se descompone como

$$x \in E \quad \text{o} \quad x \in F.$$

No aparecen dos testigos diferentes.

En cambio, $y \in f(E) \cap f(F)$ significa simultáneamente

$$\exists u \in E \quad f(u) = y$$

y

$$\exists v \in F \quad f(v) = y.$$

Los cuantificadores existenciales pueden producir testigos distintos. Para fusionarlos necesitamos saber que una misma salida no puede proceder de dos entradas diferentes. Esa es exactamente la función estructural de la inyectividad.

Podemos condensarlo así:

la unión no obliga a identificar testigos;
la intersección de imágenes puede exigir identificar dos testigos.

Las preimágenes se comportan mejor. Pasemos ahora al sentido contrario. Sea $S \subseteq B$. Recordemos la equivalencia elemental $x \in f^{-1}(S) \iff f(x) \in S$. Aquí no necesitamos buscar un antecedente: **el punto x ya está fijado**. Solo comprobamos si su imagen satisface una condición de pertenencia. Esta diferencia hace que las operaciones de conjuntos se transporten con mucha más fidelidad mediante preimágenes.

Proposición 2.5. **Leyes de preimagen.** Sea $f\colon A \to B$ y sean $S, T \subseteq B$.

1. Si $S \subseteq T$, entonces
$$f^{-1}(S) \subseteq f^{-1}(T).$$
2. La preimagen conserva uniones:
$$f^{-1}(S \cup T) = f^{-1}(S) \cup f^{-1}(T).$$
3. La preimagen conserva intersecciones:
$$f^{-1}(S \cap T) = f^{-1}(S) \cap f^{-1}(T).$$
4. La preimagen conserva diferencias:
$$f^{-1}(S \setminus T) = f^{-1}(S) \setminus f^{-1}(T).$$
5. En particular, para el complemento relativo al codominio,
$$f^{-1}(B \setminus S) = A \setminus f^{-1}(S).$$
6. Si $g\colon B \to C$ y $R \subseteq C$, entonces
$$(g \circ f)^{-1}(R) = f^{-1}\big(g^{-1}(R)\big).$$

El apartado 1 ya fue demostrado al comienzo de la sección. Los demás pueden probarse de manera muy uniforme siguiendo la pertenencia de un punto arbitrario.

Demostración de la unión. Sea $x \in A$. Entonces

$$\begin{aligned} x \in f^{-1}(S \cup T) &\iff f(x) \in S \cup T \\ &\iff \big(f(x) \in S \text{ o } f(x) \in T\big) \\ &\iff \big(x \in f^{-1}(S) \text{ o } x \in f^{-1}(T)\big) \\ &\iff x \in f^{-1}(S) \cup f^{-1}(T). \end{aligned}$$

Como la equivalencia vale para todo $x \in A$, $f^{-1}(S \cup T) = f^{-1}(S) \cup f^{-1}(T)$.

Demostración de la intersección. Del mismo modo,

$$\begin{aligned} x \in f^{-1}(S \cap T) &\iff f(x) \in S \cap T \\ &\iff \big(f(x) \in S \text{ y } f(x) \in T\big) \\ &\iff \big(x \in f^{-1}(S) \text{ y } x \in f^{-1}(T)\big) \\ &\iff x \in f^{-1}(S) \cap f^{-1}(T). \end{aligned}$$

Por tanto, $f^{-1}(S \cap T) = f^{-1}(S) \cap f^{-1}(T)$.

Observa la diferencia con las imágenes: aquí el mismo x aparece durante toda la cadena. No surge ningún problema de «dos antecedentes posiblemente distintos».

Demostración de la diferencia. Para $x \in A$,

$$\begin{aligned} x \in f^{-1}(S \setminus T) &\iff f(x) \in S \setminus T \\ &\iff \big(f(x) \in S \text{ y } f(x) \notin T\big) \\ &\iff \big(x \in f^{-1}(S) \text{ y } x \notin f^{-1}(T)\big) \\ &\iff x \in f^{-1}(S) \setminus f^{-1}(T). \end{aligned}$$

Luego $f^{-1}(S \setminus T) = f^{-1}(S) \setminus f^{-1}(T)$.

Tomando $S = B$ y recordando que $f^{-1}(B) = A$, obtenemos inmediatamente

$$\boxed{f^{-1}(B \setminus T) = A \setminus f^{-1}(T).}$$

Demostración de la composición de preimágenes. Sean

$$f\colon A \to B, \qquad g\colon B \to C,$$

y sea $R \subseteq C$. Para $x \in A$,

$$\begin{aligned} x \in (g \circ f)^{-1}(R) &\iff (g \circ f)(x) \in R \\ &\iff g(f(x)) \in R \\ &\iff f(x) \in g^{-1}(R) \\ &\iff x \in f^{-1}\big(g^{-1}(R)\big). \end{aligned}$$

Por tanto,

$$\boxed{(g \circ f)^{-1}(R) = f^{-1}\big(g^{-1}(R)\big).}$$

□

Lectura estructural
La fórmula anterior tiene el mismo orden inverso que apareció en §2.5: $A \xrightarrow{f} B \xrightarrow{g} C$. Para decidir qué puntos de A terminan dentro de $R \subseteq C$, primero preguntamos qué puntos de B caen en R bajo g y luego qué puntos de A caen en ese conjunto bajo f:

$$R \overset{g^{-1}}{\longmapsto} g^{-1}(R) \overset{f^{-1}}{\longmapsto} f^{-1}(g^{-1}(R)).$$

Aquí los símbolos g^{-1} y f^{-1} designan **preimágenes de conjuntos**; no se supone que f ni g sean biyectivas.

Por qué la preimagen conserva mejor el álgebra de conjuntos. Podemos explicar todas estas identidades sin volver a demostrarlas una por una. La afirmación $x \in f^{-1}(S)$ es simplemente otra manera de escribir $f(x) \in S$. Por eso, cuando combinamos conjuntos mediante unión, intersección o complemento, las operaciones lógicas correspondientes pasan directamente a través de la preimagen:

$$\begin{aligned} \cup &\quad\longleftrightarrow\quad \text{«o»},\\ \cap &\quad\longleftrightarrow\quad \text{«y»},\\ \text{complemento} &\quad\longleftrightarrow\quad \text{«no»}. \end{aligned}$$

La preimagen no tiene que elegir ni comparar antecedentes. Evalúa siempre el mismo punto x y pregunta dónde cae $f(x)$. La imagen directa tiene una estructura lógica distinta. Para afirmar $y \in f(E)$, necesitamos un testigo:

$$\exists x \in E \quad f(x) = y.$$

En una intersección de imágenes pueden aparecer dos testigos distintos para una misma salida. Precisamente ahí puede perderse la igualdad. La diferencia conceptual puede resumirse así:

preimagen: sustituir $f(x)$ en una condición de pertenencia;
imagen: afirmar la existencia de algún antecedente que produzca la salida.

Un mapa de las leyes obtenidas. Para $f\colon A \to B$, $E, F \subseteq A$ y $S, T \subseteq B$:

$$\begin{aligned} f(E \cup F) &= f(E) \cup f(F),\\ f(E \cap F) &\subseteq f(E) \cap f(F),\\ f(E \cap F) &= f(E) \cap f(F) \quad \text{si } f \text{ es inyectiva},\\ f^{-1}(S \cup T) &= f^{-1}(S) \cup f^{-1}(T),\\ f^{-1}(S \cap T) &= f^{-1}(S) \cap f^{-1}(T),\\ f^{-1}(S \setminus T) &= f^{-1}(S) \setminus f^{-1}(T),\\ f^{-1}(B \setminus S) &= A \setminus f^{-1}(S). \end{aligned}$$

El cuadro reúne una misma diferencia lógica: **la imagen introduce existencia de antecedentes; la preimagen no**. En la primera debemos controlar los testigos; en la segunda comprobamos una condición sobre la imagen de una entrada ya fijada. Y, sobre todo, la sección nos deja una técnica reutilizable de investigación matemática:

conjetura
→ doble inclusión
→ prueba que se atasca
→ diagnóstico del paso faltante
→ hipótesis correcta
→ contraejemplo sin ella.

En §2.7 cambiaremos de representación: estudiaremos la **gráfica** de una función como subconjunto del producto cartesiano. Allí veremos qué información estructural puede leerse de un dibujo y, con igual importancia, qué información no está contenida en la curva por sí sola.

2.7. Gráficas: una representación, no la definición completa

Una gráfica reúne entradas y salidas en pares ordenados. Esa representación permite leer propiedades de la asignación, aunque necesitaremos distinguir el conjunto exacto de pares de un dibujo aproximado y de los datos que acompañan a la función. Cuando una función real se dibuja en el plano, cada punto de la curva registra una entrada y su salida correspondiente. Si $f\colon A \to B$, entonces al valor $x \in A$ le corresponde el par ordenado $(x, f(x))$. La primera coordenada conserva la entrada; la segunda conserva la salida. Reunir todos esos pares produce la gráfica.

Definición 2.7. **Gráfica de una función.** Sea

$$f\colon A \to B.$$

La **gráfica** de f es el conjunto

$$\Gamma_f = \{(x, f(x)) : x \in A\} \subseteq A \times B.$$

Equivalentemente,

$$(x, y) \in \Gamma_f \iff x \in A \text{ y } y = f(x).$$

La gráfica queda definida como un **conjunto de pares ordenados**; un dibujo en el plano será una representación de ese conjunto. Si $A, B \subseteq \mathbb{R}$, podemos representar esos pares como puntos del plano cartesiano. Entonces el conjunto Γ_f adquiere una apariencia geométrica: puede ser una curva, varios arcos separados, un conjunto discreto de puntos o una figura mucho menos regular. Pero el objeto matemático sigue siendo el conjunto de pares.

Por ejemplo, para

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = x^2,$$

la gráfica es $\Gamma_f = \{(x, x^2) : x \in \mathbb{R}\}$.

El punto $(2, 4)$ pertenece a la gráfica porque $f(2) = 4$; el punto $(2, 5)$ no pertenece porque $f(2) \neq 5$.

Qué codifica cada coordenada. La gráfica permite recuperar inmediatamente dos conjuntos que ya conocemos. Los valores que aparecen como **primeras coordenadas** son exactamente los elementos del dominio: $A = \{x : \exists y\ ((x, y) \in \Gamma_f)\}$. Los valores que aparecen como **segundas coordenadas** son exactamente los valores efectivamente alcanzados por la función, es decir, su imagen: $f(A) = \{y : \exists x\ ((x, y) \in \Gamma_f)\}$. Así, al mirar una gráfica exacta, podemos leer horizontalmente qué entradas aparecen y verticalmente qué salidas se alcanzan. Sin embargo, hay que advertir desde ahora una asimetría:

la gráfica determina la imagen, pero no determina por sí sola un codominio mayor que la imagen.

Volveremos sobre este punto con un ejemplo explícito. Es la misma distinción que ya encontramos entre imagen y codominio en §2.1.

Cuándo un conjunto de puntos puede ser la gráfica de una función. Supongamos ahora que todavía no sabemos si cierto conjunto $R \subseteq A \times B$ es la gráfica de alguna función $A \to B$.

¿Qué tendría que ocurrir?

La definición de función exige dos cosas para **cada** $x \in A$:

1. debe existir una salida $y \in B$ asociada con x;
2. esa salida debe ser única.

Traducido al conjunto de pares, esto significa que para cada primera coordenada permitida x debe haber exactamente un par de la forma (x, y) dentro de R.

Proposición 2.6. **Caracterización de las gráficas funcionales.** Sean A y B conjuntos y sea

$$R \subseteq A \times B.$$

Entonces R es la gráfica de una única función

$$f\colon A \to B$$

si y solo si para cada $x \in A$ existe un único $y \in B$ tal que

$$(x, y) \in R.$$

Idea de la prueba. No necesitamos descubrir una técnica nueva. La afirmación es exactamente la definición de función escrita en el lenguaje de pares ordenados.

Si R ya es la gráfica de una función, cada entrada x debe aparecer emparejada con $f(x)$ y con ninguna otra salida. En la dirección inversa, si cada x aparece emparejado con un único y, podemos definir $f(x)$ como ese único valor.

Demostración. Supongamos primero que $R = \Gamma_f$ para alguna función $f\colon A \to B$. Sea $x \in A$. Como f es una función, existe el valor $f(x) \in B$, y por definición de gráfica $(x, f(x)) \in R$.

Esto demuestra existencia.

Para la unicidad, supongamos que $y_1, y_2 \in B$ satisfacen

$$(x, y_1) \in R \qquad \text{y} \qquad (x, y_2) \in R.$$

Como $R = \Gamma_f$, ambos pares deben tener como segunda coordenada el valor de f en x. Por tanto, $y_1 = f(x) = y_2$, y así $y_1 = y_2$. Recíprocamente, supongamos que para cada $x \in A$ existe un único $y \in B$ tal que $(x, y) \in R$. Definimos $f\colon A \to B$ asignando a cada x ese único y. La hipótesis de existencia garantiza que la asignación está definida para todo $x \in A$, y la hipótesis de unicidad garantiza que ninguna entrada recibe dos valores incompatibles. Por tanto, f es una función. Además, por construcción, $(x, y) \in R \iff y = f(x)$, de modo que $R = \Gamma_f$.

Finalmente, la propia construcción muestra que no puede existir otra función $g\colon A \to B$ con la misma gráfica: para cada $x \in A$, tanto $f(x)$ como $g(x)$ tendrían que ser el único y asociado con x en R. Luego $f(x) = g(x)$ para todo $x \in A$. □

Después de la prueba
La proposición no añade una condición externa a la definición de función. La descompone geométricamente:

existencia: cada entrada aparece en algún par;
unicidad: cada entrada aparece con una sola segunda coordenada.

El llamado «criterio de la recta vertical» será simplemente esta misma afirmación cuando A y B sean subconjuntos de $\mathbb{R}$ y dibujemos R en el plano.

De existencia y unicidad al criterio de la recta vertical. Supongamos que $A, B \subseteq \mathbb{R}$ y que un conjunto de puntos $R \subseteq A \times B$ está dibujado en el plano. Fijar una entrada $x = a$ significa mirar la recta vertical $x = a$. Los puntos donde esa vertical corta a R son precisamente los pares $(a, y) \in R$.

Por tanto:

- si la vertical no corta a R, la entrada a no tiene salida;
- si la vertical corta a R en dos o más puntos, la entrada a tendría dos o más salidas;
- si la vertical corta a R exactamente una vez, existe una única salida asociada con a.

Así obtenemos el criterio geométrico:

R es la gráfica de una función A
$\to B$
$\iff$ cada vertical $x = a$ con $a \in A$ corta a R exactamente una vez.

La precisión «con $a \in A$» importa. Una función cuyo dominio sea, por ejemplo, $[0, 1]$ no tiene obligación de producir puntos sobre las verticales $x = 2$ o $x = -5$.

Error frecuente
La «prueba de la recta vertical» no es una regla geométrica independiente que mágicamente decida qué es una función. Es la traducción al plano de

$$\forall x \in A\ \exists! y \in B \quad \text{tal que} \quad (x, y) \in \Gamma_f.$$

Si olvidamos el dominio, o si olvidamos que se requieren **existencia y unicidad**, el criterio pierde su significado matemático.

Una curva que no es la gráfica de una función de x. Consideremos el círculo unitario $R = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$.

¿Puede ser la gráfica de una función $f\colon [-1, 1] \to \mathbb{R}$?

Tomemos $x = 0$. La ecuación del círculo se convierte en $y^2 = 1$, por lo que aparecen dos valores:

$$y = 1 \qquad \text{y} \qquad y = -1.$$

La vertical $x = 0$ corta el círculo en los puntos

$$(0, 1) \qquad \text{y} \qquad (0, -1).$$

Por tanto, el círculo completo no es la gráfica de una función de x en $[-1, 1]$. Para esa primera coordenada faltaría unicidad de la segunda: una misma entrada recibiría dos salidas.

En cambio, si conservamos solo la semicircunferencia superior, obtenemos

$$y = \sqrt{1 - x^2}, \qquad -1 \le x \le 1,$$

y cada vertical correspondiente a un $x \in [-1, 1]$ encuentra un único punto. La semicircunferencia inferior representa análogamente $y = -\sqrt{1 - x^2}$. Una misma relación geométrica puede, por tanto, necesitar ser separada en partes para que cada parte represente una función de la coordenada x.

Leer dominio e imagen desde la gráfica. Consideremos ahora

$$f\colon [-2, 2] \to \mathbb{R}, \qquad f(x) = x^2.$$

Su gráfica es $\Gamma_f = \{(x, x^2) : -2 \le x \le 2\}$.

Las primeras coordenadas recorren exactamente $[-2, 2]$, que es el dominio. Las segundas coordenadas recorren exactamente $[0, 4]$, que es la imagen. Geométricamente podemos expresar ambas lecturas como proyecciones:

$$\boxed{\begin{aligned} \mathrm{Dom}(f) &= \{x : \exists y \ ((x, y) \in \Gamma_f)\}, \\ f(A) &= \{y : \exists x \ ((x, y) \in \Gamma_f)\}. \end{aligned}}$$

En un dibujo ordinario solemos describir esto diciendo que el dominio es la «sombra» de la gráfica sobre el eje horizontal y la imagen su «sombra» sobre el eje vertical. La metáfora es útil siempre que recordemos que detrás de ella hay cuantificadores de existencia.

Las horizontales preguntan por antecedentes. Las verticales fijaban una **entrada** y preguntaban cuántas salidas tenía. Las rectas horizontales hacen la pregunta inversa: fijan una posible **salida** $y = b$ y preguntan cuántas entradas producen ese valor. Los puntos donde la horizontal $y = b$ corta a la gráfica son exactamente los pares $(x, b) \in \Gamma_f$, es decir, las soluciones de $f(x) = b$. Esta observación conecta inmediatamente la geometría con §2.3.

Inyectividad. La función f es inyectiva si una misma salida no puede provenir de dos entradas distintas. Por tanto,

$$\boxed{f \text{ es inyectiva } \iff \text{ cada recta horizontal corta } \Gamma_f \text{ en a lo sumo un punto.}}$$

«A lo sumo uno» permite cero intersecciones: un valor b que no pertenece a la imagen simplemente no es producido por ninguna entrada.

Para

$$f(x) = x^2 \qquad (x \in [-2, 2]),$$

la horizontal $y = 1$ corta la gráfica en

$$(-1, 1) \qquad \text{y} \qquad (1, 1).$$

Eso permite leer inmediatamente la no inyectividad:

$$f(-1) = f(1) = 1, \qquad -1 \neq 1.$$

Sobreyectividad. La sobreyectividad hace una pregunta diferente. Dado el codominio declarado B, exige que **cada** $b \in B$ sea alcanzado.

Geométricamente,

$$\boxed{f\colon A \to B \text{ es sobreyectiva } \iff \text{ cada horizontal } y = b \text{ con } b \in B \text{ corta } \Gamma_f \text{ al menos una vez.}}$$

Aquí no basta mirar la curva. Necesitamos saber cuál es el conjunto B respecto del cual estamos formulando la pregunta.

Finalmente, una función es biyectiva exactamente cuando, para cada $b \in B$, la horizontal $y = b$ corta la gráfica **exactamente una vez**. Podemos reunir las cuatro lecturas:

ser función sobre A	$\longleftrightarrow$	una intersección vertical por cada $x \in A$,
inyectiva	$\longleftrightarrow$	a lo sumo una intersección horizontal,
sobreyectiva sobre B	$\longleftrightarrow$	al menos una intersección horizontal por cada $y \in B$,
biyectiva	$\longleftrightarrow$	exactamente una intersección horizontal por cada $y \in B$.

Las palabras «por cada $x \in A$» y «por cada $y \in B$» son parte esencial del enunciado, no anotaciones secundarias.

La misma curva puede corresponder a funciones con distinta sobreyectividad. Llegamos ahora a una limitación fundamental de la gráfica considerada solo como conjunto de puntos del plano.

Ejemplo 2.8. **Misma gráfica, distinto codominio.** Consideremos

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = x^2,$$

y

$$h\colon \mathbb{R} \to [0, \infty), \qquad h(x) = x^2.$$

Como ambas funciones tienen el mismo dominio y asignan a cada x el mismo valor, sus conjuntos de pares ordenados son idénticos: $\Gamma_f = \{(x, x^2) : x \in \mathbb{R}\} = \Gamma_h$. Como subconjunto del plano, vemos exactamente la misma parábola. Sin embargo, las funciones no son iguales bajo la convención adoptada en este libro, porque sus codominios son distintos.

Además,

$$f \text{ no es sobreyectiva sobre } \mathbb{R},$$

pues ningún número negativo aparece como segunda coordenada, mientras que

$$h \text{ sí es sobreyectiva sobre } [0, \infty),$$

porque todo $y \geq 0$ posee una raíz cuadrada real y, por tanto, aparece en la gráfica. La curva dibujada no cambió. Cambió la pregunta matemática porque cambió el codominio.

Las dos funciones comparten el conjunto de pares de la gráfica, pero tienen codominios diferentes. El conjunto de puntos permite recuperar la regla punto a punto y el dominio mediante las primeras coordenadas. También permite recuperar la imagen mediante las segundas coordenadas. Pero si el codominio contiene elementos que la función no alcanza, esos elementos **no aparecen como puntos adicionales en la gráfica**. Por eso, a partir de la parábola sola no podemos decidir si estamos mirando $\mathbb{R} \to \mathbb{R}$ o $\mathbb{R} \to [0, \infty)$. Y precisamente por eso la sobreyectividad nunca debe diagnosticarse solo por la forma de la curva sin conocer el codominio declarado.

Una precisión sobre el ambiente
Si escribimos explícitamente $\Gamma_f \subseteq A \times B$, y declaramos A y B, entonces el **ambiente** aporta información adicional. Pero el mero conjunto de puntos $\{(x, f(x)) : x \in A\}$ no registra qué elementos extra hemos decidido incluir en un codominio mayor que la imagen. La distinción es entre **la gráfica como conjunto de pares** y **la gráfica junto con todos los datos de la función que la acompañan**.

Gráfica, fórmula, tabla y algoritmo. Ya vimos en §2.1 que una función no necesita venir dada por una fórmula cerrada. La gráfica es otra forma de representar la misma clase de objetos, no una nueva definición limitada a curvas suaves o expresiones familiares. Una función puede presentarse mediante:

- una fórmula;
- una tabla de valores;
- una descripción verbal;
- un algoritmo;
- una gráfica;
- o una combinación de esas representaciones.

Cada representación hace visibles unas propiedades y oculta otras. Una fórmula puede facilitar cálculos exactos, pero puede ocultar el comportamiento global. Una tabla muestra valores concretos, pero normalmente solo una parte del dominio. Un algoritmo puede definir perfectamente una función aunque no produzca una expresión cerrada sencilla. Una gráfica vuelve visibles relaciones entre entradas y salidas, pero un dibujo aproximado no sustituye una demostración y, como acabamos de ver, tampoco revela por sí solo el codominio completo. Por eso conviene adoptar la siguiente regla de lectura:

$$\boxed{\text{representación} \neq \text{objeto representado.}}$$

La función es el objeto matemático completo; fórmula, tabla, algoritmo y gráfica son maneras de acceder a él.

Qué podemos y qué no podemos concluir de una gráfica. Si la gráfica es exacta y conocemos el dominio y el codominio declarados, podemos usarla para formular y, en muchos ejemplos, decidir preguntas estructurales:

- qué entradas pertenecen al dominio;
- qué valores pertenecen a la imagen;
- si alguna entrada tendría más de una salida;
- si dos entradas comparten una salida;
- si cada elemento del codominio es alcanzado.

Pero debemos evitar varios saltos ilegítimos. Un dibujo de resolución finita no demuestra por sí solo que dos puntos coincidan exactamente, que no exista una intersección demasiado pequeña para verse o que una propiedad se mantenga en todos los puntos. Cuando la afirmación exige exactitud, la gráfica orienta la conjetura y la estructura de la prueba; no reemplaza el argumento. Tampoco debemos atribuir al dibujo información que pertenece a la declaración de la función. En particular,

$$\boxed{\text{la sobreyectividad depende de la gráfica y del codominio.}}$$

Las diferencias entre objeto, descripción y representación quedan reunidas en

$$\boxed{\text{fórmula} \neq \text{función completa} \neq \text{gráfica como representación.}}$$

En §2.8 utilizaremos ya esta lectura estructural para trabajar con **funciones definidas por tramos** y con transformaciones elementales de gráficas. La regla seguirá siendo la misma: primero entender qué operación se ha realizado sobre la función; después interpretar geométricamente qué le ocurre a su representación.

2.8. Funciones por tramos y transformaciones de gráficas

Una función puede utilizar reglas diferentes en distintas partes del dominio. También podemos construir una nueva función modificando las entradas o las salidas de otra. Al seguir los pares ordenados de sus gráficas podremos justificar qué efecto geométrico produce cada modificación. En ambos casos mantendremos la misma disciplina que ha guiado todo el capítulo:

$$\boxed{\text{primero identificamos la función; después interpretamos su gráfica.}}$$

Las reglas de desplazamiento se obtendrán siguiendo un punto de la gráfica original y calculando sus nuevas coordenadas. Ese cálculo permitirá reconstruir tanto el movimiento como sus condiciones de dominio, incluso al combinar varias transformaciones.

Una función puede usar reglas diferentes en partes diferentes del dominio. No existe ninguna exigencia matemática de que una función deba estar descrita por una única fórmula en todo su dominio. Supongamos que queremos definir una función $f\colon A \to \mathbb{R}$ mediante subconjuntos $A_1, \dots, A_n \subseteq A$ y reglas $r_1, \dots, r_n$. Esquemáticamente escribiríamos

$$f(x) = \begin{cases} r_1(x), & x \in A_1, \\ r_2(x), & x \in A_2, \\ \vdots & \vdots \\ r_n(x), & x \in A_n. \end{cases}$$

Para que esta escritura determine realmente una función sobre **todo** A, debemos volver a las dos exigencias de §2.1: existencia y unicidad.

Existencia: ningún punto del dominio puede quedar sin regla. Necesitamos $A = A_1 \cup \cdots \cup A_n$.

Si algún $x \in A$ no pertenece a ninguno de los conjuntos A_i, la receta no dice qué valor debe tener $f(x)$. En ese punto fallaría la existencia de una salida.

Unicidad: un punto cubierto por varias reglas no puede recibir valores incompatibles. No es necesario que los conjuntos A_i sean disjuntos. Pueden solaparse, pero entonces las reglas deben ser compatibles en cada solapamiento.

Si $x \in A_i \cap A_j$, debemos tener $r_i(x) = r_j(x)$.

De lo contrario, la misma entrada recibiría dos valores distintos y la receta no definiría una función. Así, la definición por tramos no introduce una noción nueva de función. Solo reutiliza la definición ya conocida:

los tramos deben cubrir el dominio y deben concordar allí donde se superponen.

El valor absoluto como primer modelo. La función valor absoluto puede escribirse como

$$|x| = \begin{cases} -x, & x < 0, \\ x, & x \geq 0. \end{cases}$$

Los dos tramos cubren todo $\mathbb{R}$ y no se superponen, de modo que no existe ambigüedad.

También podríamos escribir

$$|x| = \begin{cases} -x, & x \leq 0, \\ x, & x \geq 0. \end{cases}$$

Ahora los tramos se superponen en $x = 0$, pero las dos fórmulas producen allí el mismo valor: $-0 = 0$. Por tanto, la segunda receta también define exactamente la misma función.

En cambio,

$$h(x) = \begin{cases} -x + 1, & x \leq 0, \\ x, & x \geq 0 \end{cases}$$

no define una función sobre $\mathbb{R}$, porque en el punto de solapamiento obtenemos dos instrucciones incompatibles:

$$-0 + 1 = 1, \qquad 0 = 0.$$

La entrada 0 tendría dos salidas.

Lectura estructural
Cuando veas una definición por tramos, no empieces dibujando. Comprueba primero:

1. cuál es el dominio declarado;
2. si los tramos lo cubren por completo;
3. si hay puntos que pertenecen a más de un tramo;
4. si las fórmulas coinciden en esos puntos.

Solo después la receta está auditada como función.

Fórmulas familiares y dominios naturales. Las funciones por tramos amplían el repertorio, pero también conviene ordenar las familias de expresiones que ya conocemos. Una función **polinómica**, por ejemplo $p(x) = x^3 - 2x + 1$, puede evaluarse para todo $x \in \mathbb{R}$. Si no se declara otra cosa, su dominio natural es $\mathbb{R}$.

Una expresión **racional**

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

solo puede evaluarse donde $q(x) \neq 0$.

Por ejemplo, $r(x) = \frac{1}{x^2-1}$ tiene como dominio natural $\mathbb{R} \setminus \{-1, 1\}$. Una expresión **algebraica** que contiene una raíz cuadrada real exige que el radicando sea no negativo. Por ejemplo, $s(x) = \sqrt{4 - x^2}$ requiere $4 - x^2 \geq 0$, es decir, $-2 \leq x \leq 2$. Por tanto, su dominio natural es $[-2, 2]$. Entre las funciones periódicas construidas aquí, $\sigma(x)$ y $\kappa(x)$ están definidas para todo real, mientras que una expresión como

$$q(x) = \frac{\sigma(x)}{\kappa(x)}$$

solo está definida donde $\kappa(x) \neq 0$.

Estas observaciones conducen a una convención útil, pero debemos formularla con cuidado:

> cuando se proporciona **solo una expresión** y se pide su dominio natural, buscamos el mayor subconjunto de $\mathbb{R}$ en el que todas las operaciones indicadas están definidas.

Esto no reemplaza el dominio como parte de los datos de una función. Podemos restringir deliberadamente una expresión a un dominio menor. Por ejemplo,

$$p\colon [0, 1] \to \mathbb{R}, \qquad p(x) = x^3 - 2x + 1,$$

es una función perfectamente legítima aunque la fórmula polinómica tenga sentido fuera de $[0, 1]$.

La distinción permanece:

dominio natural de una expresión $\neq$ dominio declarado de toda función que use esa expresión.

Transformar una gráfica significa transformar sus puntos. Fijemos ahora una función real $f\colon D \to \mathbb{R}$ y un punto cualquiera de su gráfica: $(u, f(u)) \in \Gamma_f$. Para entender una transformación, seguiremos ese punto. Preguntaremos qué punto de la nueva gráfica contiene **el mismo valor**, una versión modificada de ese valor o una entrada modificada. Seguir las coordenadas permitirá comprobar el dominio y el sentido de la transformación a partir de su definición.

Desplazamiento vertical: $f(x)+c$. Definamos $g(x)=f(x)+c$.

El dominio de g es el mismo D, porque no hemos cambiado la entrada que introducimos en f.

Si $y=f(u)$, entonces $g(u)=f(u)+c=y+c$.

Por tanto, el punto (u,y) de la gráfica original se convierte en $(u,y+c)$. La primera coordenada permanece fija y la segunda aumenta en c. Si $c>0$, todos los puntos se desplazan c unidades hacia arriba; si $c<0$, se desplazan $|c|$ unidades hacia abajo.

En símbolos,

$$\boxed{(u,y)\longmapsto(u,y+c).}$$

Desplazamiento horizontal: $f(x-c)$. Definamos ahora $h(x)=f(x-c)$.

Aquí sí hemos cambiado la entrada de f. Antes de hablar de la gráfica debemos auditar el dominio: $x\in\mathrm{Dom}(h)\iff x-c\in D$. Tomemos nuevamente un punto original $(u,f(u))$.

¿En qué entrada de h reaparece el valor $f(u)$? Debemos resolver $x-c=u$.

Por tanto, $x=u+c$. y $h(u+c)=f((u+c)-c)=f(u)$.

Así,

$$\boxed{(u,y)\longmapsto(u+c,y).}$$

Si $c>0$, la gráfica se desplaza c unidades **hacia la derecha**. El signo puede parecer contrario a la intuición superficial porque dentro de la fórmula aparece $x-c$. Pero el desplazamiento no se adivina mirando el signo: se obtiene resolviendo $x-c=u$.

Error frecuente: el signo del desplazamiento horizontal
No confundas $f(x)+c$ con $f(x-c)$.
La primera operación modifica la **salida**: $(u,y)\mapsto(u,y+c)$. La segunda modifica la **entrada necesaria para obtener el mismo valor**: $(u,y)\mapsto(u+c,y)$. En particular, para $c>0$, $f(x-c)$ se desplaza hacia la **derecha**, no hacia la izquierda. Si dudas, no memorices: fija un punto u de la función original y resuelve $x-c=u$.

Un ejemplo especialmente transparente es

$$f(x)=\sqrt{x},\qquad x\geq 0.$$

Entonces $h(x)=\sqrt{x-3}$ solo está definida cuando $x-3\geq 0$, es decir, cuando $x\geq 3$. El dominio $[0,\infty)$ se ha desplazado a $[3,\infty)$, exactamente como predice la transformación horizontal.

Reflexión respecto del eje horizontal: $-f(x)$. Sea $g(x)=-f(x)$.

Para cada $u\in D$, $g(u)=-f(u)$.

Por tanto,

$$\boxed{(u,y)\longmapsto(u,-y).}$$

La primera coordenada permanece fija y cambia el signo de la segunda. Geométricamente, la gráfica se refleja respecto del eje horizontal.

Reflexión respecto del eje vertical: $f(-x)$. Sea ahora $h(x) = f(-x)$.
Su dominio es $\{x \in \mathbb{R} : -x \in D\}$.
Si $u \in D$, buscamos la entrada x que satisface $-x = u$.
Entonces $x = -u$ y $h(-u) = f(u)$.
Así,

$$\boxed{(u, y) \longmapsto (-u, y).}$$

Cambia el signo de la primera coordenada y se conserva la segunda: obtenemos una reflexión respecto del eje vertical. Las dos reflexiones muestran otra vez la diferencia entre modificar una salida y modificar una entrada:

$$\boxed{\begin{array}{rcl} -f(x) & : & (u, y) \mapsto (u, -y), \\ f(-x) & : & (u, y) \mapsto (-u, y). \end{array}}$$

Escalas verticales y horizontales: $af(bx)$. Consideremos $g(x) = af(bx)$.
Supongamos primero que

$$a \neq 0, \qquad b \neq 0.$$

Si (u, y) pertenece a la gráfica de f, entonces queremos elegir x de manera que $bx = u$.
Por tanto, $x = \frac{u}{b}$.
En esa entrada, $g(\frac{u}{b}) = af(u) = ay$.
Así, la transformación completa de los puntos es

$$\boxed{(u, y) \longmapsto \left(\frac{u}{b}, ay\right).}$$

El factor a actúa directamente sobre las alturas:

- si $|a| > 1$, las distancias verticales al eje horizontal se multiplican por $|a|$;
- si $0 < |a| < 1$, se reducen por ese factor;
- si $a < 0$, además aparece la reflexión respecto del eje horizontal.

El factor b actúa de modo recíproco sobre las coordenadas horizontales:

- si $|b| > 1$, las coordenadas horizontales se dividen por un número mayor que 1 y la gráfica se comprime horizontalmente;
- si $0 < |b| < 1$, se dilata horizontalmente;
- si $b < 0$, además aparece la reflexión respecto del eje vertical.

La reciprocidad $u \mapsto u/b$ no es una convención gráfica: procede directamente de resolver $bx = u$. También el dominio debe transformarse. Si D es el dominio de f, entonces $\mathrm{Dom}(g) = \{x \in \mathbb{R} : bx \in D\}$. Los casos $a = 0$ o $b = 0$ son degenerados y deben analizarse por la definición, no por prolongación automática de la regla geométrica. Si $a = 0$, todos los valores de una función ya definida se convierten en 0. Si $b = 0$, la expresión exige evaluar $f(0)$; solo tiene sentido si $0 \in D$, y entonces la dependencia respecto de x desaparece en la parte interior.

La composición unifica las transformaciones. La composición de §2.4 reúne estas transformaciones según actúen antes o después de la función.

Definamos, para $c, a \in \mathbb{R}$,

$$\tau_c(x) = x + c, \qquad \sigma_a(x) = ax.$$

Siempre que los dominios permitan las composiciones correspondientes, tenemos $f(x) + c = (\tau_c \circ f)(x)$,

$$f(x - c) = (f \circ \tau_{-c})(x),$$

$$-f(x) = (\sigma_{-1} \circ f)(x),$$

$$f(-x) = (f \circ \sigma_{-1})(x),$$

y $af(bx) = (\sigma_a \circ f \circ \sigma_b)(x)$.

Esto explica la diferencia fundamental entre transformaciones «de afuera» y «de adentro»:

- componer **después** de f modifica las salidas;
- componer **antes** de f modifica las entradas que llegan a f.

La geometría refleja ese orden algebraico.

Un ejemplo completo: construir por tramos y transformar.

Ejemplo 2.9. **Una función por tramos bajo varias transformaciones.** Definamos $p\colon [-1, 2] \to \mathbb{R}$ mediante

$$p(x) = \begin{cases} x + 1, & -1 \le x \le 0, \\ 1 - \dfrac{x}{2}, & 0 \le x \le 2. \end{cases}$$

Los dos tramos cubren $[-1, 2]$. Se superponen únicamente en $x = 0$, y allí coinciden: $0 + 1 = 1 = 1 - \frac{0}{2}$. Por tanto, la receta define una función. La primera rama recorre valores desde 0 hasta 1, y la segunda vuelve desde 1 hasta 0. Así, $p([-1, 2]) = [0, 1]$.

Ahora sigamos varias transformaciones.

1. Desplazamiento vertical. Sea $q(x) = p(x) + 2$.

El dominio permanece $[-1, 2]$ y

$$q(x) = \begin{cases} x + 3, & -1 \le x \le 0, \\ 3 - \dfrac{x}{2}, & 0 \le x \le 2. \end{cases}$$

Cada punto (u, y) pasa a $(u, y + 2)$, y la imagen pasa de $[0, 1]$ a $[2, 3]$.

2. Desplazamiento horizontal. Sea $r(x) = p(x - 3)$.

La condición $x - 3 \in [-1, 2]$ equivale a $2 \le x \le 5$.

Por tanto, el dominio nuevo es $[2, 5]$. Sustituyendo cuidadosamente en cada tramo,

$$r(x) = \begin{cases} x - 2, & 2 \le x \le 3, \\ \dfrac{5 - x}{2}, & 3 \le x \le 5. \end{cases}$$

La gráfica completa se ha desplazado tres unidades hacia la derecha: $(u, y) \longmapsto (u + 3, y)$.

3. Dos reflexiones. Sea $s(x) = -p(-x)$.

La sustitución interior $-x$ refleja primero las entradas; el signo exterior refleja después las salidas. El dominio es $[-2, 1]$ y, desarrollando los tramos,

$$s(x) = \begin{cases} -1 - \dfrac{x}{2}, & -2 \le x \le 0, \\ x - 1, & 0 \le x \le 1. \end{cases}$$

Cada punto de la gráfica original se transforma según $(u, y) \longmapsto (-u, -y)$.

4. Cambio simultáneo de escalas. Finalmente, sea $t(x) = 2p(2x)$.

Como $2x \in [-1, 2]$, el dominio es

$$\left[-\frac{1}{2}, 1\right].$$

Además,

$$t(x) = \begin{cases} 4x + 2, & -\dfrac{1}{2} \le x \le 0, \\ 2 - 2x, & 0 \le x \le 1. \end{cases}$$

Los puntos obedecen a $(u, y) \longmapsto \left(\frac{u}{2}, 2y\right)$.

La gráfica se comprime horizontalmente por un factor 2 y las alturas se duplican. La imagen es ahora $[0, 2]$. En los cuatro casos no hemos necesitado adivinar la nueva gráfica: cada cambio se dedujo de la nueva regla y de su dominio.

Un protocolo seguro para transformar funciones. Cuando aparezca una expresión construida a partir de una función conocida, conviene seguir siempre este orden:

1. **Escribe la nueva función.** No trabajes todavía con el dibujo.
2. **Audita su dominio.** Las modificaciones interiores pueden cambiarlo.
3. **Fija un punto original** $(u, f(u))$**.**
4. **Resuelve qué nueva entrada produce el valor relacionado con** $f(u)$**.**
5. **Calcula la nueva salida.**
6. **Obtén la transformación de coordenadas.**
7. **Solo entonces interpreta desplazamiento, reflexión o cambio de escala.**

Este procedimiento es algo más lento que memorizar una tabla la primera vez, pero es mucho más robusto. Además, funciona también cuando varias transformaciones aparecen combinadas. Podemos resumir las transformaciones básicas, para los casos en que las expresiones estén definidas, del siguiente modo:

$$\begin{array}{rcll} f(x) + c & : & (u, y) \mapsto (u, y + c), & \\ f(x - c) & : & (u, y) \mapsto (u + c, y), & \\ -f(x) & : & (u, y) \mapsto (u, -y), & \\ f(-x) & : & (u, y) \mapsto (-u, y), & \\ af(bx) & : & (u, y) \mapsto \left(\dfrac{u}{b}, ay\right), & ab \neq 0. \end{array}$$

La tabla es ahora una **consecuencia** de los cálculos anteriores, no un sustituto de ellos. En §2.9 añadiremos una propiedad cualitativa nueva que depende del orden del dominio: estudiaremos cómo comparar las salidas correspondientes a entradas ordenadas y por qué ciertas funciones pueden invertirse después de restringir adecuadamente su dominio. Hasta entonces, todo lo que hemos hecho en esta sección depende solo de la definición de función, de su dominio, de la composición y de la representación mediante pares ordenados.

2.9. Monotonía e inversas sobre intervalos

La recta real permite ordenar tanto las entradas como las salidas de una función. Podemos estudiar si la asignación conserva ese orden o lo invierte, y utilizar esa información para decidir cuándo dos entradas distintas tienen que producir valores distintos. Si el dominio es un subconjunto ordenado de $\mathbb{R}$, podemos comparar dos entradas $x_1 < x_2$ y preguntar qué relación existe entre $f(x_1)$ y $f(x_2)$. Una función puede conservar el orden, invertirlo, conservarlo solo débilmente o no obedecer ningún patrón global de este tipo. Esta pregunta conduce a la noción de **monotonía**.

Comparar entradas para comparar salidas. Consideremos primero dos ejemplos sencillos. Para $f(x) = 2x + 1$, si $x_1 < x_2$, entonces $2x_1 + 1 < 2x_2 + 1$, de modo que el orden de las entradas se conserva estrictamente.

En cambio, para $g(x) = -3x + 4$, si $x_1 < x_2$, al multiplicar por -3 se invierte la desigualdad: $-3x_1 > -3x_2$, y por tanto $g(x_1) > g(x_2)$. Aquí el orden se invierte estrictamente. Hay también situaciones en las que el orden solo se conserva de manera débil. Una función puede tener dos entradas distintas con la misma salida y, aun así, no disminuir nunca. Por eso necesitamos distinguir cuidadosamente las versiones débiles de las estrictas.

Definición 2.8. **Monotonía.** Sea $D \subseteq \mathbb{R}$ y sea

$$f \colon D \to \mathbb{R}.$$

Diremos que f es:

1. **no decreciente** en D si, para cualesquiera $x_1, x_2 \in D$,
$$x_1 < x_2 \Longrightarrow f(x_1) \leq f(x_2);$$
2. **estrictamente creciente** en D si
$$x_1 < x_2 \Longrightarrow f(x_1) < f(x_2);$$
3. **no creciente** en D si
$$x_1 < x_2 \Longrightarrow f(x_1) \geq f(x_2);$$
4. **estrictamente decreciente** en D si
$$x_1 < x_2 \Longrightarrow f(x_1) > f(x_2).$$

Una función se llama **monótona** cuando es no decreciente o no creciente, y **estrictamente monótona** cuando es estrictamente creciente o estrictamente decreciente.

La terminología varía ligeramente entre libros: algunos autores usan «creciente» para la versión débil y otros para la estricta. Para evitar ambigüedades, en este libro escribiremos siempre **no decreciente**, **no creciente**, **estrictamente creciente** o **estrictamente decreciente** cuando la distinción sea relevante.

El dominio debe estar ordenado. La definición anterior no depende de que D sea un intervalo. Solo necesita que tenga sentido comparar dos elementos mediante $<$. Por ejemplo, podríamos estudiar monotonía en $D = \{-3, -1, 2, 7\}$. Sin embargo, en cálculo trabajaremos con especial frecuencia sobre **intervalos**, porque allí no hemos eliminado artificialmente puntos intermedios del dominio. Por esa razón el título de esta sección habla de monotonía sobre intervalos. Conviene separar desde ahora dos afirmaciones:

la definición de monotonía necesita orden;
muchos teoremas posteriores necesitarán además que el dominio sea un intervalo.

El primer teorema de esta sección pertenece al primer grupo: ni siquiera necesita que el dominio sea un intervalo.

Por qué la monotonía estricta impide repeticiones. En §2.3 definimos la inyectividad mediante $f(x_1) = f(x_2) \implies x_1 = x_2$. La monotonía estricta da una manera de obtener esa conclusión: al ordenar dos entradas distintas, sus salidas quedan separadas por una desigualdad estricta y, por tanto, no pueden coincidir.

Teorema 2.2. **Toda función estrictamente monótona es inyectiva.** Sea $D \subseteq \mathbb{R}$ y sea

$$f\colon D \to \mathbb{R}.$$

Si f es estrictamente creciente o estrictamente decreciente en D, entonces f es inyectiva.

Idea de la prueba. Queremos demostrar que entradas distintas no pueden compartir una salida. Como $\mathbb{R}$ está ordenado, si $x_1 \neq x_2$, entonces necesariamente ocurre una de las dos posibilidades

$$x_1 < x_2 \qquad \text{o} \qquad x_2 < x_1.$$

La monotonía estricta convierte cualquiera de esas desigualdades entre entradas en una desigualdad estricta entre salidas. Por tanto, las salidas no pueden ser iguales.

Demostración. Supongamos primero que f es estrictamente creciente. Sean $x_1, x_2 \in D$ con $x_1 \neq x_2$.

Por el orden total de $\mathbb{R}$, o bien $x_1 < x_2$ o bien $x_2 < x_1$. Si $x_1 < x_2$, la monotonía estricta da $f(x_1) < f(x_2)$, y por tanto $f(x_1) \neq f(x_2)$. Si $x_2 < x_1$, obtenemos análogamente $f(x_2) < f(x_1)$, de nuevo con $f(x_1) \neq f(x_2)$. Así, entradas distintas producen salidas distintas y f es inyectiva. Si f es estrictamente decreciente, el argumento es idéntico salvo por el sentido de las desigualdades: de $x_1 < x_2$ se sigue $f(x_1) > f(x_2)$, y de $x_2 < x_1$ se sigue $f(x_2) > f(x_1)$. En ambos casos las salidas son diferentes. Por tanto, f también es inyectiva. □

Después de la prueba
La demostración no utilizó completitud, intervalos, continuidad, límites ni derivadas. Solo necesitó:

- que el dominio esté contenido en un conjunto totalmente ordenado;
- la definición de monotonía estricta.

La razón estructural es muy simple:

$$\boxed{\text{entradas distintas} \Longrightarrow \text{entradas comparables} \Longrightarrow \text{salidas estrictamente distintas.}}$$

Por qué «estrictamente» no puede borrarse. Una función no decreciente puede conservar el orden débilmente y, sin embargo, repetir valores. El ejemplo más simple es la función constante

$$c\colon \mathbb{R} \to \mathbb{R}, \qquad c(x) = 5.$$

Si $x_1 < x_2$, entonces $c(x_1) = 5 \leq 5 = c(x_2)$, de modo que c es no decreciente. También es no creciente. Pero no es inyectiva, porque $c(0) = c(1) = 5$.

Por tanto,

$$\boxed{\text{no decreciente} \not\Longrightarrow \text{inyectiva.}}$$

Lo mismo ocurre con «no creciente». La desigualdad estricta del teorema no es una comodidad estilística: es la condición que impide los tramos horizontales y, algebraicamente, impide que dos entradas ordenadas tengan la misma salida.

Una vía natural hacia la inversa. Supongamos ahora que $I \subseteq \mathbb{R}$ es un intervalo y $f\colon I \to \mathbb{R}$ es estrictamente monótona. El teorema anterior nos da inmediatamente inyectividad. Pero para aplicar el criterio de §2.5 necesitamos también sobreyectividad. Aquí aparece una maniobra que ya usamos con x^2: en lugar de mantener un codominio posiblemente demasiado grande, consideramos la misma regla con codominio igual a su imagen: $f\colon I \to f(I)$. Esta función es sobreyectiva **por definición de** $f(I)$. Y es inyectiva por el teorema de monotonía estricta. Por tanto, es biyectiva y posee función inversa. La monotonía nos permite decir todavía más: la inversa conserva el mismo tipo de orden.

Corolario 2.2. **Inversa sobre la imagen y monotonía de la inversa.** Sea $I \subseteq \mathbb{R}$ un intervalo y sea

$$f\colon I \to \mathbb{R}$$

estrictamente monótona. Entonces la función

$$f\colon I \to f(I)$$

es biyectiva y, por tanto, admite una única inversa

$$f^{-1}\colon f(I) \to I.$$

Además:

1. si f es estrictamente creciente, entonces f^{-1} es estrictamente creciente;
2. si f es estrictamente decreciente, entonces f^{-1} es estrictamente decreciente.

Por qué existe la inversa. La sobreyectividad de $f\colon I \to f(I)$ es automática, porque cada elemento de $f(I)$ es, precisamente, un valor $f(x)$ con $x \in I$. La inyectividad proviene de Teorema 2.2. Por Teorema 2.1, la función es biyectiva y admite una única inversa.

Queda demostrar el comportamiento del orden.

Caso 1: f estrictamente creciente. Sean

$$y_1, y_2 \in f(I) \qquad \text{con} \qquad y_1 < y_2.$$

Definamos

$$x_1 = f^{-1}(y_1), \qquad x_2 = f^{-1}(y_2).$$

Entonces

$$f(x_1) = y_1, \qquad f(x_2) = y_2.$$

Queremos demostrar $x_1 < x_2$.

Como $y_1 < y_2$, tenemos $y_1 \neq y_2$ y por tanto $x_1 \neq x_2$. Si ocurriera $x_2 < x_1$, la monotonía estrictamente creciente de f produciría $f(x_2) < f(x_1)$, es decir, $y_2 < y_1$, contradicción. La única posibilidad restante es $x_1 < x_2$.

Así, $y_1 < y_2 \Longrightarrow f^{-1}(y_1) < f^{-1}(y_2)$, y f^{-1} es estrictamente creciente.

Caso 2: f estrictamente decreciente. Tomemos de nuevo $y_1 < y_2$ y pongamos

$$x_1 = f^{-1}(y_1), \qquad x_2 = f^{-1}(y_2).$$

Si ocurriera $x_1 < x_2$, como f es estrictamente decreciente tendríamos $f(x_1) > f(x_2)$, es decir, $y_1 > y_2$, contradicción.

Como $x_1 \neq x_2$, necesariamente $x_1 > x_2$.

Por tanto, $y_1 < y_2 \Longrightarrow f^{-1}(y_1) > f^{-1}(y_2)$, y la inversa es estrictamente decreciente. □

Dos inversiones distintas

En el caso decreciente aparecen dos ideas de «inversión» que no deben confundirse.

- f **invierte el orden**: $x_1 < x_2 \Longrightarrow f(x_1) > f(x_2)$.
- f^{-1} **invierte la asignación** de entradas y salidas.

La inversa funcional de una función estrictamente decreciente sigue siendo estrictamente decreciente: al recorrer la relación en sentido contrario, el orden vuelve a invertirse.

Lo que todavía no podemos concluir sobre $f(I)$. El dominio I es un intervalo, pero de lo demostrado hasta aquí **no se sigue** que la imagen $f(I)$ sea también un intervalo.

Consideremos $f\colon \mathbb{R} \to \mathbb{R}$,

$$f(x) = \begin{cases} x, & x < 0, \\ x + 1, & x \geq 0. \end{cases}$$

Esta función es estrictamente creciente. En efecto:

- si $x_1 < x_2 < 0$, entonces $f(x_1) = x_1 < x_2 = f(x_2)$;
- si $0 \leq x_1 < x_2$, entonces $f(x_1) = x_1 + 1 < x_2 + 1 = f(x_2)$;
- si $x_1 < 0 \leq x_2$, entonces

$$f(x_1) = x_1 < 0 < 1 \leq x_2 + 1 = f(x_2).$$

Sin embargo, $f(\mathbb{R}) = (-\infty, 0) \cup [1, \infty)$, que no es un intervalo. Esto marca con precisión la frontera actual del capítulo. La monotonía estricta basta para obtener **inyectividad** y, al ajustar el codominio a la imagen, una **inversa**. Para garantizar que ciertos valores intermedios no se omitan necesitaremos una propiedad adicional que todavía no hemos introducido.

El modelo fundamental: cuadrado y raíz. Volvamos ahora a una función que nos acompaña desde el comienzo del capítulo.

EJEMPLO 2.10. **El cuadrado sobre $[0,\infty)$ y su inversa.** Consideremos

$$q\colon [0,\infty) \to [0,\infty), \qquad q(x) = x^2.$$

En §2.3 demostramos que q es biyectiva y en §2.5 identificamos su inversa como $q^{-1}(y) = \sqrt{y}$. Ahora podemos comprender esa invertibilidad desde el orden.

Sean $0 \le x_1 < x_2$.

Entonces $x_2^2 - x_1^2 = (x_2 - x_1)(x_2 + x_1)$.

El primer factor es positivo: $x_2 - x_1 > 0$. Además, como $x_2 > x_1 \ge 0$, tenemos $x_2 > 0$, y por tanto $x_2 + x_1 > 0$.

Así, $x_2^2 - x_1^2 > 0$, de donde $x_1^2 < x_2^2$.

Por consiguiente, q es estrictamente creciente en $[0,\infty)$. El corolario anterior implica entonces que su inversa $\sqrt{\ }\colon [0,\infty) \to [0,\infty)$ es también estrictamente creciente. En particular,

$$0 \le y_1 < y_2 \Longrightarrow \sqrt{y_1} < \sqrt{y_2}.$$

Esta propiedad no necesita una nueva demostración mediante cuadrados: es una consecuencia estructural de que x^2 es estrictamente creciente en el dominio donde lo hemos hecho invertible.

El ejemplo muestra por qué la restricción del dominio a $[0,\infty)$ era tan natural. En todo $\mathbb{R}$, la función $x \mapsto x^2$ no es estrictamente monótona: por ejemplo, $-2 < -1$, pero

$$(-2)^2 = 4 > 1 = (-1)^2,$$

mientras que $0 < 1 \Longrightarrow 0^2 < 1^2$.

El comportamiento cambia al atravesar 0. Al restringirnos a $[0,\infty)$ obtenemos una región del dominio en la que el orden se conserva estrictamente y, con ello, la inyectividad necesaria para invertir la función.

La gráfica de la inversa y la diagonal $y = x$. En §2.7 definimos $\Gamma_f = \{(x, f(x)) : x \in I\}$.

Si $f\colon I \to f(I)$ es biyectiva, entonces $y = f(x) \Longleftrightarrow x = f^{-1}(y)$.

Por tanto, $(x,y) \in \Gamma_f \Longleftrightarrow (y,x) \in \Gamma_{f^{-1}}$.

La gráfica de la inversa se obtiene, pues, **intercambiando las dos coordenadas** de cada punto:

$$\boxed{(x,y) \longmapsto (y,x).}$$

Geométricamente, ese intercambio es la reflexión respecto de la diagonal $y = x$. No necesitamos apelar al dibujo para justificarlo. La simetría gráfica es la traducción geométrica de las equivalencias $y = f(x) \Longleftrightarrow x = f^{-1}(y)$. En el ejemplo anterior, los puntos

$$(4,16), \qquad (9,81)$$

de la gráfica de $q(x) = x^2$ sobre $[0,\infty)$ corresponden a los puntos

$$(16,4), \qquad (81,9)$$

de la gráfica de $q^{-1}(x) = \sqrt{x}$.

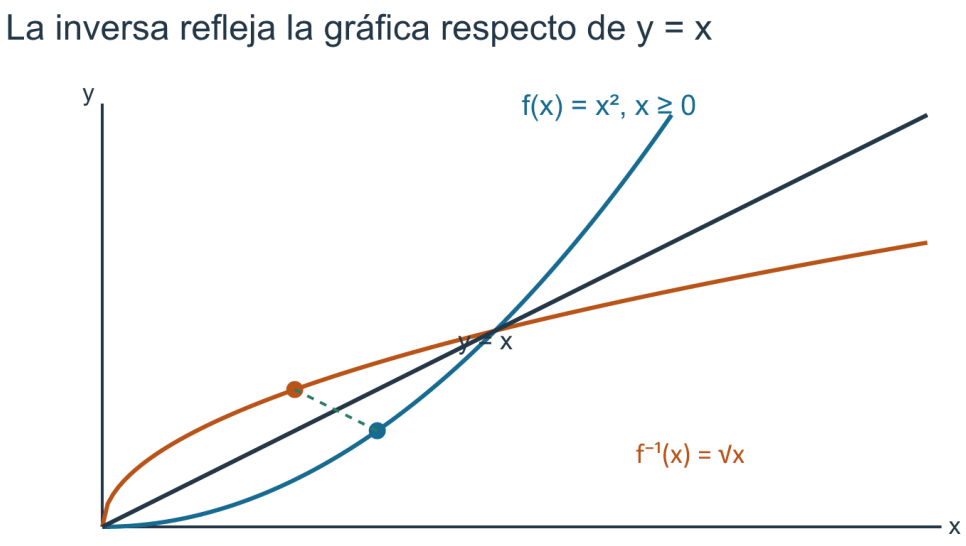


FIGURA 2.1. Para $f(x) = x^2$ en $[0, \infty)$, la gráfica de $f^{-1}(x) = \sqrt{x}$ se obtiene intercambiando las coordenadas. La diagonal $y = x$ permite ver esa simetría.

Qué aporta realmente la monotonía. Podemos reunir la cadena lógica construida en esta sección:

$$\boxed{\text{monotonía estricta} \Longrightarrow \text{inyectividad} \Longrightarrow \text{biyección sobre la imagen} \Longrightarrow \text{existencia de inversa.}}$$

Y además,

$$\boxed{\begin{array}{c}\text{estrictamente creciente} \Longrightarrow \text{inversa estrictamente creciente,} \\ \text{estrictamente decreciente} \Longrightarrow \text{inversa estrictamente decreciente.}\end{array}}$$

La importancia de esta cadena es metodológica. Para construir una inversa no siempre conviene comenzar intentando despejar una fórmula. A menudo es más útil preguntar primero:

1. ¿en qué dominio la función tiene un comportamiento ordenado simple?;
2. ¿es estrictamente monótona allí?;
3. ¿qué imagen produce ese dominio?;
4. ¿podemos tomar esa imagen como codominio?

Si las respuestas encajan, la existencia de la inversa queda resuelta antes de calcular su fórmula. En un problema concreto, identificar una región de monotonía puede facilitar la elección de una rama, mientras que calcular su imagen permite fijar el dominio de la inversa. Las herramientas se combinan según las condiciones que falten por demostrar.

2.10. Laboratorio de estructura funcional

Una fórmula puede imponer restricciones sin fijar todos los datos de una función; una composición puede fallar antes de que empiece el cálculo; una elección de rama puede hacer posible una inversa. Los problemas siguientes combinan esas situaciones. Para abordarlos conviene identificar primero los datos y la condición que debemos demostrar:

identificar la función
→ auditar tipos y conjuntos
→ elegir la estructura pertinente
→ resolver
→ revisar qué hipótesis fueron esenciales

En cada solución, los datos de la función determinarán qué operaciones están permitidas y qué resultado puede aplicarse. El cálculo deberá responder después a la condición planteada en el problema.

Laboratorio 1 — Qué sabemos y qué todavía falta. Alguien escribe solamente $p(x) = \sqrt{4 - x^2}$.

Problema. Determinar qué información puede inferirse de esa expresión y qué información falta todavía para tener una función completamente especificada. Después, construir dos funciones distintas que utilicen exactamente la misma fórmula.

Diagnóstico. La expresión contiene información algebraica. Como estamos trabajando con valores reales, la raíz cuadrada exige $4 - x^2 \geq 0$. Pero una **restricción impuesta por la fórmula** no es todavía lo mismo que un dominio declarado. Tampoco aparece un codominio. Debemos separar, por tanto, tres preguntas:

1. ¿para qué reales puede evaluarse la expresión?;
2. ¿qué conjunto elegimos efectivamente como dominio?;
3. ¿en qué conjunto declaramos que viven las salidas?

Solución. La condición $4 - x^2 \geq 0$ equivale a $x^2 \leq 4$, y por tanto a $-2 \leq x \leq 2$. Así, el **dominio real máximo de la expresión** es $[-2, 2]$.

Además, para todo $x \in [-2, 2]$, $0 \leq \sqrt{4 - x^2} \leq 2$.

Eso permite definir, por ejemplo,

$$f\colon [-2, 2] \to [0, 2], \qquad f(x) = \sqrt{4 - x^2},$$

y también

$$g\colon [0, 2] \to \mathbb{R}, \qquad g(x) = \sqrt{4 - x^2}.$$

Ambas funciones utilizan la misma fórmula, pero son funciones distintas porque sus dominios y codominios no coinciden. Todavía podríamos construir muchas otras. Si $D \subseteq [-2, 2]$, la misma expresión define una función sobre D; y cualquier codominio que contenga las salidas efectivas puede formar parte de una declaración válida. Lo que **no** podemos hacer a partir de la fórmula aislada es decidir, por ejemplo, si «p es sobreyectiva». La sobreyectividad depende del codominio declarado. Tampoco podemos preguntar sin más si «p tiene inversa», porque antes debemos conocer la función completa y su inyectividad/sobreyectividad.

Lectura de la solución. Este primer laboratorio recupera la pregunta central del capítulo:

una expresión restringe posibilidades, pero no especifica por sí sola toda la función.

El dominio natural de una expresión es una información que podemos **deducir**. El dominio y el codominio de una función son datos que debemos **declarar**.

Laboratorio 2 — Una fórmula, cuatro clasificaciones. Consideremos ahora la regla $x \longmapsto x^2 - 1$.

Definamos

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = x^2 - 1,$$

$$g\colon [0,\infty) \to [-1,\infty), \qquad g(x) = x^2 - 1,$$

$$h\colon \mathbb{R} \to [-1,\infty), \qquad h(x) = x^2 - 1,$$

y

$$k\colon [0,\infty) \to \mathbb{R}, \qquad k(x) = x^2 - 1.$$

Problema. Determinar la imagen y clasificar cada función como inyectiva, sobreyectiva, biyectiva o ninguna de ellas.

Diagnóstico. La fórmula es idéntica en las cuatro declaraciones. Por tanto, el cálculo de valores no distingue las funciones. Las diferencias deben provenir de dos lugares:

- qué entradas están permitidas;
- qué conjunto se ha declarado como codominio.

El cuadrado sobre todo $\mathbb{R}$ identifica x y $-x$, mientras que sobre $[0,\infty)$ preserva estrictamente el orden. Además, $x^2 - 1 \geq -1$. Estas dos observaciones controlan casi toda la clasificación.

Solución. Para f y h, el dominio es todo $\mathbb{R}$. Como $f(1) = f(-1) = 0$ y $h(1) = h(-1) = 0$, ninguna de las dos funciones es inyectiva. La imagen de la regla sobre $\mathbb{R}$ es $[-1,\infty)$. En efecto, $x^2 - 1 \geq -1$ para todo real x. Recíprocamente, si $y \geq -1$, entonces $y + 1 \geq 0$ y podemos tomar $x = \sqrt{y+1}$, para obtener $x^2 - 1 = y$.

Por tanto, $\operatorname{Im}(f) = \operatorname{Im}(h) = [-1,\infty)$.

Como el codominio de f es $\mathbb{R}$, los valores menores que -1 no son alcanzados. Así,

f no es inyectiva ni sobreyectiva.

En cambio, el codominio de h es exactamente $[-1,\infty)$, de modo que

h es sobreyectiva pero no inyectiva.

Consideremos ahora g y k. Si $0 \leq x_1 < x_2$, entonces, como vimos en §2.9, $x_1^2 < x_2^2$, y al restar 1 obtenemos $x_1^2 - 1 < x_2^2 - 1$. La regla es, por tanto, estrictamente creciente sobre $[0,\infty)$ y las dos funciones son inyectivas. Su imagen sigue siendo $[-1,\infty)$, porque para cada $y \geq -1$ el número $\sqrt{y+1} \in [0,\infty)$ satisface $(\sqrt{y+1})^2 - 1 = y$.

Así,

$$\boxed{g \text{ es biyectiva,}}$$

mientras que

$$\boxed{k \text{ es inyectiva pero no sobreyectiva.}}$$

Podemos resumir:

Función	Imagen	Inyectiva	Sobreyectiva	Biyectiva
$f\colon \mathbb{R} \to \mathbb{R}$	$[-1, \infty)$	no	no	no
$g\colon [0, \infty) \to [-1, \infty)$	$[-1, \infty)$	sí	sí	sí
$h\colon \mathbb{R} \to [-1, \infty)$	$[-1, \infty)$	no	sí	no
$k\colon [0, \infty) \to \mathbb{R}$	$[-1, \infty)$	sí	no	no

Lectura de la solución. La tabla muestra dos sensibilidades diferentes:

$$\boxed{\text{la inyectividad responde al dominio}}$$

y

$$\boxed{\text{la sobreyectividad responde al codominio y a la imagen.}}$$

Cambiar el dominio puede eliminar entradas que producían la misma salida. Cambiar el codominio puede convertir una función en sobreyectiva sin alterar ninguno de sus valores punto a punto.

Laboratorio 3 — Componer empieza por verificar que se puede. Sean

$$f\colon \mathbb{R} \to [0, \infty), \qquad f(x) = x^2,$$

$$g\colon [0, \infty) \to \mathbb{R}, \qquad g(u) = u + 1,$$

y

$$h\colon (0, \infty) \to \mathbb{R}, \qquad h(u) = \frac{1}{u}.$$

Problema. Decidir cuáles de las composiciones

$$g \circ f, \qquad f \circ g, \qquad h \circ f, \qquad f \circ h, \qquad h \circ g$$

están definidas con las funciones tal como han sido declaradas. Calcular las que existan y explicar por qué el orden de composición importa.

Diagnóstico. Antes de sustituir fórmulas debemos auditar tipos. Para que $v \circ u$ esté definida sobre todo el dominio de u necesitamos $\operatorname{Im}(u) \subseteq \operatorname{Dom}(v)$. Esta condición es anterior a cualquier simplificación algebraica.

Solución. Como $\mathrm{Im}(f) = [0,\infty)$ y ese conjunto es exactamente el dominio de g, existe $g \circ f \colon \mathbb{R} \to \mathbb{R}$, con $(g \circ f)(x) = g(x^2) = x^2 + 1$.

También existe $f \circ g$. En efecto, $\mathrm{Im}(g) = [1,\infty) \subseteq \mathbb{R} = \mathrm{Dom}(f)$.

Por tanto, $f \circ g \colon [0,\infty) \to [0,\infty)$ y $(f \circ g)(u) = f(u+1) = (u+1)^2$.

Estas dos composiciones ya muestran que el orden importa. Bajo nuestra convención ni siquiera son la misma función, porque tienen dominios y codominios distintos. Incluso si comparamos un valor perteneciente a ambos dominios, por ejemplo 2, obtenemos $(g \circ f)(2) = 5$ mientras que $(f \circ g)(2) = 9$. Examinemos ahora $h \circ f$. Para que existiera sobre todo $\mathbb{R}$ necesitaríamos $f(\mathbb{R}) \subseteq (0,\infty)$.

Pero $f(0) = 0 \notin (0,\infty)$.

Así,

$$\boxed{h \circ f \text{ no está definida como composición sobre todo } \mathbb{R}.}$$

Es cierto que la expresión formal $\frac{1}{x^2}$ puede evaluarse cuando $x \neq 0$. Pero eso produciría una función sobre $\mathbb{R} \setminus \{0\}$, no la composición $h \circ f$ con el dominio original de f.

En cambio, $f \circ h$ sí existe porque $\mathrm{Im}(h) = (0,\infty) \subseteq \mathbb{R} = \mathrm{Dom}(f)$.

Entonces

$$(f \circ h)(u) = \left(\frac{1}{u}\right)^2 = \frac{1}{u^2}, \qquad u > 0.$$

Finalmente,

$$\mathrm{Im}(g) = [1,\infty) \subseteq (0,\infty) = \mathrm{Dom}(h),$$

por lo que también existe $h \circ g$ y

$$(h \circ g)(u) = \frac{1}{u+1}, \qquad u \geq 0.$$

Lectura de la solución. El error que debemos evitar es

$$\boxed{\text{ver dos fórmulas y componerlas antes de mirar sus tipos.}}$$

Una expresión algebraica simplificada puede tener sentido en ciertos puntos aun cuando la composición de las funciones originalmente declaradas no exista sobre todo el dominio requerido.

Laboratorio 4 — Cuando una prueba fallida descubre el teorema correcto. Sea $f \colon A \to B$ y sean $E, F \subseteq A$.

Una igualdad tentadora es

$$\boxed{f(E \setminus F) = f(E) \setminus f(F).}$$

Problema. Investigar si la igualdad es siempre verdadera. No debemos comenzar buscando un contraejemplo al azar: intentemos primero demostrarla por doble inclusión y observemos dónde se detiene la prueba.

Diagnóstico. En §2.6 aprendimos que las imágenes directas pueden mezclar testigos distintos. Aquí aparece una dificultad parecida: que un punto x no pertenezca a F no impide, en principio, que otro punto $z \in F$ tenga la misma imagen. La prueba debe revelar exactamente dónde importa esa posibilidad.

Primera inclusión. Tomemos $y \in f(E) \setminus f(F)$.

Como $y \in f(E)$, existe $x \in E$ tal que $f(x) = y$.

Además, $y \notin f(F)$.

Si x perteneciera a F, entonces $f(x) = y$ pertenecería a $f(F)$, contradicción. Por tanto, $x \notin F$.

Así, $x \in E \setminus F$, y por ello $y = f(x) \in f(E \setminus F)$.

Hemos probado siempre que

$$\boxed{f(E) \setminus f(F) \subseteq f(E \setminus F).}$$

Intento de la inclusión inversa. Tomemos ahora $y \in f(E \setminus F)$.

Existe entonces $x \in E \setminus F$ tal que $f(x) = y$.

Sabemos que $x \in E$, así que $y \in f(E)$.

Para concluir que $y \in f(E) \setminus f(F)$ necesitamos demostrar $y \notin f(F)$. Aquí se atasca la prueba. Del hecho $x \notin F$

no podemos deducir automáticamente $f(x) \notin f(F)$.

Podría existir otro punto $z \in F$ con $f(z) = f(x)$. Ese es el paso ilegítimo que habría que introducir para completar la demostración sin hipótesis adicionales.

La hipótesis que repara la prueba. Supongamos ahora que f es inyectiva. Si $y = f(x)$ perteneciera también a $f(F)$, existiría $z \in F$ con $f(z) = y = f(x)$.

Por inyectividad, $z = x$.

Pero $z \in F$ y $x \notin F$, contradicción. Por tanto, $y \notin f(F)$, y obtenemos la inclusión inversa.

Así, bajo inyectividad,

$$\boxed{f(E \setminus F) = f(E) \setminus f(F).}$$

Contraejemplo sin inyectividad. Tomemos

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = x^2,$$

$$E = \{-1\}, \qquad F = \{1\}.$$

Como $E \cap F = \varnothing$, $E \setminus F = \{-1\}$, y por tanto $f(E \setminus F) = \{1\}$.

Pero $f(E) = \{1\} = f(F)$, de modo que $f(E) \setminus f(F) = \varnothing$.

Así,

$$\boxed{\varnothing = f(E) \setminus f(F) \subsetneq f(E \setminus F) = \{1\}.}$$

Lectura de la solución. El laboratorio reproduce un patrón de investigación matemática muy útil:

conjetura
→ intento de prueba
→ punto de atasco
→ hipótesis necesaria
→ enunciado corregido
→ contraejemplo sin la hipótesis

Una prueba que se atasca no es necesariamente tiempo perdido. Si sabemos leer el atasco, puede revelar la estructura exacta del resultado verdadero.

Laboratorio 5 — Fabricar una inversa eligiendo la rama adecuada. Consideremos

$$q\colon \mathbb{R} \to \mathbb{R}, \qquad q(x) = (x-2)^2 + 1.$$

Problema. La función q no es biyectiva. Construir dos restricciones naturales que sí sean biyectivas sobre su imagen y determinar sus inversas.

Diagnóstico. Hay dos fallos en la declaración original. Primero, $q(1) = 2 = q(3)$, así que q no es inyectiva. Segundo, $q(x) = (x-2)^2 + 1 \geq 1$, por lo que ningún real menor que 1 pertenece a la imagen. Así, q tampoco es sobreyectiva sobre $\mathbb{R}$. La gráfica algebraica tiene un punto central en $x = 2$. A la derecha de ese punto la regla es estrictamente creciente; a la izquierda es estrictamente decreciente. Eso sugiere dos restricciones posibles.

Primera rama: $x \geq 2$. Definamos

$$q_+\colon [2,\infty) \to [1,\infty), \qquad q_+(x) = (x-2)^2 + 1.$$

Si $2 \leq x_1 < x_2$, entonces

$$q_+(x_2) - q_+(x_1) = (x_2 - x_1)(x_1 + x_2 - 4).$$

El primer factor es positivo. Además, $x_2 > 2$ y $x_1 \geq 2$, así que $x_1 + x_2 - 4 > 0$.

Por tanto, $q_+(x_1) < q_+(x_2)$.

La función es estrictamente creciente y, en consecuencia, inyectiva. Para demostrar sobreyectividad sobre $[1,\infty)$, sea $y \geq 1$. Entonces $y - 1 \geq 0$ y podemos tomar $x = 2 + \sqrt{y-1} \geq 2$.

Se cumple $q_+(x) = (\sqrt{y-1})^2 + 1 = y$.

Así, q_+ es biyectiva. Resolviendo $y = (x-2)^2 + 1$ con la condición $x \geq 2$, obtenemos $x - 2 = \sqrt{y-1}$, y por tanto

$$\boxed{q_+^{-1}(y) = 2 + \sqrt{y-1}, \qquad y \geq 1.}$$

Segunda rama: $x \leq 2$. Definamos ahora

$$q_-\colon (-\infty, 2] \to [1,\infty), \qquad q_-(x) = (x-2)^2 + 1.$$

Si $x_1 < x_2 \leq 2$, entonces

$$q_-(x_2) - q_-(x_1) = (x_2 - x_1)(x_1 + x_2 - 4).$$

El primer factor es positivo, pero ahora $x_1 + x_2 - 4 < 0$, porque $x_1 < 2$ y $x_2 \leq 2$. Por tanto, $q_-(x_2) < q_-(x_1)$, de modo que q_- es estrictamente decreciente y, por ello, inyectiva. Dado $y \geq 1$, el número $x = 2 - \sqrt{y-1} \leq 2$ satisface $q_-(x) = y$. Así, q_- también es sobreyectiva sobre $[1, \infty)$ y por tanto biyectiva. Esta vez la condición $x \leq 2$ selecciona el signo negativo:

$$\boxed{q_-^{-1}(y) = 2 - \sqrt{y-1}, \qquad y \geq 1.}$$

Lectura de la solución. Las dos fórmulas son inversas de **dos funciones distintas**, obtenidas mediante restricciones diferentes del dominio: $q_+ eq_-$. Cada una posee su propia inversa única. El patrón general es importante:

función no invertible
→ localizar una región inyectiva
→ restringir dominio
→ ajustar codominio a la imagen
→ invertir

Laboratorio 6 — Una sucesión antes de hablar de límites. Definamos

$$a\colon \mathbb{N} \to \mathbb{R}, \qquad a(n) = \frac{n}{n+1},$$

donde $\mathbb{N} = \{0, 1, 2, \dots\}$.

Problema. Analizar esta regla utilizando únicamente el lenguaje funcional del capítulo: dominio, codominio, imagen, preimágenes puntuales, monotonía, inyectividad, sobreyectividad y gráfica. No debemos hablar todavía de convergencia.

Diagnóstico. La notación $a(n)$ puede hacer que el objeto parezca nuevo, pero estructuralmente no lo es. Tenemos simplemente una función cuyo dominio es $\mathbb{N}$. La primera pregunta, por tanto, no es «¿a qué se acerca?», sino

$$\boxed{\text{¿qué función es y qué propiedades funcionales tiene?}}$$

Solución. El dominio es $\mathbb{N}$, y el codominio declarado es $\mathbb{R}$. Como $n \geq 0$, el denominador $n+1$ es siempre positivo, de modo que la regla está definida para cada natural.

Podemos reescribir $a(n) = 1 - \frac{1}{n+1}$.

En particular, $0 \leq a(n) < 1$ para todo $n \in \mathbb{N}$. Sin embargo, la imagen no es todo el intervalo $[0, 1)$. Es el conjunto discreto

$$\boxed{\operatorname{Im}(a) = \left\{ \frac{n}{n+1} : n \in \mathbb{N} \right\} = \left\{ 0, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \dots \right\}.}$$

Por ejemplo,

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

$$a^{-1}\left(\left\{\frac{2}{3}\right\}\right) = \{2\},$$

y $a^{-1}(\{1\}) = \varnothing$.

Estudiemos ahora la monotonía. Para $n \in \mathbb{N}$, $a(n+1) - a(n) = \frac{n+1}{n+2} - \frac{n}{n+1}$.

Llevando a común denominador,

$$a(n+1) - a(n) = \frac{(n+1)^2 - n(n+2)}{(n+1)(n+2)} = \frac{1}{(n+1)(n+2)}.$$

Como el denominador es positivo, $a(n+1) > a(n)$. Más generalmente, si $m < n$, entonces $\frac{m}{m+1} < \frac{n}{n+1}$, pues, al multiplicar por los denominadores positivos, $m(n+1) < n(m+1)$ equivale a $m < n$. Así, a es estrictamente creciente sobre $\mathbb{N}$. Por el teorema de §2.9,

$$\boxed{a \text{ es inyectiva.}}$$

No es sobreyectiva como función $\mathbb{N} \to \mathbb{R}$, porque, por ejemplo, $2 \notin \operatorname{Im}(a)$. Si en cambio consideramos $a\colon \mathbb{N} \to \operatorname{Im}(a)$, la función es biyectiva. Para $y \in \operatorname{Im}(a)$, de $y = \frac{n}{n+1}$ obtenemos $y(n+1) = n$,

$$y = n(1-y),$$

y por tanto

$$\boxed{n = \frac{y}{1-y}.}$$

Así, sobre la imagen, $a^{-1}(y) = \frac{y}{1-y}$.

No hay conflicto con el hecho de que la expresión de la derecha parezca real: si y pertenece realmente a $\operatorname{Im}(a)$, ese cociente es precisamente el natural que produjo y.

La gráfica es

$$\Gamma_a = \left\{ \left(n, \frac{n}{n+1} \right) : n \in \mathbb{N} \right\}.$$

A diferencia de muchas gráficas anteriores, no forma una curva completa sobre un intervalo: solo contiene los puntos cuyas primeras coordenadas son naturales.

Lectura de la solución. Este objeto es lo que en el capítulo siguiente llamaremos una **sucesión real**. Pero todo lo que acabamos de hacer pertenece ya a la teoría de funciones:

$$\boxed{\text{sucesión} = \text{función con dominio } \mathbb{N}.}$$

Todavía no hemos preguntado si los valores «se acercan» a algún número ni hemos definido qué significaría rigurosamente esa frase. Esa será precisamente la nueva cuestión de `T1-C04`.

Qué estructuras se combinaron. En las seis estaciones, las mismas definiciones intervienen en combinaciones diferentes:

- El Laboratorio 1 separó **restricciones de una expresión** de **datos de una función**.
- El Laboratorio 2 mostró simultáneamente la sensibilidad de inyectividad al dominio y de sobreyectividad al codominio.
- El Laboratorio 3 obligó a verificar **tipos** antes de componer y distinguió composición funcional de mera sustitución algebraica.
- El Laboratorio 4 utilizó una **prueba fallida** como instrumento para descubrir la hipótesis correcta.
- El Laboratorio 5 combinó restricción, monotonía, imagen, biyectividad e inversión.

- El Laboratorio 6 mostró que el siguiente objeto del curso, una sucesión, entra naturalmente en el lenguaje funcional ya construido.

La arquitectura general puede verse así:

función completa
→ imagen/preimagen
→ inyectividad/sobreyectividad
→ composición
→ inversa,
función + producto cartesiano
→ gráfica,
función + orden
→ monotonía
→ inyectividad.

Un mapa de decisión antes del banco final. Cuando aparezca un problema funcional, conviene formular preguntas en un orden parecido al siguiente:

¿qué función es? → dominio/codominio → imagen/preimagen
→ inyectiva/sobreyectiva? → ¿se compone? → ¿se puede invertir?

Cada problema selecciona las preguntas pertinentes del diagrama. Especificar primero la función permite decidir si necesitamos calcular una imagen, comprobar una composición o demostrar que un antecedente existe y es único. A ese eje principal podemos añadir dos preguntas transversales:

- **¿la gráfica aporta información útil?** Entonces conviene traducir entre puntos $(x, f(x))$ y propiedades funcionales, sin confundir imagen con codominio;
- **¿el dominio está ordenado?** Entonces puede ser útil estudiar monotonía, especialmente si buscamos inyectividad o una restricción invertible.

Antes de seguir.

Antes de seguir
1. En el Laboratorio 1, ¿por qué $[-2, 2]$ no debe llamarse automáticamente «el dominio de p»?
Respuesta. Porque $[-2, 2]$ es el dominio real máximo permitido por la expresión $\sqrt{4 - x^2}$. Una función puede declararse sobre cualquier subconjunto adecuado de ese conjunto. El dominio forma parte de los datos de la función, no es siempre sinónimo de «todos los puntos donde la fórmula puede evaluarse».

Antes de seguir
2. En el Laboratorio 4, ¿qué inclusión sobre diferencias de imágenes es válida sin suponer inyectividad?
Respuesta. Siempre se cumple $f(E) \setminus f(F) \subseteq f(E \setminus F)$.
La inclusión inversa puede fallar porque un punto de $E \setminus F$ y otro punto de F pueden compartir la misma imagen. La inyectividad elimina exactamente esa posibilidad.

Antes de seguir
3. ¿Por qué las dos ramas del Laboratorio 5 no contradicen la unicidad de la función inversa?
Respuesta. Porque q_+ y q_- son funciones distintas: tienen dominios distintos. Cada una posee una única inversa, y esas inversas son diferentes porque revierten asignaciones diferentes.

Antes de seguir
4. En el Laboratorio 6, ¿por qué no podemos escribir $\mathrm{Im}(a) = [0, 1)$?
Respuesta. Porque a solo toma valores correspondientes a entradas naturales. Aunque todos sus valores pertenecen a $[0, 1)$, la mayoría de los números de ese intervalo no son alcanzados. Por ejemplo, no existe $n \in \mathbb{N}$ con $n/(n + 1) = 1/3$.

2.11. Ejercicios y soluciones

Llegamos al banco final de `T1-C03`. En las secciones anteriores las herramientas funcionales fueron apareciendo de manera progresiva y, en §2.10, se combinaron todavía con una guía explícita de estrategia. Aquí la ayuda disminuye deliberadamente: antes de calcular, habrá que decidir qué estructura es relevante y qué datos de la función deben controlarse. El banco contiene exactamente cuarenta ejercicios, distribuidos en siete niveles:

$$7A + 7B + 7C + 6D + 5E + 5F + 3G = 40.$$

Los niveles no miden solamente dificultad algebraica. Un ejercicio puede resultar corto y, sin embargo, exigir una distinción conceptual importante. En particular, se evaluará de manera recurrente la capacidad de separar:

- fórmula, función completa y gráfica;
- dominio, codominio e imagen;
- preimagen de un conjunto y función inversa;
- existencia de una inversa y cálculo de una fórmula candidata;
- propiedades de una función y propiedades de una composición.

Todos los problemas pueden resolverse con el capítulo precedente sobre los reales y con el presente capítulo. No se necesitan límites, continuidad ni derivadas.

Nivel A — Reconocimiento y comprensión. Ejercicio A1. Dominio, codominio e imagen. Sea

$$f\colon [-2, 3] \to \mathbb{R}, \qquad f(x) = x^2.$$

Identifica el dominio, el codominio y la imagen de f. Decide si imagen y codominio coinciden.

Ejercicio A2. ¿Es realmente una función? Considera la relación

$$R = \{(0,1), (1,2), (2,2), (2,3)\}$$

desde $A = \{0,1,2\}$ hacia $B = \{1,2,3\}$.

1. ¿Define R una función $A \to B$?
2. Identifica exactamente qué parte de la condición «a cada entrada corresponde exactamente una salida» falla.
3. Elimina el menor número posible de pares para obtener una función $A \to B$.

Ejercicio A3. Misma fórmula, ¿misma función? Sean

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = x^2,$$

y

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

Según la convención adoptada en este libro, decide si $f = g$ y justifica tu respuesta. Después compara sus imágenes.

Ejercicio A4. Imagen y preimagen. Sea

$$q\colon \mathbb{R} \to [0,\infty), \qquad q(x) = x^2.$$

Calcula $q([-2,1])$ y $q^{-1}([1,4))$.
Recuerda que en la segunda expresión q^{-1} designa una **preimagen de conjunto**.

Ejercicio A5. Leer una función desde una tabla. Sea $p\colon \{-2,-1,0,1\} \to \{1,2,3,4\}$ la función dada por

x	-2	-1	0	1
$p(x)$	3	1	1	4

Determina:

1. $\mathrm{Im}(p)$;
2. si p es inyectiva;
3. si p es sobreyectiva.

Ejercicio A6. Una restricción cambia la función. Sea

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = x^2,$$

y considera la restricción $f|_{[0,\infty)}\colon [0,\infty) \to \mathbb{R}$.

1. ¿Cuál es su imagen?
2. ¿Es inyectiva?
3. ¿Es sobreyectiva?
4. ¿Qué cambio adicional permitiría convertir esta restricción en una biyección?

Ejercicio A7. Leer la notación como una afirmación lógica. Sea

$$a\colon \{0,1,2,3\} \to \mathbb{R}, \qquad a(n) = \frac{n}{n+1}.$$

1. Traduce la declaración $a\colon \{0,1,2,3\} \to \mathbb{R}$ a una frase que haga explícitas existencia y unicidad de la salida.
2. Calcula $\mathrm{Im}(a)$.
3. Explica por qué el hecho de que la imagen no sea todo $\mathbb{R}$ no impide que a sea una función.

Nivel B — Aplicación directa. Ejercicio B1. Dos órdenes de composición. Sean

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = 2x - 1,$$

y

$$g\colon \mathbb{R} \to \mathbb{R}, \qquad g(x) = x^2.$$

Calcula $g \circ f$ y $f \circ g$. Decide si son la misma función.

Ejercicio B2. Antes de componer, controla los tipos. Sean

$$u\colon [-1,\infty) \to [0,\infty), \qquad u(x) = \sqrt{x+1},$$

y

$$v\colon [0,\infty) \setminus \{2\} \to \mathbb{R}, \qquad v(t) = \frac{1}{t-2}.$$

1. Explica por qué $v \circ u$ no está definida como composición sobre todo $[-1,\infty)$ con las funciones tal como fueron declaradas.
2. Encuentra el mayor subconjunto $E \subseteq [-1,\infty)$ tal que $v \circ (u|_E)$ esté bien definida.
3. Escribe su fórmula.

Ejercicio B3. Intervalos bajo una función afín. Sea

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = 3x - 2.$$

Calcula $f([1,4))$ y $f^{-1}((-5,4])$.

Ejercicio B4. La preimagen puede tener dos ramas. Sea

$$q\colon \mathbb{R} \to [0,\infty), \qquad q(x) = x^2.$$

Calcula $q^{-1}((4,9])$ y explica por qué el resultado no es un solo intervalo.

Ejercicio B5. Clasificar y reparar. Considera

$$r\colon [0,\infty) \to \mathbb{R}, \qquad r(x) = x^2.$$

1. Clasifica r respecto de inyectividad y sobreyectividad.
2. Modifica solamente el codominio para obtener una biyección.
3. Escribe la inversa de la función resultante.

Ejercicio B6. Función por tramos. Sea

$$p\colon \mathbb{R} \to \mathbb{R}, \qquad p(x) = \begin{cases} x+1, & x < 0, \\ x^2, & x \geq 0. \end{cases}$$

1. Calcula $p(-2)$, $p(0)$ y $p(2)$.
2. Determina $\mathrm{Im}(p)$.
3. Decide si p es inyectiva y si es sobreyectiva.

Ejercicio B7. Transformar una gráfica sin memorizar recetas. Sea $f\colon \mathbb{R} \to \mathbb{R}$ una función y define $h(x) = 2f(x-3) - 1$.

Supón que (a, b) pertenece a la gráfica de f.

1. Determina qué punto de la gráfica de h corresponde a (a, b).
2. Describe, en orden, las transformaciones geométricas que llevan la gráfica de f a la de h.

Nivel C — Combinación estructural. Ejercicio C1. Inyectividad de una función afín. Sea

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = ax + b,$$

donde $a \neq 0$. Demuestra que f es inyectiva.

Ejercicio C2. Biyectividad del recíproco. Sea

$$r\colon (0,\infty) \to (0,\infty), \qquad r(x) = \frac{1}{x}.$$

Demuestra que r es biyectiva y determina r^{-1}.

Ejercicio C3. Probar primero, invertir después. Sea

$$f\colon [0,\infty) \to [1,\infty), \qquad f(x) = x^2 + 1.$$

1. Demuestra que f es inyectiva.
2. Demuestra que f es sobreyectiva construyendo un antecedente de un $y \in [1,\infty)$ arbitrario.
3. Solo después de lo anterior, determina f^{-1}.

Ejercicio C4. La composición conserva la inyectividad. Sean

$$f\colon A \to B, \qquad g\colon B \to C$$

funciones inyectivas. Demuestra directamente, a partir de la definición, que $g \circ f$ es inyectiva.

Ejercicio C5. Álgebra de preimágenes. Sea $f\colon A \to B$ y sean $U, V \subseteq B$. Demuestra: $f^{-1}(U \cap V) = f^{-1}(U) \cap f^{-1}(V)$ y $f^{-1}(B \setminus U) = A \setminus f^{-1}(U)$.

Haz la demostración mediante pertenencia de un elemento arbitrario.

Ejercicio C6. El cuadrado cambia de monotonía. Demuestra que $q(x) = x^2$ es estrictamente creciente en $[0, \infty)$ y estrictamente decreciente en $(-\infty, 0]$. Deduce que su restricción a cualquiera de esos intervalos es inyectiva.

Ejercicio C7. Gráfica de una función y de su inversa. Sea

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = 2x - 3.$$

1. Demuestra que f es biyectiva y calcula f^{-1}.
2. Prueba que
$$(x, y) \in \Gamma_f \quad \Longleftrightarrow \quad (y, x) \in \Gamma_{f^{-1}}.$$
3. Interpreta geométricamente la equivalencia anterior.

Nivel D — Inversión y diagnóstico. Ejercicio D1. ¿Qué hereda una composición inyectiva? Sean

$$f\colon A \to B, \qquad g\colon B \to C.$$

Supón que $g \circ f$ es inyectiva.

1. Demuestra que f debe ser inyectiva.
2. Decide si g debe ser inyectiva. Si la respuesta es negativa, construye un contraejemplo explícito.

Ejercicio D2. ¿Qué hereda una composición sobreyectiva? Sean

$$f\colon A \to B, \qquad g\colon B \to C.$$

Supón que $g \circ f$ es sobreyectiva.

1. Demuestra que g debe ser sobreyectiva.
2. Decide si f debe ser sobreyectiva. Si la respuesta es negativa, construye un contraejemplo explícito.

Ejercicio D3. Auditoría de dominio y codominio. Considera las cuatro funciones

$$p\colon \mathbb{R} \to \mathbb{R}, \qquad p(x) = x^2,$$

$$q\colon \mathbb{R} \to [0, \infty), \qquad q(x) = x^2,$$

$$r\colon [0, \infty) \to \mathbb{R}, \qquad r(x) = x^2,$$

y

$$s\colon [0, \infty) \to [0, \infty), \qquad s(x) = x^2.$$

1. Decide qué pares, si alguno, son funciones iguales según la convención del libro.
2. Clasifica cada función como inyectiva, sobreyectiva, biyectiva o ninguna de las anteriores.
3. Explica qué propiedad se modifica al cambiar solo el dominio y qué propiedad puede modificarse al cambiar solo el codominio.

Ejercicio D4. Una «inversa» demasiado rápida. Un estudiante escribe:

Si $f(x) = x^2$, entonces $y = x^2$, de donde $x = \sqrt{y}$. Por tanto, $f^{-1}(y) = \sqrt{y}$.

Analiza el argumento para la función

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

1. Localiza el error conceptual.
2. Explica por qué resolver una ecuación no basta para demostrar que existe una función inversa.
3. Repara la situación modificando el dominio de f de dos maneras distintas y escribe las inversas correspondientes.

Ejercicio D5. Una igualdad plausible que puede fallar. Sea $f\colon A \to B$ y $E, F \subseteq A$.

1. Demuestra que siempre se cumple $f(E \cap F) \subseteq f(E) \cap f(F)$.
2. Explica dónde se atascaría un intento de demostrar la inclusión inversa sin hipótesis adicionales.
3. Indica una hipótesis suficiente sobre f que convierta la inclusión en igualdad.

Ejercicio D6. Reparar una función por tramos mal definida. Se pretende definir

$$h\colon \mathbb{R} \to \mathbb{R}, \qquad h(x) = \begin{cases} x + 1, & x \geq 0, \\ -x, & x \leq 0. \end{cases}$$

1. Explica por qué esta regla no define una función.
2. Propón dos reparaciones distintas que cambien únicamente las condiciones de los tramos, no las fórmulas $x + 1$ y $-x$.
3. ¿Las dos reparaciones obtenidas definen necesariamente la misma función? Justifica.

Nivel E — Contraejemplos. Ejercicio E1. Misma fórmula, funciones distintas. Construye dos funciones distintas que utilicen la misma fórmula $1/x$. Deben diferir por su dominio, por su codominio o por ambos. Justifica formalmente que no son iguales.

Ejercicio E2. Separar inyectividad y sobreyectividad. Construye:

1. una función inyectiva que no sea sobreyectiva;
2. una función sobreyectiva que no sea inyectiva.

Puedes usar conjuntos finitos o subconjuntos de $\mathbb{R}$, pero debes declarar dominio, codominio y regla.

Ejercicio E3. La composición no es conmutativa. Construye funciones $f, g\colon \mathbb{R} \to \mathbb{R}$ tales que $g \circ f \neq f \circ g$.

Calcula ambas composiciones y exhibe al menos un valor de x en el que difieran.

Ejercicio E4. Intersecciones e imágenes. Construye una función $f\colon A \to B$ y subconjuntos $E, F \subseteq A$ tales que $f(E \cap F) \subsetneq f(E) \cap f(F)$.

Explica qué rasgo de f hace posible la inclusión estricta.

Ejercicio E5. No decreciente no significa inyectiva. Construye una función $m\colon \mathbb{R} \to \mathbb{R}$ que sea no decreciente pero no inyectiva. Verifica ambas afirmaciones.

Nivel F — Descubrimiento guiado. Ejercicio F1. Reconstruir una función desde su gráfica como conjunto. Se sabe que $f\colon A \to B$ tiene $B = \{1, 2, 4\}$ y que su gráfica es exactamente $\Gamma_f = \{(-2, 1), (0, 1), (3, 4)\}$.

1. Reconstruye A.
2. Escribe todos los valores de f.
3. Determina $\mathrm{Im}(f)$.
4. Decide si f es inyectiva y si es sobreyectiva.
5. Explica por qué conocer solo la gráfica como conjunto de pares no habría bastado, bajo la convención de este libro, para reconstruir el codominio si B no hubiese sido declarado.

Ejercicio F2. Descubrir las leyes de preimagen. Sea $f\colon A \to B$ y sean $U, V \subseteq B$.

Completa una cadena de equivalencias empezando por un elemento arbitrario $x \in A$ para demostrar: $f^{-1}(U \cup V) = f^{-1}(U) \cup f^{-1}(V)$,

$$f^{-1}(U \cap V) = f^{-1}(U) \cap f^{-1}(V),$$

y $f^{-1}(B \setminus U) = A \setminus f^{-1}(U)$.

Después explica por qué ninguna de estas identidades requiere inyectividad ni sobreyectividad.

Ejercicio F3. Fabricar dos biyectividades a partir del cuadrado. Parte de la regla $x \mapsto x^2$.

1. Encuentra una restricción del dominio y un codominio que produzcan una biyección cuya inversa sea $y \mapsto \sqrt{y}$.
2. Encuentra otra restricción, distinta de la anterior, que produzca una biyección cuya inversa sea $y \mapsto -\sqrt{y}$.
3. Demuestra en ambos casos inyectividad y sobreyectividad.

Ejercicio F4. Descubrir el criterio exacto de existencia de una inversa. Sea $f\colon A \to B$. Para cada $y \in B$, considera la ecuación $f(x) = y$.

1. Explica qué propiedad de f garantiza que para cada y exista **al menos una** solución $x \in A$.
2. Explica qué propiedad garantiza que exista **a lo sumo una** solución.
3. Deduce que existe una función inversa $f^{-1}\colon B \to A$ si y solo si f es biyectiva.
4. Indica cómo se define $f^{-1}(y)$ a partir de la solución única.

Ejercicio F5. Reconstruir una transformación como composición. Define

$$s(x) = x^2, \qquad q(x) = -2(x - 3)^2 + 1.$$

Para $a, c \in \mathbb{R}$, escribe

$$T_a(x) = x + a, \qquad M_c(x) = cx.$$

1. Expresa q como una composición de s, traslaciones T_a y una dilatación/reflexión M_c.

2. Usa la composición obtenida para describir, en orden, cómo se transforma la gráfica de s en la de q.
3. Determina el punto al que se transforma el vértice $(0, 0)$.

Nivel G — Desafíos. Ejercicio G1. Caracterización de la inyectividad mediante intersecciones. Sea $f\colon A \to B$. Demuestra que las siguientes afirmaciones son equivalentes:

1. f es inyectiva;
2. para todos $E, F \subseteq A$, $f(E \cap F) = f(E) \cap f(F)$.

En la dirección menos inmediata, elige subconjuntos apropiados para recuperar la inyectividad a partir de la igualdad de imágenes.

Ejercicio G2. Inversas laterales y cancelación. Sean

$$f\colon A \to B, \qquad g\colon B \to A.$$

1. Demuestra que $g \circ f = \mathrm{id}_A$ implica que f es inyectiva y g es sobreyectiva.
2. Demuestra que $f \circ g = \mathrm{id}_B$ implica que f es sobreyectiva y g es inyectiva.
3. Construye contraejemplos que muestren que, en general, las conclusiones de 1 o de 2 no bastan por sí solas para recuperar la identidad de composición correspondiente.

Ejercicio G3. Composición biyectiva: qué puede concluirse. Sean

$$f\colon A \to B, \qquad g\colon B \to C,$$

y supón que $g \circ f$ es biyectiva.

1. Demuestra que f es inyectiva.
2. Demuestra que g es sobreyectiva.
3. Decide, mediante prueba o contraejemplo, si f y g deben ser biyectivas.
4. Encuentra una hipótesis adicional sobre f que sea suficiente para forzar que **ambas** funciones sean biyectivas.
5. Encuentra una hipótesis adicional sobre g que también sea suficiente para forzar que **ambas** sean biyectivas.

Soluciones. Las soluciones siguen exactamente el orden de los ejercicios. En los niveles iniciales se privilegia la lectura precisa de definiciones; en los niveles D–G se explicitan los puntos donde una implicación puede fallar y se construyen contraejemplos cuando son necesarios.

Soluciones del nivel A. **Solución A1.**

El dominio es el conjunto declarado a la izquierda de la flecha: $\mathrm{Dom}(f) = [-2, 3]$.

El codominio es $\mathbb{R}$.

Para hallar la imagen observamos que $x^2 \geq 0$ y que, dentro de $[-2, 3]$, el valor mínimo 0 se alcanza en $x = 0$, mientras que el mayor cuadrado es $3^2 = 9$. Todo valor entre 0 y 9 aparece: si $0 \leq y \leq 9$, entonces $\sqrt{y} \in [0, 3] \subseteq [-2, 3]$ y $f(\sqrt{y}) = y$.

Por tanto, $\mathrm{Im}(f) = [0, 9]$.

Imagen y codominio no coinciden, pues $[0, 9] \subsetneq \mathbb{R}$.

Solución A2.

La relación no define una función $A \to B$. Las entradas 0 y 1 tienen una salida única, pero la entrada 2 aparece en dos pares:

$$(2,2), \qquad (2,3).$$

Así, para $x = 2$ falla la **unicidad** de la salida. No falla la existencia: cada elemento de A aparece al menos una vez como primera coordenada. Para reparar la relación basta eliminar uno de los dos pares que parten de 2. Por ejemplo, $\{(0,1),(1,2),(2,2)\}$ define una función $A \to B$. También habría servido conservar $(2,3)$ y eliminar $(2,2)$.

Solución A3.

No se cumple $f = g$ según la convención del libro. Aunque tienen el mismo dominio y satisfacen $f(x) = g(x) = x^2$ para todo $x \in \mathbb{R}$, sus codominios son distintos:

$$\mathrm{Cod}(f) = \mathbb{R}, \qquad \mathrm{Cod}(g) = [0,\infty).$$

El codominio forma parte de los datos de la función. Sin embargo, sus imágenes sí coinciden: $\mathrm{Im}(f) = \mathrm{Im}(g) = [0,\infty)$. Esto muestra nuevamente que imagen y codominio son conceptos diferentes.

Solución A4.

Para la imagen de $[-2,1]$, los cuadrados van desde 0 hasta 4. El valor 4 se alcanza en $x = -2$, y 0 se alcanza en $x = 0$. Por tanto, $q([-2,1]) = [0,4]$. Ahora buscamos los x tales que $1 \le x^2 < 4$. La primera desigualdad equivale a $|x| \ge 1$ y la segunda a $|x| < 2$. Combinándolas, $1 \le |x| < 2$.

Por consiguiente, $q^{-1}([1,4)) = (-2,-1] \cup [1,2)$.

No hemos utilizado una función inversa: hemos calculado el conjunto de entradas cuya imagen pertenece a $[1,4)$.

Solución A5.

Los valores que aparecen en la segunda fila son $3,1,1,4$, así que $\mathrm{Im}(p) = \{1,3,4\}$. La función no es inyectiva, porque $p(-1) = p(0) = 1$ con $-1 \neq 0$. Tampoco es sobreyectiva sobre el codominio $\{1,2,3,4\}$, porque el valor 2 no es alcanzado. En símbolos, $2 \notin \mathrm{Im}(p)$.

Solución A6.

La restricción conserva la misma regla, pero solo admite entradas no negativas: $f|_{[0,\infty)}(x) = x^2$.

Su imagen es $[0,\infty)$.

Es inyectiva. En efecto, si $x, y \ge 0$ y $x^2 = y^2$, entonces $(x-y)(x+y) = 0$. Como $x + y \ge 0$, la única posibilidad compatible con $x, y \ge 0$ y cuadrados iguales es $x = y$; equivalentemente, el cuadrado es estrictamente creciente en $[0,\infty)$. No es sobreyectiva como función hacia $\mathbb{R}$, porque ningún número negativo es alcanzado. Si cambiamos el codominio a $[0,\infty)$, obtenemos

$$[0,\infty) \to [0,\infty), \qquad x \mapsto x^2,$$

que es biyectiva.

Solución A7.

La declaración significa: para cada $n \in \{0,1,2,3\}$ existe exactamente un número real y tal que $y = \frac{n}{n+1}$, y ese número se denota por $a(n)$.

Los cuatro valores son

$$a(0) = 0, \qquad a(1) = \frac{1}{2}, \qquad a(2) = \frac{2}{3}, \qquad a(3) = \frac{3}{4}.$$

Así,

$$\mathrm{Im}(a) = \left\{0, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}\right\}.$$

No es necesario que una función alcance todo su codominio. Eso sería la condición adicional de sobreyectividad. Para ser función basta que cada entrada del dominio tenga exactamente una salida perteneciente al codominio.

Soluciones del nivel B. **Solución B1.**

Por definición, $(g \circ f)(x) = g(f(x)) = (2x-1)^2$.

Por otro lado, $(f \circ g)(x) = f(g(x)) = 2x^2 - 1$.

No son la misma función. Por ejemplo, en $x = 1$,

$$(g \circ f)(1) = 1, \qquad (f \circ g)(1) = 1,$$

pero este valor no basta para distinguirlas. En $x = 0$,

$$(g \circ f)(0) = 1, \qquad (f \circ g)(0) = -1.$$

Luego $g \circ f \neq f \circ g$.

Solución B2.

La función u tiene imagen $[0, \infty)$, pero v no admite la entrada 2. Como $u(3) = \sqrt{4} = 2$, no se cumple $u([-1,\infty)) \subseteq \mathrm{Dom}(v)$. Por tanto, con las funciones tal como fueron declaradas, la composición no está definida sobre todo el dominio de u. Debemos excluir exactamente las entradas que u envía a 2. Resolver $\sqrt{x+1} = 2$ da $x = 3$. Así, el mayor conjunto posible es $E = [-1,\infty) \setminus \{3\}$.

En ese dominio, $(v \circ (u|_E))(x) = \frac{1}{\sqrt{x+1}-2}$.

La exclusión de $x = 3$ no es un detalle algebraico posterior: es la condición que hace posible la composición como función.

Solución B3.

Como $f(x) = 3x - 2$ es estrictamente creciente, $f([1,4)) = [f(1), f(4)) = [1, 10)$.

Para la preimagen resolvemos $-5 < 3x - 2 \leq 4$.

Sumando 2, $-3 < 3x \leq 6$, y dividiendo por $3 > 0$, $-1 < x \leq 2$.

Por tanto, $f^{-1}((-5,4]) = (-1,2]$.

Solución B4.

Buscamos $4 < x^2 \leq 9$.

Esto equivale a $2 < |x| \leq 3$.

Por tanto, $q^{-1}((4,9]) = [-3,-2) \cup (2,3]$.

Aparecen dos intervalos porque la función cuadrado toma el mismo valor positivo en dos puntos simétricos x y $-x$. La preimagen registra ambas ramas.

Solución B5.

En $[0,\infty)$ la función $x \mapsto x^2$ es inyectiva. Sin embargo, como el codominio es todo $\mathbb{R}$, no es sobreyectiva: ningún número negativo pertenece a su imagen. Si reemplazamos el codominio por $[0,\infty)$, obtenemos

$$\widetilde{r}\colon [0,\infty) \to [0,\infty), \qquad \widetilde{r}(x) = x^2.$$

Esta función es biyectiva. Su inversa asigna a cada $y \geq 0$ la única raíz cuadrada no negativa: $\widetilde{r}^{-1}(y) = \sqrt{y}$.

Solución B6.

Los valores pedidos son

$$p(-2) = -1, \qquad p(0) = 0, \qquad p(2) = 4.$$

Para $x < 0$, la expresión $x + 1$ toma todos los valores de $(-\infty, 1)$. Para $x \geq 0$, la expresión x^2 toma todos los valores de $[0,\infty)$. Por tanto, $\mathrm{Im}(p) = (-\infty,1) \cup [0,\infty) = \mathbb{R}$.

Así, p es sobreyectiva.

No es inyectiva. Por ejemplo, $p(-1) = 0 = p(0)$, aunque $-1 \neq 0$.

Solución B7.

Si $(a,b) \in \Gamma_f$, entonces $b = f(a)$.

Para obtener un punto de la gráfica de h, necesitamos elegir x de modo que $x - 3 = a$, es decir, $x = a + 3$.

Entonces $h(a+3) = 2f(a) - 1 = 2b - 1$.

Por tanto, $(a,b) \longmapsto (a+3, 2b-1)$.

Geométricamente:

1. $x \mapsto x - 3$ desplaza la gráfica 3 unidades hacia la derecha;
2. multiplicar la salida por 2 duplica las alturas;
3. restar 1 desplaza el resultado una unidad hacia abajo.

La transformación se ha deducido de la fórmula, no de una regla memorizada.

Soluciones del nivel C. **Solución C1.**

Sean $x_1, x_2 \in \mathbb{R}$ y supongamos $f(x_1) = f(x_2)$.

Entonces $ax_1 + b = ax_2 + b$.

Restando b, $ax_1 = ax_2$.

Como $a \neq 0$, podemos dividir por a y obtener $x_1 = x_2$. Por definición, f es inyectiva. La hipótesis $a \neq 0$ es esencial: si $a = 0$, la función sería constante.

Solución C2.

Para la inyectividad, supongamos $r(x_1) = r(x_2)$.

Entonces $\frac{1}{x_1} = \frac{1}{x_2}$.

Como $x_1, x_2 > 0$, ambos son no nulos y podemos multiplicar por $x_1 x_2$, obteniendo $x_2 = x_1$.

Luego r es inyectiva.

Para la sobreyectividad, tomemos un $y \in (0, \infty)$ arbitrario. Elegimos $x = \frac{1}{y}$.

Como $y > 0$, también $x > 0$, y $r(x) = \frac{1}{1/y} = y$.

Así, todo elemento del codominio tiene antecedente. Por tanto, r es biyectiva. Además, la misma fórmula deshace la asignación: $r^{-1}(y) = \frac{1}{y}$.

En particular, $r^{-1} = r$.

Solución C3.

Sean $x_1, x_2 \geq 0$ y supongamos $x_1^2 + 1 = x_2^2 + 1$.

Entonces $x_1^2 = x_2^2$.

Como ambos números son no negativos, se sigue $x_1 = x_2$. Por tanto, f es inyectiva. Para la sobreyectividad, sea $y \in [1, \infty)$. Entonces $y - 1 \geq 0$, y podemos definir $x = \sqrt{y-1}$.

Este x pertenece a $[0, \infty)$ y $f(x) = (\sqrt{y-1})^2 + 1 = y$.

Por tanto, f es sobreyectiva y, en consecuencia, biyectiva. Ahora sí podemos definir la inversa:

$$f^{-1}(y) = \sqrt{y-1}, \qquad y \geq 1.$$

Solución C4.

Tomemos $x_1, x_2 \in A$ y supongamos $(g \circ f)(x_1) = (g \circ f)(x_2)$.

Por definición de composición, $g(f(x_1)) = g(f(x_2))$.

Como g es inyectiva, $f(x_1) = f(x_2)$.

Y como f es inyectiva, $x_1 = x_2$.

Por tanto, $g \circ f$ es inyectiva.

El argumento usa las dos hipótesis en un orden determinado: primero cancelamos g y después cancelamos f.

Solución C5.

Sea $x \in A$. Entonces

$$\begin{aligned} x \in f^{-1}(U \cap V) &\iff f(x) \in U \cap V \\ &\iff f(x) \in U \text{ y } f(x) \in V \\ &\iff x \in f^{-1}(U) \text{ y } x \in f^{-1}(V) \\ &\iff x \in f^{-1}(U) \cap f^{-1}(V). \end{aligned}$$

Como la equivalencia vale para todo $x \in A$, $f^{-1}(U \cap V) = f^{-1}(U) \cap f^{-1}(V)$.

Para el complemento relativo a B,

$$\begin{aligned} x \in f^{-1}(B \setminus U) &\iff f(x) \in B \setminus U \\ &\iff f(x) \notin U \\ &\iff x \notin f^{-1}(U) \\ &\iff x \in A \setminus f^{-1}(U). \end{aligned}$$

Por tanto, $f^{-1}(B \setminus U) = A \setminus f^{-1}(U)$.

No fue necesario suponer inyectividad ni sobreyectividad.

Solución C6.

Sean $0 \leq x < y$. Entonces $y^2 - x^2 = (y - x)(y + x)$.

Aquí $y - x > 0$ y $y + x > 0$, por lo que $y^2 - x^2 > 0$.

Así, $x^2 < y^2$, y q es estrictamente creciente en $[0, \infty)$.

Ahora sean $x < y \leq 0$. De nuevo, $y^2 - x^2 = (y - x)(y + x)$.

Tenemos $y - x > 0$, pero $y + x < 0$ porque $x < y \leq 0$ implica $x < 0$ y, por tanto, $x + y < 0$. Luego $y^2 - x^2 < 0$, de modo que $y^2 < x^2$. Por tanto, q es estrictamente decreciente en $(-\infty, 0]$. Toda función estrictamente monótona es inyectiva. Así, las dos restricciones son inyectivas.

Solución C7.

La función afín tiene pendiente $2 \neq 0$, así que es inyectiva. Para cualquier $y \in \mathbb{R}$, la ecuación $y = 2x - 3$ tiene la solución $x = \frac{y+3}{2}$, que es real. Por tanto, f es sobreyectiva y, en consecuencia, biyectiva. Su inversa es $f^{-1}(y) = \frac{y+3}{2}$.

Ahora, $(x, y) \in \Gamma_f$ significa exactamente $y = f(x)$.

Como f es biyectiva, esta igualdad equivale a $x = f^{-1}(y)$, que a su vez significa $(y, x) \in \Gamma_{f^{-1}}$.

Por tanto, $(x, y) \in \Gamma_f \iff (y, x) \in \Gamma_{f^{-1}}$.

Intercambiar las coordenadas (x, y) por (y, x) corresponde geométricamente a reflejar respecto de la recta $y = x$. Por eso las gráficas de una función biyectiva y de su inversa son simétricas respecto de esa recta.

Soluciones del nivel D. **Solución D1.**

Supongamos que $g \circ f$ es inyectiva. Si $f(x_1) = f(x_2)$, entonces, aplicando g a ambos lados, $g(f(x_1)) = g(f(x_2))$, es decir, $(g \circ f)(x_1) = (g \circ f)(x_2)$. Como la composición es inyectiva, $x_1 = x_2$.

Por tanto, f es inyectiva.

En cambio, g no tiene por qué ser inyectiva fuera de la imagen de f. Consideremos

$$A = \{0, 1\}, \quad B = \{0, 1, 2\}, \quad C = \{0, 1\},$$

con

$$f(0) = 0, \qquad f(1) = 1,$$

y

$$g(0) = 0, \qquad g(1) = 1, \qquad g(2) = 1.$$

Entonces

$$(g \circ f)(0) = 0, \qquad (g \circ f)(1) = 1,$$

así que $g \circ f$ es inyectiva. Sin embargo, $g(1) = g(2) = 1$ con $1 \neq 2$, de modo que g no es inyectiva.

Solución D2.

Sea $c \in C$. Como $g \circ f$ es sobreyectiva, existe $a \in A$ tal que $g(f(a)) = c$. Si llamamos $b = f(a) \in B$, entonces $g(b) = c$. Así, todo $c \in C$ tiene antecedente bajo g, por lo que g es sobreyectiva. Pero f no tiene por qué ser sobreyectiva sobre todo B. Tomemos

$$A = \{0, 1\}, \quad B = \{0, 1, 2\}, \quad C = \{0, 1\},$$

con

$$f(0) = 0, \qquad f(1) = 1,$$

y

$$g(0) = 0, \qquad g(1) = 1, \qquad g(2) = 0.$$

La composición satisface

$$(g \circ f)(0) = 0, \qquad (g \circ f)(1) = 1,$$

así que es sobreyectiva sobre C. Sin embargo, $2 \in B$ no pertenece a la imagen de f, de modo que f no es sobreyectiva.

Solución D3.

Ninguna de las cuatro funciones es igual a otra según la convención del libro. Aunque todas usan la regla $x \mapsto x^2$, al menos el dominio o el codominio difiere en cada par.

La clasificación es:

- $p\colon \mathbb{R} \to \mathbb{R}$: no es inyectiva y no es sobreyectiva;
- $q\colon \mathbb{R} \to [0, \infty)$: es sobreyectiva, pero no inyectiva;
- $r\colon [0, \infty) \to \mathbb{R}$: es inyectiva, pero no sobreyectiva;
- $s\colon [0, \infty) \to [0, \infty)$: es biyectiva.

Cambiar el **dominio** puede eliminar pares de entradas distintas que producían la misma salida; por eso puede modificar la inyectividad. Cambiar el **codominio** puede hacer que la imagen coincida o deje de coincidir con el conjunto de llegada; por eso puede modificar la sobreyectividad. La regla puntual puede permanecer idéntica mientras cambian propiedades estructurales de la función.

Solución D4.

Para

$$f\colon \mathbb{R} \to [0, \infty), \qquad f(x) = x^2,$$

la función no es inyectiva, pues $f(1) = f(-1) = 1$. Por tanto, no puede poseer una función inversa definida sobre todo $[0, \infty)$.

El paso

$$y = x^2 \quad \Longrightarrow \quad x = \sqrt{y}$$

ha descartado silenciosamente la otra solución $x = -\sqrt{y}$ cuando $y > 0$. Resolver la ecuación produce posibles antecedentes; para construir una función inversa necesitamos **exactamente uno** para cada y del codominio. Podemos reparar la situación restringiendo el dominio a $[0, \infty)$:

$$f_+ \colon [0,\infty) \to [0,\infty), \qquad f_+(x) = x^2,$$

con $f_+^{-1}(y) = \sqrt{y}$.

O podemos restringirlo a $(-\infty, 0]$:

$$f_- \colon (-\infty,0] \to [0,\infty), \qquad f_-(x) = x^2,$$

con $f_-^{-1}(y) = -\sqrt{y}$.

No hay contradicción entre las dos inversas: f_+ y f_- son funciones distintas.

Solución D5.

Sea $y \in f(E \cap F)$. Entonces existe $x \in E \cap F$ tal que $y = f(x)$.

Como $x \in E$ y $x \in F$, tenemos

$$y \in f(E) \qquad \text{y} \qquad y \in f(F).$$

Por tanto, $y \in f(E) \cap f(F)$, y queda demostrada la inclusión $f(E \cap F) \subseteq f(E) \cap f(F)$. Para la inclusión inversa, si $y \in f(E) \cap f(F)$, sabemos que existen $e \in E$ y $r \in F$ tales que $f(e) = y = f(r)$. El problema es que no sabemos todavía que $e = r$. Sin esa igualdad no podemos concluir que exista un mismo punto perteneciente a $E \cap F$ que produzca y. La hipótesis suficiente es que f sea inyectiva. Entonces $f(e) = f(r) \Longrightarrow e = r$, y ese punto común pertenece a $E \cap F$. En consecuencia, bajo inyectividad, $f(E \cap F) = f(E) \cap f(F)$.

Solución D6.

En $x = 0$ se activan las dos condiciones. La primera rama asigna $0 + 1 = 1$, mientras que la segunda asigna $-0 = 0$. Por tanto, una misma entrada recibe dos salidas diferentes. Falla la unicidad y la regla no define una función.

Una reparación es

$$h_1(x) = \begin{cases} x + 1, & x \geq 0, \\ -x, & x < 0. \end{cases}$$

Otra es

$$h_2(x) = \begin{cases} x + 1, & x > 0, \\ -x, & x \leq 0. \end{cases}$$

Ambas definen funciones porque los tramos cubren $\mathbb{R}$ sin superponerse. No son la misma función: en $x = 0$,

$$h_1(0) = 1, \qquad h_2(0) = 0.$$

Así, reparar una ambigüedad exige tomar una decisión matemática; no hay una única reparación automática.

Soluciones del nivel E. **Solución E1.**

Un ejemplo es

$$f\colon \mathbb{R}\setminus\{0\}\to\mathbb{R},\qquad f(x)=\frac{1}{x},$$

y

$$g\colon (0,\infty)\to(0,\infty),\qquad g(x)=\frac{1}{x}.$$

Utilizan la misma fórmula, pero sus dominios y codominios son distintos. En particular, $-1\in\operatorname{Dom}(f)$ mientras que $-1\notin\operatorname{Dom}(g)$.

Por tanto, $f\neq g$.

Solución E2.

Para una función inyectiva no sobreyectiva podemos tomar

$$f\colon \mathbb{N}\to\mathbb{N},\qquad f(n)=n+1.$$

Es inyectiva porque $n_1+1=n_2+1\Longrightarrow n_1=n_2$.

No es sobreyectiva porque $0\in\mathbb{N}$ no pertenece a su imagen. Para una función sobreyectiva no inyectiva podemos tomar

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

Es sobreyectiva porque cada $y\geq 0$ es el cuadrado de $\sqrt{y}$, pero no es inyectiva porque $g(1)=g(-1)=1$.

Solución E3.

Tomemos

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

ambas como funciones $\mathbb{R}\to\mathbb{R}$.

Entonces $(g\circ f)(x)=(x+1)^2$, mientras que $(f\circ g)(x)=x^2+1$.

En $x=1$,

$$(g\circ f)(1)=4,\qquad (f\circ g)(1)=2.$$

Por tanto, $g\circ f\neq f\circ g$.

Solución E4.

Tomemos

$$f\colon \mathbb{R}\to[0,\infty),\qquad f(x)=x^2,$$

y

$$E=\{-1\},\qquad F=\{1\}.$$

Como $E\cap F=\varnothing$, $f(E\cap F)=\varnothing$.

Pero $f(E)=\{1\}=f(F)$, de modo que $f(E)\cap f(F)=\{1\}$.

Por tanto,

$$f(E \cap F) = \varnothing \subsetneq \{1\} = f(E) \cap f(F).$$

La inclusión puede ser estricta porque f no es inyectiva: los puntos distintos -1 y 1 comparten la misma imagen.

Solución E5.
Definamos

$$m(x) = \begin{cases} 0, & x \leq 0, \\ x, & x > 0. \end{cases}$$

La función es no decreciente. En efecto, si $x \leq y$:

- si $y \leq 0$, entonces $m(x) = m(y) = 0$;
- si $x \leq 0 < y$, entonces $m(x) = 0 < y = m(y)$;
- si $0 < x \leq y$, entonces $m(x) = x \leq y = m(y)$.

Sin embargo, no es inyectiva. Por ejemplo, $m(-2) = m(-1) = 0$ con $-2 \neq -1$. La monotonía **estricta** implica inyectividad; la monotonía no estricta no.

Soluciones del nivel F. **Solución F1.**
La gráfica de una función contiene exactamente los pares $(x, f(x))$ con x en el dominio. Por tanto, las primeras coordenadas de los pares de Γ_f reconstruyen $A = \{-2, 0, 3\}$.

Los valores son

$$f(-2) = 1, \qquad f(0) = 1, \qquad f(3) = 4.$$

Así, $\mathrm{Im}(f) = \{1, 4\}$.

La función no es inyectiva porque $f(-2) = f(0) = 1$. Tampoco es sobreyectiva sobre $B = \{1, 2, 4\}$, porque 2 no es alcanzado. Si solo conociéramos Γ_f, podríamos reconstruir dominio y valores, pero no el codominio declarado de manera única. Por ejemplo, la misma gráfica sería compatible con codominio $\{1, 4\}$, con $\{1, 2, 4\}$ o con un conjunto mayor que contuviera la imagen. Bajo la convención del libro, esas declaraciones corresponden a funciones distintas.

Solución F2.
Sea $x \in A$.

Para la unión,

$$\begin{aligned} x \in f^{-1}(U \cup V) &\iff f(x) \in U \cup V \\ &\iff f(x) \in U \text{ o } f(x) \in V \\ &\iff x \in f^{-1}(U) \text{ o } x \in f^{-1}(V) \\ &\iff x \in f^{-1}(U) \cup f^{-1}(V). \end{aligned}$$

Por tanto, $f^{-1}(U \cup V) = f^{-1}(U) \cup f^{-1}(V)$.

Para la intersección,

$$\begin{aligned} x \in f^{-1}(U \cap V) &\iff f(x) \in U \cap V \\ &\iff f(x) \in U \text{ y } f(x) \in V \\ &\iff x \in f^{-1}(U) \text{ y } x \in f^{-1}(V) \\ &\iff x \in f^{-1}(U) \cap f^{-1}(V). \end{aligned}$$

Así, $f^{-1}(U \cap V) = f^{-1}(U) \cap f^{-1}(V)$.

Para el complemento,

$$\begin{aligned} x \in f^{-1}(B \setminus U) &\iff f(x) \notin U \\ &\iff x \notin f^{-1}(U) \\ &\iff x \in A \setminus f^{-1}(U). \end{aligned}$$

Por tanto, $f^{-1}(B \setminus U) = A \setminus f^{-1}(U)$.

Todas las equivalencias dependen solamente de la definición de preimagen y de las operaciones de conjuntos. No fue necesario comparar entradas distintas ni exigir que todos los puntos del codominio se alcancen. Por eso no se requiere inyectividad ni sobreyectividad.

Solución F3.

La primera elección es

$$q_+ \colon [0,\infty) \to [0,\infty), \qquad q_+(x) = x^2.$$

Es inyectiva porque el cuadrado es estrictamente creciente en $[0,\infty)$. Es sobreyectiva porque, dado $y \geq 0$, existe la raíz no negativa $\sqrt{y}$ y $q_+(\sqrt{y}) = y$.

Por tanto, es biyectiva y $q_+^{-1}(y) = \sqrt{y}$.

La segunda elección es

$$q_- \colon (-\infty, 0] \to [0,\infty), \qquad q_-(x) = x^2.$$

Es inyectiva porque el cuadrado es estrictamente decreciente en $(-\infty, 0]$. Para todo $y \geq 0$, el número $-\sqrt{y} \leq 0$ satisface $q_-(-\sqrt{y}) = y$, así que q_- es sobreyectiva. Por tanto, es biyectiva y $q_-^{-1}(y) = -\sqrt{y}$. La misma regla de asignación produce dos biyectividades distintas al cambiar el dominio.

Solución F4.

Fijemos $y \in B$ y estudiemos $f(x) = y$.

La afirmación «para cada $y \in B$ existe al menos un $x \in A$ con $f(x) = y$» es exactamente la definición de **sobreyectividad**. La afirmación «para cada $y \in B$ existe a lo sumo un $x \in A$ con $f(x) = y$» equivale a **inyectividad**. En efecto, decir que una salida no puede tener dos antecedentes distintos es otra forma de decir $f(x_1) = f(x_2) \Longrightarrow x_1 = x_2$. Por tanto, para cada $y \in B$ existe **exactamente un** antecedente si y solo si f es simultáneamente sobreyectiva e inyectiva, es decir, biyectiva. Cuando esto ocurre, definimos $f^{-1}(y) = x$, donde x es el único elemento de A que satisface $f(x) = y$. La sobreyectividad garantiza la existencia de $f^{-1}(y)$; la inyectividad garantiza su unicidad.

Solución F5.

Tenemos $q(x) = -2(x-3)^2 + 1$.

Con las funciones dadas, $T_{-3}(x) = x - 3$,

$$s(T_{-3}(x)) = (x-3)^2,$$

$$M_{-2}(s(T_{-3}(x))) = -2(x-3)^2,$$

y finalmente

$$T_1(M_{-2}(s(T_{-3}(x)))) = -2(x-3)^2 + 1.$$

Por tanto, $q = T_1 \circ M_{-2} \circ s \circ T_{-3}$.

Leída desde la entrada hacia la salida, la composición indica:

1. trasladar la gráfica de s tres unidades hacia la derecha;
2. multiplicar las alturas por -2, lo que combina reflexión respecto del eje horizontal y dilatación vertical por factor 2;
3. trasladar una unidad hacia arriba.

El vértice $(0,0)$ de s pasa primero a $(3,0)$ y después a $(3,1)$. Ese es el vértice de la parábola descrita por q.

Soluciones del nivel G. **Solución G1.**

Supongamos primero que f es inyectiva. En §2.6 ya sabemos que para cualesquiera $E, F \subseteq A$ siempre se cumple $f(E \cap F) \subseteq f(E) \cap f(F)$. Para la inclusión inversa, sea $y \in f(E) \cap f(F)$.

Entonces existen $e \in E$ y $r \in F$ tales que $f(e) = y = f(r)$.

Como f es inyectiva, $e = r$.

Por tanto, ese mismo punto pertenece a $E \cap F$, y $y \in f(E \cap F)$.

Así, $f(E \cap F) = f(E) \cap f(F)$.

Recíprocamente, supongamos que para todos $E, F \subseteq A$ se cumple la igualdad anterior. Queremos recuperar la inyectividad.

Sean $x_1, x_2 \in A$ y supongamos $f(x_1) = f(x_2)$.

Tomemos

$$E = \{x_1\}, \qquad F = \{x_2\}.$$

Si $x_1 \neq x_2$, entonces $E \cap F = \varnothing$, de modo que $f(E \cap F) = \varnothing$.

Pero, como $f(x_1) = f(x_2)$, $f(E) \cap f(F) = \{f(x_1)\} \neq \varnothing$.

Esto contradice la igualdad supuesta para todos E, F. Por tanto, necesariamente $x_1 = x_2$.

Luego f es inyectiva.

Hemos demostrado la equivalencia.

Solución G2.

Supongamos primero $g \circ f = \mathrm{id}_A$.

Para demostrar que f es inyectiva, sean $x_1, x_2 \in A$ y supongamos $f(x_1) = f(x_2)$.

Aplicando g, $g(f(x_1)) = g(f(x_2))$.

Como $g \circ f = \mathrm{id}_A$, $x_1 = x_2$.

Así, f es inyectiva.

Para demostrar que g es sobreyectiva sobre A, tomemos cualquier $a \in A$. El elemento $f(a) \in B$ satisface $g(f(a)) = a$. Por tanto, cada a tiene antecedente bajo g, y g es sobreyectiva.

Ahora supongamos $f \circ g = \mathrm{id}_B$.

Para cada $b \in B$, el elemento $g(b) \in A$ satisface $f(g(b)) = b$, de modo que f es sobreyectiva. Para la inyectividad de g, si $g(b_1) = g(b_2)$, aplicamos f y obtenemos $f(g(b_1)) = f(g(b_2))$, de donde $b_1 = b_2$.

Así, g es inyectiva.

Las conversas generales fallan. Consideremos

$$A = \{0, 1\}, \qquad B = \{0, 1, 2\}.$$

Definamos

$$f(0) = 0, \qquad f(1) = 1,$$

y

$$g(0) = 0, \qquad g(1) = 0, \qquad g(2) = 1.$$

Aquí f es inyectiva y g es sobreyectiva sobre A, pero $(g \circ f)(1) = g(1) = 0 \neq 1$, por lo que $g \circ f \neq \mathrm{id}_A$.

De manera análoga, tomemos

$$A = \{0, 1, 2\}, \qquad B = \{0, 1\},$$

y definamos

$$f(0) = 0, \qquad f(1) = 0, \qquad f(2) = 1,$$

$$g(0) = 0, \qquad g(1) = 1.$$

Entonces f es sobreyectiva y g es inyectiva, pero $(f \circ g)(1) = f(1) = 0 \neq 1$, de modo que $f \circ g \neq \mathrm{id}_B$. La existencia de una inversa lateral es, por tanto, una condición más fuerte que la mera combinación de inyectividad y sobreyectividad de funciones distintas.

Solución G3.

Supongamos que $g \circ f \colon A \to C$ es biyectiva.

Como toda función biyectiva es inyectiva, $g \circ f$ es inyectiva. Por el argumento del Ejercicio D1, esto obliga a que f sea inyectiva. Como toda función biyectiva es sobreyectiva, $g \circ f$ es sobreyectiva. Por el argumento del Ejercicio D2, esto obliga a que g sea sobreyectiva. Sin embargo, no se sigue en general que f sea sobreyectiva ni que g sea inyectiva. Un solo ejemplo muestra simultáneamente ambas fallas. Tomemos

$$A = C = \{0\}, \qquad B = \{0, 1\},$$

con $f(0) = 0$, y

$$g(0) = 0, \qquad g(1) = 0.$$

La composición $g \circ f \colon \{0\} \to \{0\}$ es biyectiva. Pero f no es sobreyectiva sobre B, porque 1 no es alcanzado, y g no es inyectiva, porque $g(0) = g(1)$. Ahora añadamos la hipótesis de que f es sobreyectiva. Ya sabemos que f es inyectiva, así que f es biyectiva. También sabemos que g es sobreyectiva. Falta probar que g es inyectiva.

Sean $b_1, b_2 \in B$ y supongamos $g(b_1) = g(b_2)$.

Como f es sobreyectiva, existen $a_1, a_2 \in A$ tales que

$$f(a_1) = b_1, \qquad f(a_2) = b_2.$$

Entonces

$$(g \circ f)(a_1) = g(b_1) = g(b_2) = (g \circ f)(a_2).$$

La composición es inyectiva, de modo que $a_1 = a_2$, y por tanto

$$b_1 = f(a_1) = f(a_2) = b_2.$$

Así, g es inyectiva y, como ya era sobreyectiva, es biyectiva. Por consiguiente, una hipótesis adicional suficiente es:

$$\boxed{f \text{ sobreyectiva.}}$$

Existe una condición dual igualmente suficiente. Supongamos ahora que g es inyectiva. Ya sabemos que g es sobreyectiva, así que g es biyectiva. Falta probar que f es sobreyectiva. Sea $b \in B$. Entonces $g(b) \in C$. Como $g \circ f$ es sobreyectiva, existe $a \in A$ tal que $g(f(a)) = g(b)$.

Por la inyectividad de g, $f(a) = b$.

Así, todo $b \in B$ tiene antecedente bajo f, por lo que f es sobreyectiva. Como ya sabíamos que f es inyectiva, también f es biyectiva. Por tanto, otra hipótesis adicional suficiente es:

$$\boxed{g \text{ inyectiva.}}$$

El resultado completo puede resumirse así:

$$\boxed{\begin{array}{c} g \circ f \text{ biyectiva} \Longrightarrow f \text{ inyectiva y } g \text{ sobreyectiva,} \\ \text{pero no necesariamente } f, g \text{ biyectivas;} \\ f \text{ sobreyectiva o } g \text{ inyectiva} \Longrightarrow f \text{ y } g \text{ biyectivas.} \end{array}}$$

Cierre del capítulo. Los ejemplos de este capítulo muestran una limitación de las expresiones aisladas:

$$\boxed{\text{una fórmula no es todavía una función.}}$$

El dominio decide qué entradas podemos comparar; el codominio fija el conjunto de llegada respecto del cual preguntamos por la sobreyectividad. La asignación permite estudiar imágenes, componer etapas y, cuando cada salida tiene un único antecedente, construir una inversa. Las propiedades de una función dependen de cómo se reúnen esos datos.

La función del último laboratorio, cuyo dominio es $\mathbb{N}$, conserva esta misma estructura aunque su gráfica sea un conjunto de puntos separados. Estudiar su imagen o demostrar que es estrictamente creciente exige las definiciones ya construidas; describir qué ocurre con sus valores cuando el índice crece requerirá además una noción de aproximación.

Capítulo 3

Sucesiones y la primera noción rigurosa de límite

En los ejemplos periódicos usaremos las funciones triangulares σ y κ definidas en C03, con $\omega = 2$. Sus cotas $|\sigma|, |\kappa| \leq 1$ y los valores $\sigma(\omega/2 + n\omega) = (-1)^n$ proceden directamente de aquella definición por tramos; no se presupone trigonometría.

Una sucesión es una función con índices naturales y valores reales. Podemos calcular tantos términos como permita su regla, pero ninguna lista finita nos dice por sí sola qué ocurrirá con todos los términos posteriores. La estructura ordenada y completa de $\mathbb{R}$ permitirá responder algunas de esas preguntas sin tener que recorrer la lista entera.

La noción de **límite** describe cómo se aproximan los términos a un número cuando el índice crece. Al escribir una sucesión $a_0, a_1, a_2, a_3, \ldots$, la lista no termina. Podemos calcular cien términos, mil o un millón, pero siempre quedan términos posteriores. Por eso una teoría rigurosa del límite necesita responder una pregunta distinta de las que aparecen en un cálculo finito:

¿cómo expresamos matemáticamente lo que ocurre a partir de algún momento y para siempre después?

Como los índices son naturales, podemos separar con facilidad un tramo inicial y conservar todos los términos que lo siguen. Esa parte de la sucesión recibe el nombre de **cola**. Para expresar que los términos se aproximan a un número necesitaremos controlar una cola completa, cualquiera que sea la precisión exigida; empezar por las colas nos permitirá entender qué debe decir la definición de límite.

3.1. Una sucesión es una función, pero el límite vive en la cola

Dos sucesiones que parecen iguales durante muchísimo tiempo. Consideremos primero

$$a_n = \frac{1}{n+1}, \qquad n \in \mathbb{N}.$$

Como en este libro usamos $\mathbb{N} = \{0, 1, 2, \ldots\}$, sus primeros términos son $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \ldots$

A medida que el índice crece, los términos se hacen pequeños. No estamos formulando todavía qué significa rigurosamente «hacerse pequeños», pero la tendencia visual y aritmética resulta familiar.

Construyamos ahora otra sucesión:

$$b_n = \begin{cases} \dfrac{1}{n+1}, & 0 \leq n \leq 10^6, \\ 1, & n > 10^6. \end{cases}$$

Durante todo el tramo $n = 0, 1, 2, \ldots, 10^6$, las dos sucesiones son exactamente iguales: $a_n = b_n$. Esto significa que podemos inspeccionar más de un millón de términos sin detectar diferencia

alguna. Sin embargo, después del índice 10^6 sus comportamientos se separan por completo. La primera sigue obedeciendo $a_n = \frac{1}{n+1}$, y sus términos continúan disminuyendo en tamaño. La segunda, en cambio, queda fijada para siempre en $b_n = 1$. Si nuestro criterio dependiera solo de una muestra finita de términos, por enorme que fuera, no podría distinguir estas dos sucesiones. Podemos hacer que el tramo común tenga diez términos, un millón, un billón o cualquier cantidad finita previamente elegida. Siempre es posible modificar lo que ocurre después.

Este ejemplo contiene la primera dificultad conceptual del capítulo:

> ninguna inspección de una cantidad finita de términos determina por sí sola el comportamiento final de una sucesión.

El ejemplo obliga a dirigir la atención hacia **todos los términos suficientemente tardíos**. Para precisar esa exigencia empezaremos por la función que asigna un término a cada índice.

La variable ya no recorre un intervalo real. En `T1-C03` escribíamos funciones como $f\colon \mathbb{R} \to \mathbb{R}$ o $g\colon [0,\infty) \to \mathbb{R}$. Una sucesión real es también una función, pero ahora el dominio es discreto. En lugar de introducir cualquier número real permitido, introducimos índices naturales.

Por ejemplo,

$$a\colon \mathbb{N} \to \mathbb{R}, \qquad a(n) = \frac{1}{n+1}.$$

La notación tradicional para sucesiones reemplaza habitualmente $a(n)$ por a_n.

Así,

$$a_0 = 1, \qquad a_1 = \frac{1}{2}, \qquad a_2 = \frac{1}{3}, \qquad \dots$$

El subíndice no crea un nuevo tipo de objeto: simplemente indica que estamos evaluando una función en el índice n. La notación conserva, por tanto, la estructura de una función:

$$\boxed{a_n = a(n).}$$

La escritura con subíndices es conveniente, pero la estructura subyacente sigue siendo funcional.

Índice y término no son lo mismo. En $a_n = \frac{1}{n+1}$, el número n es el **índice** y a_n es el **término correspondiente**.

Para $n = 4$ tenemos $a_4 = \frac{1}{5}$.

Aquí 4 y 1/5 desempeñan papeles completamente distintos. El primero pertenece al dominio de la función; el segundo pertenece a su codominio. Conviene hacer esta separación desde el principio porque las afirmaciones sobre sucesiones mezclan con frecuencia ambas variables. Una frase como

> para índices suficientemente grandes, los términos son menores que 0.01

habla simultáneamente de

- **índices** n que deben ser suficientemente grandes;
- **términos** a_n que deben satisfacer una cierta desigualdad.

Al leer una definición de límite tendremos que distinguir el umbral impuesto a los índices de la condición que deben satisfacer los valores correspondientes.

Las sucesiones pueden comenzar después de cero. No toda fórmula natural está definida en $n = 0$. Por ejemplo, $c_n = \frac{1}{n}$ no tiene sentido para $n = 0$. No necesitamos alterar nuestra convención sobre $\mathbb{N}$. Simplemente podemos declarar

$$c\colon \{n \in \mathbb{N} : n \geq 1\} \to \mathbb{R}, \qquad c_n = \frac{1}{n}.$$

En general, será conveniente permitir sucesiones que comiencen en un índice $n_0 \in \mathbb{N}$. El conjunto $\{n \in \mathbb{N} : n \geq n_0\}$ es una cola de $\mathbb{N}$: contiene n_0 y todos los naturales posteriores. Permitir este dominio nos deja estudiar el comportamiento final de expresiones que solo están definidas a partir de cierto índice, sin asignarles valores artificiales en los índices anteriores.

Definición 3.1. **Sucesión real, término, cola y propiedad eventual.** Sea $n_0 \in \mathbb{N}$.

Una **sucesión real a partir de** n_0 es una función

$$a\colon \{n \in \mathbb{N} : n \geq n_0\} \to \mathbb{R}.$$

El valor

$$a_n := a(n)$$

se llama **término de índice** n. La sucesión puede denotarse por

$$(a_n)_{n \geq n_0}.$$

Cuando $n_0 = 0$, escribiremos con frecuencia simplemente

$$(a_n).$$

Si $N \geq n_0$, la **cola de la sucesión a partir de** N es la restricción

$$a|_{\{n \in \mathbb{N} : n \geq N\}},$$

que escribiremos también como

$$(a_n)_{n \geq N}.$$

Diremos que una propiedad $P(n)$ de los términos de la sucesión se cumple **eventualmente** si existe un índice $N \geq n_0$ tal que

$$P(n) \qquad \text{para todo } n \geq N.$$

En símbolos,

$$\exists N \geq n_0 \; \forall n \geq N : \; P(n).$$

La palabra **eventualmente** condensa por tanto una estructura lógica muy concreta:

$$\boxed{\text{eventualmente} \quad \Longleftrightarrow \quad \exists N \; \forall n \geq N.}$$

No significa «muchas veces», «para índices grandes en general» ni «casi siempre» en un sentido informal. Significa que podemos encontrar un umbral después del cual **no aparece ninguna excepción**.

Una sucesión no es solo el conjunto de sus valores. Consideremos $u_n = (-1)^n$.

La sucesión es $1, -1, 1, -1, 1, -1, \ldots$

El conjunto de valores que toma es simplemente $\{-1, 1\}$. Pero la sucesión contiene información que ese conjunto ha perdido: cuál valor aparece en cada índice y en qué orden aparecen los términos. La función $v_n = 1$ cuando n es par y $v_n = -1$ cuando n es impar describe la misma sucesión que $u_n = (-1)^n$. En cambio, una sucesión que vale 1 durante los primeros cien índices y luego vale siempre -1 tiene también como conjunto de valores $\{-1, 1\}$, pero su comportamiento es completamente distinto. Por eso no identificaremos (a_n) con $\{a_n : n \geq n_0\}$. La segunda expresión es un conjunto; la primera es una función indexada. En un conjunto desaparecen el orden y las repeticiones. En una sucesión, ambos forman parte de la información.

Distintas maneras de especificar una sucesión. Como cualquier función, una sucesión puede quedar determinada de distintas maneras. Una fórmula explícita es solo una posibilidad.

Ejemplo 3.1. **Cuatro maneras de especificar sucesiones.**

1. **Mediante una fórmula explícita.**

$$a_n = \frac{(-1)^n}{n+1}, \qquad n \geq 0.$$

   Cada índice determina directamente su término.

2. **Mediante una regla por casos.**

$$b_n = \begin{cases} 0, & n \text{ par}, \\ 1, & n \text{ impar}. \end{cases}$$

   Aquí la paridad del índice decide qué valor se asigna.

3. **Mediante una regla verbal inequívoca.**

   Sea $c_n = 1$ si n es un número primo y $c_n = 0$ en caso contrario. La regla determina un único número real para cada $n \in \mathbb{N}$.

4. **Mediante una recurrencia y datos iniciales.**

$$d_0 = 3, \qquad d_{n+1} = \frac{d_n + 1}{2}.$$

   No disponemos aquí de una fórmula explícita para d_n desde el comienzo, pero los datos determinan sucesivamente

$$d_1 = 2, \qquad d_2 = \frac{3}{2}, \qquad d_3 = \frac{5}{4}, \qquad \ldots$$

En los cuatro casos el objeto final es del mismo tipo: una función cuyo dominio es un conjunto de índices naturales.

Este ejemplo permite corregir otra intuición engañosa: una sucesión no necesita venir acompañada por una fórmula cerrada del tipo $a_n = F(n)$. Una recurrencia puede definirla perfectamente bien si los datos iniciales y la regla determinan de manera única cada término posterior. También debemos tener cuidado con las tablas. Una tabla **infinita** de pares índice–valor podría especificar una sucesión completa, del mismo modo que podría especificar cualquier función. Pero una tabla impresa con, por ejemplo, los valores

n	0	1	2	3	4
a_n	1	2	4	8	16

solo fija cinco términos. Por sí sola no determina qué ocurre en $n = 5, 6, 7, \ldots$. Podemos conjeturar que la regla es $a_n = 2^n$, pero también existen infinitas sucesiones diferentes que comienzan con $1, 2, 4, 8, 16$ y después se comportan de otra manera. Una muestra finita puede sugerir una ley; no la sustituye.

Qué significa mirar una cola. Sea $a_n = \frac{(-1)^n}{n+1}$.

Sus primeros términos son $1, -\frac{1}{2}, \frac{1}{3}, -\frac{1}{4}, \frac{1}{5}, \ldots$

La cola a partir de $N = 4$ es

$$\left(\frac{1}{5}, -\frac{1}{6}, \frac{1}{7}, -\frac{1}{8}, \ldots\right).$$

La cola a partir de $N = 100$ comienza en a_{100} y descarta todos los términos anteriores. En lenguaje de funciones, tomar esta cola consiste en restringir el dominio de a a $\{n \in \mathbb{N} : n \geq N\}$. La idea de cola sirve para separar dos tipos de información:

- lo que ocurre en una cantidad finita de índices iniciales;
- lo que persiste después de algún índice.

El segundo tipo será el relevante para los límites.

Propiedades que llegan para quedarse. Consideremos la afirmación $|a_n| < \frac{1}{100}$ para la sucesión $a_n = \frac{(-1)^n}{n+1}$. Como $|a_n| = \frac{1}{n+1}$, si $n \geq 100$, entonces $n + 1 \geq 101$ y por tanto

$$|a_n| = \frac{1}{n+1} \leq \frac{1}{101} < \frac{1}{100}.$$

Así, la propiedad $|a_n| < \frac{1}{100}$ se cumple eventualmente. Un índice que funciona como umbral es $N = 100$. No afirmamos que la desigualdad se cumpla para todos los índices. De hecho, $|a_0| = 1$. Lo que importa es que, una vez superado cierto umbral, no vuelve a fallar.

Comparemos con otra propiedad: $a_n > 0$.

Esta no se cumple eventualmente, porque para todo N existen índices impares $n \geq N$, y en ellos $a_n < 0$. Tampoco se cumple eventualmente $a_n < 0$, porque siempre reaparecen índices pares con $a_n > 0$. El contraste muestra que «eventualmente» no quiere decir que una afirmación sea frecuente. En esta sucesión hay infinitos términos positivos e infinitos términos negativos, pero ninguna de las dos propiedades de signo llega a mantenerse para siempre.

El orden de los cuantificadores importa. La definición de eventualidad contiene $\exists N\ \forall n \geq N$. Primero debemos encontrar **un solo** umbral N. Después, ese mismo umbral debe controlar **todos** los índices posteriores.

No sería suficiente decir:

> para cada índice n puedo encontrar algún $N \leq n$ para el cual la propiedad se cumple en n.

Eso permitiría cambiar el umbral cada vez y no impondría ningún comportamiento persistente. La estructura correcta tiene una dirección temporal lógica:

elegimos un umbral $\longrightarrow$ controlamos toda la cola posterior.

Para expresar convergencia añadiremos una exigencia a esta estructura: dada una tolerancia, tendremos que encontrar un umbral a partir del cual todos los términos la satisfagan. El umbral podrá cambiar con la tolerancia; una vez elegido, deberá servir para toda la cola.

Un prefijo enorme sigue siendo un prefijo finito. Volvamos ahora a las sucesiones con las que abrimos la sección.

Ejemplo 3.2. **Coincidir durante un millón de índices no determina el comportamiento final.** Sean $a_n = \frac{1}{n+1}$ y

$$b_n = \begin{cases} \dfrac{1}{n+1}, & 0 \le n \le 10^6, \\ 1, & n > 10^6. \end{cases}$$

Para todo índice del intervalo $0 \le n \le 10^6$ tenemos $a_n = b_n$. Sin embargo, consideremos la propiedad

$$P(n): \quad a_n < \frac{1}{2}.$$

Para la sucesión (a_n), esta propiedad se cumple eventualmente. Por ejemplo, si $n \ge 2$, entonces

$$a_n = \frac{1}{n+1} \le \frac{1}{3} < \frac{1}{2}.$$

En cambio, para (b_n) la propiedad $b_n < \frac{1}{2}$ no se cumple eventualmente. En efecto, para todo $n > 10^6$ tenemos $b_n = 1$. Las dos sucesiones comparten un prefijo extraordinariamente largo, pero sus colas satisfacen propiedades eventuales diferentes.

Podríamos reemplazar 10^6 por cualquier número natural M. La conclusión sería la misma. No existe una longitud finita privilegiada a partir de la cual una inspección por sí sola garantice el comportamiento futuro. La dificultad explica por qué una afirmación cuantificada debe sustituir impresiones como

> «miramos muchos términos y parece que…».

La experimentación numérica puede sugerir una conjetura. La definición rigurosa debe controlar una cantidad infinita de índices mediante una afirmación finita.

Cambiar una cantidad finita de términos no altera una propiedad eventual. La noción de cola produce inmediatamente un principio que utilizaremos de manera constante. Supongamos que una propiedad $P(n)$ se cumple eventualmente para una sucesión (a_n). Entonces existe algún N tal que

$$P(n) \qquad \text{para todo } n \ge N.$$

Ahora modifiquemos los términos $a_0, a_1, \dots, a_M$ para algún índice finito M, dejando intactos todos los términos posteriores. Si $M < N$, la cola a partir de N ni siquiera ha sido tocada. Si $M \ge N$, basta comenzar a mirar después de ambas fronteras. Por ejemplo, podemos elegir $K = M + 1$. Entonces, para todo $n \ge K$, los términos modificados y los originales coinciden. Como además $K > N$, la propiedad $P(n)$ sigue cumpliéndose en toda esa cola.

En forma esquemática,

una propiedad eventual depende de alguna cola, no de un prefijo finito.

Modificar el prefijo puede cambiar el primer término, el valor máximo o la validez de una identidad exigida en cada índice. La estabilidad que acabamos de demostrar corresponde a las propiedades formuladas como $\exists N\ \forall n \geq N$ son insensibles a cualquier modificación que afecte solo a una cantidad finita de índices.

«Para índices grandes» no significa que exista un índice mágico. El lenguaje informal «cuando n es grande» puede inducir a pensar que existe un número universal a partir del cual todas las propiedades interesantes se vuelven verdaderas. No es así. Para $a_n = \frac{1}{n+1}$, la propiedad $a_n < \frac{1}{10}$ se cumple, por ejemplo, desde $n = 10$. La propiedad más exigente $a_n < \frac{1}{1000}$ requiere avanzar mucho más. Podemos usar $N = 1000$. Y si exigimos $a_n < \frac{1}{10^9}$, el umbral deberá alejarse todavía más. Por tanto, «eventualmente» siempre debe leerse en relación con **una propiedad concreta**. El umbral que sirve para una condición puede no servir para otra. Esta dependencia será el centro lógico de la definición de límite: cuanto menor sea la tolerancia que exijamos, más lejos podría ser necesario avanzar en la sucesión.

La elección de un umbral tampoco exige que los términos avancen siempre en una misma dirección.

Lo eventual no exige monotonía. La sucesión $a_n = \frac{(-1)^n}{n+1}$ alterna de signo. Sus términos no avanzan todos en una misma dirección: uno es positivo, el siguiente negativo, y así sucesivamente. Sin embargo, como vimos, $|a_n| < \frac{1}{100}$ se cumple eventualmente. Esto muestra que una propiedad de cola puede estabilizarse aunque los términos continúen oscilando en otros aspectos. Más adelante distinguiremos cuidadosamente entre

- acercarse a un valor;
- ser creciente o decreciente;
- estar acotada;
- oscilar.

Por ejemplo, la alternancia de signos puede continuar mientras el valor absoluto de los términos se hace cada vez menor. Una descripción de la oscilación y una afirmación de aproximación pueden ser verdaderas a la vez.

La condición que debe cumplir una cola. Para formular la aproximación conservamos la estructura de una sucesión como función $a\colon \{n \in \mathbb{N} : n \geq n_0\} \to \mathbb{R}$. Sus términos están ordenados por índices. Una cola descarta una cantidad finita de términos iniciales y conserva todos los posteriores. Una propiedad es eventual cuando existe una cola completa en la que esa propiedad se cumple sin excepciones.

El esquema lógico es

$$\boxed{\exists N\ \forall n \geq N.}$$

Eso todavía no define que (a_n) tenga límite L. Pero ya contiene la parte temporal de la definición: **a partir de algún índice, todos los términos posteriores deberán obedecer una condición**. La pregunta siguiente será cuál debe ser esa condición. Si queremos expresar que los términos se acercan a un número L, decir simplemente «a_n está cerca de L» no basta. Necesitamos medir la cercanía y exigir que podamos hacerla tan estricta como queramos. La distancia entre números reales permitirá transformar la intuición

$$\text{«}a_n \text{ se acerca a } L\text{»}$$

en una afirmación cuantificada que pueda demostrarse línea por línea.

3.2. Tan cerca como queramos, desde algún momento: la definición ε–N

La eventualidad expresa que una propiedad se mantiene en toda una cola: $\exists N\ \forall n \geq N$. Quedó pendiente decidir qué propiedad deben satisfacer todos esos términos tardíos si queremos expresar que una sucesión «se acerca» a un número L. La palabra **cerca** todavía es demasiado imprecisa. Dos números pueden parecernos cercanos en una escala y muy alejados en otra. Si estamos aproximando una longitud de varios kilómetros, un error de un metro puede ser pequeño; si estamos midiendo el grosor de una célula, el mismo error sería enorme. Una definición matemática no puede depender de esa apreciación contextual. Necesitamos sustituir «cerca» por una condición numérica que podamos hacer tan exigente como queramos.

De la cercanía a una tolerancia. Sean a_n un término de la sucesión y $L \in \mathbb{R}$ un número que queremos considerar como posible límite. La distancia entre ambos es $|a_n - L|$. Si elegimos una tolerancia $\varepsilon > 0$, la condición $|a_n - L| < \varepsilon$ significa que a_n está a distancia menor que ε de L. Por las equivalencias de valor absoluto estudiadas en `T1-C02`, esto es lo mismo que $L - \varepsilon < a_n < L + \varepsilon$. Así, exigir $|a_n - L| < \varepsilon$ significa pedir que el término a_n pertenezca al intervalo abierto $(L - \varepsilon, L + \varepsilon)$. Si $\varepsilon = 1$, permitimos una franja bastante amplia alrededor de L. Si $\varepsilon = 10^{-3}$, exigimos mucha más precisión. Si $\varepsilon = 10^{-20}$, la exigencia es todavía más estricta.

Pero ninguna tolerancia positiva particular puede definir por sí sola el límite. Una sucesión podría terminar entrando en el intervalo de radio 1 alrededor de L y, sin embargo, permanecer siempre a distancia mayor que 10^{-3} de L. La exigencia debe valer para **cualquier precisión positiva**: una vez fijada, tiene que existir un umbral a partir del cual todos los términos estén a distancia de L menor que esa tolerancia. La expresión «tan cerca como queramos» se convierte así en $\forall \varepsilon > 0$. Y la expresión «desde algún momento y para siempre después» ya fue preparada en §3.1: $\exists N\ \forall n \geq N$. Al unir ambas estructuras aparece la definición fundamental.

Definición 3.2. **Convergencia, límite y divergencia.** Sea

$$(a_n)_{n \geq n_0}$$

una sucesión real y sea $L \in \mathbb{R}$.

Diremos que (a_n) **converge a** L si para todo $\varepsilon > 0$ existe un índice $N \geq n_0$ tal que

$$|a_n - L| < \varepsilon$$

para todo $n \geq N$.

En símbolos,

$$\forall \varepsilon > 0\ \exists N \geq n_0\ \forall n \geq N : |a_n - L| < \varepsilon.$$

En ese caso escribimos

$$a_n \to L,$$

$$\lim_{n \to \infty} a_n = L,$$

o, cuando no haya ambigüedad,

$$(a_n) \to L.$$

El número L se llama **límite** de la sucesión. Diremos que (a_n) **diverge** si no existe ningún número real L al cual converja.

La notación $n \to \infty$ no significa que n llegue a tomar como valor un objeto llamado ∞. El símbolo expresa que estamos estudiando índices arbitrariamente grandes, es decir, colas cada vez más alejadas del comienzo de la sucesión.

Leer la definición de izquierda a derecha. La definición contiene tres cuantificadores encadenados:

$$\boxed{\forall \varepsilon > 0 \quad \exists N \geq n_0 \quad \forall n \geq N.}$$

Su orden es parte esencial del significado. Cada cuantificador determina qué podemos elegir y qué debemos verificar:

1. **La exigencia:** se nos entrega una tolerancia arbitraria $\varepsilon > 0$.
2. **Nuestra respuesta:** debemos encontrar un umbral N que funcione para esa tolerancia.
3. **La verificación:** una vez elegido N, cualquier índice $n \geq N$ debe satisfacer $|a_n - L| < \varepsilon$.

El primer cuantificador es universal. No podemos elegir una tolerancia cómoda y detenernos allí. La afirmación debe funcionar para

$$\varepsilon = 1, \qquad \varepsilon = 10^{-2}, \qquad \varepsilon = 10^{-100},$$

y para cualquier otro real positivo. El segundo cuantificador es existencial. No necesitamos encontrar el menor N posible. Basta producir **algún** umbral que funcione. El tercer cuantificador vuelve a ser universal. Después de N no basta con encontrar muchos términos buenos: deben ser buenos **todos** los términos de la cola. Esta última condición es precisamente la noción de eventualidad de §3.1. Para cada tolerancia $\varepsilon > 0$, la propiedad $|a_n - L| < \varepsilon$ debe cumplirse eventualmente. Por eso podemos resumir la definición diciendo:

$$\boxed{\begin{array}{c} a_n \\ \to L \\ \\ \iff \\ \text{para toda tolerancia positiva, la sucesión está} \\ \text{eventualmente dentro de esa tolerancia alrededor de } L. \end{array}}$$

El umbral puede depender de la tolerancia. Una de las primeras dificultades lógicas consiste en entender la dependencia entre ε y N.

La definición permite que $N = N(\varepsilon)$.

Una tolerancia amplia puede admitir un umbral pequeño; una tolerancia más estricta puede obligarnos a avanzar mucho más. Por ejemplo, para $a_n = \frac{1}{n+1}$ y candidato $L = 0$, la condición $|a_n - 0| < \frac{1}{10}$ se cumple desde $n = 10$.

En cambio, $|a_n - 0| < \frac{1}{10^6}$ requiere un umbral mucho más lejano. No hay ningún problema en que los dos valores de N sean distintos. La definición no exige un único N capaz de responder simultáneamente a todas las tolerancias.

De hecho, si exigiéramos

$$\exists N \; \forall \varepsilon > 0 \; \forall n \geq N : |a_n - L| < \varepsilon,$$

estaríamos diciendo algo mucho más fuerte. Para un $n \geq N$ fijo, el número $|a_n - L|$ tendría que ser menor que **todo** real positivo. Eso solo es posible si $|a_n - L| = 0$, es decir, $a_n = L$. Por tanto, ese orden incorrecto de cuantificadores describiría sucesiones que a partir de cierto índice son exactamente constantes iguales a L. La convergencia es mucho más flexible: permite que los términos nunca lleguen a ser exactamente L, siempre que podamos hacer su distancia a L arbitrariamente pequeña en una cola adecuada.

Cómo se construye una prueba ε–N. La primera vez que vemos una prueba de convergencia puede parecer que el valor de N aparece por inspiración. En realidad, conviene separar dos momentos.

Búsqueda. Partimos de la desigualdad que queremos conseguir, $|a_n - L| < \varepsilon$, y la manipulamos para descubrir qué condición sobre n bastaría para garantizarla.

Demostración. Después escribimos el argumento en el orden lógico de la definición:

1. sea $\varepsilon > 0$ arbitrario;
2. elegimos un N que satisfaga la condición descubierta;
3. tomamos cualquier $n \geq N$;
4. verificamos que $|a_n - L| < \varepsilon$.

El cálculo exploratorio puede ir «hacia atrás»; la prueba final debe ir «hacia adelante».

Ejemplo 3.3. **Primera prueba completa: $1/(n+1) \to 0$.** Consideremos

$$a_n = \frac{1}{n+1}, \qquad n \geq 0.$$

Queremos demostrar directamente desde la definición que

$$\lim_{n\to\infty} \frac{1}{n+1} = 0.$$

Búsqueda del umbral. Necesitamos lograr

$$\left| \frac{1}{n+1} - 0 \right| < \varepsilon.$$

Como $n + 1 > 0$, esto equivale a pedir $\frac{1}{n+1} < \varepsilon$, y sería suficiente garantizar $n + 1 > \frac{1}{\varepsilon}$. La propiedad arquimediana demostrada en `T1-C02` asegura que existe un número natural N tal que $N > \frac{1}{\varepsilon}$. Ese N será una respuesta válida.

Demostración. Sea $\varepsilon > 0$. Por la propiedad arquimediana, elegimos $N \in \mathbb{N}$ tal que $N > \frac{1}{\varepsilon}$.

Si $n \geq N$, entonces

$$n + 1 \geq N + 1 > N > \frac{1}{\varepsilon}.$$

Como todas estas cantidades son positivas, al tomar recíprocos se invierte la desigualdad y obtenemos $\frac{1}{n+1} < \varepsilon$.

Por tanto,

$$\left| \frac{1}{n+1} - 0 \right| < \varepsilon$$

para todo $n \geq N$. Como $\varepsilon > 0$ era arbitrario, concluimos que $\frac{1}{n+1} \to 0$.

Vale la pena leer la prueba señalando dónde aparece cada cuantificador:

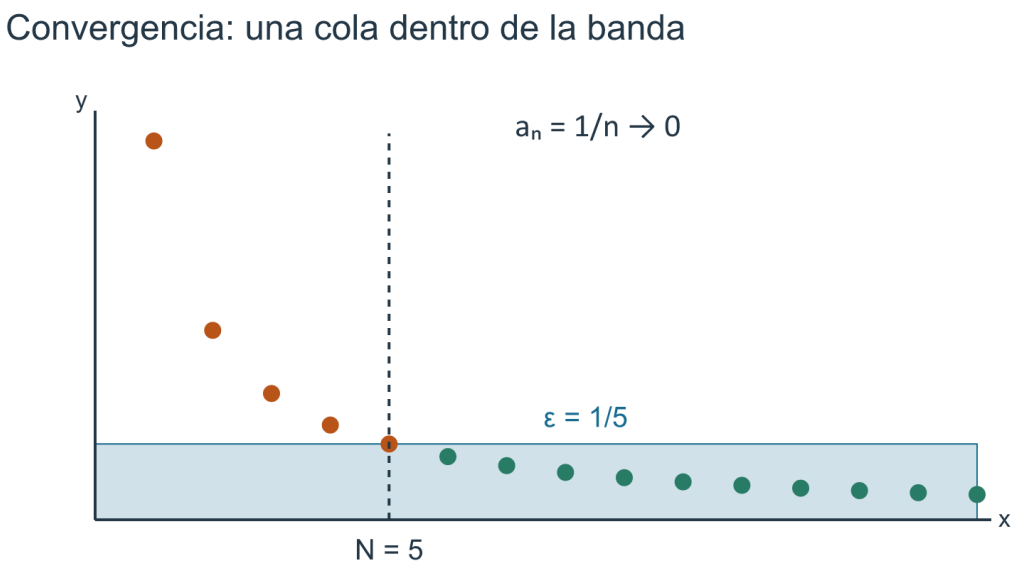


FIGURA 3.1. En $a_n = 1/n$, la banda $|a_n| < 1/5$ contiene todos los términos con $n > 5$. El término $a_5 = 1/5$ queda en el borde y no satisface la desigualdad estricta.

- **para todo** $\varepsilon > 0$: comenzamos con una tolerancia arbitraria;
- **existe** N: la propiedad arquimediana nos permite construir una respuesta;
- **para todo** $n \geq N$: la última cadena de desigualdades funciona para cualquier término de la cola.

La última desigualdad vale para cualquier índice posterior al umbral elegido. Así queda controlada una cola completa dentro de la precisión que se había fijado arbitrariamente.

Un mismo umbral no tiene que ser óptimo. En el ejemplo anterior elegimos N satisfaciendo $N > \frac{1}{\varepsilon}$. Esta elección es deliberadamente cómoda y puede estar lejos de ser mínima. La definición exige encontrar algún índice a partir del cual la desigualdad funcione para todos los términos posteriores. Un umbral mayor también sirve, porque su cola está contenida en la anterior; por ello podemos elegir una cota cómoda sin calcular el primer índice exacto.

Estar cerca muchas veces no es converger. La sucesión $a_n = (-1)^n$ toma los valores $1, -1, 1, -1, 1, -1, \ldots$

Si proponemos $L = 1$, hay infinitos términos que están exactamente en 1: todos los índices pares satisfacen $a_n = 1$. Sin embargo, eso no implica convergencia a 1. Después de cualquier índice vuelven a aparecer términos iguales a -1.

La definición no pide

> para toda tolerancia existen términos arbitrariamente tardíos que están cerca de L,

sino

para toda tolerancia existe un momento después del cual **todos** los términos están cerca de L.

En símbolos, no debemos confundir $\forall N\ \exists n \geq N$ con $\exists N\ \forall n \geq N$. La primera estructura expresa que algo vuelve a ocurrir por lejos que avancemos. La segunda expresa que, desde cierto punto, algo ocurre sin excepciones. Para la convergencia necesitamos la segunda.

Negar correctamente que $a_n \to L$. La definición de convergencia a un candidato L es

$$\forall \varepsilon > 0\ \exists N \geq n_0\ \forall n \geq N : |a_n - L| < \varepsilon.$$

Para negarla debemos invertir los cuantificadores y negar la desigualdad final. Obtenemos:

$$\boxed{a_n \not\to L \quad \Longleftrightarrow \quad \exists \varepsilon_0 > 0\ \forall N \geq n_0\ \exists n \geq N : |a_n - L| \geq \varepsilon_0.}$$

Esta fórmula merece una lectura pausada. Para demostrar que L **no** es el límite, no tenemos que fracasar ante todas las tolerancias. Basta encontrar **una** tolerancia positiva ε_0 que la sucesión sea incapaz de satisfacer eventualmente. Una vez fijada esa tolerancia, debemos demostrar que ningún umbral funciona: por lejos que alguien coloque N, todavía podemos encontrar algún término posterior que queda a distancia al menos ε_0 de L. La negación tiene así la forma de un obstáculo persistente:

existe una precisión fija que la sucesión vuelve a violar arbitrariamente lejos.

Refutar un candidato no basta para demostrar divergencia. Hay que distinguir dos afirmaciones: $a_n \not\to L$ y (a_n) diverge. La primera excluye **un** candidato concreto. La segunda afirma que **ningún** real sirve como límite. Por ejemplo, $\frac{1}{n+1} \not\to 1$, pero eso no vuelve divergente a la sucesión: acabamos de demostrar que converge a 0. Como divergencia significa ausencia de todo límite real, su forma lógica completa es

$$\forall L \in \mathbb{R}\ \exists \varepsilon_0 > 0\ \forall N \geq n_0\ \exists n \geq N : |a_n - L| \geq \varepsilon_0.$$

En muchas situaciones posteriores dispondremos de criterios más económicos para demostrar divergencia. Por ahora podemos hacerlo directamente desde esta negación.

Ejemplo 3.4. **Divergencia de $(-1)^n$ desde la definición.** Sea $a_n = (-1)^n$.

Demostraremos que la sucesión no converge a ningún número real. Sea $L \in \mathbb{R}$ un candidato arbitrario. Los dos posibles valores de la sucesión son 1 y -1. Por la desigualdad triangular,

$$2 = |1 - (-1)| \leq |1 - L| + |L - (-1)| = |1 - L| + |-1 - L|.$$

Por tanto, al menos uno de los dos números

$$|1 - L|, \qquad |-1 - L|$$

es mayor o igual que 1.

Tomemos $\varepsilon_0 = 1$.

Ahora sea $N \in \mathbb{N}$ cualquier umbral.

- Si $|1 - L| \geq 1$, elegimos un índice par $n \geq N$, por ejemplo $n = 2N$. Entonces $a_n = 1$ y $|a_n - L| = |1 - L| \geq 1 = \varepsilon_0$.
- Si $|-1 - L| \geq 1$, elegimos un índice impar $n \geq N$, por ejemplo $n = 2N + 1$. Entonces $a_n = -1$ y $|a_n - L| = |-1 - L| \geq 1 = \varepsilon_0$.

En cualquiera de los dos casos, para todo N encontramos un $n \geq N$ tal que $|a_n - L| \geq \varepsilon_0$.

Así, $a_n \not\to L$.

Como $L \in \mathbb{R}$ era arbitrario, ningún real puede ser límite de la sucesión. Por consiguiente, $((-1)^n)$ diverge.

La prueba toma un L arbitrario y muestra que la oscilación entre 1 y -1 impide encerrar toda una cola en una tolerancia suficientemente estrecha alrededor de él. Al valer para cualquier L real, el argumento excluye todos los posibles límites.

Una definición que ya contiene un método de trabajo. El orden de los cuantificadores organiza tanto la prueba de convergencia como su refutación. Para **probar convergencia hacia un candidato** L:

$$\boxed{\varepsilon > 0 \longrightarrow \text{buscar una condición suficiente sobre } n \longrightarrow \text{elegir } N(\varepsilon) \longrightarrow \forall n \geq N : \ |a_n - L| < \varepsilon.}$$

Para **refutar que L sea límite**:

$$\boxed{\text{elegir } \varepsilon_0 > 0 \longrightarrow \forall N \longrightarrow \text{encontrar } n \geq N \text{ con } |a_n - L| \geq \varepsilon_0.}$$

Estas dos arquitecturas son lógicamente opuestas. Aprender a reconocerlas será tan importante como aprender a calcular límites. En esta sección hemos definido qué significa converger, pero todavía no hemos estudiado qué consecuencias tiene hacerlo. Por ejemplo:

- ¿podría una misma sucesión converger a dos números distintos?;
- ¿puede una sucesión convergente escapar arbitrariamente lejos de todos los números?;
- ¿qué ocurre si modificamos una cantidad finita de términos de una sucesión convergente?

Estas preguntas se resuelven haciendo que una o varias condiciones de proximidad valgan en una misma cola. La elección de los umbrales será la pieza común de las pruebas.

3.3. Qué fuerza tiene converger: unicidad, colas y acotación

La definición ε–N parece, a primera vista, exigir únicamente que una sucesión se acerque a cierto número. Sin embargo, esa exigencia tiene consecuencias mucho más fuertes de lo que sugiere la intuición inicial. En esta sección responderemos tres preguntas.

1. Si una sucesión converge, ¿puede tener dos límites distintos?
2. Si alteramos solamente una cantidad finita de términos, ¿podemos cambiar su convergencia?
3. ¿Puede una sucesión convergente tener términos de tamaño arbitrariamente grande?

Las respuestas serán, respectivamente,

$$\boxed{\text{no,} \qquad \text{no,} \qquad \text{no.}}$$

Para obtenerlas tendremos que combinar umbrales, separar dos supuestos límites mediante su distancia y extender al prefijo finito una cota conocida en la cola.

Una sucesión no puede acercarse indefinidamente a dos números distintos. Supongamos que una sucesión (a_n) convergiera a dos números reales L y M. Si $L \neq M$, existe una distancia positiva entre ellos: $|L - M| > 0$. La definición de convergencia permitiría entonces exigir simultáneamente que, a partir de cierto momento, los términos de la sucesión estén muy cerca de L y muy cerca de M. Pero dos entornos suficientemente pequeños alrededor de puntos distintos no pueden compartir un mismo punto. Esta observación geométrica se convertirá en una contradicción mediante la desigualdad triangular.

TEOREMA 3.1. **Unicidad del límite secuencial.** Sea $(a_n)_{n \geq n_0}$ una sucesión real. Si

$$a_n \to L$$

y también

$$a_n \to M,$$

entonces

$$L = M.$$

Demostración. Supongamos, buscando una contradicción, que

$$L \neq M.$$

Entonces

$$d := |L - M| > 0.$$

Elegimos la tolerancia

$$\varepsilon = \frac{d}{3}.$$

Como $a_n \to L$, existe un índice $N_L \geq n_0$ tal que, para todo $n \geq N_L$,

$$|a_n - L| < \frac{d}{3}.$$

Como también $a_n \to M$, existe un índice $N_M \geq n_0$ tal que, para todo $n \geq N_M$,

$$|a_n - M| < \frac{d}{3}.$$

Tomemos

$$N = \max\{N_L, N_M\}.$$

Entonces, para todo $n \geq N$, ambas desigualdades se cumplen simultáneamente. Fijemos uno de esos índices. Por la desigualdad triangular,

$$|L - M| \leq |L - a_n| + |a_n - M|.$$

Por las estimaciones anteriores,

$$|L - M| < \frac{d}{3} + \frac{d}{3} = \frac{2d}{3}.$$

Pero $|L - M| = d$, de modo que obtenemos

$$d < \frac{2d}{3},$$

lo cual es imposible porque $d > 0$. Por tanto, la suposición $L \neq M$ era falsa y necesariamente

$$L = M.$$

Si los supuestos límites fueran distintos, su distancia positiva permitiría elegir una tolerancia que los separase. La convergencia a L y a M produce entonces dos umbrales; tomar su máximo $N = \max\{N_L, N_M\}$ permite entrar simultáneamente en ambas colas. En esa cola, un mismo término a_n tendría que estar a menos de $d/3$ de ambos extremos de un segmento de longitud d, lo que contradice la desigualdad triangular. Podríamos haber usado, por ejemplo, $\varepsilon = d/4$. Lo esencial es elegir una tolerancia tal que dos errores sumados sigan siendo estrictamente menores que d.

La elección $\varepsilon = \frac{d}{2}$ sería menos conveniente para esta prueba tal como está escrita: produciría solamente $|L - M| < \frac{d}{2} + \frac{d}{2} = d$, que sigue siendo una contradicción porque las desigualdades son estrictas, pero deja el argumento exactamente en la frontera. Elegir $d/3$ hace visible un margen positivo y prepara un hábito útil: cuando una desigualdad triangular divide un error entre varias partes, conviene reservar espacio suficiente para todas ellas. A partir de ahora, gracias a la unicidad, tiene sentido escribir

$$\lim_{n\to\infty} a_n = L$$

sin ambigüedad: si el límite real existe, hay uno solo.

La convergencia pertenece a la cola. En §3.1 vimos que una propiedad eventual no depende de un prefijo finito. La convergencia está formulada precisamente mediante propiedades eventuales:

$$\forall \varepsilon > 0 \; \exists N \; \forall n \geq N : |a_n - L| < \varepsilon.$$

Por eso modificar algunos términos iniciales no puede destruir una convergencia ya existente ni fabricar un límite diferente. Conviene formular esta estabilidad de manera algo más general.

Proposición 3.1. **Coincidencia eventual y estabilidad de la convergencia.** Sean (a_n) y (b_n) dos sucesiones reales definidas, al menos, a partir de algún índice común n_0. Supongamos que coinciden eventualmente; es decir, existe $K \geq n_0$ tal que

$$a_n = b_n \qquad \text{para todo } n \geq K.$$

Entonces, para todo $L \in \mathbb{R}$,

$$a_n \to L \quad \Longleftrightarrow \quad b_n \to L.$$

En particular, cambiar una cantidad finita de términos de una sucesión no modifica su convergencia ni su límite.

Demostración. Supongamos primero que

$$a_n \to L.$$

Sea $\varepsilon > 0$. Por convergencia existe $N_a \geq n_0$ tal que

$$n \geq N_a \quad \Longrightarrow \quad |a_n - L| < \varepsilon.$$

Definimos

$$N = \text{máx}\{N_a, K\}.$$

Si $n \geq N$, entonces simultáneamente

$$n \geq N_a$$

y

$$n \geq K.$$

La primera desigualdad da

$$|a_n - L| < \varepsilon,$$

y la segunda garantiza que

$$b_n = a_n.$$

Por tanto,

$$|b_n - L| = |a_n - L| < \varepsilon.$$

Hemos demostrado que $b_n \to L$.

El argumento inverso es idéntico, intercambiando los papeles de (a_n) y (b_n). Por consiguiente,

$$a_n \to L \quad \Longleftrightarrow \quad b_n \to L.$$

El mecanismo de la prueba puede resumirse en una sola operación:

$$\boxed{\text{umbral de convergencia} + \text{umbral de coincidencia} \longrightarrow \text{máx}\{N_a, K\}.}$$

Este uso del máximo aparecerá constantemente. Siempre que varias propiedades sean verdaderas eventualmente, podemos hacerlas verdaderas **simultáneamente** avanzando hasta el mayor de sus umbrales.

Quitar o añadir una cantidad finita de términos. La proposición anterior tiene una consecuencia inmediata. Si $a_n \to L$, entonces cualquier cola $(a_n)_{n \geq K}$ converge al mismo L. Recíprocamente, si alguna cola de la sucesión converge a L, entonces la sucesión completa también converge a L: los términos anteriores a K forman solo un prefijo finito y no intervienen en ninguna condición que deba cumplirse desde un umbral suficientemente tardío.

Así,

$$\boxed{(a_n) \text{ converge a } L \quad \Longleftrightarrow \quad \text{cualquier cola de } (a_n) \text{ converge a } L.}$$

Esta equivalencia expresa con precisión una frase que utilizaremos muchas veces:

> para estudiar un límite podemos descartar una cantidad finita de términos iniciales.

Descartar no significa borrar información del objeto original. Significa que esa información no afecta la pregunta específica de convergencia.

Desplazar los índices tampoco cambia el límite. Sea $r \in \mathbb{N}$ fijo y, para $n \geq n_0$, definamos $b_n = a_{n+r}$.

La sucesión (b_n) recorre los mismos términos que una cola de (a_n), pero vuelve a numerarlos comenzando en un índice anterior.

Demostremos directamente que

$$a_n \to L \quad \iff \quad b_n \to L.$$

Supongamos que $a_n \to L$ y sea $\varepsilon > 0$. Existe N tal que

$$m \geq N \quad \implies \quad |a_m - L| < \varepsilon.$$

Si elegimos un índice n suficientemente grande para que $n + r \geq N$, entonces $|b_n - L| = |a_{n+r} - L| < \varepsilon$.

Por ejemplo, basta tomar $N_b = \max\{n_0, N - r\}$.

En la dirección contraria, si $b_n \to L$, entonces existe un umbral N_b después del cual $|a_{n+r} - L| < \varepsilon$. Para todo $m \geq N_b + r$, podemos escribir $m = n + r$ con $n = m - r \geq N_b$, y obtenemos $|a_m - L| < \varepsilon$. Por tanto, desplazar una cantidad fija de posiciones el índice tampoco altera el límite. Este hecho será decisivo cuando estudiemos recurrencias: si $a_n \to L$, entonces también $a_{n+1} \to L$. No estamos usando continuidad ni álgebra de límites. Es una consecuencia directa de que quitar el primer término no cambia el comportamiento de la cola.

Convergencia no significa que todos los términos estén cerca del límite. La estabilidad bajo modificaciones finitas corrige una intuición frecuente. Una sucesión convergente puede comenzar con términos enormes o completamente irregulares. Por ejemplo, definamos

$$a_0 = 10^{100}, \qquad a_1 = -10^{50}, \qquad a_2 = 37,$$

y, para $n \geq 3$, $a_n = \frac{1}{n+1}$.

Como la cola coincide con la sucesión ya estudiada $1/(n+1)$, tenemos $a_n \to 0$. Los primeros términos pueden estar muy lejos de 0 sin afectar el límite.

La convergencia no dice

$$|a_n - L| < \varepsilon \qquad \text{para todo } n.$$

Dice que, para cada tolerancia, esa desigualdad termina siendo verdadera **eventualmente**. Esta distinción será importante en la siguiente prueba. La convergencia controlará automáticamente una cola, pero para obtener una afirmación sobre **todos** los términos deberemos tratar por separado el prefijo finito restante.

Acotación: encerrar todos los términos en un intervalo fijo. Hasta ahora hemos utilizado cotas para conjuntos de números reales. Como una sucesión es una función, también podemos preguntar si su conjunto de valores permanece dentro de una región finita de la recta.

Definición 3.3. **Sucesión acotada.** Sea $(a_n)_{n \geq n_0}$ una sucesión real. Diremos que (a_n) es **acotada** si existe un número $M > 0$ tal que

$$|a_n| \leq M \qquad \text{para todo } n \geq n_0.$$

Equivalentemente, existe $M > 0$ tal que todos los términos pertenecen al intervalo

$$[-M, M].$$

La constante M debe ser **una sola cota para toda la sucesión**. No sería suficiente permitir una cota distinta para cada término: para cualquier número real a_n podríamos elegir trivialmente $M_n = |a_n| + 1$, y esa condición no expresaría ningún control global. En lenguaje de funciones, la definición equivale a afirmar que la imagen $\{a_n : n \geq n_0\}$ es un subconjunto acotado de $\mathbb{R}$. También podemos separar cotas superiores e inferiores. Una sucesión es acotada si existen números $m, M \in \mathbb{R}$ tales que

$$m \leq a_n \leq M \qquad \text{para todo } n \geq n_0.$$

La forma simétrica $|a_n| \leq C$ es especialmente cómoda porque reúne ambas desigualdades en una sola expresión.

Toda sucesión convergente es acotada. La definición de convergencia ya contiene casi toda la prueba. Si $a_n \to L$, podemos escoger una tolerancia concreta, por ejemplo $\varepsilon = 1$.

Entonces, desde algún índice N, $|a_n - L| < 1$.

La desigualdad triangular da $|a_n| \leq |a_n - L| + |L| < 1 + |L|$.

Así, toda la cola está acotada por $|L|+1$. Pero la definición de sucesión acotada exige controlar **todos** los términos, incluidos los anteriores a N. Aquí aparece el segundo ingrediente: antes de N solo hay una cantidad finita de términos, y una cantidad finita de números reales siempre admite un máximo de sus valores absolutos.

Proposición 3.2. **Toda sucesión convergente es acotada.** Si

$$a_n \to L,$$

entonces la sucesión (a_n) es acotada.

Demostración. Como $a_n \to L$, podemos aplicar la definición con la tolerancia

$$\varepsilon = 1.$$

Existe entonces un índice $N_0 \geq n_0$ tal que

$$n \geq N_0 \quad \Longrightarrow \quad |a_n - L| < 1.$$

Si fuera necesario, reemplazamos N_0 por

$$N = \max\{N_0, n_0 + 1\}.$$

La propiedad anterior sigue siendo válida para $n \geq N$, y ahora existe al menos un término en el prefijo que va desde n_0 hasta $N - 1$. Para todo $n \geq N$, por la desigualdad triangular,

$$|a_n| \leq |a_n - L| + |L| < 1 + |L|.$$

Por otra parte, los términos

$$a_{n_0}, a_{n_0+1}, \ldots, a_{N-1}$$

son finitos en número. Por tanto existe el máximo

$$M_0 = \max\{|a_{n_0}|, |a_{n_0+1}|, \ldots, |a_{N-1}|\}.$$

Definamos

$$M = \max\{M_0, |L| + 1\}.$$

Entonces $M > 0$ y controla ambas regiones:

- si $n_0 \leq n < N$, tenemos $|a_n| \leq M_0 \leq M$;
- si $n \geq N$, tenemos $|a_n| < |L| + 1 \leq M$.

En consecuencia,

$$|a_n| \leq M \qquad \text{para todo } n \geq n_0.$$

Por definición, (a_n) es acotada.

Esta demostración contiene un patrón que conviene conservar:

$$\boxed{\text{controlar la cola} + \text{controlar el prefijo finito} \longrightarrow \text{control global.}}$$

La convergencia proporciona el primer control. La finitud del prefijo proporciona el segundo. Observe también que no necesitábamos la definición completa para **todas** las tolerancias. Para demostrar acotación bastó utilizar una sola, $\varepsilon = 1$. Una definición universal puede tener consecuencias que se obtienen seleccionando estratégicamente un caso particular.

Qué hipótesis usamos realmente en la prueba de acotación. Es útil auditar el argumento.

1. **Convergencia.** Se usó para encontrar una cola dentro de distancia 1 de L.
2. **Desigualdad triangular.** Transformó una cota sobre $|a_n - L|$ en una cota sobre $|a_n|$.
3. **Finitud del prefijo.** Permitió tomar el máximo de los términos no controlados por la cola.

No se utilizó completitud, monotonía, álgebra de límites ni ningún teorema posterior. La prueba es, por tanto, una consecuencia muy temprana y directa de la definición ε–N.

La recíproca es falsa. Acabamos de demostrar

$$\boxed{\text{convergente} \Longrightarrow \text{acotada.}}$$

No debemos invertir automáticamente la implicación. La sucesión $a_n = (-1)^n$ satisface $|a_n| = 1$ para todo n. Por tanto es acotada, con $M = 1$. Sin embargo, en §3.2 demostramos directamente que diverge.

Así,

$$\boxed{\text{acotada} \not\Longrightarrow \text{convergente.}}$$

Esta separación será fundamental más adelante. La acotación impide que los términos escapen indefinidamente en tamaño, pero no impide que sigan oscilando entre regiones distintas.

Convergencia tampoco significa monotonía. Otra confusión frecuente consiste en imaginar una sucesión convergente como una lista de términos que se aproxima al límite siempre desde el mismo lado y en una sola dirección. La definición no exige nada semejante.

EJEMPLO 3.5. **Dos separaciones esenciales: convergencia sin monotonía y acotación sin convergencia.**

Consideremos primero $a_n = \frac{(-1)^n}{n+1}$.

Para todo n, $|a_n - 0| = \frac{1}{n+1}$.

Dado $\varepsilon > 0$, en §3.2 ya construimos un umbral N tal que

$$n \geq N \quad \Longrightarrow \quad \frac{1}{n+1} < \varepsilon.$$

El mismo umbral produce $|a_n| < \varepsilon$.

Por tanto, $\frac{(-1)^n}{n+1} \to 0$.

Pero la sucesión no es creciente ni decreciente. Sus tres primeros términos son $1, -\frac{1}{2}, \frac{1}{3}$, de modo que $a_0 > a_1$ y también $a_1 < a_2$. Una sucesión creciente no podría presentar la primera desigualdad, y una decreciente no podría presentar la segunda. Así, (a_n) converge pero no es monótona.

Consideremos ahora $b_n = (-1)^n$.

Como $|b_n| = 1$ para todo n, la sucesión es acotada. Sin embargo, Ejemplo 3.4 demostró que no converge a ningún número real. Tenemos entonces dos contraejemplos complementarios:

$$\boxed{\text{convergente} \not\Longrightarrow \text{monótona},}$$

$$\boxed{\text{acotada} \not\Longrightarrow \text{convergente}.}$$

Estos ejemplos muestran que conviene mantener separadas tres propiedades:

- **convergencia:** toda cola suficientemente tardía queda dentro de cualquier tolerancia alrededor de un único L;
- **acotación:** todos los términos permanecen dentro de algún intervalo fijo;
- **monotonía:** los términos respetan sistemáticamente un sentido de orden entre índices consecutivos o cualesquiera índices ordenados.

Ninguna de estas frases es sinónimo de las otras. Más adelante aparecerá una interacción mucho más profunda: una sucesión **monótona y acotada** sí tendrá que converger. Esa afirmación requerirá la completitud de $\mathbb{R}$ y será uno de los resultados centrales del capítulo. Todavía no podemos usarla.

Una primera red de implicaciones. Con los resultados disponibles hasta ahora podemos escribir:

$$\boxed{\begin{array}{ccccc} & & a_n \to L & & \\ & & \Downarrow & & \\ \text{límite único} & + & \text{acotación} & + & \text{estabilidad de colas} \end{array}}$$

Pero no podemos añadir las flechas recíprocas sin hipótesis adicionales.

En particular: acotada $\not\Longrightarrow$ convergente,

$$\text{convergente} \not\Longrightarrow \text{monótona}.$$

Y la unicidad no afirma existencia: decir que **si** existe un límite entonces es único no demuestra que una sucesión dada converja. Esta distinción entre existencia y unicidad es la misma que empleamos al demostrar la unicidad de inversos y la existencia de raíces en el capítulo inicial. El Teorema 3.1 resuelve solamente la parte de «a lo sumo uno». Para establecer que una sucesión concreta tiene límite todavía necesitamos demostrar existencia mediante la definición o mediante resultados que construiremos después.

Qué hemos ganado. La definición de §3.2 ya no es una pieza aislada. Ha producido tres herramientas generales:

1. **Unicidad:** si una sucesión converge, no hay ambigüedad acerca de su límite.
2. **Estabilidad de cola:** eliminar, añadir o modificar finitos términos no afecta el límite; desplazar el índice una cantidad fija tampoco.
3. **Acotación:** toda sucesión convergente permanece globalmente dentro de algún intervalo finito.

Estas propiedades permitirán que, a partir de la próxima sección, dejemos de volver a la definición ε–N para cada cálculo elemental. El siguiente problema será demostrar que las operaciones algebraicas respetan los límites. Si $a_n \to A$ y $b_n \to B$,

¿podemos concluir que $a_n + b_n \to A + B$,

$$a_n b_n \to AB,$$

o que $\frac{a_n}{b_n} \to \frac{A}{B}$?

La respuesta será afirmativa bajo las hipótesis adecuadas, pero la prueba del producto revelará por qué la acotación que acabamos de establecer no era un resultado accesorio: será exactamente la herramienta necesaria para controlar un factor variable.

3.4. El álgebra de límites: cuándo podemos calcular sin volver a ε

Hasta ahora, cada vez que quisimos demostrar un límite tuvimos que regresar a la definición

$$\forall \varepsilon > 0 \; \exists N \; \forall n \geq N : |a_n - L| < \varepsilon.$$

Ese trabajo era necesario. Antes de utilizar reglas de cálculo debíamos demostrar que esas reglas son compatibles con la noción rigurosa de convergencia que acabamos de construir. La pregunta de esta sección es la siguiente:

si conocemos los límites de dos sucesiones, ¿qué ocurre al sumarlas, multiplicarlas o dividirlas?

Las respuestas serán las que sugiere el álgebra ordinaria, pero no las aceptaremos por analogía. Las deduciremos de la definición ε–N. Una vez demostradas estas leyes, podremos aplicarlas a expresiones nuevas comprobando sus hipótesis, sin reconstruir una prueba con tolerancias para cada expresión completa.

Las reglas algebraicas son teoremas, no convenciones. Supongamos que $a_n \to A$ y $b_n \to B$.

Es natural esperar $a_n + b_n \to A + B$,

$$a_n b_n \to AB,$$

y, cuando el denominador no plantee problemas, $\frac{a_n}{b_n} \to \frac{A}{B}$. Pero cada una de estas afirmaciones contiene una dificultad distinta. Para la suma, la distancia al límite esperado se separa inmediatamente:

$$\begin{aligned} |(a_n + b_n) - (A + B)| &= |(a_n - A) + (b_n - B)| \\ &\leq |a_n - A| + |b_n - B|. \end{aligned}$$

Así, podemos repartir la tolerancia entre dos errores que ya sabemos controlar. Para el producto aparece algo nuevo. La identidad útil es

$$\begin{aligned} a_n b_n - AB &= a_n b_n - A b_n + A b_n - AB \\ &= b_n(a_n - A) + A(b_n - B). \end{aligned}$$

Por tanto,

$$|a_n b_n - AB| \leq |b_n|\,|a_n - A| + |A|\,|b_n - B|.$$

El segundo término contiene una constante $|A|$, pero el primero contiene el factor variable $|b_n|$. Saber únicamente que $|a_n - A|$ puede hacerse pequeño no basta si no controlamos cuánto puede valer $|b_n|$. Aquí reaparece exactamente el resultado de §3.3:

> toda sucesión convergente es acotada.

La acotación permitirá reemplazar el factor variable $|b_n|$ por una constante fija. El cociente contiene una dificultad diferente: aunque $B \neq 0$, necesitamos asegurar que los términos b_n tampoco se acerquen peligrosamente a cero cuando n sea grande. Repartir el error, acotar los factores variables y separar los denominadores de cero nos permitirá transformar la convergencia de las piezas en una estimación para la expresión completa.

Proposición 3.3. **Álgebra de límites secuenciales.** Sean (a_n) y (b_n) sucesiones reales tales que

$$a_n \to A \qquad \text{y} \qquad b_n \to B.$$

Entonces:

1. para todo $c \in \mathbb{R}$,
$$c a_n \to cA;$$
2. $$a_n + b_n \to A + B;$$
3. $$a_n b_n \to AB;$$
4. si $B \neq 0$, entonces existe una cola en la que $b_n \neq 0$ y, sobre esa cola,
$$\frac{a_n}{b_n} \to \frac{A}{B}.$$

En particular, usando $c = -1$, también

$$a_n - b_n \to A - B.$$

Demostración. Demostraremos cada afirmación directamente desde la definición de convergencia.

1. Multiplicación por una constante. Si $c = 0$, entonces $ca_n = 0$ para todo n, y la sucesión constante nula converge a $0 = cA$.

Supongamos ahora $c \neq 0$. Sea $\varepsilon > 0$. Como $a_n \to A$, existe N tal que

$$n \geq N \quad \Longrightarrow \quad |a_n - A| < \frac{\varepsilon}{|c|}.$$

Entonces, para todo $n \geq N$,

$$|ca_n - cA| = |c|\,|a_n - A| < |c|\frac{\varepsilon}{|c|} = \varepsilon.$$

Por tanto,

$$ca_n \to cA.$$

2. Suma. Sea $\varepsilon > 0$. Como $a_n \to A$, existe N_1 tal que

$$n \geq N_1 \quad \Longrightarrow \quad |a_n - A| < \frac{\varepsilon}{2}.$$

Como $b_n \to B$, existe N_2 tal que

$$n \geq N_2 \quad \Longrightarrow \quad |b_n - B| < \frac{\varepsilon}{2}.$$

Tomemos

$$N = \text{máx}\{N_1, N_2\}.$$

Si $n \geq N$, ambas estimaciones son válidas simultáneamente y, por la desigualdad triangular,

$$\begin{aligned} |(a_n + b_n) - (A + B)| &\leq |a_n - A| + |b_n - B| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon. \end{aligned}$$

Luego

$$a_n + b_n \to A + B.$$

3. Producto. Esta es la parte delicada. Por Proposición 3.2, la sucesión convergente (b_n) es acotada. Por tanto existe $M > 0$ tal que

$$|b_n| \leq M$$

para todo índice de la sucesión. Sea $\varepsilon > 0$. Como $a_n \to A$, existe N_1 tal que

$$n \geq N_1 \quad \Longrightarrow \quad |a_n - A| < \frac{\varepsilon}{2M}.$$

Como $b_n \to B$, existe N_2 tal que

$$n \geq N_2 \quad \Longrightarrow \quad |b_n - B| < \frac{\varepsilon}{2(|A|+1)}.$$

Elegimos

$$N = \text{máx}\{N_1, N_2\}.$$

Para todo $n \geq N$,

$$\begin{aligned} |a_n b_n - AB| &= |b_n(a_n - A) + A(b_n - B)| \\ &\leq |b_n|\,|a_n - A| + |A|\,|b_n - B| \\ &\leq M|a_n - A| + |A|\,|b_n - B| \\ &< M\frac{\varepsilon}{2M} + |A|\frac{\varepsilon}{2(|A|+1)} \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon. \end{aligned}$$

Por consiguiente,

$$a_n b_n \to AB.$$

4. Cociente. Supongamos $B \neq 0$. Antes de estudiar a_n/b_n debemos demostrar que el denominador no se anula en una cola.

Como $b_n \to B$, podemos aplicar la definición con la tolerancia

$$\frac{|B|}{2} > 0.$$

Existe N_0 tal que

$$n \geq N_0 \quad \Longrightarrow \quad |b_n - B| < \frac{|B|}{2}.$$

Por la desigualdad triangular inversa,

$$\big||b_n| - |B|\big| \leq |b_n - B|.$$

Así, para $n \geq N_0$,

$$|b_n| > |B| - \frac{|B|}{2} = \frac{|B|}{2} > 0.$$

En particular,

$$b_n \neq 0$$

para todo $n \geq N_0$. El cociente está, por tanto, bien definido en esa cola.

Demostremos primero que

$$\frac{1}{b_n} \to \frac{1}{B}.$$

Sea $\varepsilon > 0$. Además del umbral N_0, como $b_n \to B$ existe N_1 tal que

$$n \geq N_1 \quad \Longrightarrow \quad |b_n - B| < \frac{\varepsilon |B|^2}{2}.$$

Tomemos

$$N = \max\{N_0, N_1\}.$$

Entonces, si $n \geq N$,

$$\begin{aligned} \left| \frac{1}{b_n} - \frac{1}{B} \right| &= \frac{|B - b_n|}{|B|\,|b_n|} \\ &\leq \frac{|B - b_n|}{|B|(|B|/2)} \\ &= \frac{2|b_n - B|}{|B|^2} \\ &< \varepsilon. \end{aligned}$$

Por tanto,

$$\frac{1}{b_n} \to \frac{1}{B}.$$

Finalmente,

$$\frac{a_n}{b_n} = a_n \frac{1}{b_n}$$

para todo n suficientemente grande. Aplicando la parte del producto,

$$\frac{a_n}{b_n} \to A \frac{1}{B} = \frac{A}{B}.$$

Esto completa la demostración.

Cómo leer la prueba del producto. La prueba anterior contiene una técnica que aparecerá muchas veces en cálculo.

Nuestro objetivo era controlar $|a_n b_n - AB|$.

La expresión no estaba inicialmente escrita en términos de los dos errores que la convergencia nos permite hacer pequeños:

$$|a_n - A| \qquad \text{y} \qquad |b_n - B|.$$

Para hacerlos aparecer introdujimos un término intermedio: $a_n b_n \longrightarrow A b_n \longrightarrow AB$. Algebraicamente, esto equivale a sumar y restar Ab_n:

$$a_n b_n - AB = (a_n b_n - A b_n) + (A b_n - AB).$$

Después la desigualdad triangular transformó el problema en $|b_n|\,|a_n - A| + |A|\,|b_n - B|$. Ahora sí aparecen los errores conocidos, pero uno viene multiplicado por $|b_n|$. La acotación de (b_n) convierte ese factor variable en una constante manejable. La cadena de ideas es, por tanto,

crear los errores conocidos → usar desigualdad triangular → acotar el factor variable → repartir ε.

La cota proporcionada por Proposición 3.2 es la que permite elegir una tolerancia para cada error sin que esta elección dependa del índice n.

Una observación sobre las cotas usadas en una prueba. En la demostración del producto usamos una cota global $|b_n| \leq M$. En realidad habría bastado una **cota eventual**. Como $b_n \to B$, tomando, por ejemplo, $\varepsilon = 1$ obtenemos $|b_n - B| < 1$ desde algún índice, y entonces $|b_n| \leq |b_n - B| + |B| < 1 + |B|$. Esto controla exactamente la cola que interviene en la definición de límite. La versión global obtenida en §3.3 es más fuerte y más cómoda, pero conviene reconocer qué parte de una hipótesis utiliza realmente una demostración.

El denominador no nulo no puede tratarse como un detalle. La afirmación $b_n \to B \neq 0$ contiene más información que simplemente «el límite del denominador no es cero». Garantiza que los términos quedan finalmente dentro de una región que no contiene al cero. En efecto, desde cierto índice tenemos $|b_n - B| < \frac{|B|}{2}$, y por tanto $|b_n| > \frac{|B|}{2}$. Geométricamente, los términos terminan atrapados en un intervalo centrado en B cuyo radio es menor que la distancia de B al origen. Este paso cumple dos funciones distintas:

1. garantiza que $1/b_n$ está definido en una cola;
2. proporciona la estimación $\frac{1}{|b_n|} < \frac{2}{|B|}$, necesaria para controlar el error del recíproco.

No debemos escribir simplemente «como $B \neq 0$, dividimos». El denominador de cada término es b_n, no B; la separación eventual de cero es el puente entre ambos hechos.

Sucesiones constantes y diferencias. Las leyes anteriores producen inmediatamente varias herramientas básicas. Para cualquier $c \in \mathbb{R}$, la sucesión constante $a_n = c$ converge a c, porque $|a_n - c| = 0 < \varepsilon$ para toda $\varepsilon > 0$ y para todos los índices.

Además, si

$$a_n \to A, \qquad b_n \to B,$$

entonces, aplicando primero la multiplicación por -1 y después la suma, $a_n - b_n = a_n + (-1)b_n \to A - B$. Estas consecuencias pequeñas permiten construir expresiones algebraicas cada vez más complejas sin regresar a la definición en cada paso.

El valor absoluto también respeta el límite. El valor absoluto no es una operación algebraica binaria como suma o producto, pero la desigualdad triangular inversa permite controlarlo directamente.

Corolario 3.1. **Límite del valor absoluto.** Si

$$a_n \to A,$$

entonces

$$|a_n| \to |A|.$$

Demostración. Para todo n, la desigualdad triangular inversa da

$$\big||a_n| - |A|\big| \leq |a_n - A|.$$

Sea $\varepsilon > 0$. Como $a_n \to A$, existe N tal que

$$n \geq N \quad \Longrightarrow \quad |a_n - A| < \varepsilon.$$

Para esos mismos índices,

$$\big||a_n| - |A|\big| \le |a_n - A| < \varepsilon.$$

Por definición,

$$|a_n| \to |A|.$$

La desigualdad triangular inversa, demostrada en `T1-C02`, controla el nuevo error mediante el que ya sabemos hacer pequeño. El argumento establece la propiedad para el valor absoluto con las herramientas disponibles, sin requerir todavía una teoría de continuidad.

De la prueba a la herramienta de cálculo. Una vez demostrado Proposición 3.3, ya podemos utilizarlo como un resultado previo. Esto cambia el modo de resolver muchos problemas. Si sabemos, por ejemplo, que $a_n \to A$ y queremos estudiar una expresión como $3a_n^2 - 5a_n + 7$, no necesitamos construir desde cero un $N(\varepsilon)$ para toda la expresión. Podemos razonar en capas: $a_n \to A$, luego $a_n^2 = a_n a_n \to A^2$, entonces $3a_n^2 \to 3A^2$,

$$-5a_n \to -5A,$$

y, sumando con la sucesión constante 7, $3a_n^2 - 5a_n + 7 \to 3A^2 - 5A + 7$. La definición ε–N no ha desaparecido. Está **encapsulada** dentro de los teoremas que acabamos de demostrar. Esta es una característica fundamental del trabajo matemático: una vez justificada una herramienta, podemos reutilizarla sin repetir cada vez toda su construcción interna.

Un cociente de polinomios en n. Veamos ahora cómo estas leyes convierten un límite aparentemente más complicado en una combinación de límites ya conocidos.

Ejemplo 3.6. **Un límite racional calculado mediante el álgebra de límites.** Consideremos

$$r_n = \frac{2n+1}{3n-4}, \qquad n \in \mathbb{N}.$$

El denominador nunca se anula para un índice natural, porque la ecuación $3n - 4 = 0$ tendría la solución $n = 4/3$, que no pertenece a $\mathbb{N}$. Para $n \ge 1$ podemos dividir numerador y denominador por n:

$$r_n = \frac{2 + 1/n}{3 - 4/n}.$$

Ya sabemos que $\frac{1}{n} \to 0$ para $n \ge 1$. En efecto, dada $\varepsilon > 0$, la propiedad arquimediana permite elegir N con $N > 1/\varepsilon$; entonces, si $n \ge N$,

$$0 < \frac{1}{n} \le \frac{1}{N} < \varepsilon.$$

Es la misma estimación básica que utilizamos en §3.2, ahora escrita para una sucesión que comienza en $n = 1$. Por la multiplicación por constantes, $\frac{4}{n} \to 0$. Como las sucesiones constantes 2 y 3 convergen respectivamente a 2 y 3, la ley de la suma da $2 + \frac{1}{n} \to 2$ y $3 - \frac{4}{n} \to 3$. El límite del denominador es $3 \neq 0$. Por la ley del cociente,

$$\frac{2 + 1/n}{3 - 4/n} \to \frac{2}{3}.$$

Por tanto,

$$\boxed{\frac{2n+1}{3n-4} \to \frac{2}{3}.}$$

Obsérvese qué cambió respecto de §3.2. Allí una afirmación como $\frac{1}{n+1} \to 0$ requirió abrir la definición, fijar ε y fabricar un umbral mediante la propiedad arquimediana. Aquí ya no necesitamos repetir ese mecanismo para toda la fracción. Hemos reducido el problema a piezas cuyos límites conocemos y hemos aplicado resultados generales previamente demostrados.

Qué está permitido escribir desde ahora. Después de esta sección podremos abreviar razonamientos como

$$a_n \to A, \qquad b_n \to B$$

mediante expresiones del tipo

$$\lim_{n\to\infty}(a_n + b_n) = A + B,$$

$$\lim_{n\to\infty}(a_n b_n) = AB,$$

o, si $B \neq 0$,

$$\lim_{n\to\infty}\frac{a_n}{b_n} = \frac{A}{B}.$$

Pero estas fórmulas deben leerse como aplicaciones de Proposición 3.3, no como una licencia para sustituir símbolos mecánicamente. Antes de usar una regla conviene comprobar:

- que los límites de las piezas realmente existen;
- que la expresión está definida en una cola suficiente;
- que, en un cociente, el límite del denominador es no nulo.

La última condición es esencial. Si $b_n \to 0$, la ley del cociente no autoriza ninguna conclusión del tipo $\frac{a_n}{b_n} \to \frac{A}{0}$. La expresión $A/0$ ni siquiera es un número real. Los cocientes con denominadores que tienden a cero requerirán análisis adicional según el problema concreto; no existe aquí una regla algebraica automática.

Qué hemos ganado. La teoría de límites dispone ahora de su primer cálculo interno.

A partir de

$$a_n \to A, \qquad b_n \to B,$$

sabemos transportar la convergencia a través de:

escalas, sumas y diferencias, productos, cocientes con límite denominador no nulo,

y también a través del valor absoluto. La estructura de las pruebas revela tres principios reutilizables:

sincronizar varios controles mediante máx,
convertir factores variables en cotas fijas,
separar de cero aquello que aparecerá en un denominador.

En la siguiente sección añadiremos una estructura distinta: el **orden**. Ya no preguntaremos qué ocurre al combinar sucesiones algebraicamente, sino qué información sobre desigualdades sobrevive al pasar al límite y cómo dos sucesiones pueden atrapar a una tercera.

3.5. Orden y teorema del sándwich

Las leyes algebraicas de §3.4 nos permiten transportar límites a través de sumas, productos y cocientes. Pero muchas sucesiones no se controlan mejor mediante una identidad algebraica, sino mediante una **desigualdad**. A veces no sabemos calcular directamente una sucesión complicada, pero sí podemos compararla con otras más sencillas. En otras ocasiones sabemos que una sucesión queda siempre por debajo de otra y queremos averiguar si esa relación puede invertirse cuando pasamos al límite. La pregunta de esta sección es, por tanto,

¿qué información de orden que vale en una cola sigue siendo visible en el límite?

La palabra **cola** vuelve a ser decisiva. Como el límite ignora cualquier prefijo finito, no necesitaremos exigir que una desigualdad sea verdadera desde el primer término. Bastará que lo sea eventualmente.

Una desigualdad eventual no puede invertirse en el límite. Supongamos que $a_n \le b_n$ para todos los índices suficientemente grandes y que

$$a_n \to A, \qquad b_n \to B.$$

¿Podría ocurrir que $A > B$?

Si así fuera, existiría una distancia positiva entre ambos límites. La convergencia permitiría entonces obligar a a_n a permanecer muy cerca de A y a b_n muy cerca de B. Para tolerancias suficientemente pequeñas, los términos de a_n quedarían necesariamente **por encima** de los de b_n, contradiciendo la desigualdad eventual. La intuición se convierte en la siguiente proposición.

Proposición 3.4. **Preservación eventual del orden bajo límites.** Sean (a_n) y (b_n) sucesiones reales tales que

$$a_n \to A$$

y

$$b_n \to B.$$

Supongamos además que existe un índice K tal que

$$a_n \le b_n \qquad \text{para todo } n \ge K.$$

Entonces

$$A \le B.$$

Demostración. Supongamos, buscando una contradicción, que

$$A > B.$$

Definamos

$$d := A - B > 0.$$

Elegimos la tolerancia

$$\varepsilon = \frac{d}{3}.$$

Como $a_n \to A$, existe N_a tal que, si $n \geq N_a$, entonces

$$|a_n - A| < \frac{d}{3}.$$

En particular,

$$a_n > A - \frac{d}{3}.$$

Como $A = B + d$,

$$A - \frac{d}{3} = B + \frac{2d}{3}.$$

Por tanto, para $n \geq N_a$,

$$a_n > B + \frac{2d}{3}.$$

Por otra parte, como $b_n \to B$, existe N_b tal que, si $n \geq N_b$, entonces

$$|b_n - B| < \frac{d}{3},$$

y de aquí

$$b_n < B + \frac{d}{3}.$$

Tomemos ahora

$$N = \text{máx}\{K, N_a, N_b\}.$$

Para todo $n \geq N$ se cumplen simultáneamente las tres informaciones:

$$a_n \leq b_n,$$

$$a_n > B + \frac{2d}{3},$$

y

$$b_n < B + \frac{d}{3}.$$

Las dos últimas implican

$$a_n > b_n,$$

porque

$$B + \frac{2d}{3} > B + \frac{d}{3}.$$

Esto contradice $a_n \leq b_n$.

Por consiguiente, la suposición $A > B$ es imposible y debe cumplirse

$$A \leq B.$$

La prueba utiliza una arquitectura ya conocida:

hipótesis eventual + dos controles de convergencia $\longrightarrow$ un único umbral mediante máx.

No importa que la desigualdad falle en los primeros cien, mil o un millón de términos. Si existe una cola en la que $a_n \leq b_n$, esa cola basta para impedir que los límites aparezcan en el orden contrario.

Consecuencias inmediatas. La proposición anterior contiene varias formas útiles como casos particulares. Si $a_n \geq 0$ eventualmente y $a_n \to A$, podemos comparar (a_n) con la sucesión constante 0. Como $0 \to 0$, obtenemos $A \geq 0$. Análogamente, si existe una constante $c \in \mathbb{R}$ tal que $a_n \leq c$ eventualmente y $a_n \to A$, entonces $A \leq c$.

Y si $a_n \geq c$ eventualmente, entonces $A \geq c$.

Estas versiones permiten transportar al límite cotas que solo empiezan a cumplirse después de cierto índice.

El orden estricto puede desaparecer. Hay que leer Proposición 3.4 con precisión. De $a_n < b_n$ eventualmente no podemos concluir necesariamente $A < B$. La razón es que dos sucesiones pueden mantener una separación positiva en cada índice y, sin embargo, esa separación hacerse cada vez menor. El ejemplo más sencillo ya está disponible: $0 < \frac{1}{n+1}$ para todo $n \in \mathbb{N}$.

Sin embargo, $0 \to 0$ y, por §3.2, $\frac{1}{n+1} \to 0$.

Por tanto los límites satisfacen $0 = 0$, no una desigualdad estricta.

Así,

$$\boxed{a_n < b_n \text{ eventualmente} \quad \not\Longrightarrow \quad A < B.}$$

Lo que sí preserva el teorema es la versión débil:

$$a_n < b_n \quad \Longrightarrow \quad a_n \leq b_n \quad \Longrightarrow \quad A \leq B.$$

Para conservar una separación estricta en el límite necesitamos información cuantitativa adicional. Por ejemplo, si existe $\delta > 0$ tal que eventualmente $a_n + \delta \leq b_n$, entonces, usando §3.4, $a_n + \delta \to A + \delta$. La preservación del orden produce $A + \delta \leq B$, y por tanto $A < B$. La diferencia es importante: no basta que cada desigualdad sea estricta; hace falta que exista una **separación positiva uniforme** que no se desvanezca al avanzar por la sucesión.

Atrapar una sucesión entre dos aproximaciones. La preservación del orden resuelve una pregunta cuando ya sabemos que las dos sucesiones comparadas convergen. Pero a menudo queremos utilizar desigualdades para demostrar precisamente que una tercera sucesión **converge**. Supongamos que podemos establecer, eventualmente, $a_n \leq b_n \leq c_n$, y que las sucesiones exteriores convergen al mismo número:

$$a_n \to L, \qquad c_n \to L.$$

Intuitivamente, si a_n y c_n terminan dentro de cualquier tolerancia alrededor de L, entonces b_n, atrapada entre ambas, no tiene ningún lugar fuera de ese mismo intervalo al cual escapar. Esta es la idea del teorema del sándwich.

PROPOSICIÓN 3.5. **Teorema del sándwich.** Sean (a_n), (b_n) y (c_n) sucesiones reales. Supongamos que existe un índice K tal que

$$a_n \leq b_n \leq c_n \qquad \text{para todo } n \geq K.$$

Si

$$a_n \to L$$

y

$$c_n \to L,$$

entonces

$$b_n \to L.$$

Demostración. Sea $\varepsilon > 0$.
Como $a_n \to L$, existe N_a tal que

$$n \geq N_a \quad \Longrightarrow \quad |a_n - L| < \varepsilon.$$

Esto implica

$$L - \varepsilon < a_n < L + \varepsilon.$$

Como $c_n \to L$, existe N_c tal que

$$n \geq N_c \quad \Longrightarrow \quad |c_n - L| < \varepsilon,$$

y por tanto

$$L - \varepsilon < c_n < L + \varepsilon.$$

Tomemos

$$N = \max\{K, N_a, N_c\}.$$

Si $n \geq N$, entonces

$$a_n \leq b_n \leq c_n,$$

mientras que

$$L - \varepsilon < a_n$$

y

$$c_n < L + \varepsilon.$$

Encadenando las desigualdades,

$$L - \varepsilon < a_n \leq b_n \leq c_n < L + \varepsilon.$$

En particular,

$$L - \varepsilon < b_n < L + \varepsilon.$$

Por la caracterización de valor absoluto,

$$|b_n - L| < \varepsilon.$$

Como $\varepsilon > 0$ era arbitrario, concluimos que

$$b_n \to L.$$

Hay un detalle lógico importante en esta prueba. **No** podemos justificar el teorema diciendo simplemente:

> como $a_n \leq b_n \leq c_n$, pasamos al límite y obtenemos $L \leq \lim b_n \leq L$.

Esa escritura presupone que $\lim b_n$ ya existe, que es precisamente lo que estamos intentando demostrar. El teorema del sándwich debe producir simultáneamente dos cosas:

1. la **existencia** del límite de (b_n);
2. la **identificación** de ese límite como L.

La prueba ε–N anterior hace ambas tareas sin circularidad.

Los extremos deben converger al mismo número. La coincidencia de los límites exteriores es esencial.

Por ejemplo, consideremos

$$a_n = 0, \qquad c_n = 1,$$

y $b_n = \frac{1+(-1)^n}{2}$.

Para todo n, $0 \leq b_n \leq 1$.

Las sucesiones exteriores convergen:

$$a_n \to 0, \qquad c_n \to 1.$$

Pero sus límites son distintos. La sucesión intermedia alterna entre 1 y 0, de modo que no converge. Así, una mera cota $a_n \leq b_n \leq c_n$ no basta. El sándwich funciona porque las dos paredes exteriores **se cierran sobre el mismo número**.

La forma absoluta: una herramienta especialmente útil. En muchos problemas no resulta natural buscar dos sucesiones exteriores por separado. Es más eficiente controlar directamente la distancia del término al candidato.

Supongamos que $|u_n| \leq v_n$ eventualmente y que $v_n \to 0$.

De la desigualdad de valor absoluto obtenemos $-v_n \leq u_n \leq v_n$. Por el álgebra de límites, $-v_n \to 0$, mientras que, por hipótesis, $v_n \to 0$. El teorema del sándwich da entonces $u_n \to 0$. Podemos registrar esta forma operativa como

$$\boxed{|u_n| \leq v_n \text{ eventualmente}, \qquad v_n \to 0 \quad \Longrightarrow \quad u_n \to 0.}$$

Esta versión será una de las herramientas más reutilizadas del capítulo. Su lógica es particularmente transparente: si el **tamaño** de u_n está dominado por una cantidad que puede hacerse arbitrariamente pequeña, entonces u_n debe acercarse a cero, aunque cambie de signo, oscile o carezca de monotonía.

Una oscilación que desaparece por estar dominada.

Ejemplo 3.7. **Una sucesión oscilante controlada por sándwich.** Consideremos

$$a_n = \frac{\sigma n}{n+1}, \qquad n \in \mathbb{N}.$$

Los valores de σn oscilan y no necesitamos describir su comportamiento con mayor precisión. Basta la desigualdad elemental $|\sigma n| \leq 1$.

Por tanto,

$$\left|\frac{\sigma n}{n+1}\right| = \frac{|\sigma n|}{n+1} \leq \frac{1}{n+1}.$$

Ya sabemos que $\frac{1}{n+1} \to 0$.

Aplicando la forma absoluta del teorema del sándwich,

$$\boxed{\frac{\sigma n}{n+1} \to 0.}$$

Equivalentemente, podríamos escribir $-\frac{1}{n+1} \leq \frac{\sigma n}{n+1} \leq \frac{1}{n+1}$, y observar que las dos sucesiones exteriores convergen a 0.

Este ejemplo muestra por qué el sándwich es más que una regla para «calcular límites». Permite ignorar una parte complicada de una sucesión cuando esa parte está uniformemente controlada. Aquí no necesitamos averiguar si σn se acerca a algún número. Ni siquiera necesitamos que converja. Lo único relevante es que su tamaño nunca excede 1, mientras que el factor $\frac{1}{n+1}$ se hace arbitrariamente pequeño.

El patrón general es

$$\boxed{\text{oscilación acotada} \times \text{factor que tiende a } 0 \longrightarrow \text{producto que tiende a } 0,}$$

siempre que podamos justificar la estimación correspondiente.

Cómo buscar un sándwich. En la práctica, rara vez nos entregan las tres sucesiones ya preparadas. La parte creativa consiste en **fabricar** buenas cotas. Una estrategia especialmente eficaz para demostrar $a_n \to L$ es comenzar por el error $|a_n - L|$. Si podemos encontrar una sucesión sencilla (r_n) tal que $|a_n - L| \le r_n$ eventualmente y $r_n \to 0$, la forma absoluta del sándwich termina la prueba. El trabajo se desplaza entonces desde

calcular exactamente a_n

hacia

encontrar una estimación suficientemente buena del error.

Este cambio de perspectiva es central en cálculo. Una cota puede contener menos información que una fórmula exacta y, sin embargo, ser exactamente la información necesaria para demostrar convergencia.

Tres errores que conviene evitar. 1. Una sola cota no determina un límite.
De $a_n \le b_n$ y $a_n \to L$ no se sigue que $b_n \to L$. Por ejemplo, $0 \le 1$ para todos los índices, pero las sucesiones constantes 0 y 1 tienen límites distintos.

2. Estar entre dos sucesiones convergentes no basta si sus límites son distintos.
El ejemplo $0 \le \frac{1+(-1)^n}{2} \le 1$ muestra que una sucesión intermedia puede divergir cuando las cotas exteriores convergen a números diferentes.

3. La desigualdad estricta no tiene por qué permanecer estricta.
Aunque $0 < \frac{1}{n+1}$ para todo n, ambos lados convergen a 0. Las tres advertencias comparten una misma lección: al trabajar con orden no basta mirar el símbolo de desigualdad; hay que auditar **qué límites poseen las sucesiones comparadas y qué separación cuantitativa conservan**.

Qué hemos ganado. A la estructura algebraica de §3.4 añadimos ahora una estructura de orden.

Sabemos que:

$$a_n \le b_n \text{ eventualmente}, \qquad a_n \to A, \qquad b_n \to B \quad \Longrightarrow \quad A \le B;$$

y que

$$a_n \le b_n \le c_n \text{ eventualmente}, \qquad a_n, c_n \to L \quad \Longrightarrow \quad b_n \to L.$$

Además, disponemos de la forma especialmente útil

$$|u_n| \le v_n \text{ eventualmente}, \qquad v_n \to 0 \quad \Longrightarrow \quad u_n \to 0.$$

Con estas herramientas podemos pasar de la definición y las leyes generales a un pequeño repertorio de límites fundamentales. En §3.6 utilizaremos el sándwich junto con desigualdades elementales e inducción para demostrar, entre otros resultados, que las potencias recíprocas tienden a cero y que $q^n \to 0$ cuando $|q| < 1$.

3.6. Un pequeño repertorio de límites fundamentales

Las secciones anteriores construyeron herramientas generales. Ya sabemos trabajar directamente con la definición ε–N, combinar límites mediante operaciones algebraicas y utilizar desigualdades para atrapar una sucesión entre otras más sencillas. Ahora conviene transformar esas herramientas en un pequeño repertorio de resultados que podamos reutilizar sin reconstruir

cada prueba desde cero. Conocer la demostración de cada resultado permitirá reconocer cuándo una expresión nueva se reduce a estas piezas y qué hipótesis deben verificarse para combinarlas. El criterio será siempre el mismo:

un límite fundamental debe quedar apoyado en resultados anteriores y no en teoría todavía no construida.

Por eso en esta sección no utilizaremos continuidad de funciones, exponenciales reales, logaritmos, series ni regla de l'Hôpital. Todo saldrá de arquimedianidad, inducción, álgebra de límites y teorema del sándwich.

Las potencias recíprocas también desaparecen. Ya demostramos que $\frac{1}{n+1} \to 0$.

¿Qué ocurre si aumentamos la potencia del denominador?

Sea $p \in \mathbb{N}_{>0}$. Para todo $n \in \mathbb{N}$ tenemos $n+1 \geq 1$.

Como $p \geq 1$, $(n+1)^p \geq n+1$, y, al tomar recíprocos positivos, $0 < \frac{1}{(n+1)^p} \leq \frac{1}{n+1}$.

Los extremos satisfacen $0 \to 0$ y $\frac{1}{n+1} \to 0$.

Por el teorema del sándwich,

$$\boxed{\frac{1}{(n+1)^p} \to 0 \qquad \text{para todo } p \in \mathbb{N}_{>0}.}$$

Este pequeño argumento ya muestra una ventaja importante de disponer de resultados generales. No necesitamos volver a fabricar un $N(\varepsilon)$ diferente para cada potencia. Una sola comparación con una sucesión cuyo límite conocemos resuelve toda la familia. También podemos multiplicar por una constante. Si $c \in \mathbb{R}$, el álgebra de límites da $\frac{c}{(n+1)^p} \to 0$. Así obtenemos un primer patrón reutilizable:

$$\boxed{\text{constante} \times \text{potencia recíproca positiva} \longrightarrow 0.}$$

Un problema diferente: ¿por qué una potencia geométrica tiende a cero? Consideremos ahora una sucesión de la forma q^n. La intuición es familiar cuando $|q| < 1$.

Por ejemplo,

$$\left(\frac{1}{2}\right)^n, \qquad \left(\frac{2}{3}\right)^n, \qquad \left(-\frac{3}{4}\right)^n$$

parecen hacerse cada vez menores en valor absoluto. Pero aquí aparece una dificultad distinta de la que encontramos con $1/(n+1)$. La variable n está en el **exponente**, de modo que una desigualdad como $q^n < \varepsilon$ no se despeja mediante el álgebra elemental que hemos utilizado hasta ahora sin introducir logaritmos. Y los logaritmos todavía no forman parte de nuestra teoría. Necesitamos, por tanto, otra estrategia. Para $0 < q < 1$ podemos escribir $q = \frac{1}{1+h}$ con algún $h > 0$. Basta tomar $h = \frac{1-q}{q}$.

Entonces $q^n = \frac{1}{(1+h)^n}$.

Si conseguimos una **cota inferior sencilla** para $(1+h)^n$, obtendremos una cota superior para q^n. La herramienta adecuada es una desigualdad elemental que puede demostrarse por inducción.

La desigualdad de Bernoulli como puente entre crecimiento multiplicativo y crecimiento lineal. Sea $h > 0$. Afirmamos que para todo $n \in \mathbb{N}$,

$$\boxed{(1+h)^n \geq 1 + nh.}$$

La demostración será inductiva: indicaremos explícitamente el caso base y el paso que transmite la desigualdad de n a $n+1$.

Caso base. Para $n = 0$, $(1+h)^0 = 1 = 1 + 0h$.

Paso inductivo. Supongamos que $(1+h)^n \geq 1 + nh$.

Como $1 + h > 0$, podemos multiplicar la desigualdad por ese número sin cambiar su sentido: $(1+h)^{n+1} \geq (1+nh)(1+h)$.

Desarrollando el lado derecho,

$$(1+nh)(1+h) = 1 + (n+1)h + nh^2.$$

Como $nh^2 \geq 0$, tenemos $(1+h)^{n+1} \geq 1 + (n+1)h$.

La inducción queda cerrada.

Para controlar el recíproco de $(1+h)^n$ basta una estimación inferior que crezca con n. La cota lineal $1 + nh$ es muy rudimentaria comparada con la potencia. Pero es suficiente para nuestro objetivo. Al tomar recíprocos positivos obtenemos $0 < \frac{1}{(1+h)^n} \leq \frac{1}{1+nh}$. Hemos convertido una potencia en una expresión lineal en n que sí podemos controlar directamente.

Un caso completo antes del resultado general.

Ejemplo 3.8. **Construcción detallada del límite geométrico: $(2/3)^n \to 0$.** Consideremos

$$a_n = \left(\frac{2}{3}\right)^n, \qquad n \geq 0.$$

No utilizaremos logaritmos ni una función exponencial real.

Escribimos $\frac{2}{3} = \frac{1}{1+1/2}$.

Por la desigualdad de Bernoulli con $h = 1/2$,

$$\left(1 + \frac{1}{2}\right)^n \geq 1 + \frac{n}{2}.$$

Tomando recíprocos,

$$0 < \left(\frac{2}{3}\right)^n \leq \frac{1}{1 + n/2} = \frac{2}{n+2}.$$

Ahora debemos justificar que la sucesión de la derecha tiende a cero. Podemos hacerlo con el álgebra ya disponible:

$$\frac{2}{n+2} = \frac{2/(n+1)}{(n+2)/(n+1)} = \frac{2/(n+1)}{1 + 1/(n+1)}.$$

Como $\frac{1}{n+1} \to 0$, tenemos $\frac{2}{n+1} \to 0$ y $1 + \frac{1}{n+1} \to 1$.

Por la ley del cociente, $\frac{2}{n+2} \to 0$.

Por tanto,

$$0 \leq \left(\frac{2}{3}\right)^n \leq \frac{2}{n+2},$$

con ambos extremos convergiendo a 0. El teorema del sándwich da

$$\boxed{\left(\frac{2}{3}\right)^n \to 0.}$$

Lectura de la estrategia. La potencia no se resolvió despejando n. La reemplazamos por una cota más sencilla:

$$\left(\frac{2}{3}\right)^n = \frac{1}{(1+1/2)^n} \leq \frac{1}{1+n/2}.$$

Bernoulli produjo la comparación y el sándwich convirtió esa comparación en convergencia.

El ejemplo contiene ya toda la estructura necesaria para el caso general.

COROLARIO 3.2. **Límite de una sucesión geométrica con razón de módulo menor que uno.** Sea $q \in \mathbb{R}$ con

$$|q| < 1.$$

Entonces, considerando la sucesión $(q^n)_{n\geq 1}$,

$$\boxed{q^n \to 0.}$$

Demostración. Distinguimos casos.

Si $q = 0$, entonces

$$q^n = 0$$

para todo $n \geq 1$, y la conclusión es inmediata.

Supongamos ahora

$$0 < |q| < 1.$$

Definimos

$$r = |q|.$$

Entonces $0 < r < 1$. Existe $h > 0$ tal que

$$r = \frac{1}{1+h};$$

por ejemplo,

$$h = \frac{1-r}{r} > 0.$$

Por Bernoulli,

$$(1+h)^n \geq 1 + nh.$$

Por tanto,

$$0 < r^n = \frac{1}{(1+h)^n} \leq \frac{1}{1+nh}.$$

Veamos directamente que

$$\frac{1}{1+nh} \to 0.$$

Sea $\varepsilon > 0$. Por la propiedad arquimediana podemos elegir $N \in \mathbb{N}$ con

$$N > \frac{1}{h\varepsilon}.$$

En particular $N > 0$. Si $n \geq N$, entonces

$$1 + nh > nh \geq Nh > \frac{1}{\varepsilon}.$$

Tomando recíprocos positivos,

$$0 < \frac{1}{1+nh} < \varepsilon.$$

Así,

$$\frac{1}{1+nh} \to 0.$$

El sándwich produce

$$r^n \to 0.$$

Finalmente,

$$|q^n| = |q|^n = r^n.$$

Como $r^n \to 0$, la forma absoluta del teorema del sándwich implica

$$q^n \to 0.$$

Esta prueba merece una segunda lectura porque combina casi todo lo construido hasta ahora:

1. **álgebra:** $|q| = 1/(1+h)$;
2. **inducción:** Bernoulli;
3. **arquimedianidad:** producir un umbral para $1/(1+nh)$;
4. **orden:** convertir una desigualdad en una cota útil;
5. **sándwich:** transportar el límite de la cota a la potencia geométrica;
6. **valor absoluto:** reducir el caso negativo al positivo.

No hemos utilizado ninguna teoría de la función exponencial. El resultado pertenece enteramente a la teoría secuencial ya disponible.

Por qué la hipótesis $|q| < 1$ es esencial. Conviene auditar la frontera del corolario. Si $q = 1$, entonces $q^n = 1$ para todo n, de modo que el límite es 1, no 0. Si $q = -1$, obtenemos $(-1)^n$, cuya divergencia ya demostramos en §3.2. Por tanto, no podemos reemplazar la condición $|q| < 1$ por $|q| \leq 1$. Cuando $|q| > 1$, los tamaños $|q|^n$ crecen en lugar de contraerse. Todavía no formalizaremos ese comportamiento mediante límites infinitos; esa será tarea de §3.9.

Cocientes de polinomios: dividir por la potencia dominante. Los límites anteriores permiten sistematizar una familia importante de ejemplos.

Consideremos $a_n = \frac{3n^2-5n+7}{2n^2+n-4}$.

La dificultad aparente es que numerador y denominador aumentan de tamaño. Pero para un límite real no necesitamos estudiar ambos por separado. Dividimos toda la fracción por la mayor potencia presente, aquí n^2.

Para $n \geq 1$,

$$\frac{3n^2-5n+7}{2n^2+n-4} = \frac{3-5/n+7/n^2}{2+1/n-4/n^2}.$$

Ya sabemos que $\frac{1}{n} \to 0$ y $\frac{1}{n^2} \to 0$.

Por el álgebra de límites, $3 - \frac{5}{n} + \frac{7}{n^2} \to 3$ y $2 + \frac{1}{n} - \frac{4}{n^2} \to 2$. Como el límite del denominador es distinto de cero, la ley del cociente da

$$\boxed{\frac{3n^2-5n+7}{2n^2+n-4} \to \frac{3}{2}.}$$

El procedimiento general ya es visible. Si numerador y denominador son polinomios del mismo grado, dividir por la potencia dominante hace desaparecer todos los términos de grado inferior mediante factores del tipo $1/n^p$. El límite queda determinado por el cociente de los coeficientes principales. Si el grado del numerador es menor que el del denominador, la misma operación deja en el numerador únicamente términos que contienen potencias recíprocas. Por ejemplo,

$$\frac{4n+1}{n^3+2} = \frac{4/n^2+1/n^3}{1+2/n^3} \to 0.$$

Así obtenemos, para los casos de límite real que necesitamos ahora, la regla estructural:

$$\boxed{\begin{aligned} \deg P < \deg Q \quad &\Longrightarrow \frac{P(n)}{Q(n)} \to 0, \\ \deg P = \deg Q \quad &\Longrightarrow \frac{P(n)}{Q(n)} \to \frac{\text{coeficiente principal de } P}{\text{coeficiente principal de } Q}, \end{aligned}}$$

siempre que el denominador esté definido en la cola correspondiente. No hemos creado una nueva regla independiente. Todo proviene de $\frac{1}{n^p} \to 0$ y del álgebra de límites. El caso $\deg P > \deg Q$ suele producir crecimiento sin límite real finito. Lo estudiaremos cuando dispongamos de la definición rigurosa de $+\infty$ y $-\infty$ en §3.9.

Oscilaciones amortiguadas: lo complicado puede quedar encerrado en una cota. El ejemplo de §3.5 con $\frac{\sigma n}{n+1}$ mostró un fenómeno general: una parte de la sucesión puede oscilar sin converger, mientras otro factor reduce su tamaño hasta volver irrelevante esa oscilación para el límite.

Consideremos, por ejemplo,

$$a_n = \frac{3(-1)^n - 2\sigma n}{(n+1)^2}.$$

No necesitamos comprender la interacción exacta entre $(-1)^n$ y σn. Basta estimar:

$$\begin{aligned}|3(-1)^n - 2\sigma n| &\leq 3|(-1)^n| + 2|\sigma n| \\ &\leq 3 + 2 \\ &= 5.\end{aligned}$$

Por tanto, $|a_n| \leq \frac{5}{(n+1)^2}$.

Como $\frac{5}{(n+1)^2} \to 0$, la forma absoluta del sándwich da

$$\boxed{\frac{3(-1)^n - 2\sigma n}{(n+1)^2} \to 0.}$$

Aquí aparece un principio que conviene conservar:

numerador uniformemente acotado
+denominador cuyo recíproco tiende a 0
$\Longrightarrow$
cociente que tiende a 0,

siempre que la estimación se formule rigurosamente. Más generalmente, si (u_n) es una sucesión acotada y $v_n \to 0$, entonces $u_n v_n \to 0$. Podemos justificarlo de dos maneras con herramientas ya conocidas. Por un lado, como (u_n) es acotada, existe $M > 0$ tal que $|u_n| \leq M$.

Entonces $|u_n v_n| \leq M|v_n|$.

Como $|v_n| \to 0$ y por álgebra $M|v_n| \to 0$, el sándwich da $u_n v_n \to 0$. Por otro lado, si además supiéramos que (u_n) converge, el resultado sería un caso particular de la ley del producto. La estimación anterior es más fuerte para este propósito: **no exige que** (u_n) **converja**, solo que esté acotada. Esto explica por qué podemos controlar factores como

$$(-1)^n, \qquad \sigma n, \qquad \kappa n,$$

cuando aparecen multiplicados por una sucesión que tiende a cero.

Una red mínima de límites que ya podemos reutilizar. Podemos reunir lo construido hasta ahora sin convertirlo en una lista aislada de recetas. Sabemos que $\frac{1}{n+1} \to 0$, que para cada $p \in \mathbb{N}_{>0}$, $\frac{1}{(n+1)^p} \to 0$, y que si $|q| < 1$, $q^n \to 0$. Además, mediante álgebra y sándwich podemos obtener inmediatamente límites de expresiones construidas con estas piezas.

Por ejemplo, $\frac{7}{(n+1)^3} \to 0$,

$$\left(-\frac{4}{5}\right)^n \to 0,$$

y $\frac{2+\sigma n}{n+1} \to 0$, porque $|2 + \sigma n| \leq 3$.

La meta de esta sección no era acumular ejemplos, sino disponer de unas pocas **unidades de construcción** cuyo origen conocemos.

Qué no hemos utilizado. Es importante hacer explícita la frontera lógica de los argumentos. Para obtener los límites de esta sección usamos:

- la propiedad arquimediana;
- inducción matemática;
- desigualdades elementales;
- álgebra de límites;
- preservación del orden;
- teorema del sándwich;
- valor absoluto.

No utilizamos:

- límites de funciones;
- continuidad;
- función exponencial real;
- logaritmos;
- derivadas;
- regla de l'Hôpital;
- series;
- subsucesiones;
- Bolzano–Weierstrass;
- criterio de Cauchy.

Esta auditoría importa porque el siguiente gran paso del capítulo será cualitativamente distinto. Hasta ahora, en todos los límites que hemos demostrado, **ya conocíamos o sospechábamos el candidato** y después verificábamos que los términos se acercaban a él. En §3.7 preguntaremos algo más profundo:

> ¿puede la estructura de $\mathbb{R}$ garantizar que una sucesión converge aunque todavía no conozcamos de antemano su límite?

La respuesta será afirmativa para sucesiones monótonas y acotadas. Allí aparecerá por primera vez en este capítulo la completitud de $\mathbb{R}$ como mecanismo directo de existencia de límites.

3.7. Convergencia de sucesiones monótonas y acotadas

Hasta ahora, para demostrar que una sucesión converge, hemos seguido casi siempre este esquema:

1. proponemos o descubrimos un candidato L;
2. demostramos que $a_n \to L$ mediante la definición, el álgebra o el sándwich.

Ese procedimiento responde bien a preguntas del tipo

«¿cuál es el límite de esta sucesión?»

cuando ya disponemos de una expresión que permite adivinar el candidato. Pero existe otra pregunta, más estructural:

¿podemos demostrar que un límite existe sin conocer primero su valor?

La respuesta será afirmativa cuando una sucesión combine dos tipos de control que hasta ahora hemos mantenido separados:

- **monotonía**, que impide retrocesos en una dirección de orden;
- **acotación**, que impide avanzar indefinidamente en esa misma dirección.

La afirmación decisiva no será una propiedad puramente algebraica. Para convertir esos dos controles en la existencia de un número real que actúe como límite necesitaremos la **completitud de** $\mathbb{R}$. Esta es la cadena que debemos hacer visible:

$$\boxed{\text{monotonía} + \text{cota} + \text{completitud} \longrightarrow \sup / \inf \longrightarrow \text{límite.}}$$

Monotonía para sucesiones. Como una sucesión es una función cuyo dominio está ordenado, podemos trasladar directamente el lenguaje de monotonía desarrollado en `T1-C03`. Conviene, sin embargo, fijar la notación específica que usaremos a partir de ahora.

Definición 3.4. **Sucesiones crecientes, decrecientes y monótonas.** Sea $(a_n)_{n\geq n_0}$ una sucesión real.

Diremos que (a_n) es **creciente** si

$$n < m \quad \Longrightarrow \quad a_n \leq a_m.$$

Diremos que es **decreciente** si

$$n < m \quad \Longrightarrow \quad a_n \geq a_m.$$

La sucesión es **monótona** si es creciente o decreciente. Diremos que es **estrictamente creciente** si

$$n < m \quad \Longrightarrow \quad a_n < a_m,$$

y **estrictamente decreciente** si

$$n < m \quad \Longrightarrow \quad a_n > a_m.$$

En este libro, por tanto, las palabras **creciente** y **decreciente** se usan en sentido débil: se permite la igualdad entre términos distintos. Cuando necesitemos excluirla diremos explícitamente **estrictamente** creciente o decreciente. Por ejemplo, una sucesión constante es a la vez creciente y decreciente. Esta convención es importante porque el teorema que demostraremos no necesita monotonía estricta. Para verificar monotonía suele bastar comparar términos consecutivos. Si $a_n \leq a_{n+1}$ para todo $n \geq n_0$, entonces, dados $n < m$, podemos encadenar

$$a_n \leq a_{n+1} \leq \cdots \leq a_m,$$

y por tanto $a_n \leq a_m$. El argumento para sucesiones decrecientes es análogo. Así, en la práctica, muchas pruebas de monotonía comenzarán estudiando el signo de $a_{n+1} - a_n$ o comparando directamente a_{n+1} con a_n.

Dos controles distintos: dirección y barrera. Una sucesión creciente satisface

$$a_{n_0} \leq a_{n_0+1} \leq a_{n_0+2} \leq \cdots .$$

Esto controla la **dirección**, pero no impide que los términos crezcan sin límite real finito. La sucesión $a_n = n$ es creciente y no está acotada superiormente. Una cota superior, en cambio, proporciona una **barrera**. Si existe $M \in \mathbb{R}$ tal que $a_n \leq M$ para todo n, los términos no pueden

atravesar M. Pero la acotación por sí sola no controla la dirección: $(-1)^n$ permanece entre -1 y 1 y, sin embargo, oscila para siempre.

Por separado, entonces,

$$\text{monotonía} \qquad \text{y} \qquad \text{acotación}$$

no garantizan convergencia.

Cuando los términos avanzan en una sola dirección y una cota les impide seguir creciendo en ella, la completitud de $\mathbb{R}$ proporciona un límite.

¿De dónde saldrá el candidato a límite? Supongamos que (a_n) es creciente y está acotada superiormente. Consideremos el conjunto de todos sus valores: $S = \{a_n : n \geq n_0\}$. El conjunto S es no vacío y está acotado superiormente. Por la completitud de $\mathbb{R}$, existe $L = \sup S$. Aquí aparece el candidato que antes no teníamos que adivinar. La elección no es arbitraria. Como L es una cota superior, $a_n \leq L$ para todo n. Y como es la **menor** cota superior, cualquier número un poco menor que L deja de ser cota superior. En particular, para todo $\varepsilon > 0$, $L - \varepsilon$ no puede ser una cota superior de S. Por la caracterización aproximativa del supremo demostrada en `T1-C02`, existe entonces algún término a_N tal que $L - \varepsilon < a_N \leq L$.

Hasta aquí hemos encontrado **un** término suficientemente próximo a L. Eso todavía no es convergencia: necesitamos controlar todos los términos posteriores. Aquí entra exactamente la monotonía. Si $n \geq N$, entonces $a_N \leq a_n$.

Por tanto,

$$L - \varepsilon < a_N \leq a_n \leq L < L + \varepsilon.$$

Y ahora sí toda la cola queda atrapada en el intervalo $(L - \varepsilon, L + \varepsilon)$. La prueba completa no necesita nada más.

Teorema 3.2. **Teorema de convergencia monótona.** Sea $(a_n)_{n \geq n_0}$ una sucesión real.

1. Si (a_n) es creciente y está acotada superiormente, entonces converge y
$$\lim_{n \to \infty} a_n = \sup\{a_n : n \geq n_0\}.$$
2. Si (a_n) es decreciente y está acotada inferiormente, entonces converge y
$$\lim_{n \to \infty} a_n = \text{ínf}\{a_n : n \geq n_0\}.$$

En particular, toda sucesión monótona y acotada converge.

Demostración. Demostremos primero el caso creciente.

Sea

$$S = \{a_n : n \geq n_0\}.$$

Como la sucesión está acotada superiormente, S es un conjunto no vacío y acotado superiormente. Por completitud de $\mathbb{R}$, existe

$$L = \sup S.$$

Queremos demostrar que $a_n \to L$.

Sea $\varepsilon > 0$. Como L es el supremo de S, el número

$$L - \varepsilon$$

no es una cota superior de S. Por tanto existe algún índice $N \geq n_0$ tal que

$$L - \varepsilon < a_N.$$

Como L sí es una cota superior de S,

$$a_n \leq L$$

para todo $n \geq n_0$.
Además, como la sucesión es creciente, para todo $n \geq N$ tenemos

$$a_N \leq a_n.$$

Combinando las tres desigualdades,

$$L - \varepsilon < a_N \leq a_n \leq L < L + \varepsilon.$$

Así, para todo $n \geq N$,

$$L - \varepsilon < a_n < L + \varepsilon,$$

lo cual equivale a

$$|a_n - L| < \varepsilon.$$

Como $\varepsilon > 0$ era arbitrario,

$$a_n \to L.$$

Por tanto,

$$\lim_{n\to\infty} a_n = \sup\{a_n : n \geq n_0\}.$$

Consideremos ahora el caso decreciente. Sea nuevamente

$$S = \{a_n : n \geq n_0\}.$$

Como la sucesión está acotada inferiormente, S es no vacío y acotado inferiormente. La completitud de $\mathbb{R}$, mediante la existencia de ínfimos ya deducida en `T1-C02`, garantiza que existe

$$L = \text{ínf}\, S.$$

Sea $\varepsilon > 0$. Como L es el ínfimo de S, el número

$$L + \varepsilon$$

no es una cota inferior de S. Por tanto existe algún índice $N \geq n_0$ tal que

$$a_N < L + \varepsilon.$$

Como L es una cota inferior,

$$L \leq a_n$$

para todo $n \geq n_0$.

Y como la sucesión es decreciente, si $n \geq N$ entonces

$$a_n \leq a_N.$$

En consecuencia,

$$L - \varepsilon < L \leq a_n \leq a_N < L + \varepsilon.$$

Por tanto,

$$|a_n - L| < \varepsilon$$

para todo $n \geq N$. Como $\varepsilon > 0$ era arbitrario,

$$a_n \to L.$$

Así,

$$\lim_{n\to\infty} a_n = \text{ínf}\{a_n : n \geq n_0\}.$$

Quedan demostradas ambas afirmaciones.

Dónde se usa exactamente cada hipótesis. Las hipótesis cumplen funciones diferentes en la prueba: una permite definir el supremo y otra extiende a toda una cola la aproximación obtenida en un término.

En el caso creciente:

1. **Acotación superior.** Garantiza que el conjunto $S = \{a_n : n \geq n_0\}$ tiene cotas superiores.
2. **Completitud de $\mathbb{R}$.** Garantiza que entre esas cotas existe una menor: $L = \sup S \in \mathbb{R}$.
3. **Propiedad aproximativa del supremo.** Para cada $\varepsilon > 0$ produce un índice N con $L - \varepsilon < a_N$.
4. **Monotonía creciente.** Propaga esa aproximación puntual a toda la cola:
$$n \geq N \quad \Longrightarrow \quad a_N \leq a_n.$$
5. **Propiedad de cota superior del supremo.** Impide que los términos atraviesen L: $a_n \leq L$.

La arquitectura es, por tanto,

$$\boxed{\begin{array}{c} \text{acotación superior} \\ + \\ \text{completitud} \\ \Longrightarrow L = \sup S \\ \\ \Longrightarrow \begin{array}{c} \text{un término entra cerca de } L \\ + \\ \text{monotonía} \end{array} \\ \\ \Longrightarrow \text{toda la cola entra cerca de } L. \end{array}}$$

El caso decreciente reproduce el mismo mecanismo con ínfimos y orden invertido.

La completitud no está escondida: es el motor de existencia. En las primeras pruebas ε–N del capítulo, la propiedad arquimediana nos permitía fabricar umbrales. En el teorema anterior ocurre algo distinto: antes de buscar un umbral necesitamos **producir el propio candidato a límite**.

Ese candidato aparece como $\sup S$ o $\inf S$.

La afirmación de que esos números existen en $\mathbb{R}$ no procede de la monotonía. Procede de la completitud. Esto permite localizar con precisión el punto donde fallaría la misma demostración si trabajáramos únicamente en $\mathbb{Q}$. Un conjunto no vacío y acotado de racionales puede carecer de supremo racional, como vimos en `T1-C02`. En ese caso, el paso

$$S \text{ acotado superiormente} \quad \Longrightarrow \quad \sup S \text{ existe en el mismo sistema numérico}$$

no estaría disponible.

Por eso el teorema de convergencia monótona no es simplemente una consecuencia del orden: es una manifestación secuencial de la completitud de los números reales. Podemos condensar esta segunda gran aparición operativa de completitud en el libro como

$$\boxed{\text{completitud} \longrightarrow \sup / \inf \longrightarrow \text{convergencia monótona.}}$$

Qué falla si quitamos una hipótesis. El teorema necesita las dos piezas de control.

Sin acotación suficiente. La sucesión $a_n = n$ es creciente, pero no está acotada superiormente. Tampoco puede converger a un número real: si convergiera, Proposición 3.2 implicaría que sería acotada, contradiciendo la propiedad arquimediana.

Así,

$$\boxed{\text{creciente} \not\Longrightarrow \text{convergente.}}$$

La versión decreciente presenta el mismo problema si falta una cota inferior.

Sin monotonía. La sucesión $b_n = (-1)^n$ es acotada, pero Ejemplo 3.4 demostró que diverge.

Por tanto,

$$\boxed{\text{acotada} \not\Longrightarrow \text{convergente.}}$$

Lo que sí obtenemos al combinar ambas condiciones es

$$\boxed{\text{monótona} + \text{acotada} \Longrightarrow \text{convergente.}}$$

Los dos contraejemplos muestran qué aporta cada hipótesis: la monotonía controla la dirección y la cota impide que el avance en esa dirección supere cualquier número real.

Una cota en la dirección correcta es suficiente. La formulación «monótona y acotada» es cómoda, pero la demostración revela una versión más precisa. Para una sucesión **creciente** solo necesitamos una cota **superior**. Ya existe automáticamente una cota inferior: el primer término satisface $a_{n_0} \leq a_n$ para todo $n \geq n_0$. Del mismo modo, para una sucesión **decreciente** solo necesitamos una cota **inferior**, porque $a_n \leq a_{n_0}$ para todo $n \geq n_0$. Así, las hipótesis realmente ajustadas son

$$\boxed{\begin{aligned} \text{creciente} + \text{acotada superiormente} &\Longrightarrow \text{convergente,} \\ \text{decreciente} + \text{acotada inferiormente} &\Longrightarrow \text{convergente.} \end{aligned}}$$

Esta precisión será útil en sucesiones recursivas: no necesitaremos encerrar siempre todos los términos entre dos cotas nuevas; bastará demostrar monotonía y encontrar una barrera en la dirección hacia la que avanzan.

El teorema demuestra existencia antes de identificar un número sencillo. Consideremos una sucesión creciente y acotada superiormente. El teorema garantiza que $L = \sup\{a_n : n \geq n_0\}$ es su límite, incluso si todavía no disponemos de una fórmula elemental para L. Este cambio de perspectiva es importante. Hasta ahora el candidato solía aparecer antes que la prueba de convergencia. A partir de ahora podremos invertir el orden lógico:

primero demostrar que el límite existe $\longrightarrow$ después identificarlo.

En una sucesión recursiva, por ejemplo, la monotonía y una cota podrán darnos existencia mediante el teorema de convergencia monótona. Solo después tendrá sentido utilizar la recurrencia y el álgebra de límites para deducir una ecuación satisfecha por el límite. Resolver primero esa ecuación y declarar que la sucesión converge sería invertir indebidamente la lógica.

Monotonía eventual también basta. La convergencia depende de una cola. Por ello, si una sucesión no es monótona desde su primer término pero existe un índice N_0 tal que la cola $(a_n)_{n\geq N_0}$ es creciente y está acotada superiormente, el teorema de convergencia monótona se aplica a esa cola. La cola converge y, por Proposición 3.1, la sucesión completa converge al mismo límite. Análogamente, una cola decreciente y acotada inferiormente basta. Esta observación evita exigir regularidad innecesaria a una cantidad finita de términos iniciales.

La herramienta que necesitábamos para las recurrencias. Con §3.7 disponemos por primera vez de un criterio de **existencia de límite** que no requiere conocer de antemano el valor del límite. El protocolo que emerge es:

demostrar monotonía +encontrar una cota en la dirección correcta $\Longrightarrow$ obtener convergencia por completitud.

En §3.8 aplicaremos exactamente este mecanismo a una sucesión definida recursivamente. Allí habrá que separar cinco preguntas:

1. ¿la recurrencia está bien definida?;
2. ¿qué región permanece invariante?;
3. ¿la sucesión es monótona?;
4. ¿por qué existe un límite?;
5. una vez establecida la existencia, ¿cómo identificamos ese límite sin invocar continuidad?

El cuarto paso será una aplicación directa del Teorema 3.2. El quinto utilizará el desplazamiento de índices y el álgebra de límites ya demostrados.

3.8. Sucesiones recursivas: fabricar y controlar un límite

Las sucesiones estudiadas hasta ahora estaban dadas, en su mayoría, por una fórmula que permitía calcular directamente el término de índice n. Por ejemplo, $a_n = \frac{1}{n+1}$ nos entrega a_n sin necesidad de conocer a_{n-1}. Una **regla recursiva** funciona de otra manera. En lugar de decirnos directamente cuánto vale el término de índice n, nos indica cómo fabricar el término siguiente a partir de uno o varios términos anteriores.

Por ejemplo,

$$a_0 = 1, \qquad a_{n+1} = \sqrt{2 + a_n}$$

produce sucesivamente $a_1 = \sqrt{3}$,

$$a_2 = \sqrt{2 + \sqrt{3}},$$

$$a_3 = \sqrt{2 + \sqrt{2 + \sqrt{3}}},$$

y así indefinidamente, siempre que cada paso esté bien definido. Esta diferencia cambia también la estrategia de estudio. Una fórmula explícita puede permitirnos comparar directamente a_n con un candidato de límite. En una recurrencia, en cambio, la información principal suele estar escondida en la relación entre términos consecutivos. La pregunta deja de ser solamente

> ¿qué expresión tiene a_n?,

y pasa a ser

> **¿qué propiedades conserva el procedimiento que transforma a_n en a_{n+1}?**

La información sobre términos consecutivos permitirá buscar cotas, probar monotonía y decidir si se cumplen las hipótesis del teorema de convergencia monótona.

Una regla recursiva no garantiza por sí sola una sucesión real infinita. Es fácil escribir una recurrencia sintácticamente correcta y olvidar que cada término posterior debe existir dentro del sistema numérico en el que estamos trabajando. Por ejemplo, si intentáramos definir

$$c_0 = -3, \qquad c_{n+1} = \sqrt{2 + c_n},$$

el primer paso exigiría $c_1 = \sqrt{-1}$, que no es un número real. Por tanto, la expresión $c_{n+1} = \sqrt{2 + c_n}$ no basta por sí sola para garantizar una sucesión real definida para todo $n \in \mathbb{N}$. Debemos verificar que, cada vez que aplicamos la regla, el nuevo término vuelve a quedar dentro del dominio donde el siguiente paso tiene sentido. Esta observación es una versión dinámica de una pregunta ya conocida desde `T1-C03`:

antes de usar una regla, debemos comprobar que está bien definida sobre los objetos que realmente aparecen.

En una recurrencia esa comprobación suele hacerse encontrando una **región invariante**: un conjunto de valores del cual la regla no consigue sacar a la sucesión. Podemos justificar la invariancia mediante inducción:

> si demostramos que a_n pertenece a cierto intervalo y que esa pertenencia obliga a que a_{n+1} permanezca en el mismo intervalo, la inducción controla todos los términos.

Cinco preguntas para auditar una recurrencia. A partir de ahora organizaremos muchos problemas recursivos mediante cinco preguntas.

1. **¿Está bien definida?**
   ¿Cada paso produce un número real permitido?
2. **¿Qué región permanece invariante?**
   ¿Podemos encerrar todos los términos en un intervalo o encontrar una barrera útil?
3. **¿Es monótona?**
   ¿Los términos avanzan sistemáticamente en una dirección?
4. **¿Por qué existe un límite?**
   Si obtenemos monotonía y una cota en la dirección correcta, podemos invocar Teorema 3.2.
5. **Una vez demostrada la existencia, ¿cómo identificamos el límite?**
   Aquí utilizaremos desplazamiento de índices y el álgebra de límites de §3.4.

La separación entre las preguntas 4 y 5 es esencial:

$$\boxed{\text{demostrar existencia del límite} \quad \neq \quad \text{resolver una ecuación que un eventual límite tendría que satisfacer}}.$$

Una ecuación puede producir **candidatos**. No produce por sí sola convergencia.

El ejemplo rector: un radical construido término a término. Analicemos completamente la sucesión

$$a_0 = 1, \qquad a_{n+1} = \sqrt{2 + a_n}.$$

Los primeros términos son aproximadamente

$$1, \quad 1.732, \quad 1.932, \quad 1.983, \quad \dots$$

Esta inspección sugiere dos conjeturas:

- los términos parecen crecer;
- parecen acercarse a 2 sin superarlo.

Pero una tabla finita no demuestra ninguna de las dos afirmaciones. Nuestro objetivo será convertir esas observaciones en una cadena deductiva completa.

Ejemplo 3.9. **Una recurrencia controlada desde el principio hasta el límite.** Sea (a_n) la sucesión definida por

$$a_0 = 1, \qquad a_{n+1} = \sqrt{2 + a_n}.$$

Demostraremos, en este orden, que

1. la sucesión está bien definida para todo n;
2. $1 \le a_n < 2$ para todo n;
3. (a_n) es estrictamente creciente;
4. (a_n) converge por el teorema de convergencia monótona;

5. su límite es 2.

Paso 1 y paso 2: bien definida y región invariante.

Demostraremos por inducción la afirmación

$$P(n): \qquad 1 \leq a_n < 2.$$

Caso base. Como $a_0 = 1$, se cumple inmediatamente $1 \leq a_0 < 2$.

Paso inductivo. Supongamos que para algún $n \in \mathbb{N}$ tenemos $1 \leq a_n < 2$.

Entonces $3 \leq 2 + a_n < 4$.

En particular, $2 + a_n > 0$, de modo que la raíz cuadrada real $a_{n+1} = \sqrt{2 + a_n}$ está bien definida.

Además, como $a_{n+1} \geq 0$ y $a_{n+1}^2 = 2 + a_n$, las desigualdades

$$1 < 3 \leq a_{n+1}^2 < 4 = 2^2$$

implican $1 < a_{n+1} < 2$.

Por tanto, $1 \leq a_{n+1} < 2$.

La inducción queda cerrada. Hemos demostrado simultáneamente que todos los pasos de la recurrencia tienen sentido en $\mathbb{R}$ y que

$$\boxed{1 \leq a_n < 2 \qquad \text{para todo } n \in \mathbb{N}.}$$

El intervalo $[1, 2)$ es, en este sentido, una región invariante para la recurrencia. En particular, 2 es una cota superior de la sucesión.

Paso 3: monotonía.

Queremos demostrar que $a_{n+1} > a_n$ para todo n. Ya sabemos que ambos números son no negativos. Por tanto, podemos comparar sus cuadrados.

Usando la recurrencia, $a_{n+1}^2 - a_n^2 = (2 + a_n) - a_n^2$.

Factorizamos:

$$\begin{aligned} 2 + a_n - a_n^2 &= -(a_n^2 - a_n - 2) \\ &= -(a_n - 2)(a_n + 1) \\ &= (2 - a_n)(a_n + 1). \end{aligned}$$

Por el paso anterior, $a_n < 2$ y $a_n \geq 1$, así que $2 - a_n > 0$ y $a_n + 1 > 0$.

Por consiguiente, $a_{n+1}^2 - a_n^2 = (2 - a_n)(a_n + 1) > 0$.

Luego $a_{n+1}^2 > a_n^2$.

Como $a_{n+1} \geq 0$ y $a_n \geq 0$, se sigue que $a_{n+1} > a_n$.

Así, (a_n) es estrictamente creciente.

Paso 4: existencia del límite.

Hemos demostrado dos propiedades: (a_n) es creciente y

$$a_n < 2 \qquad \text{para todo } n.$$

Por tanto, la sucesión es creciente y está acotada superiormente. Aplicando Teorema 3.2, existe un número real L tal que $a_n \to L$. Además, el teorema identifica abstractamente ese número

como $L = \sup\{a_n : n \in \mathbb{N}\}$. Como cada término satisface $1 \leq a_n < 2$, tenemos en particular $1 \leq L \leq 2$. Hasta aquí hemos demostrado **existencia**. Todavía no hemos demostrado que $L = 2$.

Paso 5: identificación del límite sin continuidad.

Como $a_n \to L$, la invariancia por desplazamiento de índices establecida en §3.3 da $a_{n+1} \to L$. La recurrencia satisface, para todo n, $a_{n+1}^2 = 2 + a_n$. Ahora utilizamos únicamente el álgebra de límites ya demostrada. Como $a_{n+1} \to L$, la regla del producto implica $a_{n+1}^2 = a_{n+1}a_{n+1} \to L^2$. Y como $a_n \to L$, las reglas para constantes y sumas dan $2 + a_n \to 2 + L$. Pero las sucesiones (a_{n+1}^2) y $(2 + a_n)$ son idénticas término a término. Por unicidad del límite, $L^2 = 2 + L$.

Reordenando, $L^2 - L - 2 = 0$, y factorizando, $(L - 2)(L + 1) = 0$.

Los candidatos algebraicos son, por tanto, $L = 2$ o $L = -1$. Sin embargo, ya sabemos que $1 \leq L \leq 2$. La segunda posibilidad queda excluida. Luego

$$\boxed{\lim_{n\to\infty} a_n = 2.}$$

La demostración obtiene primero un límite mediante monotonía y acotación; después utiliza la recurrencia para determinar su valor. La ecuación $L^2 = L + 2$ solo aparece **después** de saber que existe un límite L. Antes de ese punto, la letra L no designa todavía un objeto cuya existencia esté demostrada.

Cómo se descubre la barrera 2. En una demostración final conviene presentar las ideas en orden lógico. Durante la búsqueda, sin embargo, podemos trabajar hacia atrás. Los primeros términos sugieren que 2 podría desempeñar un papel especial. Para comprobar si es una barrera estable, preguntamos:

> si $a_n < 2$, ¿la recurrencia obliga a que también $a_{n+1} < 2$?

La condición buscada es $\sqrt{2 + a_n} < 2$.

Como ambos lados son no negativos, podemos elevar al cuadrado y obtener la condición equivalente $2 + a_n < 4$, es decir, $a_n < 2$. La desigualdad que queremos conservar reaparece exactamente en el lado derecho. Este es un indicio muy fuerte de invariancia:

$$a_n < 2 \quad \Longrightarrow \quad a_{n+1} < 2.$$

La propia recurrencia permite comprobar la barrera sugerida por los primeros términos: si la desigualdad vale en un paso, vuelve a valer en el siguiente. La búsqueda de regiones invariantes suele funcionar así:

1. conjeturamos una región a partir de los primeros términos o de la forma algebraica de la regla;
2. suponemos que a_n pertenece a esa región;
3. utilizamos la recurrencia para demostrar que a_{n+1} vuelve a pertenecer;
4. cerramos el argumento por inducción.

Cómo se descubre la monotonía. También podemos reconstruir la búsqueda de la desigualdad $a_{n+1} > a_n$.

La recurrencia nos dice $a_{n+1} = \sqrt{2 + a_n}$.

Como ya sabemos que los términos son positivos, podemos comparar cuadrados:

$$a_{n+1} > a_n \quad \Longleftrightarrow \quad a_{n+1}^2 > a_n^2.$$

Sustituyendo la recurrencia, $2 + a_n > a_n^2$.

Llevando todo a un lado, $2 + a_n - a_n^2 > 0$.

Y la factorización $2+a_n-a_n^2 = (2-a_n)(a_n+1)$ revela por qué la región invariante encontrada en el paso anterior era exactamente la información que necesitábamos. Dentro de $1 \leq a_n < 2$, ambos factores son positivos. Así, la prueba de monotonía **reutiliza** la prueba de acotación. Las dos tareas no son independientes: encontrar una región adecuada puede proporcionar simultáneamente dominio, cotas y signo para la comparación entre términos consecutivos.

La ecuación de punto fijo es una condición necesaria, no una prueba de convergencia. Existe un error particularmente tentador con las recurrencias. Supongamos que escribimos $x_{n+1} = F(x_n)$ y queremos estudiar su comportamiento. Es natural pensar:

> si $x_n \to L$, entonces también $x_{n+1} \to L$; por tanto, tal vez L deba satisfacer $L = F(L)$.

La idea puede ser útil para **encontrar candidatos**, siempre que el paso algebraico utilizado para llegar a la ecuación esté justificado por resultados ya demostrados. El error consiste en invertir la implicación:

$$\boxed{L = F(L) \not\Longrightarrow x_n \to L.}$$

Veamos un ejemplo elemental. Definamos

$$y_0 = 1, \qquad y_{n+1} = 3 - y_n.$$

Si alguien supone sin demostrar que y_n tiene límite L, entonces el desplazamiento y el álgebra de límites obligarían a $L = 3 - L$, de donde $L = \frac{3}{2}$.

Pero calculemos los términos:

$$y_0 = 1, \qquad y_1 = 2, \qquad y_2 = 1, \qquad y_3 = 2, \qquad \dots$$

La sucesión oscila para siempre.

Más aún, para todo n,

$$\left| y_n - \frac{3}{2} \right| = \frac{1}{2}.$$

Tomando, por ejemplo, $\varepsilon = \frac{1}{4}$, ninguna cola satisface

$$\left| y_n - \frac{3}{2} \right| < \varepsilon.$$

Por tanto, $y_n \not\to \frac{3}{2}$.

Y como cualquier eventual límite tendría que satisfacer necesariamente $L = 3 - L$, no existe ningún otro candidato posible. La sucesión diverge. El número $3/2$ es un **punto fijo** de la regla $F(x) = 3 - x$, pero eso no obliga a que la órbita que comienza en 1 se acerque a él.

La lección lógica es:

$$\boxed{\begin{array}{ll} \text{convergencia demostrada} & \\ \quad + \text{ recurrencia} & \Longrightarrow \text{ecuación necesaria para el límite,} \\ \quad + \text{ leyes de límites} & \end{array}}$$

pero no al revés.

En nuestro ejemplo no pasamos el límite «a través de la raíz». Conviene aislar un detalle metodológico del Ejemplo 3.9. Podría resultar tentador escribir $a_{n+1} = \sqrt{2 + a_n}$ y, como $a_n \to L$, declarar inmediatamente $L = \sqrt{2 + L}$. Ese paso sería precisamente utilizar, sin haberlo demostrado todavía, que la función $x \mapsto \sqrt{x}$ preserva límites. Para justificar ese paso haría falta demostrar que la raíz preserva límites. Podemos resolver el problema con una identidad que solo involucre operaciones algebraicas ya estudiadas. En lugar de ello utilizamos la identidad algebraica $a_{n+1}^2 = 2 + a_n$. Cada operación que aparece ahora —producto, suma y constante— ya tiene una ley de límites demostrada en §3.4. Por eso la cadena legítima es

$$\boxed{a_n \to L \Longrightarrow a_{n+1} \to L \Longrightarrow a_{n+1}^2 \to L^2}$$

y, simultáneamente,

$$\boxed{a_n \to L \Longrightarrow 2 + a_n \to 2 + L.}$$

La igualdad término a término entre ambos lados obliga entonces a $L^2 = 2 + L$. No hemos apelado a continuidad de la raíz ni a un límite funcional.

Existencia e identificación son dos problemas distintos. El ejemplo rector puede resumirse en dos bloques.

Bloque A: demostrar existencia.

$$\boxed{\begin{array}{l} 1 \le a_n < 2 \\ a_{n+1} > a_n \end{array} \Longrightarrow \text{creciente + cota superior} \Longrightarrow a_n \to L.}$$

Aquí entra la completitud a través de Teorema 3.2.

Bloque B: identificar el número.

$$\boxed{a_n \to L \Longrightarrow a_{n+1} \to L \Longrightarrow L^2 = L + 2 \Longrightarrow L \in \{-1, 2\} \Longrightarrow L = 2.}$$

Aquí entran la estabilidad bajo desplazamiento, el álgebra de límites y la información de orden obtenida antes. Confundir ambos bloques produce uno de los errores más frecuentes en problemas de recurrencias.

El protocolo reutilizable. Ante una sucesión definida por $a_{n+1} = F(a_n)$, no buscaremos automáticamente una fórmula cerrada para a_n. Muchas recurrencias útiles pueden estudiarse sin ella.

El protocolo será:

$$\boxed{\begin{array}{c} \text{1. bien definida} \\ \downarrow \\ \text{2. región invariante / cota} \\ \downarrow \\ \text{3. monotonía} \\ \downarrow \\ \text{4. existencia por convergencia monótona} \\ \downarrow \\ \text{5. identificación mediante desplazamiento + álgebra} \end{array}}$$

No todos los problemas seguirán exactamente esta ruta. Algunas sucesiones no serán monótonas; otras necesitarán otra técnica. Pero cuando la recurrencia conserva una región y empuja los términos siempre en una misma dirección, este esquema debe ser una de las primeras estrategias que examinemos.

Qué hemos ganado. La recurrencia $a_{n+1} = \sqrt{2 + a_n}$ nos permitió reunir por primera vez casi todas las herramientas del capítulo en una sola demostración:

- inducción para conservar una región;
- orden para construir cotas;
- comparación algebraica para demostrar monotonía;
- completitud, a través del teorema de convergencia monótona, para obtener existencia;
- estabilidad de colas y desplazamientos para pasar de (a_n) a (a_{n+1});
- álgebra de límites para identificar el valor final;
- control lógico para no confundir un punto fijo con una prueba de convergencia.

Esta arquitectura aparecerá muchas veces en análisis numérico, ecuaciones y procesos iterativos, aunque más adelante dispongamos de herramientas mucho más potentes. Hasta ahora, sin embargo, todas las sucesiones convergentes del capítulo tenían un límite real finito. La siguiente pregunta será qué lenguaje utilizar cuando los términos crecen más allá de toda cota real o disminuyen por debajo de toda cota real. Para describir ese comportamiento habrá que exigir que una cola completa supere cualquier barrera fijada. La exigencia distinguirá los **límites infinitos** de la mera ausencia de una cota.

3.9. Cuando no hay límite real: límites infinitos y divergencia

Hasta ahora, cuando una sucesión convergía, buscábamos un número real L alrededor del cual terminara concentrándose toda la cola. Pero hay sucesiones cuyo comportamiento final es ordenado y describible aunque no se acerquen a ningún número real. El ejemplo más simple es $a_n = n$. Los términos $0, 1, 2, 3, \ldots$ no se aproximan a un real fijo. Al contrario: sobrepasan cualquier barrera real que fijemos, siempre que avancemos lo suficiente. Podemos precisar ese crecimiento mediante una afirmación cuantificada: **dada cualquier barrera real, existe una cola completa cuyos términos quedan por encima de ella**.

Este patrón es análogo a la definición ε–N, pero cambia el tipo de desafío. Allí alguien fijaba una tolerancia alrededor de un candidato L; aquí alguien fija una barrera real M que debemos superar de manera permanente.

Superar cualquier barrera.

Definición 3.5. **Límites infinitos de una sucesión.** Sea $(a_n)_{n \geq n_0}$ una sucesión real. Diremos que

$$a_n \to +\infty$$

si para todo número real M existe un índice $N \geq n_0$ tal que

$$a_n > M$$

para todo $n \geq N$. En símbolos,

$$\forall M \in \mathbb{R}\ \exists N \geq n_0\ \forall n \geq N : \quad a_n > M.$$

Diremos que

$$a_n \to -\infty$$

si para todo número real M existe un índice $N \geq n_0$ tal que

$$a_n < M$$

para todo $n \geq N$. En símbolos,

$$\forall M \in \mathbb{R}\ \exists N \geq n_0\ \forall n \geq N: \quad a_n < M.$$

La estructura lógica debe leerse con el mismo cuidado que en la convergencia ordinaria:

$$\boxed{\text{barrera arbitraria} \longrightarrow \text{umbral adecuado} \longrightarrow \text{toda la cola al otro lado de la barrera.}}$$

Para $+\infty$, la barrera se supera por arriba. Para $-\infty$, se supera por abajo.

El símbolo ∞ no nombra un número real. La notación

$$\lim_{n\to\infty} a_n = +\infty$$

es útil, pero debe interpretarse mediante Definición 3.5. No significa que exista un número real llamado $+\infty$ al cual se acerquen los términos.

En particular,

$$+\infty \notin \mathbb{R}, \qquad -\infty \notin \mathbb{R}.$$

Por eso, si $a_n \to +\infty$ o $a_n \to -\infty$, la sucesión **no converge en el sentido real ordinario** de Definición 3.2. Hay también una razón estructural inmediata. Si $a_n \to +\infty$, entonces la sucesión no está acotada superiormente: dada cualquier supuesta cota B, la definición con $M = B$ produce términos tardíos mayores que B. Análogamente, si $a_n \to -\infty$, no está acotada inferiormente. Pero Proposición 3.2 demostró que toda sucesión convergente a un número real es acotada. Por tanto,

$$\boxed{a_n \to \pm\infty \Longrightarrow (a_n) \text{ diverge como sucesión real.}}$$

La palabra **diverge**, sin embargo, es más amplia. Una sucesión puede divergir sin escapar ordenadamente hacia ninguno de los dos infinitos.

El ejemplo básico: $n \to +\infty$. Probemos la afirmación directamente desde la propiedad arquimediana de $\mathbb{R}$. Sea $M \in \mathbb{R}$ arbitrario. Por la propiedad arquimediana existe un número natural N tal que $N > M$.

Entonces, para todo $n \geq N$, $n \geq N > M$.

Por consiguiente,

$$\forall M \in \mathbb{R}\ \exists N \in \mathbb{N}\ \forall n \geq N: \quad n > M,$$

y por Definición 3.5,

$$\boxed{n \to +\infty.}$$

El argumento muestra con precisión qué papel desempeña la arquimedianidad: garantiza que ninguna barrera real puede quedar por encima de todos los naturales. De la misma manera,

$-n \to -\infty$, porque dada una barrera real M, basta elegir $N > -M$; entonces, si $n \geq N$, $-n \leq -N < M$.

No acotada no significa tender a infinito. Una sucesión que tiende a $+\infty$ es necesariamente no acotada superiormente. Pero la recíproca es falsa. Ser no acotada solo afirma que podemos encontrar términos arbitrariamente grandes; no dice que **todos los términos suficientemente tardíos** sean grandes. La diferencia vuelve a estar en los cuantificadores. No estar acotada superiormente significa

$$\forall M \in \mathbb{R}\ \exists n \geq n_0 : \quad a_n > M.$$

Tender a $+\infty$ exige mucho más:

$$\forall M \in \mathbb{R}\ \exists N \geq n_0\ \forall n \geq N : \quad a_n > M.$$

En la primera afirmación encontramos **algún término** que supera la barrera. En la segunda encontramos **una cola completa** que queda por encima de ella.

Ejemplo 3.10. **No acotación frente a límite infinito.** Consideremos $a_n = (-1)^n n$.

La sucesión toma la forma $0, -1, 2, -3, 4, -5, \ldots$

Es no acotada superiormente y no acotada inferiormente. Sin embargo, $a_n \not\to +\infty$ y $a_n \not\to -\infty$.

No está acotada superiormente. Sea $B \in \mathbb{R}$. Por arquimedianidad podemos elegir un número natural m tan grande que $2m > B$.

Para el índice par $n = 2m$, $a_{2m} = 2m > B$.

Así, ninguna cota superior real puede controlar todos los términos.

No está acotada inferiormente. Dado $B \in \mathbb{R}$, elijamos m con $2m + 1 > -B$.

Entonces $a_{2m+1} = -(2m + 1) < B$.

Por tanto, tampoco existe una cota inferior real.

No tiende a $+\infty$. Fijemos la barrera $M = 0$.

Para que $a_n \to +\infty$, debería existir un índice N después del cual todos los términos fueran positivos. Pero, dado cualquier N, podemos elegir un índice impar $n \geq N$. Para ese índice, $a_n = -n < 0$. Así, ninguna cola queda enteramente por encima de 0.

No tiende a $-\infty$. De nuevo usemos la barrera $M = 0$. Dado cualquier N, podemos elegir un índice par $n \geq N$. Entonces $a_n = n \geq 0$, de modo que ninguna cola queda enteramente por debajo de 0.

La sucesión escapa a cotas arbitrariamente grandes en ambos sentidos, pero sigue cruzando una y otra vez de un lado al otro. No existe una dirección unilateral de escape. El ejemplo separa tres ideas:

$$\boxed{\text{no acotada} \not\Longrightarrow +\infty, \qquad \text{no acotada} \not\Longrightarrow -\infty.}$$

El recíproco convierte crecimiento infinito en pequeñez. Hay una conexión muy útil entre los límites infinitos y los límites reales ya estudiados. Para números positivos, afirmar que a_n termina superando cualquier barrera equivale a afirmar que su recíproco termina siendo tan pequeño como queramos.

Proposición 3.6. **Recíprocos y crecimiento hacia $+\infty$.** Sea $(a_n)_{n \geq n_0}$ una sucesión real que es eventualmente positiva. Entonces

$$a_n \to +\infty \quad \Longleftrightarrow \quad \frac{1}{a_n} \to 0.$$

Demostración. Supongamos primero que $a_n \to +\infty$.

Sea $\varepsilon > 0$. Como $1/\varepsilon$ es un número real, la definición de límite infinito proporciona un índice N_1 tal que $a_n > \frac{1}{\varepsilon}$ para todo $n \geq N_1$. Además, por positividad eventual existe N_0 tal que $a_n > 0$ para todo $n \geq N_0$.

Si $N = \text{máx}\{N_0, N_1\}$, entonces, para $n \geq N$, $0 < \frac{1}{a_n} < \varepsilon$.

Por tanto,

$$\left|\frac{1}{a_n} - 0\right| < \varepsilon,$$

y concluimos $\frac{1}{a_n} \to 0$.

Recíprocamente, supongamos que $\frac{1}{a_n} \to 0$ y que $a_n > 0$ eventualmente. Queremos demostrar que toda barrera real termina siendo superada.

Sea $M \in \mathbb{R}$.

Si $M \leq 0$, la positividad eventual basta: existe N_0 tal que, para $n \geq N_0$, $a_n > 0 \geq M$.

Si $M > 0$, tomemos $\varepsilon = \frac{1}{M}$.

Como $1/a_n \to 0$, existe N_1 tal que, para todo $n \geq N_1$,

$$\left|\frac{1}{a_n}\right| < \frac{1}{M}.$$

Tomando además una cola donde $a_n > 0$, obtenemos $0 < \frac{1}{a_n} < \frac{1}{M}$. Como ambos lados son positivos, el orden de los recíprocos se invierte y resulta $a_n > M$. En ambos casos existe un umbral después del cual $a_n > M$. Como M era arbitrario, $a_n \to +\infty$.

■

La positividad eventual determina la dirección del crecimiento. Al retirar esa hipótesis, la condición $\frac{1}{a_n} \to 0$ solo informa que $|a_n|$ se hace grande en magnitud; no selecciona hacia qué lado de la recta escapan los términos. El ejemplo $a_n = (-1)^n n$ lo muestra de inmediato: su recíproco, a partir de $n = 1$, satisface $\frac{1}{a_n} = \frac{(-1)^n}{n} \to 0$, pero (a_n) no tiende ni a $+\infty$ ni a $-\infty$. Por simetría de signo, una versión análoga relaciona $a_n \to -\infty$ con $1/a_n \to 0$ cuando la sucesión es eventualmente negativa. No necesitamos introducir un nuevo principio: basta aplicar el resultado anterior a $(-a_n)$.

Divergir tiene varias formas. A esta altura conviene ordenar el vocabulario.

Una sucesión puede:

- converger a un número real L;
- tender a $+\infty$;
- tender a $-\infty$;
- divergir sin tener ninguno de esos comportamientos, como $(-1)^n$ o $(-1)^n n$.

Así, decir simplemente

> «la sucesión diverge»

no especifica **cómo** falla la convergencia real.

En particular,

$$\boxed{\text{divergente} \not\Longrightarrow \text{límite infinito.}}$$

Esta distinción será importante cuando más adelante clasifiquemos comportamientos de funciones y series.

No haremos álgebra formal con el símbolo ∞. Las leyes de Proposición 3.3 fueron demostradas para límites **reales**. No hemos demostrado todavía un cálculo general para expresiones que involucren $+\infty$ o $-\infty$. Y no podemos obtenerlo tratando esos símbolos como números. Por ejemplo, las tres parejas

$$a_n = n, \qquad b_n = -n,$$

$$c_n = 2n, \qquad b_n = -n,$$

y

$$a_n = n, \qquad d_n = -2n$$

tienen siempre un término que tiende a $+\infty$ y otro que tiende a $-\infty$. Sin embargo, $a_n + b_n = 0$,

$$c_n + b_n = n \to +\infty,$$

y $a_n + d_n = -n \to -\infty$.

Por tanto, la escritura informal $+\infty + (-\infty)$ no determina por sí sola un resultado. Más adelante aprenderemos a reconocer y resolver distintas **formas indeterminadas** bajo hipótesis adicionales. Por ahora basta fijar una disciplina:

$$\boxed{\begin{array}{ll} \text{los símbolos } \pm\infty & \text{se interpretan mediante definiciones cuantificadas, no} \\ & \text{mediante aritmética ficticia.} \end{array}}$$

Qué hemos ganado. Ya podemos distinguir cuatro niveles de comportamiento final:

$$\boxed{\begin{array}{c} \text{convergencia a un real} \\ \text{escape a } +\infty \text{ o } -\infty \\ \text{no acotación sin dirección unilateral} \\ \text{otras formas de divergencia} \end{array}}$$

La herramienta común sigue siendo la misma: controlar **toda una cola** mediante cuantificadores. Para un límite real, dada una tolerancia $\varepsilon > 0$, toda la cola debe entrar en un intervalo alrededor de L. Para $+\infty$, dada una barrera M, toda la cola debe quedar por encima de ella. Para $-\infty$, dada una barrera M, toda la cola debe quedar por debajo. Al resolver un problema tendremos que reconocer cuál de estas exigencias corresponde al comportamiento que buscamos demostrar y qué información permite controlar la cola. Los ejemplos siguientes reúnen esas decisiones con las herramientas algebraicas y de orden.

3.10. Laboratorio de límites secuenciales

Hasta aquí hemos construido las herramientas del capítulo una por una. Sabemos traducir «acercarse» a una afirmación ε–N; negar correctamente una convergencia; usar unicidad, acotación, álgebra y orden; reconocer cuándo conviene aplicar el teorema del sándwich; obtener convergencia a partir de monotonía y completitud; analizar recurrencias; y distinguir un límite real de un escape hacia $+\infty$ o $-\infty$. Las seis estaciones reúnen problemas en los que debemos elegir una herramienta, comprobar sus hipótesis y justificar los pasos que una solución demasiado rápida podría ocultar. Conviene adoptar este protocolo general:

leer el objetivo
→ identificar la estructura de la cola
→ elegir la herramienta
→ probar
→ auditar hipótesis y dependencias

La pregunta ya no será solamente

> ¿puedo calcular este límite?,

sino también

> **¿qué teorema o definición convierte la información disponible en control de toda una cola?**

Estación 1 — Convertir un candidato en una prueba ε–N. Consideremos

$$a_n = \frac{4n-1}{2n+3}, \qquad n \geq 0.$$

Una manipulación informal de los términos dominantes sugiere $\frac{4n-1}{2n+3} \approx \frac{4n}{2n} = 2$.

Problema. Demostrar directamente desde la definición que $a_n \to 2$, fabricando un umbral $N(\varepsilon)$ y señalando dónde aparece cada cuantificador.

Diagnóstico. El candidato está dado. Por tanto, la definición nos exige controlar $|a_n - 2|$. Antes de escribir «sea $\varepsilon > 0$», conviene hacer el cálculo de búsqueda: simplificar el error hasta descubrir qué condición sobre n bastaría para hacerlo menor que ε.

Búsqueda del umbral. Tenemos

$$\begin{aligned}\left|\frac{4n-1}{2n+3} - 2\right| &= \left|\frac{4n-1-2(2n+3)}{2n+3}\right| \\ &= \left|\frac{-7}{2n+3}\right| \\ &= \frac{7}{2n+3}.\end{aligned}$$

Queremos garantizar $\frac{7}{2n+3} < \varepsilon$.

Sería suficiente imponer una condición un poco más fuerte y más cómoda: $2n > \frac{7}{\varepsilon}$, es decir, $n > \frac{7}{2\varepsilon}$. La propiedad arquimediana nos permite producir un natural que supere esa cantidad real.

Demostración. Sea $\varepsilon > 0$ arbitrario. Por la propiedad arquimediana, elegimos $N \in \mathbb{N}$ tal que $N > \frac{7}{2\varepsilon}$.

Sea ahora $n \geq N$. Entonces $2n + 3 > 2n \geq 2N > \frac{7}{\varepsilon}$.

Como todas las cantidades son positivas, $\frac{7}{2n+3} < \varepsilon$.

Por consiguiente,

$$\left|\frac{4n-1}{2n+3} - 2\right| < \varepsilon$$

para todo $n \geq N$. Como $\varepsilon > 0$ era arbitrario,

$$\boxed{\frac{4n-1}{2n+3} \to 2.}$$

Lectura de la solución. La prueba contiene exactamente la arquitectura $\forall\varepsilon > 0\ \exists N\ \forall n \geq N$.

- El adversario lógico elige ε.
- Nosotros respondemos con un N que puede depender de ε.
- Después ese mismo N debe controlar **todos** los índices posteriores.

La desigualdad preliminar $n > \frac{7}{2\varepsilon}$ no es todavía la demostración: es la **búsqueda** que nos permite diseñar el testigo existencial N. Podríamos obtener el mismo límite más rápidamente mediante el álgebra de límites de Proposición 3.3. Aquí hemos evitado deliberadamente ese atajo porque el objetivo de la estación era reconstruir la lógica elemental sobre la que descansa toda la teoría.

Estación 2 — Negar una convergencia sin cambiar el orden de los cuantificadores. Sea $b_n = (-1)^n + \frac{1}{n+1}$.

Algunos términos son positivos y otros negativos, y el término adicional $1/(n+1)$ se hace pequeño. Alguien podría mirar esa pequeñez y conjeturar erróneamente que $b_n \to 0$.

Problema. Refutar ese candidato directamente desde la negación de la definición de convergencia.

Diagnóstico. Para probar $b_n \not\to 0$ no debemos demostrar que «muchos términos» quedan lejos de cero. Necesitamos una tolerancia fija $\varepsilon_0 > 0$ que ninguna cola consiga satisfacer por completo:

$$\exists\varepsilon_0 > 0\ \forall N\ \exists n \geq N : |b_n| \geq \varepsilon_0.$$

La alternancia sugiere elegir siempre un índice par suficientemente tardío.

Solución. Tomemos $\varepsilon_0 = \frac{1}{2}$.

Sea $N \in \mathbb{N}$ arbitrario. Elegimos $n = 2N$.

Entonces n es par y además $n \geq N$. Por tanto, $(-1)^n = 1$, y $b_n = 1 + \frac{1}{2N+1} > 1$.

En particular,

$$|b_n| > 1 > \frac{1}{2} = \varepsilon_0.$$

Así, para todo umbral N encontramos un índice $n \geq N$ que queda fuera del intervalo

$$\left(-\frac{1}{2}, \frac{1}{2}\right).$$

Por la negación de la definición,

$$\boxed{b_n \not\to 0.}$$

Qué demuestra y qué no demuestra. La conclusión excluye **el candidato** 0. Por sí sola no demuestra todavía que (b_n) diverja, porque podría existir otro número real como límite. Esta separación reproduce una distinción de §3.2:

$$\boxed{\text{refutar un candidato} \neq \text{refutar todos los candidatos.}}$$

La estación se concentra en una habilidad lógica más básica: leer correctamente $\forall N\ \exists n \geq N$ como «por lejos que avancemos, todavía reaparece una violación», y no confundirla con $\exists N\ \forall n \geq N$, que expresa estabilidad de toda una cola.

Estación 3 — Abrir una prueba algebraica y encontrar los controles ocultos. Supongamos que

$$x_n \to 2 \qquad \text{y} \qquad y_n \to -1.$$

El álgebra de límites permite concluir inmediatamente $x_n y_n \to -2$ y $\frac{x_n}{y_n} \to -2$. Pero ahora queremos mirar **dentro** de esas dos conclusiones.

Problema. Reconstruir qué controles hacen posibles las pruebas del producto y del cociente. En particular:

1. explicar por qué la prueba del producto necesita una cota para uno de los factores;
2. explicar por qué la prueba del cociente necesita separar y_n de cero.

Primera reparación: controlar el producto. La diferencia respecto del límite esperado puede escribirse como

$$\begin{aligned} x_n y_n - (-2) &= x_n y_n + 2 \\ &= x_n(y_n + 1) - (x_n - 2). \end{aligned}$$

Por la desigualdad triangular,

$$|x_n y_n + 2| \leq |x_n|\,|y_n + 1| + |x_n - 2|.$$

Aquí aparece el punto delicado. Sabemos que $|y_n + 1| \to 0$, pero ese factor está multiplicado por $|x_n|$. Para convertir la pequeñez de $|y_n + 1|$ en pequeñez del producto necesitamos impedir que $|x_n|$ crezca sin control. Como $x_n \to 2$, existe una cola en la que $|x_n - 2| < 1$.

En esa cola, $|x_n| \leq |x_n - 2| + 2 < 3$.

Ya tenemos la cota que faltaba. Sea ahora $\varepsilon > 0$. Elegimos índices suficientemente grandes para que simultáneamente $|x_n| < 3$,

$$|y_n + 1| < \frac{\varepsilon}{6},$$

y $|x_n - 2| < \frac{\varepsilon}{2}$.

Tomando el máximo de los tres umbrales correspondientes, todas las condiciones valen en una misma cola. Entonces

$$\begin{aligned} |x_n y_n + 2| &\le |x_n|\,|y_n + 1| + |x_n - 2| \\ &< 3\frac{\varepsilon}{6} + \frac{\varepsilon}{2} \\ &= \varepsilon. \end{aligned}$$

La convergencia de (x_n) proporciona aquí una cota para $|x_n|$. Con esa cota fija podemos elegir la precisión exigida a $|y_n + 1|$ y controlar el producto en toda una cola.

Segunda reparación: controlar el cociente. Ahora examinemos $\frac{x_n}{y_n}$.

Antes de estimar el error debemos garantizar que el denominador no se acerque a cero. Como $y_n \to -1$, podemos exigir $|y_n + 1| < \frac{1}{2}$.

Entonces $-\frac{3}{2} < y_n < -\frac{1}{2}$, y por tanto $|y_n| > \frac{1}{2}$.

En esa cola el cociente está bien definido y además $\frac{1}{|y_n|} < 2$.

Ahora calculamos

$$\begin{aligned} \left|\frac{x_n}{y_n} + 2\right| &= \frac{|x_n + 2y_n|}{|y_n|} \\ &= \frac{|(x_n - 2) + 2(y_n + 1)|}{|y_n|} \\ &\le \frac{|x_n - 2| + 2|y_n + 1|}{|y_n|}. \end{aligned}$$

Sea $\varepsilon > 0$. En una cola común podemos exigir $|y_n + 1| < \frac{1}{2}$,

$$|x_n - 2| < \frac{\varepsilon}{4},$$

y $|y_n + 1| < \frac{\varepsilon}{8}$.

Entonces

$$\begin{aligned} \left|\frac{x_n}{y_n} + 2\right| &< 2\left(\frac{\varepsilon}{4} + 2\frac{\varepsilon}{8}\right) \\ &= \varepsilon. \end{aligned}$$

Por tanto, $\frac{x_n}{y_n} \to -2$.

Lectura de la solución. Las dos pruebas tienen obstáculos diferentes:

producto: controlar un factor que multiplica el error;
cociente: impedir que el denominador haga explotar el error.

El teorema algebraico de Proposición 3.3 comprime estos razonamientos para reutilizarlos. Esta estación muestra qué información permanece escondida cuando escribimos simplemente «por el álgebra de límites».

Estación 4 — Cuando la oscilación no necesita tener límite. Consideremos $u_n = \frac{2\sigma n - \kappa n}{n+1}$.

Problema. Determinar el límite sin suponer que las sucesiones (σn) o (κn) tengan límite.

Diagnóstico. La expresión contiene dos factores oscilantes. Intentar aplicar el álgebra de límites término a término nos obligaría a conocer primero límites para σn y κn. Pero no necesitamos esa información. Lo que sí sabemos es

$$|\sigma n| \leq 1, \qquad |\kappa n| \leq 1.$$

La pregunta correcta no es, por tanto,

> ¿a qué número se acercan las funciones σ y κ?,

sino

> ¿podemos encerrar toda la expresión entre cantidades cuya magnitud tienda a cero?

Eso apunta al teorema del sándwich.

Solución. Por desigualdad triangular, $|2\sigma n - \kappa n| \leq 2|\sigma n| + |\kappa n| \leq 3$.

Por tanto, $|u_n| = \frac{|2\sigma n - \kappa n|}{n+1} \leq \frac{3}{n+1}$.

Ya sabemos que $\frac{1}{n+1} \to 0$, así que, por multiplicación por el escalar 3, $\frac{3}{n+1} \to 0$. La forma absoluta del teorema del sándwich Proposición 3.5 da entonces

$$\boxed{u_n \to 0.}$$

Lectura de la solución. La estimación controla la magnitud de la oscilación en todos los índices de la cola:

$$\boxed{\text{factor oscilante acotado} \times \text{factor que tiende a } 0 \longrightarrow \text{producto que tiende a } 0.}$$

Esta arquitectura es más importante que el ejemplo periódico concreto. Cuando una parte de una expresión se resiste a tener un comportamiento simple, conviene preguntar si basta controlar su magnitud.

Estación 5 — Detectar cuándo la completitud produce el límite. Definamos

$$a_n = \sqrt{2 - \frac{1}{n+1}}, \qquad n \geq 0.$$

Problema. Demostrar que (a_n) converge y determinar su límite **sin utilizar continuidad de la raíz cuadrada**.

Diagnóstico. Es tentador escribir $2 - \frac{1}{n+1} \to 2$ y después afirmar inmediatamente

$$\sqrt{2 - \frac{1}{n+1}} \to \sqrt{2}.$$

Pero esa segunda inferencia sería precisamente una aplicación de continuidad de $x \mapsto \sqrt{x}$, noción que todavía no pertenece a este capítulo. Debemos separar las dos preguntas:

1. ¿por qué existe un límite real para (a_n)?;
2. una vez asegurada su existencia, ¿cómo identificamos su valor usando únicamente el álgebra ya demostrada?

La forma de la sucesión sugiere que sus términos crecen y permanecen por debajo de $\sqrt{2}$.

Monotonía y cota. Para todo $n \geq 0$, $0 < \frac{1}{n+1} \leq 1$, de modo que $1 \leq 2 - \frac{1}{n+1} < 2$.

Por tanto, $1 \leq a_n < \sqrt{2}$.

Así, (a_n) está acotada superiormente por $\sqrt{2}$.

Además, $\frac{1}{n+2} < \frac{1}{n+1}$, por lo que $2 - \frac{1}{n+2} > 2 - \frac{1}{n+1}$.

La raíz cuadrada es estrictamente creciente sobre $[0, \infty)$, de modo que $a_{n+1} > a_n$. La sucesión es, pues, creciente y acotada superiormente.

Existencia del límite. Por el teorema de convergencia monótona Teorema 3.2, existe $L \in \mathbb{R}$ tal que $a_n \to L$. Aquí está el paso en el que actúa la completitud de $\mathbb{R}$: el teorema de convergencia monótona se demostró tomando el supremo de la imagen de la sucesión. Además, como todos los términos satisfacen $a_n \geq 1$, el límite producido por el teorema es no negativo; de hecho, $1 \leq L \leq \sqrt{2}$.

Identificación sin continuidad. La definición de a_n nos da la identidad exacta $a_n^2 = 2 - \frac{1}{n+1}$. Como $a_n \to L$, el álgebra de límites implica $a_n^2 \to L^2$.

Por otra parte, $\frac{1}{n+1} \to 0$, y por tanto $2 - \frac{1}{n+1} \to 2$. Las dos sucesiones de ambos lados de la identidad son la misma término a término, así que sus límites deben coincidir. Por unicidad, $L^2 = 2$. Como $L \geq 0$, la unicidad de la raíz no negativa demostrada en `T1-C02` obliga a $L = \sqrt{2}$.

Concluimos

$$\boxed{\sqrt{2 - \frac{1}{n+1}} \to \sqrt{2}.}$$

Dónde estaba la completitud. La prueba tiene dos capas claramente separadas:

$$\boxed{\text{monotonía + cota + completitud} \to \text{existencia de } L \to \text{álgebra} \to \text{identificación de } L.}$$

La completitud asegura primero que existe un límite real. La identidad algebraica y la información de signo permiten después identificarlo, utilizando únicamente resultados ya demostrados.

Estación 6 — Auditar una recurrencia antes de resolver su ecuación de punto fijo. Sea

$$u_0 = 0, \qquad u_{n+1} = \frac{u_n + 3}{2}.$$

Una solución apresurada podría decir:

> «Si $u_n \to L$, entonces $L = (L+3)/2$, de donde $L = 3$. Por tanto, $u_n \to 3$.»

Problema. Localizar el fallo lógico y reparar completamente la demostración sin utilizar continuidad.

Diagnóstico del argumento defectuoso. La primera frase comienza con una hipótesis condicional: si $u_n \to L$. La ecuación $L = \frac{L+3}{2}$ solo puede servir para **identificar un candidato después de saber que el límite existe**. Resolverla no demuestra que la sucesión converja. La prueba correcta debe seguir el protocolo de §3.8:

$$\boxed{\text{bien definida} \to \text{región invariante} \to \text{monotonía} \to \text{existencia} \to \text{identificación}.}$$

Paso 1: la recurrencia está bien definida y preserva una región. Demostraremos por inducción que $0 \leq u_n < 3$ para todo $n \geq 0$. Para $n = 0$, $u_0 = 0$, de modo que la afirmación es verdadera.

Supongamos ahora $0 \leq u_n < 3$.

Entonces $3 \leq u_n + 3 < 6$, y al dividir por 2, $\frac{3}{2} \leq u_{n+1} < 3$.

En particular, $0 \leq u_{n+1} < 3$.

Por inducción, la región $[0, 3)$ es invariante bajo la recurrencia.

Paso 2: monotonía. Para todo n,

$$\begin{aligned} u_{n+1} - u_n &= \frac{u_n + 3}{2} - u_n \\ &= \frac{3 - u_n}{2}. \end{aligned}$$

Como $u_n < 3$, $u_{n+1} - u_n > 0$.

Así, (u_n) es estrictamente creciente.

Además, la región invariante ya nos dio la cota superior $u_n < 3$.

Paso 3: existencia por completitud. La sucesión es creciente y acotada superiormente. Por Teorema 3.2, $u_n \to L$ para algún $L \in \mathbb{R}$. Solo ahora tenemos derecho a buscar una ecuación para L.

Paso 4: identificación mediante desplazamiento y álgebra. Por la estabilidad bajo desplazamiento de índices Proposición 3.1, $u_{n+1} \to L$.

La recurrencia satisface $2u_{n+1} = u_n + 3$.

Aplicando únicamente el álgebra de límites, $2L = L + 3$.

Por tanto, $L = 3$.

Concluimos

$$\boxed{u_n \to 3.}$$

Lectura de la solución. La ecuación de punto fijo identifica el único candidato posible. Para que ese candidato sea el límite de la sucesión, la prueba tuvo que establecer antes la convergencia.

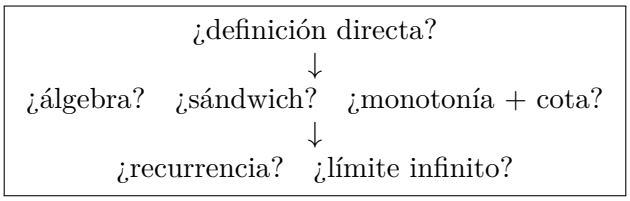
$$\boxed{\text{ecuación del candidato} \not\Longrightarrow \text{existencia del límite.}}$$

El orden correcto es

$$\boxed{\text{probar que el límite existe} \longrightarrow \text{usar la recurrencia para identificarlo.}}$$

Esta distinción será especialmente importante más adelante, cuando las recurrencias sean suficientemente complejas como para poseer varios puntos fijos o para no converger a ninguno de ellos.

Un mapa de decisiones para problemas de sucesiones. Las seis estaciones muestran que «calcular un límite» no designa una única técnica. Antes de hacer álgebra conviene diagnosticar qué tipo de información tenemos. Un primer mapa puede organizarse así:



Podemos convertirlo en preguntas operativas.

Si nos dan un candidato real L, preguntamos primero si podemos estimar directamente $|a_n - L|$.

Cuando el error se reduce a una expresión elemental en n, una prueba ε–N puede ser la ruta más transparente.

Si la sucesión está construida a partir de otras cuyos límites ya conocemos, el álgebra de Proposición 3.3 puede comprimir el trabajo. Pero debemos auditar sus hipótesis: en productos aparece acotación y en cocientes el denominador debe quedar eventualmente separado de cero.

Si aparece una oscilación acotada multiplicada por algo pequeño, conviene buscar una desigualdad en valor absoluto y aplicar el sándwich en vez de intentar asignar un límite a cada factor.

Si no conocemos el candidato pero la sucesión tiene una dirección de movimiento, preguntamos si es monótona y está acotada en la dirección adecuada. Si ambas respuestas son afirmativas, la completitud puede producir el límite mediante Teorema 3.2.

Si la sucesión está dada por recurrencia, no comenzamos resolviendo una ecuación de punto fijo. Primero debemos demostrar que la regla está bien definida, encontrar una región invariante o cotas, estudiar monotonía y obtener existencia. Solo después usamos desplazamiento de índices y álgebra para identificar el límite.

Si los términos parecen crecer sin cota, debemos distinguir no acotación de límite infinito. Para probar $a_n \to +\infty$ o $a_n \to -\infty$, la definición exige superar o descender cualquier barrera con **toda una cola**, no solo con términos arbitrariamente grandes en magnitud.

Si queremos refutar una convergencia, buscamos una tolerancia fija que reaparezca violada arbitrariamente lejos. La estructura correcta es $\exists \varepsilon_0 > 0\ \forall N\ \exists n \geq N$.

Qué debe poder hacer ahora el lector. Al terminar este laboratorio, una solución completa ya no debería consistir únicamente en una cadena de manipulaciones simbólicas. Antes de calcular, debemos ser capaces de decir:

- qué afirmación cuantificada intentamos probar o negar;
- qué información pertenece solo a un prefijo y qué información controla una cola;
- si estamos usando una definición o un teorema ya demostrado;
- cuáles son las hipótesis delicadas de ese teorema;
- si la existencia del límite está demostrada o solo hemos encontrado un candidato;
- dónde entra la completitud, cuando entra;
- y qué herramientas posteriores estamos evitando para no razonar circularmente.

El capítulo comenzó preguntando cómo expresar rigurosamente lo que ocurre «a partir de algún momento y para siempre después». Después de estas seis estaciones, esa pregunta ya dispone de un repertorio de respuestas diferentes pero coordinadas:

ε–N,
álgebra y orden,
sándwich,
monotonía + completitud,
recurrencias,
M–N para límites infinitos.

En los ejercicios se combinan estas dificultades. Para elegir una herramienta habrá que reconocer la información del enunciado, formular la afirmación que se quiere demostrar y comprobar cómo esa información permite controlar todos los términos de una cola.

3.11. Ejercicios y soluciones

Llegamos al banco final de `T1-C04`. Hasta aquí las técnicas aparecieron primero de manera aislada y después, en §3.10, como estaciones guiadas. En esta sección desaparece esa señalización: antes de comenzar una prueba habrá que decidir si el problema pide trabajar directamente con cuantificadores, usar álgebra de límites, estimar por sándwich, explotar monotonía y completitud, controlar una recurrencia o interpretar un límite infinito. El banco contiene exactamente cuarenta ejercicios, distribuidos en siete niveles:

$$7A + 7B + 7C + 6D + 5E + 5F + 3G = 40.$$

Los niveles no miden solo dificultad de cálculo. En particular, los niveles D y E obligan a distinguir afirmaciones verdaderas de recíprocas falsas, mientras que los niveles F y G exigen reconstruir por qué funcionan las herramientas del capítulo. Todos los problemas pueden resolverse con los capítulos publicados desde los números reales hasta el presente. No se necesitan límites de funciones, continuidad, derivadas, subsucesiones, Bolzano–Weierstrass, criterio de Cauchy ni $\limsup$ / $\liminf$.

Nivel A — Reconocimiento y comprensión. Ejercicio A1. Índices y términos. Sea

$$a_n = \frac{2n+1}{n+1}, \qquad n \geq 0.$$

1. Escribe los cuatro primeros términos.
2. Identifica el término de índice 3.
3. Explica la diferencia entre el índice 3 y el término a_3.

Ejercicio A2. Cola y eventualidad. Sea $a_n = \frac{(-1)^n}{n+1}$.
Decide cuáles de las siguientes propiedades se cumplen eventualmente:

1. $|a_n| < 1/10$;
2. $a_n > 0$;
3. $-1/10 < a_n < 1/10$.

Cuando la respuesta sea afirmativa, proporciona un umbral N que funcione.

Ejercicio A3. Tres comportamientos distintos. Considera

$$p_n = 2 - \frac{1}{n+1}, \qquad q_n = (-1)^n, \qquad r_n = n.$$

Para cada sucesión, indica si es acotada, si es monótona y cuál de las siguientes descripciones corresponde:

- converge a un número real;
- tiende a $+\infty$;
- diverge sin límite infinito unilateral.

Ejercicio A4. Leer ε–N. Traduce a lenguaje ordinario la afirmación

$$\forall \varepsilon > 0 \; \exists N \in \mathbb{N} \; \forall n \geq N : |a_n - 5| < \varepsilon.$$

Después responde:

1. ¿puede N depender de ε?;
2. ¿puede N depender de un índice n elegido después?;
3. ¿qué número afirma la fórmula que es el límite?

Ejercicio A5. Un cambio finito. Sean $a_n = \frac{1}{n+1}$ y

$$b_n = \begin{cases} 1000 - n, & 0 \leq n \leq 50, \\ \dfrac{1}{n+1}, & n > 50. \end{cases}$$

Sin volver a demostrar el límite desde la definición, determina el límite de (b_n) y explica qué resultado del capítulo estás usando.

Ejercicio A6. Real, infinito o ninguno. Clasifica cada sucesión:

$$a_n = \frac{1}{n+1}, \qquad b_n = n, \qquad c_n = -n, \qquad d_n = (-1)^n.$$

Para cada una indica si tiene límite real, límite $+\infty$, límite $-\infty$ o ninguna de esas tres posibilidades.

Ejercicio A7. Infinitas visitas no son una cola. Define

$$a_n = \begin{cases} 0, & n \text{ par}, \\ 1, & n \text{ impar}. \end{cases}$$

1. Explica por qué hay términos arbitrariamente tardíos exactamente iguales a 0.
2. Decide si $a_n \to 0$.
3. Señala qué orden de cuantificadores distingue ambas afirmaciones.

Nivel B — Aplicación directa. Ejercicio B1. Álgebra de límites. Supón que

$$a_n \to 2, \qquad b_n \to -1.$$

Calcula

$$\lim_{n\to\infty} \frac{3a_n - 2b_n}{a_n + b_n}.$$

Justifica por qué el cociente puede tratarse con el álgebra de límites.

Ejercicio B2. Cociente de polinomios. Para $n \geq 1$, calcula

$$\lim_{n\to\infty} \frac{5n^3 - 2n}{2n^3 + n^2 + 1}.$$

Ejercicio B3. Oscilación amortiguada. Demuestra que $\frac{2+\sigma n}{n+1} \to 0$.

Ejercicio B4. Geométrica dentro de una expresión. Calcula

$$\lim_{n\to\infty} \left(2 + 4\left(-\frac{3}{5}\right)^n\right).$$

Ejercicio B5. Qué preserva el orden. Sean (a_n) y (b_n) sucesiones convergentes tales que $a_n < b_n$ para todo n. Si

$$a_n \to A, \qquad b_n \to B,$$

¿qué relación entre A y B está garantizada? Da un ejemplo que muestre por qué no puede exigirse siempre una desigualdad estricta entre los límites.

Ejercicio B6. Valor absoluto y cociente. Supón que

$$a_n \to -3, \qquad b_n \to 2.$$

Calcula

$$\lim_{n\to\infty} \frac{|a_n|}{b_n + 1}.$$

Ejercicio B7. Una barrera cualquiera. Demuestra desde la definición M–N que $2n - 5 \to +\infty$.

Nivel C — Combinación estructural. Ejercicio C1. Una prueba ε–N completa. Demuestra directamente desde la definición que $\frac{3n+1}{n+1} \to 3$.

Tu prueba debe exhibir una elección explícita de $N(\varepsilon)$.

Ejercicio C2. Diseñar el umbral. Demuestra directamente que $\frac{1}{2n+3} \to 0$.

Antes de escribir la prueba final, determina qué condición suficiente sobre n garantiza $\frac{1}{2n+3} < \varepsilon$.

Ejercicio C3. Convergencia y acotación visible. Para $a_n = \frac{(-1)^n}{n+1}$, demuestra que $a_n \to 0$ mediante sándwich y encuentra además una cota global M tal que $|a_n| \leq M$ para todo n.

Ejercicio C4. Sándwich racional. Para $n \geq 1$, demuestra que $\frac{n}{n^2+1} \to 0$ sin dividir numerador y denominador por n^2. Busca directamente una sucesión dominante sencilla.

Ejercicio C5. Separar un denominador de cero. Supón que $b_n \to 4$.

1. Demuestra que existe N tal que $b_n > 2$ para todo $n \geq N$.
2. Deduce que $b_n \neq 0$ eventualmente.
3. Concluye que $\frac{1}{b_n} \to \frac{1}{4}$.

Ejercicio C6. Una geométrica desde Bernoulli. Demuestra que

$$\left(\frac{3}{4}\right)^n \to 0$$

usando únicamente la desigualdad de Bernoulli y el teorema del sándwich.
Sugerencia: escribe $\frac{3}{4} = \frac{1}{1+h}$ con $h > 0$.

Ejercicio C7. Un límite infinito desde la definición. Demuestra que $n^2 \to +\infty$ usando directamente la definición M–N.

Nivel D — Inversión y diagnóstico. En cada ejercicio decide si la afirmación es verdadera. Si es falsa, construye un contraejemplo y verifica las propiedades relevantes.

Ejercicio D1. «Acotada implica convergente». Decide si toda sucesión real acotada converge.

Ejercicio D2. «Convergente implica monótona». Decide si toda sucesión real convergente es monótona.

Ejercicio D3. «No acotada implica límite infinito». Decide si toda sucesión no acotada satisface $a_n \to +\infty$ o $a_n \to -\infty$.

Ejercicio D4. Cuadrados convergentes. Decide si

$$a_n^2 \text{ converge} \quad \Longrightarrow \quad a_n \text{ converge.}$$

Ejercicio D5. Diferencias sucesivas pequeñas. Decide si

$$a_{n+1} - a_n \to 0 \quad \Longrightarrow \quad a_n \text{ converge.}$$

Ejercicio D6. El punto fijo no garantiza convergencia. Una sucesión satisface

$$y_0 = 0, \qquad y_{n+1} = 3 - y_n.$$

La ecuación de punto fijo $L = 3 - L$ tiene la solución $L = 3/2$. Decide si (y_n) converge y explica qué error lógico cometeríamos al concluir convergencia solo a partir de esa ecuación.

Nivel E — Contraejemplos construidos. Ejercicio E1. Acotada y divergente. Construye una sucesión que permanezca en el intervalo $[0, 2]$ y no converja a ningún número real. Verifica ambas afirmaciones.

Ejercicio E2. Convergente pero no monótona. Construye una sucesión que converja a 1 y no sea monótona. Demuestra las dos propiedades.

Ejercicio E3. No acotada sin escape unilateral. Construye una sucesión no acotada que no tienda ni a $+\infty$ ni a $-\infty$. Justifica las tres afirmaciones.

Ejercicio E4. Un prefijo arbitrariamente largo no decide el límite. Sea $M \in \mathbb{N}$ arbitrario. Construye dos sucesiones (a_n) y (b_n) tales que

$$a_n = b_n \qquad \text{para } 0 \le n \le M,$$

pero cuyos límites reales sean distintos.

Ejercicio E5. Infinitamente cerca y, sin embargo, divergente. Construye una sucesión que tenga términos arbitrariamente tardíos exactamente iguales a 0, pero que no converja a 0. Verifica la negación ε–N.

Nivel F — Descubrimiento guiado. Ejercicio F1. Reconstruir $1/(n+1) \to 0$. Completa el siguiente programa.

1. Parte de $\frac{1}{n+1} < \varepsilon$. ¿Qué desigualdad suficiente sobre n aparece al invertir?
2. Usa la propiedad arquimediana para producir un natural que satisfaga esa desigualdad.
3. Reescribe el argumento en el orden

$$\varepsilon \to N \to n \ge N \to |a_n| < \varepsilon.$$

Ejercicio F2. Descubrir la estimación del producto. Sean

$$a_n \to A, \qquad b_n \to B.$$

Reconstruye una prueba de $a_n b_n \to AB$ siguiendo estas pistas:

1. suma y resta $a_n B$;
2. demuestra primero que (a_n) está eventualmente acotada;
3. obtiene una estimación de $|a_n b_n - AB|$ mediante dos errores;
4. elige tolerancias para que cada error contribuya menos que $\varepsilon/2$.

Ejercicio F3. Descubrir el límite geométrico. Sea q real con $0 < q < 1$.

1. Escribe $q = \frac{1}{1+h}$ con $h > 0$.
2. Usa Bernoulli para demostrar $0 < q^n \le \frac{1}{1+nh}$.
3. Demuestra que el término de la derecha tiende a 0.
4. Concluye $q^n \to 0$ por sándwich.
5. Explica cómo se extiende el resultado a $-1 < q < 0$.

Ejercicio F4. Una recurrencia completa. Define

$$a_0 = 0, \qquad a_{n+1} = \frac{a_n + 4}{2}.$$

Demuestra, en este orden:

1. que $0 \le a_n < 4$ para todo n;
2. que (a_n) es creciente;
3. que converge por el teorema de convergencia monótona;
4. que su límite es 4, usando desplazamiento de índices y álgebra de límites.

Ejercicio F5. Reconstruir convergencia monótona. Sea (a_n) creciente y acotada superiormente, y define

$$S = \{a_n : n \ge n_0\}, \qquad L = \sup S.$$

Usando únicamente la caracterización aproximativa del supremo y la monotonía, demuestra que $a_n \to L$. Señala exactamente dónde se utilizó la completitud de $\mathbb{R}$.

Nivel G — Desafíos de síntesis. Ejercicio G1. Intervalos encajados y sándwich secuencial. Sea

$$I_n = [a_n, b_n], \qquad n \ge 0,$$

una sucesión de intervalos cerrados no vacíos tal que $I_{n+1} \subseteq I_n$ para todo n y $b_n - a_n \to 0$.

1. Demuestra que existe un único $L \in \mathbb{R}$ perteneciente a todos los intervalos.
2. Demuestra que
$$a_n \to L, \qquad b_n \to L.$$
3. Si (x_n) satisface $x_n \in [a_n, b_n]$ para todo n, demuestra que $x_n \to L$.

Ejercicio G2. Método de Herón sin derivadas ni continuidad. Sea $a > 0$ y elige $x_0 > \sqrt{a}$.
Define

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

Demuestra que:

1. la sucesión está bien definida;
2. para todo n, $\sqrt{a} \le x_{n+1} \le x_n$;
3. (x_n) converge;
4. su límite es $\sqrt{a}$.

La identificación final debe utilizar únicamente desplazamiento de índices y álgebra de límites.

Ejercicio G3. Diferencias pequeñas no bastan. Define $a_n = \sqrt{n+1}$.

1. Demuestra que $a_n \to +\infty$.
2. Demuestra que $a_{n+1} - a_n = \frac{1}{\sqrt{n+2}+\sqrt{n+1}} \to 0$.
3. Explica con precisión qué afirmación falsa destruye este ejemplo.

Soluciones. Las soluciones siguen exactamente el orden de los ejercicios. En los niveles A–C se hace visible la herramienta utilizada; en D–G se insiste en la arquitectura lógica, la verificación de contraejemplos y la frontera de no circularidad del capítulo.

Soluciones del nivel A. **Solución A1.**

Sustituyendo $n = 0, 1, 2, 3$ obtenemos

$$a_0 = 1, \qquad a_1 = \frac{3}{2}, \qquad a_2 = \frac{5}{3}, \qquad a_3 = \frac{7}{4}.$$

Por tanto, el término de índice 3 es $a_3 = \frac{7}{4}$. El número 3 es una **entrada** de la función que define la sucesión; pertenece al dominio de índices. El número $7/4$ es la **salida** asociada a esa entrada. En símbolos, $a(3) = a_3 = \frac{7}{4}$.

Solución A2.

Como $|a_n| = \frac{1}{n+1}$, si $n \geq 10$, entonces

$$|a_n| = \frac{1}{n+1} \leq \frac{1}{11} < \frac{1}{10}.$$

Así, la primera propiedad se cumple eventualmente; por ejemplo, $N = 10$ funciona. La segunda propiedad no se cumple eventualmente. Para cualquier N existe un índice impar $n \geq N$, y entonces $a_n < 0$. La tercera propiedad equivale a $|a_n| < \frac{1}{10}$, de modo que también se cumple eventualmente con $N = 10$.

Solución A3.

Para $p_n = 2 - \frac{1}{n+1}$, el término $1/(n+1)$ disminuye hacia 0, de modo que (p_n) es creciente. Además, $1 \leq p_n < 2$. Por tanto es acotada y, usando el álgebra de límites, $p_n \to 2$. La sucesión $q_n = (-1)^n$ está acotada por $|q_n| = 1$, pero alterna entre 1 y -1, así que no es monótona ni converge a un real. Tampoco tiende a $+\infty$ ni a $-\infty$, pues continúa tomando ambos valores.

Finalmente, $r_n = n$ es creciente, no está acotada superiormente y satisface $r_n \to +\infty$.

En resumen:

Sucesión	Acotada	Monótona	Comportamiento
(p_n)	sí	creciente	$p_n \to 2$
(q_n)	sí	no	divergente, sin límite infinito
(r_n)	no	creciente	$r_n \to +\infty$

Solución A4.

La fórmula dice:

> cualquiera que sea la tolerancia positiva ε, existe un índice N tal que todos los términos con índice $n \geq N$ están a distancia menor que ε de 5.

Sí, N puede depender de la tolerancia: $N = N(\varepsilon)$. No puede depender de un índice n elegido **después**, porque el orden de los cuantificadores es $\forall\varepsilon\ \exists N\ \forall n \geq N$. Primero se fija ε, después se elige un único N, y ese mismo N debe controlar todos los índices posteriores. La fórmula afirma exactamente que $a_n \to 5$.

Solución A5.

Para todo $n > 50$, $a_n = b_n$.

Las dos sucesiones coinciden eventualmente. Por Proposición 3.1, la convergencia y el límite no cambian al modificar una cantidad finita de términos. Como $a_n = \frac{1}{n+1} \to 0$, se sigue inmediatamente que

$$\boxed{b_n \to 0.}$$

Los valores iniciales de (b_n) pueden ser grandes e irregulares; no intervienen en el comportamiento de su cola.

Solución A6.

Tenemos: $\frac{1}{n+1} \to 0$, así que (a_n) tiene límite real 0. Por la propiedad arquimediana, $n \to +\infty$, de modo que (b_n) tiene límite infinito positivo. Análogamente, $-n \to -\infty$, de modo que (c_n) tiene límite infinito negativo.

Finalmente, $d_n = (-1)^n$ alterna entre 1 y -1. No converge a un real y tampoco puede superar eventualmente toda barrera ni quedar eventualmente por debajo de toda barrera. Por tanto no tiene límite real ni límite infinito unilateral.

Solución A7.

Para todo N podemos encontrar un índice par $n \geq N$. En ese índice, $a_n = 0$.

Por tanto, $\forall N\ \exists n \geq N :\ a_n = 0$.

Sin embargo, $a_n \not\to 0$. Tomemos, por ejemplo, $\varepsilon_0 = \frac{1}{2}$. Para todo N existe un índice impar $n \geq N$, y entonces $|a_n - 0| = 1 \geq \frac{1}{2}$.

Así,

$$\exists\varepsilon_0 > 0\ \forall N\ \exists n \geq N : |a_n| \geq \varepsilon_0,$$

que es precisamente la negación de la convergencia a 0. La diferencia lógica es $\forall N\ \exists n \geq N$ frente a $\exists N\ \forall n \geq N$. La primera expresa reaparición arbitrariamente tardía; la segunda, control de toda una cola.

Soluciones del nivel B. **Solución B1.**

Por el álgebra de límites, $3a_n - 2b_n \to 3 \cdot 2 - 2(-1) = 8$, y $a_n + b_n \to 2 + (-1) = 1$. Como el límite del denominador es distinto de cero, el denominador queda eventualmente separado de cero y podemos aplicar la ley del cociente. Por tanto,

$$\boxed{\frac{3a_n - 2b_n}{a_n + b_n} \to 8.}$$

Solución B2.
Dividimos numerador y denominador por n^3:

$$\frac{5n^3-2n}{2n^3+n^2+1}=\frac{5-\dfrac{2}{n^2}}{2+\dfrac{1}{n}+\dfrac{1}{n^3}}.$$

Las potencias recíprocas satisfacen

$$\frac{1}{n}\to 0, \qquad \frac{1}{n^2}\to 0, \qquad \frac{1}{n^3}\to 0.$$

Así, el numerador tiende a 5 y el denominador a $2\neq 0$. Por la ley del cociente,

$$\boxed{\frac{5n^3-2n}{2n^3+n^2+1}\to\frac{5}{2}.}$$

Solución B3.
Como $-1\leq \sigma n\leq 1$, tenemos $1\leq 2+\sigma n\leq 3$.
Para todo $n\geq 0$, $0\leq \frac{2+\sigma n}{n+1}\leq \frac{3}{n+1}$.
Y $\frac{3}{n+1}\to 0$.
Por el teorema del sándwich,

$$\boxed{\frac{2+\sigma n}{n+1}\to 0.}$$

Solución B4.
Como

$$\left|-\frac{3}{5}\right|<1,$$

el resultado geométrico Corolario 3.2 da

$$\left(-\frac{3}{5}\right)^n\to 0.$$

Por el álgebra de límites,

$$2+4\left(-\frac{3}{5}\right)^n\to 2+4\cdot 0=2.$$

Por tanto,

$$\boxed{\lim_{n\to\infty}\left(2+4\left(-\frac{3}{5}\right)^n\right)=2.}$$

Solución B5.
De $a_n < b_n$ para todo n se sigue, en particular, $a_n \leq b_n$ eventualmente. Por Proposición 3.4,

$$\boxed{A \leq B.}$$

No podemos concluir necesariamente $A < B$. Por ejemplo, tomemos

$$a_n = 0, \qquad b_n = \frac{1}{n+1}.$$

Entonces $a_n < b_n$ para todo n, pero

$$a_n \to 0, \qquad b_n \to 0.$$

La desigualdad estricta de los términos puede colapsar a igualdad en el límite.

Solución B6.
Por Corolario 3.1, $|a_n| \to |-3| = 3$.
Además, $b_n + 1 \to 2 + 1 = 3$.
Como $3 \neq 0$, podemos aplicar la ley del cociente:

$$\boxed{\frac{|a_n|}{b_n + 1} \to \frac{3}{3} = 1.}$$

Solución B7.
Debemos demostrar que para todo $M \in \mathbb{R}$ existe N tal que

$$n \geq N \quad \Longrightarrow \quad 2n - 5 > M.$$

Sea $M \in \mathbb{R}$. Por la propiedad arquimediana podemos elegir $N \in \mathbb{N}$ tal que $N > \frac{M+5}{2}$.
Si $n \geq N$, entonces $2n - 5 \geq 2N - 5 > M$.
Por tanto,

$$\forall M \in \mathbb{R}\ \exists N\ \forall n \geq N : 2n - 5 > M,$$

y concluimos

$$\boxed{2n - 5 \to +\infty.}$$

Soluciones del nivel C. **Solución C1.**
Queremos controlar

$$\left| \frac{3n+1}{n+1} - 3 \right|.$$

Simplificando,

$$\left| \frac{3n + 1 - 3n - 3}{n+1} \right| = \frac{2}{n+1}.$$

Sea $\varepsilon > 0$. Por la propiedad arquimediana elegimos $N \in \mathbb{N}$ tal que $N > \frac{2}{\varepsilon}$.
Si $n \geq N$, entonces $n + 1 > N > \frac{2}{\varepsilon}$, y por tanto

$$\left|\frac{3n+1}{n+1} - 3\right| = \frac{2}{n+1} < \varepsilon.$$

Como $\varepsilon > 0$ era arbitrario,

$$\boxed{\frac{3n+1}{n+1} \to 3.}$$

Solución C2.

En la fase de búsqueda queremos $\frac{1}{2n+3} < \varepsilon$. Sería suficiente exigir $2n > \frac{1}{\varepsilon}$, es decir, $n > \frac{1}{2\varepsilon}$. Ahora escribimos la prueba en el orden lógico correcto. Sea $\varepsilon > 0$. Por la propiedad arquimediana elegimos $N \in \mathbb{N}$ tal que $N > \frac{1}{2\varepsilon}$.

Si $n \geq N$, entonces $2n + 3 > 2N > \frac{1}{\varepsilon}$.

Todas las cantidades son positivas, de modo que al tomar recíprocos obtenemos $\frac{1}{2n+3} < \varepsilon$.

Por consiguiente,

$$\boxed{\frac{1}{2n+3} \to 0.}$$

Solución C3.

Para todo n, $|a_n| = \frac{1}{n+1}$.

Así, $-\frac{1}{n+1} \leq a_n \leq \frac{1}{n+1}$.

Como

$$-\frac{1}{n+1} \to 0 \qquad \text{y} \qquad \frac{1}{n+1} \to 0,$$

el teorema del sándwich da

$$\boxed{a_n \to 0.}$$

Además, $|a_n| = \frac{1}{n+1} \leq 1$ para todo $n \geq 0$. Por tanto, una cota global válida es

$$\boxed{M = 1.}$$

Este ejemplo hace visible que una sucesión puede oscilar y, sin embargo, converger.

Solución C4.

Para $n \geq 1$ tenemos $n^2 + 1 \geq n^2 > 0$.

Por tanto,

$$0 \leq \frac{n}{n^2+1} \leq \frac{n}{n^2} = \frac{1}{n}.$$

Y $\frac{1}{n} \to 0$.

Por el teorema del sándwich,

$$\boxed{\frac{n}{n^2+1} \to 0.}$$

La ventaja de esta ruta es que no necesitamos reescribir el cociente completo: basta reconocer una cota dominante cuyo límite ya conocemos.

Solución C5.

Como $b_n \to 4$, podemos utilizar en la definición la tolerancia $\varepsilon = 2$. Existe entonces N tal que, para todo $n \geq N$, $|b_n - 4| < 2$.

Esto equivale a $2 < b_n < 6$.

En particular, $b_n > 2$ para todo $n \geq N$. Por tanto, b_n queda eventualmente separado de cero y, en particular, $b_n \neq 0$ eventualmente. Ahora podemos aplicar la ley del cociente a la sucesión constante 1 y a (b_n): $\frac{1}{b_n} \to \frac{1}{4}$.

Así,

$$\boxed{\frac{1}{b_n} \to \frac{1}{4}.}$$

Solución C6.

Escribimos

$$\frac{3}{4} = \frac{1}{1 + \frac{1}{3}}.$$

Así, con $h = \frac{1}{3} > 0$, tenemos

$$\left(\frac{3}{4}\right)^n = \frac{1}{(1+h)^n}.$$

Por la desigualdad de Bernoulli, $(1+h)^n \geq 1 + nh$.

Por tanto,

$$0 < \left(\frac{3}{4}\right)^n \leq \frac{1}{1 + n/3} = \frac{3}{n+3}.$$

Ahora, $0 \leq \frac{3}{n+3} \leq \frac{3}{n+1}$, y $\frac{3}{n+1} \to 0$.

De nuevo por sándwich, $\frac{3}{n+3} \to 0$.

Aplicando una vez más el sándwich,

$$\boxed{\left(\frac{3}{4}\right)^n \to 0.}$$

Solución C7.

Sea $M \in \mathbb{R}$. Definimos $R = \max\{1, M\}$.

Por la propiedad arquimediana elegimos $N \in \mathbb{N}$ tal que $N > R$. Si $n \geq N$, entonces $n \geq 1$, de modo que

$$n^2 \geq n \geq N > R \geq M.$$

Así, $\forall M \in \mathbb{R}\ \exists N\ \forall n \geq N : n^2 > M$.

Por definición,

$$\boxed{n^2 \to +\infty.}$$

Soluciones del nivel D. **Solución D1.**

La afirmación es falsa.

Tomemos $a_n = (-1)^n$.

La sucesión está acotada porque $|a_n| = 1$ para todo n. Sin embargo, no converge. Como se demostró en Ejemplo 3.4, la alternancia entre 1 y -1 impide que una cola completa quede dentro de una tolerancia suficientemente pequeña alrededor de cualquier real.

Por tanto,

$$\boxed{\text{acotada} \not\Longrightarrow \text{convergente.}}$$

Solución D2.

La afirmación es falsa.

Consideremos $a_n = \frac{(-1)^n}{n+1}$.

Como $|a_n| \leq \frac{1}{n+1}$ y $1/(n+1) \to 0$, el teorema del sándwich da $a_n \to 0$. Pero la sucesión no es monótona. Por ejemplo,

$$a_0 = 1, \qquad a_1 = -\frac{1}{2}, \qquad a_2 = \frac{1}{3}.$$

Primero disminuye y luego aumenta. Por tanto no es creciente ni decreciente.

Así,

$$\boxed{\text{convergente} \not\Longrightarrow \text{monótona.}}$$

Solución D3.

La afirmación es falsa.

Sea $a_n = (-1)^n n$.

La sucesión no está acotada: en los índices pares toma valores positivos arbitrariamente grandes y en los impares valores negativos de magnitud arbitrariamente grande. No tiende a $+\infty$. Basta elegir la barrera $M = 0$. Después de cualquier umbral N existe un índice impar $n \geq N$, y entonces

$$a_n = -n < 0 = M.$$

Tampoco tiende a $-\infty$. Con la barrera $M = 0$ en la definición correspondiente, después de cualquier N existe un índice par $n \geq N$ con $a_n = n > 0$.

Por tanto,

$$\boxed{\text{no acotada} \not\Longrightarrow a_n \to +\infty \text{ o } a_n \to -\infty.}$$

Solución D4.

La afirmación es falsa.

Tomemos $a_n = (-1)^n$.

Entonces $a_n^2 = 1$ para todo n, de modo que $a_n^2 \to 1$. Sin embargo, (a_n) diverge porque alterna entre 1 y -1.

Así,

$$\boxed{a_n^2 \text{ convergente} \not\Longrightarrow a_n \text{ convergente.}}$$

El cuadrado ha borrado la información de signo responsable de la oscilación.

Solución D5.

La afirmación es falsa.

Tomemos $a_n = \sqrt{n+1}$.

Primero observemos que $a_{n+1} - a_n = \sqrt{n+2} - \sqrt{n+1}$.

Racionalizando, $a_{n+1} - a_n = \frac{1}{\sqrt{n+2}+\sqrt{n+1}}$.

Como el denominador es al menos $\sqrt{n+1}$, $0 < a_{n+1} - a_n \leq \frac{1}{\sqrt{n+1}}$. Además, $\sqrt{n+1} \to +\infty$, y por Proposición 3.6, $\frac{1}{\sqrt{n+1}} \to 0$.

Por sándwich, $a_{n+1} - a_n \to 0$.

Pero la propia sucesión satisface $a_n = \sqrt{n+1} \to +\infty$, así que no converge a ningún real.

Por tanto,

$$\boxed{a_{n+1} - a_n \to 0 \not\Longrightarrow a_n \text{ convergente.}}$$

Solución D6.

Calculamos los primeros términos:

$$y_0 = 0, \qquad y_1 = 3, \qquad y_2 = 0, \qquad y_3 = 3, \qquad \dots$$

Por inducción,

$$y_{2k} = 0, \qquad y_{2k+1} = 3.$$

La sucesión no converge: por ejemplo, respecto del candidato $3/2$, todos sus términos están a distancia

$$\left|0 - \frac{3}{2}\right| = \left|3 - \frac{3}{2}\right| = \frac{3}{2}.$$

La ecuación $L = 3 - L$ solo responde a la pregunta:

> si la sucesión convergiera y pudiéramos identificar el límite mediante la recurrencia, ¿qué valor tendría que satisfacer?

La respuesta es $L = 3/2$. Pero la existencia del límite nunca fue demostrada.

Así,

$$\boxed{L = F(L) \text{ identifica candidatos; no demuestra convergencia.}}$$

Soluciones del nivel E. **Solución E1.**

Una construcción sencilla es

$$a_n = \begin{cases} 0, & n \text{ par}, \\ 2, & n \text{ impar}. \end{cases}$$

Claramente, $0 \le a_n \le 2$ para todo n, así que la sucesión está acotada y permanece en $[0, 2]$. Demostremos que no converge. Supongamos que $a_n \to L$. Como hay términos pares e impares arbitrariamente tardíos, una cola tendría que contener valores 0 y 2 simultáneamente. Más directamente, por la desigualdad triangular, $2 = |2 - 0| \le |2 - L| + |L|$. Por tanto, al menos una de las cantidades $|L|$ y $|2 - L|$ es mayor o igual que 1. Tomemos $\varepsilon_0 = 1$. Después de cualquier N podemos elegir un índice de la paridad que produce ese valor y obtener $|a_n - L| \ge 1$. Así no existe límite real.

Solución E2.

Podemos tomar $a_n = 1 + \frac{(-1)^n}{n+1}$.

Como $\frac{(-1)^n}{n+1} \to 0$, el álgebra de límites da

$$\boxed{a_n \to 1.}$$

No es monótona. En efecto,

$$a_0 = 2, \qquad a_1 = \frac{1}{2}, \qquad a_2 = \frac{4}{3}.$$

De $a_0 > a_1 < a_2$ vemos que la dirección cambia ya en los primeros pasos. Por tanto, la sucesión converge a 1 pero no es monótona.

Solución E3.

Definamos

$$a_n = \begin{cases} n, & n \text{ par}, \\ -n, & n \text{ impar}. \end{cases}$$

La sucesión no está acotada superiormente, porque para cualquier M podemos elegir un índice par suficientemente grande con $n > M$. Tampoco está acotada inferiormente, porque los índices impares producen valores $-n$ arbitrariamente negativos. No tiende a $+\infty$: con la barrera $M = 0$, después de todo umbral reaparece un índice impar con $a_n < 0$. No tiende a $-\infty$: después de todo umbral reaparece un índice par con $a_n > 0$. Así, la sucesión es no acotada pero no posee límite infinito unilateral.

Solución E4.

Fijemos un $M \in \mathbb{N}$ arbitrario y definamos $a_n = 0$ para todo n, mientras que

$$b_n = \begin{cases} 0, & 0 \le n \le M, \\ 1, & n > M. \end{cases}$$

Entonces $a_n = b_n$ para todos los índices del prefijo $0, \dots, M$. Sin embargo, $a_n \to 0$, porque es constante igual a 0, mientras que $b_n \to 1$, porque a partir de $M + 1$ es constante igual a 1. Como M era arbitrario, podemos hacer el prefijo común tan largo como queramos. Ninguna cantidad finita de términos iniciales determina por sí sola el límite.

Solución E5.
Tomemos

$$a_n = \begin{cases} 0, & n \text{ par}, \\ 1, & n \text{ impar}. \end{cases}$$

Para todo N existe un índice par $n \geq N$, así que hay términos arbitrariamente tardíos exactamente iguales a 0.

Pero $a_n \not\to 0$. Fijemos $\varepsilon_0 = \frac{1}{2}$.

Para todo N existe un índice impar $n \geq N$, y entonces $|a_n - 0| = 1 \geq \frac{1}{2}$.

Por tanto,

$$\exists \varepsilon_0 > 0 \; \forall N \; \exists n \geq N : |a_n| \geq \varepsilon_0.$$

Esta es la negación exacta de $a_n \to 0$.

Soluciones del nivel F. **Solución F1.**
Queremos demostrar $\frac{1}{n+1} \to 0$.

La búsqueda comienza con $\frac{1}{n+1} < \varepsilon$.

Como ambos lados son positivos, esto queda garantizado si $n + 1 > \frac{1}{\varepsilon}$. Es suficiente, por ejemplo, pedir $n > \frac{1}{\varepsilon}$. Sea ahora $\varepsilon > 0$. Por la propiedad arquimediana elegimos $N \in \mathbb{N}$ tal que $N > \frac{1}{\varepsilon}$.

Si $n \geq N$, entonces

$$n + 1 \geq N + 1 > N > \frac{1}{\varepsilon}.$$

Al tomar recíprocos, $\frac{1}{n+1} < \varepsilon$.

Así,

$$\forall \varepsilon > 0 \; \exists N \; \forall n \geq N : \left| \frac{1}{n+1} - 0 \right| < \varepsilon,$$

y por definición,

$$\boxed{\frac{1}{n+1} \to 0.}$$

La fase de búsqueda fue hacia atrás; la prueba final respeta el orden de cuantificadores.

Solución F2.
Partimos de $a_n b_n - AB$.

Sumamos y restamos $a_n B$:

$$a_n b_n - AB = a_n(b_n - B) + B(a_n - A).$$

Por la desigualdad triangular,

$$|a_n b_n - AB| \leq |a_n| \, |b_n - B| + |B| \, |a_n - A|.$$

Necesitamos controlar $|a_n|$. Como $a_n \to A$, tomando $\varepsilon = 1$ existe N_0 tal que, para $n \geq N_0$, $|a_n - A| < 1$.

Entonces $|a_n| \leq |a_n - A| + |A| < 1 + |A|$.

Definamos $K = 1 + |A| > 0$.

Dado $\varepsilon > 0$, como $b_n \to B$, existe N_1 tal que $|b_n - B| < \frac{\varepsilon}{2K}$ para $n \geq N_1$. Para el segundo término distinguimos solo una pequeña comodidad técnica. Definimos $C = |B| + 1 > 0$. Como $a_n \to A$, existe N_2 tal que $|a_n - A| < \frac{\varepsilon}{2C}$ para $n \geq N_2$.

Si $n \geq N = \max\{N_0, N_1, N_2\}$, entonces

$$\begin{aligned} |a_n b_n - AB| &\leq |a_n|\,|b_n - B| + |B|\,|a_n - A| \\ &< K\frac{\varepsilon}{2K} + |B|\frac{\varepsilon}{2C} \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \end{aligned}$$

porque $|B|/C \leq 1$.

Por tanto,

$$\boxed{a_n b_n \to AB.}$$

La cota eventual no es un adorno: es lo que impide que el factor $|a_n|$ destruya el control del error $|b_n - B|$.

Solución F3.

Sea $0 < q < 1$. Definimos $h = \frac{1}{q} - 1$.

Como $q < 1$ y $q > 0$, $h > 0$, y $q = \frac{1}{1+h}$.

Entonces $q^n = \frac{1}{(1+h)^n}$.

Por Bernoulli, $(1 + h)^n \geq 1 + nh$.

Como todas las cantidades son positivas, $0 < q^n \leq \frac{1}{1+nh}$. Para ver que el término de la derecha tiende a 0, observemos que para $n \geq 1$,

$$0 < \frac{1}{1 + nh} \leq \frac{1}{nh} = \frac{1}{h}\frac{1}{n}.$$

Como $1/n \to 0$, el sándwich da $\frac{1}{1+nh} \to 0$.

Aplicando sándwich una vez más,

$$\boxed{q^n \to 0.}$$

Si ahora $-1 < q < 0$, entonces $|q| \in (0, 1)$ y $|q^n| = |q|^n \to 0$.

La forma absoluta del sándwich implica $q^n \to 0$. Así recuperamos el resultado completo para $|q| < 1$.

Solución F4.

La recurrencia es

$$a_0 = 0, \qquad a_{n+1} = \frac{a_n + 4}{2}.$$

1. Región invariante. Demostraremos por inducción que $0 \leq a_n < 4$.

Para $n = 0$, $a_0 = 0$.

Supongamos $0 \leq a_n < 4$.

Entonces $4 \leq a_n + 4 < 8$, y al dividir por 2, $2 \leq a_{n+1} < 4$.

En particular, $0 \leq a_{n+1} < 4$.

Por inducción, la afirmación vale para todo n.

2. Monotonía. Tenemos $a_{n+1} - a_n = \frac{4-a_n}{2}$.

Como $a_n < 4$, $a_{n+1} - a_n > 0$.

Por tanto, (a_n) es estrictamente creciente.

3. Existencia del límite. La sucesión es creciente y está acotada superiormente por 4. Por Teorema 3.2 existe $L \in \mathbb{R}$ tal que $a_n \to L$.

4. Identificación. Por Proposición 3.1, $a_{n+1} \to L$.

La recurrencia puede escribirse como $2a_{n+1} = a_n + 4$. Usando solo el álgebra de límites, $2L = L + 4$.

Por tanto,

$$\boxed{L = 4.}$$

En consecuencia,

$$\boxed{a_n \to 4.}$$

Solución F5.

Sea $S = \{a_n : n \geq n_0\}$.

Como la sucesión está acotada superiormente, S es no vacío y acotado superiormente. La **completitud** de $\mathbb{R}$ garantiza la existencia de $L = \sup S \in \mathbb{R}$. Este es el punto exacto donde entra la completitud. Sea ahora $\varepsilon > 0$. Como L es el supremo, el número $L - \varepsilon$ no puede ser una cota superior de S. Por la caracterización aproximativa del supremo existe un índice N tal que $a_N > L - \varepsilon$. Como la sucesión es creciente, para todo $n \geq N$, $a_n \geq a_N$. Y como L es cota superior de S, $a_n \leq L$.

Por tanto, para todo $n \geq N$,

$$L - \varepsilon < a_N \leq a_n \leq L < L + \varepsilon.$$

Esto implica $|a_n - L| < \varepsilon$.

Como $\varepsilon > 0$ era arbitrario,

$$\boxed{a_n \to L = \sup S.}$$

La arquitectura de la prueba es

$$\boxed{\text{completitud} \to L = \sup S \to \text{un término cerca de } L \to \text{monotonía} \to \text{toda la cola cerca de } L.}$$

Soluciones del nivel G. **Solución G1.**

Como $I_{n+1} \subseteq I_n$, los intervalos están encajados. Además, cada I_n es cerrado, no vacío y está contenido en I_0, que es acotado. Por el principio de intervalos cerrados encajados demostrado en `T1-C02`, existe al menos un punto $L \in \bigcap_{n=0}^{\infty} I_n$. Demostremos que es único. Si L y K pertenecieran a todos los intervalos, entonces para todo n, $|L-K| \leq b_n - a_n$. Como $b_n - a_n \to 0$, el sándwich implicaría que la sucesión constante $|L-K|$ tiene límite 0. Por unicidad del límite, $|L-K| = 0$, y por tanto $L = K$. Así, la intersección contiene exactamente un punto.

Ahora, como $L \in [a_n, b_n]$, $0 \leq L - a_n \leq b_n - a_n$.

El extremo derecho tiende a 0, luego por sándwich $L - a_n \to 0$. Por el álgebra de límites,

$$\boxed{a_n \to L.}$$

Análogamente, $0 \leq b_n - L \leq b_n - a_n$, de modo que

$$\boxed{b_n \to L.}$$

Finalmente, si $x_n \in [a_n, b_n]$, entonces tanto x_n como L pertenecen al mismo intervalo $[a_n, b_n]$. Por tanto, $|x_n - L| \leq b_n - a_n$. Como el lado derecho tiende a 0, la forma absoluta del teorema del sándwich da

$$\boxed{x_n \to L.}$$

El argumento integra dos manifestaciones de completitud: el principio de intervalos encajados produce el punto común y el lenguaje secuencial permite demostrar que toda elección dentro de los intervalos converge hacia él.

Solución G2.

Sea $a > 0$ y $x_0 > \sqrt{a} > 0$.

Demostraremos simultáneamente que todos los términos están bien definidos y permanecen por encima de $\sqrt{a}$.

Supongamos que $x_n \geq \sqrt{a} > 0$.

Entonces $x_n \neq 0$, por lo que

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

está bien definido.

Además,

$$\begin{aligned} x_{n+1} - \sqrt{a} &= \frac{1}{2}\left(x_n + \frac{a}{x_n} - 2\sqrt{a}\right) \\ &= \frac{x_n^2 - 2x_n\sqrt{a} + a}{2x_n} \\ &= \frac{(x_n - \sqrt{a})^2}{2x_n} \geq 0. \end{aligned}$$

Por tanto, $x_{n+1} \geq \sqrt{a}$.

Como $x_0 > \sqrt{a}$, la inducción muestra que todos los términos están definidos y satisfacen $x_n \geq \sqrt{a} > 0$.

Ahora comparemos términos consecutivos:

$$\begin{aligned} x_n - x_{n+1} &= x_n - \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) \\ &= \frac{x_n^2 - a}{2x_n}. \end{aligned}$$

Como $x_n \geq \sqrt{a}$, $x_n^2 \geq a$, y por tanto $x_n - x_{n+1} \geq 0$.

Así,

$$\boxed{\sqrt{a} \leq x_{n+1} \leq x_n}$$

para todo n.

La sucesión es decreciente y está acotada inferiormente por $\sqrt{a}$. Por el teorema de convergencia monótona existe $L \in \mathbb{R}$ tal que $x_n \to L$.

Además, $L \geq \sqrt{a} > 0$.

Para identificar L no usamos continuidad. Multiplicamos la recurrencia por $2x_n$: $2x_n x_{n+1} = x_n^2 + a$.

Por estabilidad ante desplazamientos, $x_{n+1} \to L$.

El álgebra de límites da entonces $2L^2 = L^2 + a$.

Por tanto, $L^2 = a$.

Como $L \geq 0$ y $\sqrt{a}$ es la raíz no negativa de a,

$$\boxed{L = \sqrt{a}.}$$

Así,

$$\boxed{x_n \to \sqrt{a}.}$$

Solución G3.

Sea $a_n = \sqrt{n+1}$.

1. Límite infinito. Debemos probar que toda barrera real es superada por una cola.

Sea $M \in \mathbb{R}$. Definimos $M_+ = \max\{M, 0\}$.

Por la propiedad arquimediana existe $N \in \mathbb{N}$ tal que $N > M_+^2$.

Si $n \geq N$, entonces $n + 1 > N > M_+^2$.

Como ambos lados son no negativos, $\sqrt{n+1} > M_+ \geq M$.

Por definición,

$$\boxed{a_n \to +\infty.}$$

2. Diferencias sucesivas. Racionalizando,

$$\begin{aligned} a_{n+1} - a_n &= \sqrt{n+2} - \sqrt{n+1} \\ &= \frac{(n+2) - (n+1)}{\sqrt{n+2} + \sqrt{n+1}} \\ &= \frac{1}{\sqrt{n+2} + \sqrt{n+1}}. \end{aligned}$$

Como $\sqrt{n+2} + \sqrt{n+1} \geq \sqrt{n+1}$, tenemos $0 < a_{n+1} - a_n \leq \frac{1}{\sqrt{n+1}}$.

Ya demostramos que $\sqrt{n+1} \to +\infty$ y la sucesión es positiva. Por Proposición 3.6, $\frac{1}{\sqrt{n+1}} \to 0$. El teorema del sándwich da

$$\boxed{a_{n+1} - a_n \to 0.}$$

3. Diagnóstico. El ejemplo destruye la afirmación

$$\boxed{a_{n+1} - a_n \to 0 \quad \Longrightarrow \quad a_n \text{ converge.}}$$

Que los pasos consecutivos se hagan arbitrariamente pequeños solo controla la distancia **local** entre términos vecinos. No impide que la sucesión continúe avanzando indefinidamente en una misma dirección. Aquí los incrementos tienden a cero mientras la sucesión completa escapa hacia $+\infty$.

Capítulo 4

Límites de funciones

En los ejemplos periódicos usaremos las funciones triangulares σ y κ definidas en C03, con $\omega = 2$. Sus cotas $|\sigma|, |\kappa| \leq 1$ y los valores $\sigma(\omega/2 + n\omega) = (-1)^n$ proceden directamente de aquella definición por tramos; no se presupone trigonometría.

En una sucesión, afirmar que los términos se aproximan a un número exige controlar **todos los términos suficientemente tardíos**. Una muestra finita, por extensa que sea, deja sin examinar los términos posteriores; por eso la definición de límite se formula sobre una cola completa: $\exists N\ \forall n \geq N$. El índice N marcaba un umbral. Antes de ese índice podían ocurrir muchas cosas; después de él, la propiedad que nos interesaba debía mantenerse sin excepciones. Ahora queremos trasladar esa misma lógica a funciones reales. Sea

$$f\colon A \to \mathbb{R}, \qquad A \subseteq \mathbb{R}.$$

En lugar de preguntar qué ocurre cuando el índice natural n crece, comenzaremos a preguntar qué ocurre con $f(x)$ cuando la variable real x **se acerca** a un punto a. La frase parece inocente. Sin embargo, contiene una dificultad nueva. En una sucesión, el dominio es discreto y está ordenado naturalmente: sabemos qué significa pasar a índices posteriores. En una función real, en cambio, alrededor de un punto hay números a distancias arbitrariamente pequeñas. No existe un «índice siguiente» que nos lleve progresivamente hacia a. Antes de poder hablar rigurosamente de límites necesitamos entonces responder una pregunta más elemental:

¿qué significa que x pueda acercarse a a dentro del dominio de f?

Para expresar esa aproximación utilizaremos el valor absoluto como distancia. Un entorno describirá los puntos próximos a a; al retirar el centro obtendremos un entorno perforado. Necesitaremos además que el dominio tenga puntos distintos de a arbitrariamente próximos a él, condición que precisaremos mediante la noción de punto de acumulación. Hay un principio que conviene conservar desde el comienzo:

en una sucesión, el límite vive en la cola;	en una función, vivirá en los entornos perforados.

Este capítulo desarrollará esa traducción paso a paso.

4.1. Acercarse a un punto: entornos, acumulación y puntos aislados

La distancia vuelve a ser el lenguaje adecuado. En `T1-C02` vimos que el valor absoluto permite medir distancia en la recta real. La distancia entre x y a es $|x - a|$.

Por ejemplo,

$$|2.1 - 2| = 0.1, \qquad |1.99 - 2| = 0.01, \qquad |2.001 - 2| = 0.001.$$

Decir que x está cerca de a puede traducirse entonces en una desigualdad de la forma $|x-a| < \delta$, donde $\delta > 0$ indica cuánto alejamiento estamos dispuestos a permitir. Ya sabemos resolver esta desigualdad:

$$|x-a| < \delta \quad \Longleftrightarrow \quad -\delta < x-a < \delta \quad \Longleftrightarrow \quad a-\delta < x < a+\delta.$$

Por tanto, los números que están a distancia menor que δ de a forman exactamente el intervalo $(a-\delta, a+\delta)$. Este intervalo será nuestra primera herramienta local.

Si, por ejemplo, $a = 3$ y $\delta = \frac{1}{2}$, obtenemos

$$V_{1/2}(3) = \left(\frac{5}{2}, \frac{7}{2}\right).$$

Todos los puntos de ese intervalo están a menos de $1/2$ de 3. Si reducimos el radio a $\delta = \frac{1}{100}$, obtenemos el intervalo mucho más pequeño $V_{1/100}(3) = (2.99, 3.01)$. El centro no cambia. Lo que cambia es la escala con la que observamos sus alrededores.

Acercarse no es necesariamente llegar. Cuando estudiemos el comportamiento de una función **cerca** de a, necesitaremos distinguir lo que ocurre en puntos próximos de lo que ocurre exactamente en a. La desigualdad $|x-a| < \delta$ incluye el punto $x = a$, porque $|a-a| = 0 < \delta$. Pero si queremos observar únicamente puntos próximos distintos del centro, debemos imponer además $x \neq a$. Una manera equivalente de escribir ambas condiciones es $0 < |x-a| < \delta$. La desigualdad estricta de la izquierda, $0 < |x-a|$, excluye exactamente el caso $x = a$. Geométricamente estamos tomando el intervalo $(a-\delta, a+\delta)$ y retirando su centro. Ese pequeño agujero será esencial. Más adelante permitirá estudiar el comportamiento de $f(x)$ alrededor de a incluso cuando $f(a)$ no exista o cuando su valor no coincida con el comportamiento de los puntos vecinos.

Todavía no estamos definiendo el límite. Estamos construyendo el lenguaje geométrico que hará posible definirlo sin ambigüedad.

Definición 4.1. **Entorno, entorno perforado, punto de acumulación y punto aislado.** Sean $A \subseteq \mathbb{R}$, $a \in \mathbb{R}$ y $\delta > 0$.

El **entorno de radio δ centrado en** a es el intervalo

$$V_\delta(a) := (a-\delta, a+\delta) = \{x \in \mathbb{R} : |x-a| < \delta\}.$$

El **entorno perforado de radio δ centrado en** a es

$$V_\delta^*(a) := V_\delta(a) \setminus \{a\} = \{x \in \mathbb{R} : 0 < |x-a| < \delta\}.$$

Diremos que a es un **punto de acumulación** de A si todo entorno perforado de a contiene al menos un punto de A; es decir, si

$$\forall \delta > 0, \qquad V_\delta^*(a) \cap A \neq \varnothing.$$

Equivalentemente,

$$\forall \delta > 0 \; \exists x \in A : 0 < |x-a| < \delta.$$

Diremos que a es un **punto aislado** de A si $a \in A$ y existe algún radio $\delta > 0$ para el cual a es el único punto de A dentro del entorno:

$$V_\delta(a) \cap A = \{a\}.$$

Equivalentemente, $a \in A$ es aislado si existe $\delta > 0$ tal que

$$V_\delta^*(a) \cap A = \varnothing.$$

La definición contiene una asimetría importante. Para ser **punto de acumulación** de A, el punto a no necesita pertenecer a A. Lo único que exigimos es que haya puntos de A, distintos de a, arbitrariamente cerca de él. Para ser **punto aislado**, en cambio, sí exigimos $a \in A$. Un punto aislado pertenece al conjunto, pero queda separado de todos los demás puntos del conjunto por alguna distancia positiva. Esta diferencia aparecerá repetidamente en cálculo.

Primeros ejemplos: intervalos y enteros. Consideremos $A = [0, 1]$.

El punto $a = \frac{1}{2}$ es un punto de acumulación de A. En efecto, dado cualquier $\delta > 0$, podemos encontrar puntos del intervalo distintos de $1/2$ y a distancia menor que δ. Por ejemplo, basta tomar un desplazamiento positivo suficientemente pequeño. Si elegimos

$$h = \text{mín}\left\{\frac{\delta}{2}, \frac{1}{4}\right\},$$

entonces $h > 0$ y $x = \frac{1}{2} + h$ satisface

$$x \in [0, 1], \qquad 0 < \left|x - \frac{1}{2}\right| = h < \delta.$$

Los extremos 0 y 1 también son puntos de acumulación de $[0, 1]$. Esto merece atención. Para demostrar que 0 es punto de acumulación no necesitamos disponer de puntos del conjunto a ambos lados de 0. Basta con que **todo** entorno perforado contenga algún punto de A.

Dado $\delta > 0$, el número

$$x = \text{mín}\left\{\frac{\delta}{2}, \frac{1}{2}\right\}$$

pertenece a $[0, 1]$ y cumple $0 < x < \delta$.

Por tanto, $0 < |x - 0| < \delta$.

Así, 0 es punto de acumulación aunque los puntos de A que se aproximan a él lo hagan solamente desde la derecha. Esta observación será el germen de los límites laterales, pero todavía no necesitamos introducirlos formalmente.

Comparemos ahora con $A = \mathbb{Z}$.

Sea $m \in \mathbb{Z}$. Entre m y cualquier otro entero hay distancia al menos 1. Por eso, si elegimos $\delta = \frac{1}{2}$, obtenemos $V_{1/2}(m) \cap \mathbb{Z} = \{m\}$. Así, **todo entero es un punto aislado de** $\mathbb{Z}$.

La comparación es instructiva:

- en un intervalo como $[0, 1]$, podemos encontrar puntos del conjunto arbitrariamente cerca de cada uno de sus puntos;
- en $\mathbb{Z}$, alrededor de cada entero existe un pequeño intervalo que no contiene ningún otro entero.

La diferencia no es que un conjunto sea «grande» y el otro «pequeño». Es una diferencia en su estructura **local**.

Un punto de acumulación puede quedar fuera del conjunto. Consideremos ahora

$$A = \left\{ \frac{1}{n} : n \in \mathbb{N}_{>0} \right\} = \left\{ 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots \right\}.$$

El número 0 no pertenece a A: $0 \notin A$.

Sin embargo, 0 es punto de acumulación de A. Para demostrarlo debemos verificar la definición. Sea $\delta > 0$. Por la propiedad arquimediana existe $n \in \mathbb{N}_{>0}$ tal que $n > \frac{1}{\delta}$. Como ambos números son positivos, al tomar recíprocos se invierte la desigualdad: $0 < \frac{1}{n} < \delta$.

Si elegimos $x = \frac{1}{n} \in A$, entonces

$$0 < |x - 0| = \frac{1}{n} < \delta.$$

Esto funciona para **cada** radio positivo δ, por pequeño que sea. Por tanto, todos los entornos perforados de 0 contienen puntos de A.

Hemos probado que

$$0 \text{ es punto de acumulación de } A,$$

aunque $0 \notin A$.

El ejemplo muestra por qué sería incorrecto definir «punto de acumulación» exigiendo primero que el punto pertenezca al conjunto. Además conecta con `T1-C04`. Allí demostramos que $\frac{1}{n} \to 0$. Aquí estamos observando el mismo fenómeno desde otro ángulo: los **valores** $1/n$ forman un conjunto con puntos arbitrariamente próximos a 0. No confundamos, sin embargo, ambas afirmaciones. Una sucesión conserva índices y repeticiones; el conjunto A no. En esta sección solo nos interesa la geometría local del conjunto de puntos disponibles.

Los puntos $1/n$ sí son aislados. El ejemplo anterior contiene una segunda sorpresa. Aunque 0 es punto de acumulación de A, cada punto $\frac{1}{m} \in A$ es aislado. Fijemos $m \in \mathbb{N}_{>0}$. Si $m = 1$, el punto es 1 y el siguiente valor del conjunto es $1/2$, de modo que cualquier radio menor que $1/2$ aísla a 1. Si $m \geq 2$, los vecinos más próximos en la lista son

$$\frac{1}{m-1} \qquad \text{y} \qquad \frac{1}{m+1}.$$

Ambos están a distancia positiva de $1/m$. Podemos elegir, por ejemplo,

$$\delta = \frac{1}{2} \min \left\{ \frac{1}{m-1} - \frac{1}{m}, \frac{1}{m} - \frac{1}{m+1} \right\} > 0.$$

Con ese radio, ningún otro punto de A entra en $V_\delta(1/m)$. Por tanto, cada $1/m$ es aislado, mientras que 0 —que ni siquiera pertenece al conjunto— es un punto de acumulación. La figura conceptual es entonces:

muchos puntos aislados pueden acumularse alrededor de un punto que no pertenece al conjunto.

Esta posibilidad será importante cuando el dominio de una función tenga una forma menos sencilla que un intervalo.

Acercarse siempre significa acercarse dentro del dominio. Sea nuevamente $f\colon A \to \mathbb{R}$.

La expresión $f(x)$ solo tiene sentido cuando $x \in A$. Por ello, cuando más adelante escribamos informalmente que «x se acerca a a», no estaremos permitiendo que x recorra todos los números reales sin restricción. Los puntos que se aproximan deben pertenecer al dominio de la función. El objeto geométrico relevante no será por tanto solo $V_\delta^*(a)$, sino la intersección $V_\delta^*(a) \cap A$. Esta intersección responde exactamente a la pregunta:

> ¿qué puntos del dominio están disponibles a distancia menor que δ de a, sin usar el propio punto a?

Veamos tres situaciones distintas.

Si $A = \mathbb{R}$, entonces $V_\delta^*(a) \cap A = V_\delta^*(a)$, porque todos los puntos reales están permitidos.

Si

$$A = [0, \infty) \qquad \text{y} \qquad a = 0,$$

entonces los puntos disponibles cerca de 0 son únicamente los positivos: $V_\delta^*(0) \cap A = (0, \delta)$.

Si, en cambio,

$$A = \mathbb{Z} \qquad \text{y} \qquad a = 0,$$

para cualquier $0 < \delta < 1$ tenemos $V_\delta^*(0) \cap \mathbb{Z} = \varnothing$.

En las dos primeras situaciones podemos encontrar puntos del dominio arbitrariamente cerca de a. En la tercera, no. Este contraste explica por qué el dominio completo de una función, introducido en `T1-C03`, vuelve a ser decisivo ahora. La misma fórmula puede tener un comportamiento de aproximación distinto cuando cambia el dominio, porque cambian los puntos desde los cuales podemos acercarnos.

La pertenencia del centro no es lo importante. Hay cuatro posibilidades que conviene separar.

1. $a \in A$ y a es punto de acumulación de A. Por ejemplo, $a = 0$ en $A = [0, 1]$.
2. $a \notin A$ y a es punto de acumulación de A. Por ejemplo, $a = 0$ en $A = \{1/n : n \geq 1\}$.
3. $a \in A$ y a es aislado. Por ejemplo, cualquier entero en $A = \mathbb{Z}$.
4. $a \notin A$ y a no es punto de acumulación. Por ejemplo, $a = 10$ en $A = [0, 1]$.

Las dos primeras situaciones tienen algo en común: aunque el centro pertenezca o no al conjunto, existen puntos de A distintos de a arbitrariamente próximos a él. Eso es exactamente lo que necesitaremos para estudiar comportamiento **cerca** del punto. Por tanto, para la teoría de límites que viene, la pregunta relevante no será primero

> ¿pertenece a al dominio?,

sino

> ¿es posible aproximarse a a mediante puntos del dominio distintos de a?

La noción de punto de acumulación responde a esta segunda pregunta.

Por qué los puntos aislados crean un problema lógico. La condición de acumulación impide que la definición de límite quede satisfecha por falta de puntos del dominio en los que comprobarla.

Consideremos $A = \{0\} \cup [1, \infty)$.

El punto 0 pertenece a A, pero está aislado. En efecto, si $0 < \delta < 1$, entonces $V_\delta(0) \cap A = \{0\}$, y por tanto $V_\delta^*(0) \cap A = \varnothing$. No existen puntos del dominio distintos de 0 suficientemente próximos a 0.

¿Por qué debería preocuparnos esto?

Queremos que, cuando x pertenezca al dominio y esté suficientemente cerca de a sin coincidir con él, $f(x)$ quede cerca de un número L. Para que esa exigencia describa el comportamiento de la función, debe haber argumentos admisibles en cada entorno perforado. La estructura lógica básica será una implicación:

$$0 < |x - a| < \delta \quad \Longrightarrow \quad \text{una condición sobre } f(x).$$

En una implicación, si el antecedente no se cumple, no existe en ese caso un contraejemplo al enunciado condicional. Si no existe ningún $x \in A$ que satisfaga $0 < |x - a| < \delta$, entonces cualquier afirmación de la forma

$$\forall x \in A: \quad 0 < |x - a| < \delta \Longrightarrow Q(x)$$

resulta automáticamente verdadera, sea cual sea la propiedad $Q(x)$. Eso es una **verdad vacua**. La lógica de la implicación es correcta; el problema es que la afirmación ya no está controlando ningún punto real del dominio.

Ejemplo 4.1. **Un punto aislado y el peligro de un criterio vacuo.** Sea $A = \{0\} \cup [1, \infty)$ y sea $f: A \to \mathbb{R}$ cualquier función. No necesitamos especificar sus valores.

Tomemos $a = 0$. Como 0 es aislado, podemos elegir $\delta = \frac{1}{2}$. Entonces no existe ningún $x \in A$ tal que $0 < |x| < \frac{1}{2}$. Fijemos ahora cualquier $L \in \mathbb{R}$ y cualquier $\varepsilon > 0$. La afirmación

$$\forall x \in A: \quad 0 < |x| < \frac{1}{2} \Longrightarrow |f(x) - L| < \varepsilon$$

es verdadera, porque su antecedente es falso para todos los $x \in A$. Más aún: el mismo argumento funciona para **todo** número real L. Por tanto, si intentáramos caracterizar el comportamiento próximo a 0 utilizando únicamente una condición de ese tipo y sin exigir que 0 sea punto de acumulación del dominio, el criterio aceptaría simultáneamente cualquier candidato L. No habríamos descubierto que la función se aproxima a muchos números distintos. Simplemente habríamos formulado una condición que no examina ningún punto del dominio.

Este ejemplo explica una decisión que adoptaremos durante todo el capítulo:

solo estudiaremos el límite de f
cuando el punto de aproximación
sea punto de acumulación de su dominio.

La condición garantiza que cada entorno perforado contiene algún punto del dominio. Por pequeño que elijamos el radio, la condición sobre las imágenes tendrá argumentos en los que ejercer su control. También prepara un resultado esencial: la unicidad del límite. Si alrededor de a no hubiera puntos disponibles, ninguna comparación entre dos supuestos valores límite podría tener contenido.

Una condición local, no global. Ser punto de acumulación es una propiedad enteramente local. Para decidir si a es punto de acumulación de A no necesitamos comprender todo el conjunto. Solo necesitamos saber qué ocurre en entornos arbitrariamente pequeños alrededor de a.

Por ejemplo, consideremos $A = (-1, 1) \cup \{1000\}$.

El punto 0 es punto de acumulación de A. La presencia del punto 1000 no cambia absolutamente nada alrededor de 0. Del mismo modo, si definimos

$$B = (-1, 1) \cup \{10^6, 10^6 + 1, 10^6 + 2, \dots\},$$

el comportamiento local de B cerca de 0 sigue siendo exactamente el mismo que el de $(-1, 1)$. Esta insensibilidad frente a lo que ocurre lejos del punto es uno de los rasgos fundamentales de la teoría de límites. Más adelante podremos modificar una función en regiones alejadas de a sin afectar su comportamiento límite en a. Incluso podremos modificar su valor exactamente en a. Pero para justificar esas afirmaciones necesitaremos primero la definición rigurosa. Por ahora basta con reconocer la geometría que las hace plausibles:

el análisis local observa entornos cada vez más pequeños alrededor del punto.

No hace falta introducir todavía toda la topología de la recta. Las palabras **entorno**, **punto de acumulación** y **punto aislado** pertenecen al lenguaje topológico de $\mathbb{R}$. Podríamos continuar inmediatamente y definir conjuntos abiertos y cerrados, interior, clausura, frontera y otras nociones relacionadas. Todas ellas serán importantes más adelante. Sin embargo, hacerlo ahora ocultaría la razón por la que necesitamos estas primeras definiciones. Nuestro objetivo inmediato es mucho más concreto: dar significado preciso a la frase

> «x se acerca a a dentro del dominio de f».

Para ello bastan tres ingredientes:

1. medir cercanía mediante $|x - a|$;
2. excluir el centro mediante $0 < |x - a|$;
3. exigir que existan puntos del dominio en todos los entornos perforados.

En símbolos, el tercer ingrediente es

$$\forall \delta > 0 \; \exists x \in A : 0 < |x - a| < \delta.$$

Eso es exactamente decir que a es punto de acumulación de A. La teoría topológica más sistemática aparecerá cuando resuelva nuevos problemas matemáticos, no antes.

Antes de seguir. Conviene comprobar que las distinciones anteriores están firmes.

Si $A = (0, 1)$,

¿es 0 punto de acumulación de A?

Sí. Aunque $0 \notin A$, para todo $\delta > 0$ podemos escoger un número positivo suficientemente pequeño que pertenezca simultáneamente a $(0, 1)$ y a $(0, \delta)$.

Si $A = \{2, 5, 8\}$,

¿es 5 punto aislado?

Sí. Por ejemplo, con $\delta = 1$ tenemos $V_1(5) = (4, 6)$ y $V_1(5) \cap A = \{5\}$.

¿Es 4 punto aislado de ese mismo conjunto?

No, porque $4 \notin A$.

Tampoco es punto de acumulación: existe, por ejemplo, un entorno suficientemente pequeño alrededor de 4 que no contiene puntos de A.

Finalmente, si $A = \mathbb{Q}$, todo número real es punto de acumulación de A. Esto es una reformulación local de la densidad de los racionales demostrada en `T1-C02`: entre números reales arbitrariamente próximos siempre podemos encontrar racionales. En particular, incluso un irracional $a \notin \mathbb{Q}$ puede ser punto de acumulación de $\mathbb{Q}$.

De la cola al entorno perforado. Ya podemos ver con precisión qué parte de la teoría de sucesiones sobrevivirá y qué parte deberá cambiar. Para una sucesión, decir que estudiamos su comportamiento final significa ignorar un número finito de términos iniciales y pasar a una cola: $\{n \in \mathbb{N} : n \geq N\}$. Para una función, decir que estudiamos su comportamiento cerca de a significará ignorar los puntos suficientemente lejanos y concentrarnos en $V_\delta^*(a) \cap A$.

El paralelismo es:

Sucesiones	Funciones
índice n	variable real x
cola $n \geq N$	entorno perforado $0 < \lvert x - a \rvert < \delta$
umbral N	radio δ
término a_n	valor $f(x)$
dominio discreto	dominio real posiblemente continuo o irregular

Hay, sin embargo, una diferencia profunda. En la sucesión, una cola siempre contiene infinitos índices. En una función sobre un dominio arbitrario, un entorno perforado podría no contener **ningún** punto del dominio. Precisamente por eso necesitamos la hipótesis de acumulación. Al exigir que a sea punto de acumulación de A, garantizamos que $V_\delta^*(a) \cap A \neq \varnothing$ para todo $\delta > 0$. Así queda preparado el escenario lógico para la próxima sección. Hasta aquí hemos hablado únicamente de **dónde puede estar** x. Todavía no hemos impuesto ninguna condición sobre **dónde debe estar** $f(x)$. Ese será el paso siguiente.

Para una sucesión, una tolerancia ε alrededor de L exigía encontrar un umbral N que controlase todos los términos posteriores. En una función real buscaremos un radio δ alrededor de a que garantice la precisión de salida para todos los puntos del dominio situados dentro del entorno perforado correspondiente. La próxima sección convertirá esa idea en una definición cuantificada completa:

$$\boxed{\text{cola } n \geq N \quad \longrightarrow \quad \text{entorno perforado } 0 < |x - a| < \delta.}$$

Con el lenguaje local ya establecido, la definición ε–δ podrá aparecer como lo que realmente es: la versión funcional de una lógica de aproximación que ya conocemos.

4.2. De la cola al entorno perforado: la definición ε–δ

La misma pregunta, con otra geometría. En `T1-C04`, para expresar que una sucesión (a_n) se aproxima a L, fijábamos primero una tolerancia $\varepsilon > 0$ en la salida y después buscábamos un índice N a partir del cual todos los términos quedaran dentro de esa tolerancia:

$$n \geq N \quad \Longrightarrow \quad |a_n - L| < \varepsilon.$$

La estructura lógica era $\forall \varepsilon > 0\ \exists N\ \forall n \geq N$.

Ahora queremos formular la misma idea para una función

$$f\colon A \to \mathbb{R}, \qquad A \subseteq \mathbb{R},$$

cuando la variable x se aproxima a un punto a. La diferencia es geométrica. Ya no disponemos de una cola de índices. En su lugar tenemos los puntos del dominio que caen en un entorno perforado de a:

$$x \in A, \qquad 0 < |x - a| < \delta.$$

El radio δ desempeñará el papel que antes desempeñaba el umbral N. La traducción fundamental es, por tanto,

Sucesiones	Funciones
fijamos $\varepsilon > 0$	fijamos $\varepsilon > 0$
buscamos N	buscamos $\delta > 0$
$n \geq N$	$0 < \lvert x - a\rvert < \delta$
controlamos todos los términos de la cola	controlamos todos los puntos del dominio en el entorno perforado
$\lvert a_n - L\rvert < \varepsilon$	$\lvert f(x) - L\rvert < \varepsilon$

Hay una inversión intuitiva que conviene notar. En las sucesiones, para mejorar el control avanzamos hacia índices **más grandes**. En las funciones, para mejorar el control reducimos el radio y miramos puntos **más cercanos** a a. Pero la lógica profunda es la misma:

dada cualquier precisión de salida, debemos poder imponer una condición de entrada que la garantice.

Qué queremos controlar. Supongamos que esperamos que los valores de $f(x)$ se aproximen a un número L cuando x se aproxima a a. La tolerancia de salida se expresa mediante $|f(x)-L| < \varepsilon$. Esto significa que $f(x)$ debe quedar dentro del intervalo $(L-\varepsilon, L+\varepsilon)$. La condición de entrada se expresa mediante $0 < |x - a| < \delta$. Esto significa que x debe pertenecer al entorno perforado $V_\delta^*(a)$. Pero, como vimos en §4.1, los acercamientos se realizan **dentro del dominio**. Por tanto, los puntos relevantes son exactamente $A \cap V_\delta^*(a)$.

La pregunta completa es:

> dada una tolerancia $\varepsilon > 0$ alrededor de L, ¿podemos encontrar un radio $\delta > 0$ alrededor de a tal que **todo** punto permitido del dominio dentro de ese entorno perforado produzca un valor de la función dentro de la tolerancia?

Esta frase contiene ya la definición.

Definición 4.2. **Límite funcional finito.** Sea

$$f\colon A \to \mathbb{R}, \qquad A \subseteq \mathbb{R},$$

y sea $a \in \mathbb{R}$ un punto de acumulación de A. Diremos que $f(x)$ **tiende a** L **cuando** x **tiende a** a, o que f **tiene límite** L **en** a, si

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in A: \qquad 0 < |x - a| < \delta \Longrightarrow |f(x) - L| < \varepsilon.$$

En ese caso escribimos

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

Equivalentemente,

$$\forall \varepsilon > 0\ \exists \delta > 0: \qquad f\big(A \cap V_\delta^*(a)\big) \subseteq (L-\varepsilon, L+\varepsilon).$$

La negación de la afirmación

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

es

$$\exists \varepsilon_0 > 0\ \forall \delta > 0\ \exists x \in A: \qquad 0 < |x-a| < \delta \quad \text{y} \quad |f(x) - L| \geq \varepsilon_0.$$

La hipótesis de acumulación asegura que existen argumentos del dominio distintos de a arbitrariamente próximos a él. Sin ella, la implicación local podría ser verdadera porque no hubiese puntos que cumplieran su antecedente; esa verdad por vacuidad no describiría una aproximación de los valores de la función.

Leer los cuantificadores en el orden correcto. La parte esencial de la definición no es la presencia de las letras griegas. Es su orden:

$$\boxed{\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in A.}$$

Leámoslo lentamente.

Primero aparece $\forall \varepsilon > 0$.

Esto significa que no podemos elegir una precisión cómoda. La afirmación debe funcionar para **toda** tolerancia positiva: 1, 0.1, 10^{-6} o cualquier otra. Solo después de conocer ε podemos escoger $\delta > 0$. Por eso δ puede depender de ε. En una demostración es completamente legítimo escribir, por ejemplo, $\delta = \frac{\varepsilon}{3}$ o $\delta = \min\left\{1, \frac{\varepsilon}{7}\right\}$. Lo que no sería legítimo es fijar una única tolerancia ε y declarar terminado el problema. Tampoco podemos elegir δ después de conocer el punto particular x que queremos controlar.

El orden exige:

1. el adversario fija ε;
2. nosotros respondemos con un único δ;
3. ese mismo δ debe funcionar para **todos** los puntos $x \in A$ que satisfagan $0 < |x-a| < \delta$.

Por eso la última parte es universal: $\forall x \in A$. No basta con encontrar algunos puntos cercanos cuyos valores estén cerca de L. Debemos controlar todos los puntos permitidos suficientemente próximos al centro.

δ puede depender de ε, pero no del punto examinado. Esta restricción merece aislarse porque es una fuente frecuente de falsas demostraciones. Imaginemos que, después de recibir un punto x, escogemos $\delta = 2|x-a|$.

Entonces ciertamente ese x satisface $|x-a| < \delta$.

Pero hemos invertido el orden lógico. La definición no permite observar primero el punto y fabricar después un radio a su medida. El radio debe estar decidido **antes** del cuantificador universal sobre x.

Esquemáticamente,

$$\varepsilon \longmapsto \delta(\varepsilon) \longmapsto \text{control de todos los } x.$$

No $\varepsilon \longmapsto x \longmapsto \delta(\varepsilon, x)$.

En las primeras pruebas convendrá incluso escribir explícitamente la dependencia: $\delta = \delta(\varepsilon)$. Más adelante, cuando la arquitectura sea familiar, podremos comprimir esa notación.

La desigualdad de entrada tiene dos partes. La condición $0 < |x - a| < \delta$ reúne dos afirmaciones diferentes. La desigualdad $|x - a| < \delta$ dice que x está suficientemente cerca de a.

La desigualdad $0 < |x - a|$ dice que $x \neq a$.

El segundo requisito es lo que perfora el entorno. Por eso el valor exacto de la función en a no aparece en la definición. Puede ocurrir que $a \notin A$, de modo que $f(a)$ ni siquiera esté definido, y aun así la pregunta por el comportamiento de $f(x)$ alrededor de a tenga perfecto sentido siempre que a sea punto de acumulación de A. También puede ocurrir que $a \in A$ y que $f(a)$ tenga un valor completamente distinto del comportamiento de los puntos próximos. La definición está diseñada precisamente para separar:

comportamiento alrededor de a	de	valor exactamente en a.

Primer ejemplo: una función con un agujero. Consideremos la expresión $\frac{x^2-1}{x-1}$.

En $x = 1$ el denominador se anula, así que la función

$$f\colon \mathbb{R} \setminus \{1\} \to \mathbb{R}, \qquad f(x) = \frac{x^2 - 1}{x - 1},$$

no está definida en 1.

Sin embargo, para $x \neq 1$ podemos factorizar: $x^2 - 1 = (x-1)(x+1)$, y por tanto $f(x) = x + 1$. Esto sugiere que, cuando x se aproxima a 1 sin tocarlo, los valores de $f(x)$ deberían aproximarse a 2. La definición nos obliga ahora a demostrar esa intuición.

Ejemplo 4.2. **Un límite en un punto donde la función no está definida.** Sea

$$f\colon \mathbb{R} \setminus \{1\} \to \mathbb{R}, \qquad f(x) = \frac{x^2 - 1}{x - 1}.$$

Demostraremos directamente que

$$\lim_{x \to 1} f(x) = 2.$$

El punto 1 es punto de acumulación del dominio $\mathbb{R} \setminus \{1\}$, porque todo entorno perforado de 1 contiene números reales distintos de 1.

Sea $\varepsilon > 0$.

Queremos encontrar $\delta > 0$ tal que, para todo $x \neq 1$, $0 < |x - 1| < \delta$ implique

$$\left| \frac{x^2 - 1}{x - 1} - 2 \right| < \varepsilon.$$

Como $x \neq 1$, podemos simplificar: $\frac{x^2-1}{x-1} = x + 1$.

Entonces

$$\left|\frac{x^2-1}{x-1}-2\right| = |x+1-2| = |x-1|.$$

La expresión que queremos hacer menor que ε coincide exactamente con la distancia de x al punto 1.

Por ello elegimos $\delta = \varepsilon$.

Si $0 < |x-1| < \delta$, entonces

$$\left|\frac{x^2-1}{x-1}-2\right| = |x-1| < \delta = \varepsilon.$$

Esto prueba que

$$\lim_{x\to 1} \frac{x^2-1}{x-1} = 2$$

para la función definida sobre $\mathbb{R} \setminus \{1\}$.

La demostración muestra una característica típica de las pruebas ε–δ: la parte creativa suele ocurrir antes de escribir la prueba formal.

Aquí comenzamos preguntando: $|f(x)-2| =?$

La simplificación reveló que $|f(x)-2| = |x-1|$.

Una vez encontrada esa relación, la elección $\delta = \varepsilon$ resulta natural. La prueba final puede ser breve porque la búsqueda previa ya descubrió qué condición sobre la entrada controla el error de salida.

Trabajar hacia atrás y demostrar hacia delante. Para problemas algo más difíciles, será útil distinguir dos fases. En la fase de **búsqueda**, partimos del objetivo $|f(x)-L| < \varepsilon$ y manipulamos la expresión para descubrir qué condición suficiente sobre $|x-a|$ garantizaría esa desigualdad. En la fase de **demostración**, invertimos el recorrido:

1. fijamos $\varepsilon > 0$;
2. declaramos una elección concreta de δ;
3. suponemos $0 < |x-a| < \delta$;
4. deducimos $|f(x)-L| < \varepsilon$.

Esta separación ya apareció en el capítulo sobre demostraciones. Trabajar hacia atrás puede ser una excelente estrategia de descubrimiento, siempre que la prueba final verifique hacia delante que la elección realmente funciona. En el ejemplo anterior, la búsqueda fue $|f(x)-2| = |x-1| < \varepsilon$, de donde descubrimos que bastaba exigir $|x-1| < \varepsilon$. La demostración escribió entonces $\delta = \varepsilon$ y recorrió las implicaciones en el sentido correcto.

El valor puntual no participa. El ejemplo con $f(x) = \frac{x^2-1}{x-1}$ muestra que $f(1)$ no necesita existir. Podemos ir más lejos. Consideremos funciones que coinciden para todos los puntos cercanos distintos de 1, pero que tienen valores diferentes exactamente en 1.

Por ejemplo,

$$g(x) = \begin{cases} x+1, & x \neq 1, \\ 1000, & x = 1, \end{cases}$$

y $h(x) = x+1$.

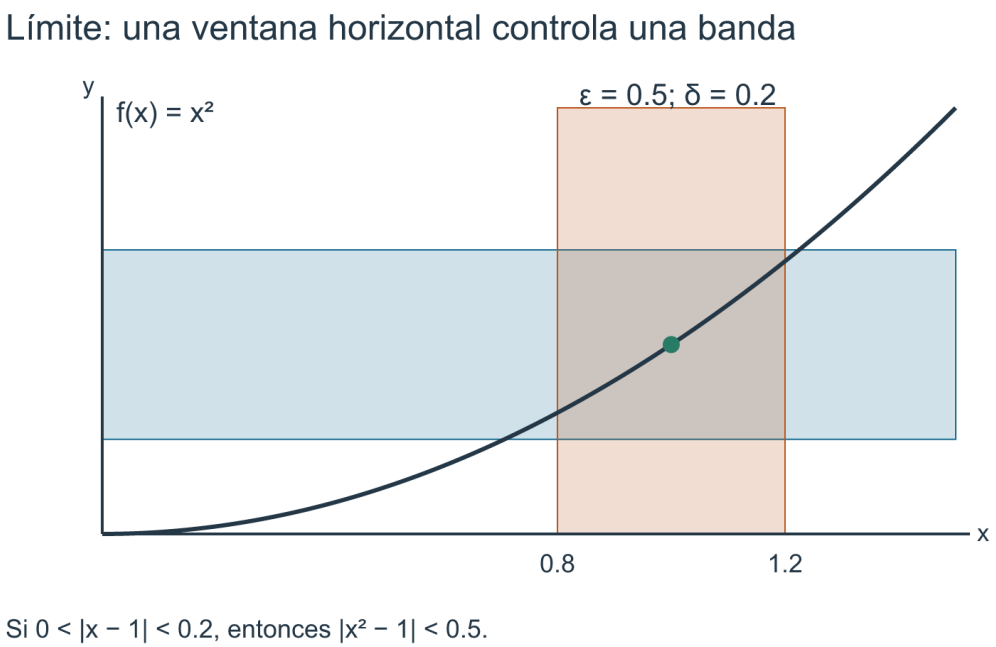


FIGURA 4.1. Para $f(x) = x^2$ cerca de 1, la ventana $0 < |x-1| < 0.2$ obliga a que $|x^2 - 1| < 0.5$. La ventana de entrada y la banda de salida representan dos tolerancias diferentes.

La primera satisface $g(1) = 1000$, mientras que la segunda satisface $h(1) = 2$. La función racional anterior ni siquiera tenía valor en 1. Sin embargo, las tres se comportan exactamente igual en todo entorno perforado de 1.

EJEMPLO 4.3. **Mismo comportamiento perforado, valores puntuales distintos.** Consideremos

$$f\colon \mathbb{R} \setminus \{1\} \to \mathbb{R}, \qquad f(x) = \frac{x^2 - 1}{x - 1},$$

$$g\colon \mathbb{R} \to \mathbb{R}, \qquad g(x) = \begin{cases} x + 1, & x \neq 1, \\ 1000, & x = 1, \end{cases}$$

y

$$h\colon \mathbb{R} \to \mathbb{R}, \qquad h(x) = x + 1.$$

Para todo $x \neq 1$,

$$f(x) = g(x) = h(x) = x + 1.$$

En particular, si $0 < |x-1| < \delta$, entonces las tres funciones tienen exactamente el mismo valor. La prueba del ejemplo anterior se aplica sin cambio a los tres casos: dado $\varepsilon > 0$, elegimos $\delta = \varepsilon$. Entonces, para todo punto del dominio correspondiente con $0 < |x-1| < \delta$, se cumple

$$|f(x) - 2| = |g(x) - 2| = |h(x) - 2| = |x - 1| < \varepsilon.$$

Por tanto,

$$\lim_{x\to 1} f(x) = \lim_{x\to 1} g(x) = \lim_{x\to 1} h(x) = 2.$$

Sin embargo,

$$f(1) \text{ no existe}, \qquad g(1) = 1000, \qquad h(1) = 2.$$

El límite describe el comportamiento en puntos próximos distintos del centro; no impone todavía ninguna relación con el valor puntual.

Esta separación será decisiva en el capítulo siguiente. Allí estudiaremos precisamente qué condición adicional conecta el comportamiento límite con el valor de la función en el punto. Por ahora debemos resistir la tentación de anticiparla.

La definición es relativa al dominio. La escritura $x \to a$ no significa que x pueda recorrer indiscriminadamente todos los números reales. Si $f\colon A \to \mathbb{R}$, el cuantificador de la definición dice explícitamente $\forall x \in A$. Solo se examinan puntos pertenecientes al dominio. Esto puede cambiar radicalmente el problema. Supongamos, por ejemplo, que $A = (0, \infty)$ y queremos estudiar el comportamiento cuando $x \to 0$. El punto 0 no pertenece a A, pero es punto de acumulación de A. Los únicos puntos disponibles cerca de 0 son positivos.

En cambio, si el dominio fuera $B = \mathbb{R} \setminus \{0\}$, habría puntos disponibles a ambos lados de 0. El dominio determina qué argumentos pueden aproximarse al centro. La definición debe conservar esa información incluso cuando los puntos admisibles se encuentran solo a un lado.

La notación abreviada

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

oculta visualmente el conjunto A, pero la definición no lo olvida.

Cómo negar correctamente un candidato a límite. La negación incluida en Definición 4.2 merece una lectura independiente.

La afirmación positiva dice:

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in A: \quad 0 < |x - a| < \delta \Longrightarrow |f(x) - L| < \varepsilon.$$

Para negarla invertimos los cuantificadores y negamos la conclusión de la implicación bajo la condición de entrada. Obtenemos:

$$\exists \varepsilon_0 > 0\ \forall \delta > 0\ \exists x \in A: \quad 0 < |x - a| < \delta \quad \text{y} \quad |f(x) - L| \geq \varepsilon_0.$$

En palabras:

> existe una tolerancia positiva fija ε_0 que no podemos garantizar, por pequeño que hagamos el entorno de entrada.

El orden vuelve a ser esencial. Para demostrar que un candidato L falla, debemos encontrar **una sola** tolerancia $\varepsilon_0 > 0$ que resista todos nuestros intentos de estrechar el radio. Después, para cada $\delta > 0$, debemos producir algún punto del dominio suficientemente próximo a a cuya imagen permanezca fuera de esa tolerancia. El punto testigo puede depender de δ: $x = x(\delta)$. La tolerancia problemática, en cambio, queda fija: ε_0. Esta estructura será especialmente útil cuando estudiemos oscilación y, más adelante, cuando conectemos los límites funcionales con sucesiones.

Una afirmación sobre todos los puntos, no sobre una muestra. La definición explica también por qué una tabla o una gráfica aproximada nunca constituyen por sí solas una demostración de límite.

Podemos calcular

$$f(0.9), \quad f(0.99), \quad f(0.999), \quad f(1.001), \quad f(1.01)$$

y observar valores muy próximos a cierto número L. Eso puede sugerir una conjetura. Pero la definición exige controlar **todos** los puntos del dominio que estén suficientemente cerca de a, no solo una lista finita o una colección escogida de ejemplos. La diferencia es la misma que aprendimos al comenzar el libro:

evidencia numérica puede sugerir; la definición cuantificada exige demostrar.

Una gráfica puede ayudarnos a descubrir qué límite esperar y qué estimaciones podrían funcionar. La prueba debe transformar esa intuición en una elección de δ válida para cualquier tolerancia ε.

Qué hemos ganado con la definición. La expresión

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

ya no significa simplemente «parece que los valores se acercan a L». Significa una afirmación precisa con cuatro piezas:

1. **dominio de aproximación:** $x \in A$;
2. **centro excluido:** $x \neq a$;
3. **control de entrada:** $|x - a| < \delta$;
4. **control de salida:** $|f(x) - L| < \varepsilon$.

Y esas piezas están enlazadas por el orden $\forall \varepsilon \; \exists \delta \; \forall x$. Podemos resumir la arquitectura en una sola línea:

por pequeña que sea la tolerancia de salida, podemos hacer suficientemente pequeño el entorno de entrada para controlar todos sus puntos permitidos.

Todavía quedan preguntas importantes.

Si dos números distintos parecieran satisfacer esta definición, ¿puede ocurrir realmente?

Si modificamos una función en un solo punto, ya vimos por ejemplos que el límite no cambia; ¿podemos convertir esa observación en un principio general?

Si un límite finito existe, ¿qué otras propiedades locales quedan forzadas?

La siguiente sección responderá estas preguntas. Allí demostraremos que el límite funcional es único y extraeremos las primeras consecuencias estructurales de la definición que acabamos de construir.

4.3. Qué fuerza tiene un límite: unicidad y estabilidad local

La definición ε–δ no se limita a decir que los valores de una función «se acercan» a cierto número. Una vez que un límite finito existe, quedan forzadas varias propiedades locales que utilizaremos repetidamente durante el resto del capítulo. En esta sección responderemos tres preguntas.

1. Si $f(x)$ tiene límite cuando $x \to a$, ¿puede haber dos números distintos que satisfagan la definición?
2. Si dos funciones son iguales suficientemente cerca de a, salvo quizá en el propio punto a, ¿pueden tener comportamientos límite distintos?
3. Si $f(x) \to L$, ¿qué control podemos obtener sobre el tamaño de $f(x)$ cerca de a? ¿Qué ocurre si además $L \neq 0$?

Las respuestas condensan una idea común:

un límite finito impone estructura en algún entorno perforado del punto.

La palabra **local** es esencial. Ninguno de los resultados de esta sección dice cómo debe comportarse la función lejos de a.

El límite funcional es único. En `T1-C04` demostramos que una sucesión convergente no puede tener dos límites distintos. El argumento utilizaba la distancia entre los dos candidatos y un término suficientemente tardío que debía quedar cerca de ambos. Para funciones aparece la misma geometría, pero con una diferencia lógica importante. Supongamos que dos números L y M fueran límites de $f(x)$ cuando $x \to a$. Si $L \neq M$, podemos tomar entornos de salida suficientemente pequeños alrededor de L y de M para que un mismo valor $f(x)$ no pueda pertenecer a ambos.

La definición ε–δ nos dará un radio para cada candidato. Tomaremos el menor de los dos radios. Pero, a diferencia del caso secuencial, todavía falta una pieza: debemos saber que **existe efectivamente algún punto del dominio distinto de a dentro de ese radio**. Esa existencia es exactamente lo que garantiza que a sea un punto de acumulación del dominio.

Teorema 4.1. **Unicidad del límite funcional.** Sea

$$f\colon A \to \mathbb{R},$$

y sea a un punto de acumulación de A. Si

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

y también

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

entonces

$$L = M.$$

Demostración. Supongamos, buscando una contradicción, que

$$L \neq M.$$

Entonces la distancia entre ambos candidatos es positiva. Definamos

$$d := |L - M| > 0$$

y elijamos

$$\varepsilon = \frac{d}{3}.$$

Como

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

existe $\delta_L > 0$ tal que, para todo $x \in A$,

$$0 < |x - a| < \delta_L \quad \Longrightarrow \quad |f(x) - L| < \frac{d}{3}.$$

Análogamente, como

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

existe $\delta_M > 0$ tal que, para todo $x \in A$,

$$0 < |x - a| < \delta_M \quad \Longrightarrow \quad |f(x) - M| < \frac{d}{3}.$$

Tomemos

$$\delta = \text{mín}\{\delta_L, \delta_M\} > 0.$$

Ahora entra una hipótesis que no podemos omitir. Como a es punto de acumulación de A, existe algún $x \in A$ tal que

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

Para este mismo punto tenemos simultáneamente

$$|f(x) - L| < \frac{d}{3}$$

y

$$|f(x) - M| < \frac{d}{3}.$$

Por la desigualdad triangular,

$$|L - M| \leq |L - f(x)| + |f(x) - M| < \frac{d}{3} + \frac{d}{3} = \frac{2d}{3}.$$

Pero $|L - M| = d$, de modo que

$$d < \frac{2d}{3},$$

lo cual es imposible porque $d > 0$. Por tanto, la suposición $L \neq M$ es falsa y necesariamente

$$L = M.$$

La prueba se parece mucho a la unicidad del límite secuencial, pero conviene aislar la diferencia. En una sucesión, después de elegir $N = \max\{N_L, N_M\}$, sabemos automáticamente que existen índices $n \geq N$ y, por tanto, términos a_n sobre los cuales aplicar simultáneamente las dos estimaciones. En una función, después de elegir $\delta = \min\{\delta_L, \delta_M\}$, no es automático que exista un punto del dominio en $V_\delta^*(a)$.

La condición

$$a \text{ es punto de acumulación de } A$$

es precisamente la que suministra ese punto. Así vemos ahora, dentro de una demostración, por qué la convención adoptada en §4.1 era necesaria. Si intentáramos hablar de límite en un punto aislado, la condición ε–δ podría hacerse verdadera por vacuidad para muchos candidatos; la prueba de unicidad se detendría exactamente en el momento en que necesitáramos elegir un x próximo distinto de a. A partir de Teorema 4.1, la notación

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

queda libre de ambigüedad: cuando el límite funcional finito existe en el sentido de Definición 4.2, existe un único número real que puede ocupar el lugar de L.

El límite depende solo de lo que ocurre localmente. En §4.2 vimos tres funciones con comportamientos idénticos alrededor de 1 y valores distintos —o inexistentes— en el propio punto. Ahora podemos elevar esa observación a un principio general. No necesitamos que dos funciones sean iguales en todo su dominio. Para comparar sus límites en a, basta con que, dentro de algún radio, presenten exactamente los mismos puntos de dominio disponibles y los mismos valores en esos puntos.

Proposición 4.1. **Estabilidad bajo coincidencia local.** Sean

$$f\colon A \to \mathbb{R}, \qquad g\colon B \to \mathbb{R},$$

y supongamos que a es punto de acumulación tanto de A como de B. Supongamos además que existe $r > 0$ tal que

$$A \cap V_r^*(a) = B \cap V_r^*(a)$$

y que, para todo punto x de ese conjunto común,

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

Entonces, para todo $L \in \mathbb{R}$,

$$\lim_{x \to a} f(x) = L \quad \Longleftrightarrow \quad \lim_{x \to a} g(x) = L.$$

En particular, añadir, quitar o cambiar el valor de una función **en el propio punto** a no modifica su límite cuando $x \to a$.

Demostración. Supongamos primero que

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

Sea $\varepsilon > 0$. Por la definición del límite de f, existe $\delta_f > 0$ tal que, para todo $x \in A$,

$$0 < |x - a| < \delta_f \quad \Longrightarrow \quad |f(x) - L| < \varepsilon.$$

Definimos

$$\delta = \text{mín}\{\delta_f, r\}.$$

Sea ahora $x \in B$ y supongamos que

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

Como $\delta \leq r$, tenemos

$$x \in B \cap V_r^*(a).$$

Por la igualdad local de los dominios,

$$x \in A \cap V_r^*(a),$$

y por la coincidencia local de los valores,

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

Además, $\delta \leq \delta_f$, de modo que

$$|f(x) - L| < \varepsilon.$$

Por tanto,

$$|g(x) - L| = |f(x) - L| < \varepsilon.$$

Hemos demostrado que

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

La implicación inversa se obtiene intercambiando los papeles de f y g. En consecuencia,

$$\lim_{x \to a} f(x) = L \quad \Longleftrightarrow \quad \lim_{x \to a} g(x) = L.$$

La operación central de la prueba es el análogo funcional del máximo que utilizábamos con sucesiones:

$$\boxed{\text{radio de control del límite} + \text{radio de coincidencia} \longrightarrow \text{mín}\{\delta_f, r\}.}$$

Para sucesiones, varias propiedades eventuales se sincronizan avanzando hasta el mayor de sus umbrales. Para funciones, varias propiedades locales se sincronizan **reduciendo el radio hasta el menor de los radios disponibles**.

Esta dualidad es útil:

$$\boxed{\begin{array}{rcl} \text{sucesiones} & : & \text{máx}\{N_1, N_2, \dots\}, \\ \text{funciones} & : & \text{mín}\{\delta_1, \delta_2, \dots\}. \end{array}}$$

En ambos casos buscamos una región donde todas las condiciones relevantes sean verdaderas simultáneamente.

Cambiar el valor puntual no cambia el límite. La última afirmación de Proposición 4.1 merece quedar completamente explícita. Supongamos que f y g tienen el mismo comportamiento en todos los puntos distintos de a y que solamente difieren en el centro. Por ejemplo, puede ocurrir que

$$f(a) = 2, \qquad g(a) = 10^{100},$$

mientras

$$f(x) = g(x) \qquad \text{para todo } x \neq a$$

en un entorno del punto. También puede ocurrir que una de las dos funciones ni siquiera esté definida en a. En cualquiera de esos casos, los dominios perforados y los valores correspondientes coinciden localmente. Como la definición de límite solo cuantifica sobre puntos que satisfacen $0 < |x - a| < \delta$, el punto a queda excluido desde el comienzo.

Por eso

el límite en a es una propiedad del comportamiento perforado alrededor de a, no del valor puntual.

Más adelante, la continuidad añadirá precisamente una condición que relaciona ambas cosas. En este capítulo todavía las mantenemos separadas.

Un límite finito obliga a la función a permanecer localmente acotada. Supongamos ahora que

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

La definición permite exigir que $f(x)$ esté tan cerca de L como queramos. Para obtener acotación local no necesitamos una tolerancia sofisticada: basta escoger una tolerancia fija, por ejemplo $\varepsilon = 1$.

Entonces, suficientemente cerca de a, $|f(x) - L| < 1$.

Geométricamente, todos esos valores quedan encerrados en el intervalo $(L - 1, L + 1)$. Por tanto, su tamaño no puede crecer sin límite dentro de ese entorno perforado. Si además $L \neq 0$, podemos pedir algo más fuerte. Como L está a una distancia positiva de 0, podemos obligar a $f(x)$ a permanecer suficientemente cerca de L para que no alcance el cero. Ambas consecuencias se reúnen en la siguiente proposición.

Proposición 4.2. **Acotación local y separación respecto de cero.** Sea

$$f \colon A \to \mathbb{R},$$

sea a un punto de acumulación de A y supongamos que

$$\lim_{x \to a} f(x) = L \in \mathbb{R}.$$

Entonces:

1. **Acotación local.** Existen $r > 0$ y $M > 0$ tales que, para todo $x \in A$,
$$0 < |x - a| < r \quad \Longrightarrow \quad |f(x)| < M.$$

2. **Separación respecto de cero.** Si $L \neq 0$, existe $r_0 > 0$ tal que, para todo $x \in A$,

$$0 < |x - a| < r_0 \quad \Longrightarrow \quad |f(x)| > \frac{|L|}{2} > 0.$$

En particular,

$$f(x) \neq 0$$

en algún entorno perforado suficientemente pequeño de a.

Demostración. Para la primera afirmación, aplicamos la definición del límite con

$$\varepsilon = 1.$$

Existe entonces $r > 0$ tal que, para todo $x \in A$,

$$0 < |x - a| < r \quad \Longrightarrow \quad |f(x) - L| < 1.$$

Por la desigualdad triangular,

$$|f(x)| = |f(x) - L + L| \leq |f(x) - L| + |L| < 1 + |L|.$$

Por tanto, basta tomar

$$M = |L| + 1.$$

Esto demuestra la acotación local.
Supongamos ahora que

$$L \neq 0.$$

Entonces

$$\frac{|L|}{2} > 0.$$

Aplicamos la definición del límite con

$$\varepsilon = \frac{|L|}{2}.$$

Existe $r_0 > 0$ tal que

$$0 < |x - a| < r_0 \quad \Longrightarrow \quad |f(x) - L| < \frac{|L|}{2}.$$

Por la desigualdad triangular inversa,

$$\big||f(x)| - |L|\big| \leq |f(x) - L|.$$

En particular,

$$|f(x)| \geq |L| - |f(x) - L| > |L| - \frac{|L|}{2} = \frac{|L|}{2}.$$

Así,

$$|f(x)| > \frac{|L|}{2} > 0,$$

y por tanto

$$f(x) \neq 0$$

para todos los puntos del dominio suficientemente próximos a a y distintos de a.

Las dos partes de Proposición 4.2 responden a necesidades distintas. La acotación local nos dice que un límite finito excluye explosiones arbitrarias de $f(x)$ **suficientemente cerca del punto**. No afirma que f sea acotada en todo su dominio. Una función puede tender a 3 cuando $x \to 0$ y, al mismo tiempo, tomar valores arbitrariamente grandes lejos de 0. La separación respecto de cero tampoco dice que f nunca se anule. Dice solamente que, si su límite en a es un número no nulo, entonces existe una región perforada alrededor de a donde los valores ya no pueden tocar el cero. Esta segunda conclusión será indispensable al estudiar cocientes. Si más adelante sabemos que

$$g(x) \to M \qquad \text{con} \qquad M \neq 0,$$

Proposición 4.2 nos permitirá asegurar primero que $g(x) \neq 0$ suficientemente cerca de a. Solo después tendrá sentido controlar rigurosamente expresiones como

$$\frac{f(x)}{g(x)}.$$

No estamos usando todavía el álgebra de límites funcionales; estamos preparando una de sus herramientas locales.

La misma idea también conserva el signo. La separación respecto de cero contiene una consecuencia geométrica adicional. Si $L > 0$, podemos tomar nuevamente $\varepsilon = \frac{L}{2}$. Suficientemente cerca de a, $|f(x) - L| < \frac{L}{2}$, de donde $-\frac{L}{2} < f(x) - L < \frac{L}{2}$.

Sumando L obtenemos $\frac{L}{2} < f(x) < \frac{3L}{2}$.

En particular, $f(x) > 0$.

Si $L < 0$, el argumento análogo muestra que $f(x) < 0$ suficientemente cerca de a.

Por tanto,

$$\boxed{L \neq 0 \quad \Longrightarrow \quad f(x) \text{ tiene localmente el mismo signo que } L.}$$

Esta observación no añade una nueva hipótesis ni una nueva teoría: es otra lectura del mismo control ε–δ.

Tres patrones de prueba que conviene conservar. Los resultados de esta sección introducen tres movimientos que reaparecerán una y otra vez.

Primer patrón: separar candidatos distintos. Si dos números L y M supuestamente describen el mismo comportamiento límite, utilizamos su distancia $|L - M|$ para construir una tolerancia incompatible con ambos a la vez.

Segundo patrón: sincronizar condiciones locales. Si una propiedad vale dentro de un radio δ_1 y otra dentro de un radio δ_2, ambas valen simultáneamente dentro de $\delta = \min\{\delta_1, \delta_2\}$.

Cuando aparezcan tres o más condiciones, usaremos del mismo modo $\delta = \min\{\delta_1, \delta_2, \delta_3, \dots\}$.

Tercer patrón: elegir una tolerancia estratégica. La definición permite usar cualquier $\varepsilon > 0$. Para probar una consecuencia estructural, muchas veces conviene fijar una tolerancia

especialmente útil: $\varepsilon = 1$ para obtener acotación local, o $\varepsilon = \frac{|L|}{2}$ para mantenernos lejos de cero cuando $L \neq 0$.

La definición no solo sirve para verificar límites concretos. También funciona como una máquina de producir información local una vez que el límite es conocido.

Qué queda preparado para el criterio secuencial. Hasta ahora todo el capítulo se ha construido directamente con la geometría de la recta y la definición ε–δ. No hemos usado el criterio secuencial ni ninguna teoría de continuidad.

Ya sabemos que:

- el punto de aproximación debe ser de acumulación del dominio;
- el límite funcional, si existe, es único;
- el valor puntual puede cambiar sin alterar el límite;
- dos funciones que coinciden localmente en un entorno perforado tienen el mismo comportamiento límite;
- un límite finito fuerza acotación local;
- un límite no nulo fuerza separación local respecto de cero.

La siguiente pregunta es de otra naturaleza. En `T1-C04` construimos una teoría completa de convergencia de sucesiones. ¿Podemos detectar el comportamiento de una función cerca de a observando **todas las sucesiones de puntos del dominio que se aproximan a** a?

La respuesta será afirmativa, pero exigirá formular con cuidado qué sucesiones están permitidas. En particular, sus términos deberán evitar el propio punto a, exactamente porque el límite funcional vive en un entorno perforado. Ese será el contenido de §4.4: el criterio secuencial del límite funcional.

4.4. El criterio secuencial: dos lenguajes para el mismo acercamiento

Hasta ahora hemos descrito el acercamiento de una variable real mediante entornos perforados. La definición de límite funcional dice que, dada una tolerancia de salida $\varepsilon > 0$, debemos encontrar un radio $\delta > 0$ tal que **todo** punto permitido del dominio suficientemente próximo a a sea enviado cerca de L. Pero `T1-C04` nos proporcionó otro lenguaje para describir aproximación: las sucesiones. Una sucesión $x_0, x_1, x_2, \ldots$ puede recorrer puntos del dominio cada vez más próximos a a. Si una función posee límite L en a, parece razonable esperar que los valores $f(x_0), f(x_1), f(x_2), \ldots$ se aproximen a L. La pregunta importante es si esta observación funciona también en sentido inverso:

¿controlar todas las sucesiones que se acercan a a equivale a controlar todo un entorno perforado?

La respuesta es sí. Esta equivalencia será una de las herramientas más útiles del capítulo porque permite pasar de la geometría local de ε–δ a toda la teoría secuencial ya construida. Hay, sin embargo, una condición que debemos vigilar desde el comienzo. El límite funcional estudia puntos que satisfacen $0 < |x - a| < \delta$. Por tanto, las sucesiones que utilicemos deberán acercarse a a **sin tomar el valor** a. El centro del entorno perforado sigue estando excluido.

El teorema de caracterización secuencial. La equivalencia puede formularse de manera compacta.

Teorema 4.2. **Criterio secuencial del límite funcional.** Sea

$$f\colon A \to \mathbb{R},$$

sea a un punto de acumulación de A y sea $L \in \mathbb{R}$. Entonces son equivalentes:

1. $$\lim_{x \to a} f(x) = L.$$
2. Para toda sucesión (x_n) de puntos de A que satisfaga
$$x_n \neq a \qquad \text{para todo } n,$$
y
$$x_n \to a,$$
se cumple
$$f(x_n) \to L.$$

Demostración. Demostraremos las dos implicaciones por separado.
(1) $\Rightarrow$ (2)**.** Supongamos que

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

Sea (x_n) una sucesión cualquiera de puntos de A tal que

$$x_n \neq a$$

para todo n y

$$x_n \to a.$$

Debemos demostrar que

$$f(x_n) \to L.$$

Sea $\varepsilon > 0$. Como $f(x) \to L$ cuando $x \to a$, la definición ε–δ proporciona un radio $\delta > 0$ tal que, para todo $x \in A$,

$$0 < |x - a| < \delta \quad \Longrightarrow \quad |f(x) - L| < \varepsilon.$$

Ahora usamos la convergencia secuencial

$$x_n \to a.$$

Aplicando su definición con la tolerancia $\delta > 0$, existe un índice N tal que, para todo $n \geq N$,

$$|x_n - a| < \delta.$$

Además, por hipótesis,

$$x_n \neq a,$$

de modo que

$$|x_n - a| > 0.$$

Por tanto, para todo $n \geq N$,

$$0 < |x_n - a| < \delta.$$

Como cada x_n pertenece a A, podemos aplicar el control funcional y obtenemos

$$|f(x_n) - L| < \varepsilon$$

para todo $n \geq N$.

Hemos probado exactamente que

$$f(x_n) \to L.$$

Como la sucesión (x_n) era arbitraria entre todas las sucesiones admisibles, queda demostrada la primera implicación.

(2) $\Rightarrow$ (1)**.** Demostraremos esta dirección por contraposición. Supongamos que **no se cumple**

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

Por la negación de la definición ε–δ establecida en Definición 4.2, existe una tolerancia fija

$$\varepsilon_0 > 0$$

tal que, para **todo** $\delta > 0$, existe un punto $x \in A$ con

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

y

$$|f(x) - L| \geq \varepsilon_0.$$

Aplicaremos esta afirmación con una sucesión de radios que tiende a cero. Para cada $n \in \mathbb{N}$, tomemos

$$\delta_n = \frac{1}{n+1}.$$

La negación del límite garantiza que podemos elegir un punto $x_n \in A$ tal que

$$0 < |x_n - a| < \frac{1}{n+1}$$

y simultáneamente

$$|f(x_n) - L| \geq \varepsilon_0.$$

Así hemos construido una sucesión (x_n) de puntos del dominio con

$$x_n \neq a$$

para todo n.

Además,

$$0 \leq |x_n - a| < \frac{1}{n+1}.$$

En `T1-C04` demostramos que

$$\frac{1}{n+1} \to 0.$$

Por el teorema del sándwich aplicado a la sucesión no negativa $|x_n - a|$,

$$|x_n - a| \to 0.$$

Esto equivale a

$$x_n \to a.$$

Por tanto, (x_n) es una sucesión admisible para la condición 2.

Pero sus imágenes satisfacen

$$|f(x_n) - L| \geq \varepsilon_0$$

para **todo** n. En particular, la sucesión $(f(x_n))$ nunca entra en la tolerancia ε_0 alrededor de L, ni siquiera eventualmente. Luego

$$f(x_n) \not\to L.$$

Hemos demostrado que, si falla el límite funcional, existe una sucesión admisible $x_n \to a$ para la cual falla la convergencia de $f(x_n)$ a L. Esta es la contraposición de (2) $\Rightarrow$ (1).

Concluimos que 1 y 2 son equivalentes.

La demostración merece ser leída como una traducción exacta entre dos sistemas de control. En la primera dirección, la definición funcional produce un radio: $\varepsilon \longrightarrow \delta$. Después, la convergencia de x_n a a convierte ese radio en un umbral secuencial: $\delta \longrightarrow N$. A partir de ese índice, todos los términos de la sucesión están dentro del entorno perforado apropiado, y por ello sus imágenes quedan dentro de la tolerancia de salida.

El flujo lógico es

$$\boxed{\varepsilon \longrightarrow \delta \longrightarrow N \longrightarrow |f(x_n) - L| < \varepsilon.}$$

La segunda dirección invierte la perspectiva. Si no existe un radio que funcione, la negación de la definición proporciona, para cada escala δ, un **testigo malo**. Elegimos escalas cada vez menores, $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$, y tomamos un testigo malo en cada una. Los testigos forman entonces una sucesión que se aproxima a a, pero cuyas imágenes siguen violando la misma tolerancia fija ε_0.

Esquemáticamente,

$$\boxed{\text{fallo } \varepsilon\text{–}\delta \longrightarrow \text{un testigo malo en cada escala} \longrightarrow \text{una sucesión testigo del fallo.}}$$

Esta construcción explica por qué el criterio secuencial no es una analogía informal: es **equivalente** a la definición de límite funcional en la recta real.

Por qué debemos excluir el propio punto. En la primera mitad de la demostración utilizamos de manera explícita $x_n \neq a$. Sin esa condición, de $|x_n - a| < \delta$ no podríamos deducir $0 < |x_n - a| < \delta$, porque algunos términos podrían ser exactamente iguales a a. Y la definición funcional no controla el valor $f(a)$. El siguiente ejemplo muestra que esta precaución no es técnica ni opcional.

Ejemplo 4.4. **Por qué las sucesiones del criterio deben evitar a.** Definamos $f\colon \mathbb{R} \to \mathbb{R}$ por

$$f(x) = \begin{cases} 1, & x = 0, \\ 0, & x \neq 0. \end{cases}$$

Afirmamos que

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

En efecto, sea $\varepsilon > 0$. Podemos tomar, por ejemplo, $\delta = 1$.

Si $0 < |x| < 1$, entonces $x \neq 0$, y por definición $f(x) = 0$.

Por tanto, $|f(x) - 0| = 0 < \varepsilon$.

Así, el límite funcional es efectivamente 0.

Consideremos ahora la sucesión constante $x_n = 0$ para todo n. Claramente, $x_n \to 0$. Pero $f(x_n) = f(0) = 1$ para todo n, de modo que $f(x_n) \to 1$, y en particular $f(x_n) \not\to 0$. No hay contradicción con Teorema 4.2, porque esta sucesión **no es admisible**: sus términos no satisfacen $x_n \neq 0$.

El ejemplo recupera una idea de §4.2: el límite funcional no ve el valor situado exactamente en el centro. Una sucesión que permanezca en ese centro estaría introduciendo información que la definición ε–δ excluye deliberadamente. Podríamos formular una variante equivalente del criterio permitiendo que $x_n = a$ ocurra una cantidad finita de veces, siempre que finalmente $x_n \neq a$. En `T1-C04` vimos que modificar o eliminar un prefijo finito no altera el límite de una sucesión. Por eso exigir $x_n \neq a$ **para todo** n es principalmente la forma más limpia de expresar el criterio; una exclusión eventual produciría una versión equivalente.

Lo que sí debemos impedir es que la sucesión toque a infinitamente muchas veces sin quedar finalmente fuera del centro, porque entonces los valores $f(a)$ podrían intervenir una y otra vez.

Una sola sucesión mala basta para refutar un candidato. El cuantificador central de Teorema 4.2 es universal:

$$\boxed{\text{para toda sucesión admisible } (x_n), \qquad f(x_n) \to L.}$$

Esto tiene una consecuencia metodológica inmediata. Para **demostrar** un límite mediante el criterio secuencial debemos controlar una sucesión arbitraria. Probar que funciona para una sucesión particular no basta.

En cambio, para **refutar** que L sea el límite basta encontrar una sola sucesión admisible (x_n) tal que $x_n \to a$ pero $f(x_n) \not\to L$. Esta asimetría es la misma que aparece siempre al negar una afirmación universal: un único contraejemplo destruye la universalidad. Consideremos, por ejemplo, la función

$$g(x) = \begin{cases} 1, & x > 0, \\ -1, & x < 0, \end{cases}$$

con dominio $A = \mathbb{R} \setminus \{0\}$.

Tomemos las sucesiones $x_n = \frac{1}{n+1}$ y $y_n = -\frac{1}{n+1}$.

Ambas pertenecen a A, evitan 0 y satisfacen

$$x_n \to 0, \qquad y_n \to 0.$$

Sin embargo, $g(x_n) = 1$ para todo n, mientras $g(y_n) = -1$ para todo n. Por tanto,

$$g(x_n) \to 1, \qquad g(y_n) \to -1.$$

Si existiera un límite finito

$$\lim_{x\to 0} g(x) = L,$$

Teorema 4.2 obligaría a ambas sucesiones de imágenes a converger al mismo número L. Pero sus límites secuenciales son distintos. Luego el límite funcional en 0 no existe.

Este argumento anticipa, sin definir todavía límites laterales, una estrategia que reaparecerá en §4.7: aproximaciones distintas al mismo punto pueden revelar comportamientos incompatibles.

Una sucesión buena no demuestra nada por sí sola. El ejemplo anterior también permite evitar un error frecuente. Si tomáramos solamente $x_n = \frac{1}{n+1}$, observaríamos que $g(x_n) = 1 \to 1$. Eso **no** demuestra que

$$\lim_{x\to 0} g(x) = 1.$$

El criterio no dice

> existe una sucesión que se aproxima a a y cuyas imágenes convergen a L.

Dice

> para **toda** sucesión del dominio que se aproxima a a evitando a, las imágenes convergen a L.

La diferencia lógica es decisiva:

$$\exists(x_n) \qquad \text{frente a} \qquad \forall(x_n).$$

Una sucesión bien elegida puede sugerir un candidato. Solo el control universal de todas las sucesiones admisibles caracteriza el límite.

La condición de acumulación vuelve a aparecer. El teorema supone, igual que la definición funcional, que a es punto de acumulación de A. Esta hipótesis desempeña ahora dos papeles relacionados. Primero, garantiza que la pregunta sobre aproximación no sea vacua: existen puntos del dominio distintos de a arbitrariamente cerca de a. Segundo, permite construir sucesiones admisibles. Para cada n podemos buscar un punto $x_n \in A$ tal que $0 < |x_n - a| < \frac{1}{n+1}$. La propia definición de punto de acumulación asegura que cada una de esas búsquedas tiene al menos una solución. Cualquier elección de un punto en cada escala produce una sucesión con $x_n \to a$. Así, la noción introducida en §4.1 enlaza directamente con el lenguaje secuencial de `T1-C04`.

Qué hemos ganado con la equivalencia. Disponemos ahora de dos puertas de entrada al mismo concepto. La definición ε–δ observa **simultáneamente todos los puntos** de un entorno perforado:

$$0 < |x - a| < \delta \quad \Longrightarrow \quad |f(x) - L| < \varepsilon.$$

El criterio secuencial observa **todas las maneras secuenciales** de aproximarse al punto:

$$x_n \to a, \quad x_n \neq a \quad \Longrightarrow \quad f(x_n) \to L.$$

Ninguno de los lenguajes es «más verdadero» que el otro. En la recta real describen exactamente el mismo fenómeno, pero pueden ser convenientes para tareas distintas. La vía ε–δ será especialmente útil cuando podamos estimar directamente $|f(x) - L|$ en función de $|x - a|$. La vía secuencial será especialmente útil cuando ya dispongamos de resultados potentes sobre sucesiones. A partir de ahora podremos tomar una sucesión arbitraria $x_n \to a$, aplicar a $(f(x_n))$ el álgebra, el orden o el teorema del sándwich construidos en `T1-C04`, y después regresar al lenguaje funcional mediante Teorema 4.2. Ese mecanismo será importante inmediatamente. En §4.5 compararemos dos estilos de demostración para el álgebra de límites funcionales: una prueba directa desde ε–δ y otra que transporta el problema al mundo secuencial.

El mapa conceptual del capítulo se ha ampliado así:

entornos perforados $\longleftrightarrow$ sucesiones que evitan el centro $\longleftrightarrow$ convergencia de las imágenes.

La siguiente sección utilizará esta equivalencia para construir el álgebra de límites funcionales y para justificar con precisión cuándo productos y cocientes pueden controlarse localmente.

4.5. Álgebra de límites funcionales y control local

Hasta ahora la definición ε–δ nos ha permitido estudiar límites uno por uno. Para cada función fijábamos una tolerancia, buscábamos un radio y verificábamos que todos los puntos del dominio suficientemente próximos al centro produjeran valores adecuados. Ese procedimiento es indispensable para construir la teoría, pero no sería razonable repetirlo desde cero cada vez que aparezca una suma, un producto o un cociente. Queremos saber si, una vez conocidos

$$\lim_{x \to a} f(x) = L \qquad \text{y} \qquad \lim_{x \to a} g(x) = M,$$

podemos calcular límites de expresiones formadas algebraicamente a partir de f y g.

La expectativa es familiar:

$$f + g \longrightarrow L + M, \qquad fg \longrightarrow LM, \qquad \frac{f}{g} \longrightarrow \frac{L}{M}.$$

Pero cada una de estas fórmulas es todavía una afirmación que debe demostrarse. En particular, el producto contiene factores variables que deben ser controlados y el cociente plantea una cuestión previa de dominio: antes de dividir por $g(x)$ necesitamos saber que $g(x)$ no se anula en los puntos suficientemente próximos a a. La sección anterior ya preparó exactamente esas dos herramientas:

- un límite finito produce **acotación local**;
- un límite no nulo produce **separación local respecto de cero**.

Veremos ahora que esas consecuencias no eran accesorios de la definición. Son las piezas que permiten construir el cálculo algebraico de límites funcionales.

El álgebra que queremos justificar. Trabajaremos primero con dos funciones definidas sobre un mismo dominio $f, g\colon A \to \mathbb{R}$, y con un punto a de acumulación de A. Si dos funciones tienen dominios distintos, las mismas leyes podrán aplicarse después de restringirlas a un dominio común en el que a siga siendo punto de acumulación.

PROPOSICIÓN 4.3. **Álgebra de límites funcionales.** Sean

$$f, g\colon A \to \mathbb{R},$$

sea a un punto de acumulación de A y supongamos que

$$\lim_{x\to a} f(x) = L, \qquad \lim_{x\to a} g(x) = M,$$

con $L, M \in \mathbb{R}$. Entonces:

1. para todo $c \in \mathbb{R}$,
$$\lim_{x\to a} cf(x) = cL;$$
2. $$\lim_{x\to a}(f(x) + g(x)) = L + M;$$
3. $$\lim_{x\to a} f(x)g(x) = LM;$$
4. si $M \neq 0$ y
$$C := \{x \in A : g(x) \neq 0\},$$
entonces a es punto de acumulación de C y, para la función cociente definida en C,
$$\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{L}{M};$$
5. $$\lim_{x\to a} |f(x)| = |L|.$$

En particular,

$$\lim_{x\to a}(f(x) - g(x)) = L - M.$$

Demostración. Demostraremos primero las leyes directamente desde la definición ε–δ y desde las consecuencias locales de §4.3.

1. Multiplicación por una constante. Si $c = 0$, entonces $cf(x) = 0$ para todo $x \in A$, de modo que el límite es $0 = cL$.

Supongamos $c \neq 0$ y sea $\varepsilon > 0$. Como $f(x) \to L$, existe $\delta > 0$ tal que

$$0 < |x - a| < \delta \quad \Longrightarrow \quad |f(x) - L| < \frac{\varepsilon}{|c|}.$$

Por tanto,

$$|cf(x) - cL| = |c|\,|f(x) - L| < \varepsilon.$$

Así,

$$\lim_{x\to a} cf(x) = cL.$$

2. Suma. Sea $\varepsilon > 0$. Como $f(x) \to L$, existe $\delta_1 > 0$ tal que

$$0 < |x-a| < \delta_1 \quad \Longrightarrow \quad |f(x) - L| < \frac{\varepsilon}{2}.$$

Como $g(x) \to M$, existe $\delta_2 > 0$ tal que

$$0 < |x-a| < \delta_2 \quad \Longrightarrow \quad |g(x) - M| < \frac{\varepsilon}{2}.$$

Tomemos

$$\delta = \text{mín}\{\delta_1, \delta_2\}.$$

Si $x \in A$ y $0 < |x-a| < \delta$, ambas estimaciones valen simultáneamente y

$$\begin{aligned} |(f(x) + g(x)) - (L+M)| &\leq |f(x) - L| + |g(x) - M| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon. \end{aligned}$$

Luego

$$\lim_{x \to a} (f(x) + g(x)) = L + M.$$

3. Producto. Esta es la primera parte en la que necesitamos algo más que repartir ε.
Por Proposición 4.2, como $g(x) \to M$ existe un radio $r > 0$ y una constante $B > 0$ tales que

$$0 < |x-a| < r \quad \Longrightarrow \quad |g(x)| < B.$$

Sea $\varepsilon > 0$. Como $f(x) \to L$, existe $\delta_1 > 0$ tal que

$$0 < |x-a| < \delta_1 \quad \Longrightarrow \quad |f(x) - L| < \frac{\varepsilon}{2B}.$$

Como $g(x) \to M$, existe $\delta_2 > 0$ tal que

$$0 < |x-a| < \delta_2 \quad \Longrightarrow \quad |g(x) - M| < \frac{\varepsilon}{2(|L|+1)}.$$

Definamos

$$\delta = \text{mín}\{r, \delta_1, \delta_2\}.$$

Para $x \in A$ con $0 < |x-a| < \delta$ podemos escribir

$$f(x)g(x) - LM = g(x)(f(x) - L) + L(g(x) - M).$$

Por la desigualdad triangular,

$$\begin{aligned} |f(x)g(x) - LM| &\leq |g(x)|\,|f(x) - L| + |L|\,|g(x) - M| \\ &< B\frac{\varepsilon}{2B} + |L|\frac{\varepsilon}{2(|L|+1)} \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon. \end{aligned}$$

Por tanto,

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

4. Cociente. Supongamos ahora que

$$M \neq 0.$$

Antes de calcular el límite debemos justificar que el cociente está definido cerca de a. Por la segunda parte de Proposición 4.2, existe $r_0 > 0$ tal que, para todo $x \in A$,

$$0 < |x-a| < r_0 \quad \Longrightarrow \quad |g(x)| > \frac{|M|}{2} > 0.$$

En particular,

$$g(x) \neq 0$$

en ese entorno perforado. Por consiguiente, si

$$C = \{x \in A : g(x) \neq 0\},$$

entonces C y A tienen exactamente los mismos puntos dentro de $V^*_{r_0}(a)$. Como a es punto de acumulación de A, también lo es de C: dado $\rho > 0$, basta buscar un punto de A en el entorno perforado de radio

$$\text{mín}\{\rho, r_0\};$$

ese punto pertenece automáticamente a C.

Consideremos ahora la función

$$h\colon C \to \mathbb{R}, \qquad h(x) = \frac{1}{g(x)}.$$

Demostraremos que

$$\lim_{x\to a} h(x) = \frac{1}{M}.$$

Sea $\varepsilon > 0$. Como $g(x) \to M$, existe $\delta_1 > 0$ tal que

$$0 < |x-a| < \delta_1 \quad \Longrightarrow \quad |g(x) - M| < \frac{\varepsilon |M|^2}{2}.$$

Tomemos

$$\delta = \text{mín}\{r_0, \delta_1\}.$$

Si $x \in C$ y $0 < |x-a| < \delta$, entonces

$$|g(x)| > \frac{|M|}{2},$$

y por tanto

$$\begin{aligned}\left|\frac{1}{g(x)}-\frac{1}{M}\right| &= \frac{|M-g(x)|}{|M|\,|g(x)|}\\ &< \frac{|M-g(x)|}{|M|(|M|/2)}\\ &= \frac{2|g(x)-M|}{|M|^2}\\ &< \varepsilon.\end{aligned}$$

Luego

$$\frac{1}{g(x)} \longrightarrow \frac{1}{M} \qquad (x \to a,\ x \in C).$$

La restricción $f|_C$ conserva el límite L, porque todo punto de C es también un punto de A y el control ε–δ de f sigue siendo válido al restringir el dominio. Aplicando ahora la ley del producto, ya demostrada, a $f|_C$ y h, obtenemos

$$\frac{f(x)}{g(x)} = f(x)h(x) \longrightarrow L\frac{1}{M} = \frac{L}{M}.$$

5. Valor absoluto. Para todo $x \in A$, la desigualdad triangular inversa da

$$\big||f(x)| - |L|\big| \le |f(x) - L|.$$

Sea $\varepsilon > 0$. Como $f(x) \to L$, existe $\delta > 0$ tal que

$$0 < |x-a| < \delta \quad \Longrightarrow \quad |f(x) - L| < \varepsilon.$$

Entonces, para esos mismos puntos,

$$\big||f(x)| - |L|\big| < \varepsilon.$$

Por tanto,

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

Finalmente, tomando $c = -1$ en la primera ley y combinándola con la suma,

$$f - g = f + (-1)g$$

tiene límite

$$L - M.$$

Esto completa la demostración.

Cómo leer la prueba del producto. La dificultad del producto no consiste en una manipulación algebraica complicada. Consiste en que el error $|f(x)g(x) - LM|$ no aparece inicialmente escrito solo en términos de los errores que sabemos hacer pequeños:

$$|f(x) - L| \qquad \text{y} \qquad |g(x) - M|.$$

La identidad

$$f(x)g(x) - LM = g(x)(f(x) - L) + L(g(x) - M)$$

crea precisamente esos dos errores. Pero aparece un factor variable $|g(x)|$. La acotación local de Proposición 4.2 permite sustituirlo, suficientemente cerca de a, por una constante fija B. La arquitectura de la prueba es

$$\boxed{\begin{array}{c}\text{crear los errores conocidos}\\ \to \text{ acotar el factor variable}\\ \to \text{ repartir } \varepsilon\\ \to \text{ sincronizar radios mediante } \min.\end{array}}$$

Esta es la versión funcional del patrón que apareció en el producto de límites secuenciales. La diferencia externa es que allí sincronizábamos colas mediante un máximo de índices; aquí sincronizamos regiones locales mediante un mínimo de radios.

Por qué el cociente necesita una auditoría de dominio. La hipótesis $M \neq 0$ no dice directamente que $g(x) \neq 0$ para cada punto del dominio. Podría haber ceros lejos de a, e incluso podría ocurrir $g(a) = 0$ si el punto a pertenece al dominio. Ninguna de esas posibilidades afecta por sí sola al límite. Lo que necesitamos es una afirmación local:

$$\boxed{g(x) \to M \neq 0 \quad \Longrightarrow \quad g(x) \neq 0 \text{ suficientemente cerca de } a, \text{ con } x \neq a.}$$

Eso es exactamente lo que proporciona la separación respecto de cero de Proposición 4.2. Por ello la ley del cociente no debe leerse como una sustitución formal

$$\frac{f(x)}{g(x)} \rightsquigarrow \frac{L}{M}.$$

Antes hay que verificar que el límite del denominador es no nulo. Esa verificación garantiza simultáneamente dos cosas:

1. el cociente está definido en algún entorno perforado del punto;
2. los recíprocos $1/g(x)$ quedan controlados porque $|g(x)|$ posee una cota inferior positiva.

La segunda propiedad es la que hace posible la estimación

$$\left|\frac{1}{g(x)} - \frac{1}{M}\right| = \frac{|g(x) - M|}{|g(x)|\,|M|}.$$

Sin una cota inferior para $|g(x)|$, el denominador de esta expresión podría volverse arbitrariamente pequeño.

Una segunda demostración: transportar el problema a sucesiones. La prueba anterior es importante porque muestra **por qué** las leyes algebraicas funcionan. Pero §4.4 nos proporciona ahora una ruta mucho más corta. Tomemos, por ejemplo, una sucesión arbitraria (x_n) de puntos de A tal que

$$x_n \neq a \qquad \text{y} \qquad x_n \to a.$$

Por Teorema 4.2,

$$f(x_n) \to L, \qquad g(x_n) \to M.$$

Pero estas son ahora dos sucesiones reales ordinarias. Podemos aplicar directamente el álgebra secuencial de `T1-C04`: $cf(x_n) \to cL$,

$$f(x_n) + g(x_n) \to L + M,$$

$$f(x_n)g(x_n) \to LM.$$

Además, por Corolario 3.1, $|f(x_n)| \to |L|$. Como esto ocurre para **toda** sucesión admisible (x_n), el criterio secuencial devuelve inmediatamente los correspondientes límites funcionales. Para el cociente procedemos sobre su dominio natural $C = \{x \in A : g(x) \neq 0\}$. Ya sabemos por la separación local respecto de cero que a es punto de acumulación de C. Si (x_n) es una sucesión arbitraria de puntos de C con $x_n \neq a$ y $x_n \to a$, entonces

$$f(x_n) \to L, \qquad g(x_n) \to M \neq 0.$$

La ley secuencial del cociente de Proposición 3.3 da

$$\frac{f(x_n)}{g(x_n)} \to \frac{L}{M}.$$

Aplicando nuevamente Teorema 4.2,

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \frac{L}{M}.$$

Esta segunda demostración es mucho más breve porque todo el trabajo técnico —acotación de factores, separación de denominadores, distribución de tolerancias— ya quedó encapsulado en los teoremas sobre sucesiones. Las dos rutas no compiten. Cumplen funciones distintas:

ε–δ : expone el mecanismo local de la prueba,
criterio secuencial : reutiliza una teoría ya construida y comprime el argumento.

Esta comparación muestra una ventaja real de haber demostrado Teorema 4.2: los resultados de `T1-C04` pueden transportarse al mundo funcional cuando sus hipótesis están correctamente traducidas.

Del álgebra general a polinomios y funciones racionales. Las leyes anteriores permiten construir límites de expresiones más complejas sin volver a abrir la definición para cada término. Primero necesitamos una pieza elemental. Para la función identidad $i(x) = x$, tenemos

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

En efecto, dada $\varepsilon > 0$, basta tomar $\delta = \varepsilon$, porque

$$0 < |x - a| < \delta \quad \Longrightarrow \quad |x - a| < \varepsilon.$$

A partir de este límite y de Proposición 4.3 podemos elevar a potencias, multiplicar por constantes y sumar términos.

Corolario 4.1. **Límites de polinomios y funciones racionales.** Sea

$$p(x) = c_0 + c_1 x + \cdots + c_m x^m$$

un polinomio real. Entonces, para todo $a \in \mathbb{R}$,

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

Si además

$$q(x) = d_0 + d_1 x + \cdots + d_n x^n$$

es otro polinomio y

$$q(a) \neq 0,$$

entonces, para la función racional

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

definida en su dominio natural

$$D = \{x \in \mathbb{R} : q(x) \neq 0\},$$

se cumple

$$\lim_{x \to a} r(x) = \frac{p(a)}{q(a)}.$$

Demostración. Ya hemos probado que

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

Aplicando repetidamente la ley del producto obtenemos, para cada entero $k \geq 1$,

$$\lim_{x \to a} x^k = a^k.$$

Las funciones constantes satisfacen

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

y por multiplicación escalar,

$$\lim_{x\to a} c_k x^k = c_k a^k.$$

La aplicación repetida de la ley de la suma produce

$$\lim_{x\to a} p(x) = c_0 + c_1 a + \cdots + c_m a^m = p(a).$$

Del mismo modo,

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

Si $q(a) \neq 0$, la ley del cociente de Proposición 4.3 se aplica en el dominio natural D y da

$$\lim_{x\to a} \frac{p(x)}{q(x)} = \frac{p(a)}{q(a)}.$$

Es importante describir correctamente qué acabamos de usar. No hemos invocado todavía un teorema de continuidad de los polinomios ni de las funciones racionales. Esa teoría aparecerá en `T1-C06`. Aquí el razonamiento es anterior y más elemental:

$$\boxed{\lim_{x\to a} x = a \quad + \quad \text{álgebra de límites} \quad \Longrightarrow \quad \lim_{x\to a} p(x) = p(a).}$$

Más adelante la continuidad permitirá encapsular este resultado en un lenguaje nuevo. Por ahora, su fundamento es exclusivamente el álgebra de límites que acabamos de demostrar.

Un cálculo racional con auditoría del denominador.

Ejemplo 4.5. **Calcular no es solo sustituir: hay que auditar el denominador.** Consideremos $R(x) = \frac{2x^2+3x-1}{x^2-5x+5}$ sobre su dominio natural y estudiemos

$$\lim_{x\to 1} R(x).$$

Definamos

$$p(x) = 2x^2 + 3x - 1, \qquad q(x) = x^2 - 5x + 5.$$

Por Corolario 4.1,

$$\lim_{x\to 1} p(x) = p(1) = 2 + 3 - 1 = 4$$

y

$$\lim_{x\to 1} q(x) = q(1) = 1 - 5 + 5 = 1.$$

El dato decisivo para poder usar la ley del cociente es $q(1) = 1 \neq 0$. Por Proposición 4.2, como $q(x) \to 1$, existe un radio $r > 0$ tal que

$$0 < |x - 1| < r \quad \Longrightarrow \quad |q(x)| > \frac{1}{2}.$$

En particular, $q(x) \neq 0$ para todos los puntos suficientemente próximos a 1 distintos de 1. Esto confirma que los posibles ceros que el denominador pueda tener en otras regiones no interfieren con el límite que estamos estudiando. Ahora sí aplicamos la ley del cociente:

$$\lim_{x\to 1} R(x) = \frac{\lim_{x\to 1} p(x)}{\lim_{x\to 1} q(x)} = \frac{4}{1} = 4.$$

Por tanto,

$$\boxed{\lim_{x\to 1} \frac{2x^2+3x-1}{x^2-5x+5} = 4.}$$

El orden del razonamiento importa. No escribimos primero «sustituimos $x = 1$ en la fracción». Hicimos tres verificaciones separadas:

1. calculamos el límite del numerador;
2. calculamos el límite del denominador;
3. comprobamos que el límite del denominador es no nulo.

Solo entonces la ley del cociente quedó autorizada.

Qué ocurre si el límite del denominador es cero. Si $g(x) \to 0$,
Proposición 4.3 **no** permite concluir nada automáticamente sobre

$$\frac{f(x)}{g(x)}.$$

No existe una regla $\frac{L}{0}$ dentro de los números reales, y distintas situaciones con denominador tendiendo a cero pueden producir comportamientos completamente diferentes. Por ejemplo, cuando $x \to 0$, $\frac{x}{x} = 1$ para $x \neq 0$, mientras que $\frac{1}{x}$ no posee un límite real finito en 0. En ambos casos el denominador tiende a 0, pero eso no determina por sí solo el comportamiento del cociente. No desarrollaremos todavía la teoría de límites infinitos funcionales; está reservada para §4.8. El punto actual es más básico:

la ley algebraica del cociente exige un límite denominador distinto de cero.

Cuando esa hipótesis falla, debemos analizar el problema por otros métodos.

Qué podemos escribir desde ahora. Después de Proposición 4.3 y Corolario 4.1, expresiones como

$$\lim_{x\to a}(f(x)+g(x)) = L + M,$$

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

o

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{L}{M}$$

pueden utilizarse como herramientas, siempre que sus hipótesis hayan sido verificadas. La definición ε–δ no ha desaparecido. Ha quedado **encapsulada** dentro de resultados generales cuya demostración ya conocemos. La sección deja además tres hábitos de auditoría:

para productos: buscar una cota local para el factor variable,
para cocientes: separar localmente el denominador de cero,
para varios controles: tomar el mínimo de los radios disponibles.

Con el álgebra disponible, el siguiente paso ya no será combinar valores mediante igualdades, sino transportar **desigualdades** al límite. En §4.6 estudiaremos preservación del orden y el teorema del sándwich funcional, que permitirán controlar funciones cuyo límite es difícil de calcular directamente pero que pueden quedar atrapadas entre funciones más sencillas.

4.6. Orden, sándwich y oscilación local

En §4.5 aprendimos a transportar al límite operaciones algebraicas. Si conocemos los límites de dos funciones, podemos sumar, multiplicar o, bajo la hipótesis adecuada, dividir. Pero una gran cantidad de problemas de análisis no se resuelven mediante una identidad exacta. A veces la función cuyo límite buscamos es difícil de describir, mientras que sí podemos encerrarla entre funciones más sencillas. Otras veces sabemos que una función permanece localmente por debajo de otra y queremos averiguar qué queda de esa relación al pasar al límite. La pregunta de esta sección es, por tanto,

¿qué información de orden válida cerca de a sigue siendo visible en el límite?

La expresión **cerca de** a es decisiva. Como el límite funcional solo observa entornos perforados, no necesitaremos exigir desigualdades globales. Bastará que sean verdaderas en algún entorno perforado del punto de aproximación.

Una desigualdad local no puede invertirse en el límite. Sean $f, g\colon A \to \mathbb{R}$, donde a es punto de acumulación de A. Supongamos que

$$\lim_{x\to a} f(x) = L, \qquad \lim_{x\to a} g(x) = M.$$

Imaginemos además que existe un radio $r > 0$ tal que $f(x) \le g(x)$ para todo $x \in A$ con $0 < |x - a| < r$. La desigualdad puede fallar completamente fuera de ese entorno. Eso no importa: para estudiar el límite solo interesa lo que sucede a escalas suficientemente pequeñas alrededor de a.

¿Podría ocurrir, sin embargo, que $L > M$?

La teoría secuencial ya construida permite responder de manera especialmente limpia.

Proposición 4.4. **Preservación local del orden bajo límites funcionales.** Sean

$$f, g\colon A \to \mathbb{R},$$

y sea a un punto de acumulación de A. Supongamos que

$$\lim_{x\to a} f(x) = L, \qquad \lim_{x\to a} g(x) = M.$$

Si existe $r > 0$ tal que

$$f(x) \le g(x)$$

para todo $x \in A$ que satisfaga

$$0 < |x - a| < r,$$

entonces

$$L \leq M.$$

Demostración. Como a es punto de acumulación de A, en §4.4 vimos que existen sucesiones de puntos de A que se aproximan a a evitando el centro. Tomemos una de ellas, (x_n), tal que

$$x_n \neq a \qquad \text{para todo } n,$$

y

$$x_n \to a.$$

Como $x_n \to a$, aplicado el criterio de convergencia secuencial a la tolerancia $r > 0$, existe un índice N_r tal que

$$n \geq N_r \quad \Longrightarrow \quad |x_n - a| < r.$$

Además, $x_n \neq a$, de modo que para $n \geq N_r$ tenemos

$$0 < |x_n - a| < r.$$

La hipótesis local de orden da entonces

$$f(x_n) \leq g(x_n) \qquad \text{para todo } n \geq N_r.$$

Por Teorema 4.2,

$$f(x_n) \to L$$

y

$$g(x_n) \to M.$$

Ahora estamos ante dos sucesiones reales con una desigualdad eventual. La preservación del orden para sucesiones, demostrada en Proposición 3.4, implica

$$L \leq M.$$

Como la conclusión ya no depende de la sucesión elegida, queda demostrado el resultado.

Esta prueba hace visible una traducción estructural muy precisa:

desigualdad en un entorno perforado
↓
desigualdad eventual sobre toda sucesión admisible
↓
orden de los límites secuenciales
↓
orden de los límites funcionales.

No hemos necesitado reconstruir desde cero la prueba de orden. El criterio secuencial permite transportar exactamente el resultado de `T1-C04` al contexto funcional.

La hipótesis es local, no global. Conviene aislar una consecuencia conceptual de Proposición 4.4. Para concluir $L \leq M$, no hace falta que $f(x) \leq g(x)$ para todo $x \in A$. Basta encontrar **un solo radio** $r > 0$ tal que la desigualdad se cumpla en $V_r^*(a) \cap A$. Por ejemplo, dos funciones podrían satisfacer $f(x) > g(x)$ para todos los puntos muy alejados de a y, sin embargo, cumplir $f(x) \leq g(x)$ cada vez que $0 < |x - a| < 10^{-3}$. Para el límite en a, la segunda información es la relevante. Este es el análogo funcional de una idea ya conocida para sucesiones: allí bastaba una desigualdad **eventual**; aquí basta una desigualdad **local perforada**.

cola suficientemente tardía $\quad\longleftrightarrow\quad$ entorno perforado suficientemente pequeño.

El orden estricto puede colapsar. Hay que leer Proposición 4.4 con la misma precaución que su versión secuencial. De $f(x) < g(x)$ para todos los puntos suficientemente próximos a a, no se sigue necesariamente $L < M$.

Un ejemplo elemental es

$$f(x) = 0, \qquad g(x) = (x - a)^2.$$

Para todo $x \neq a$, $0 < (x - a)^2$, de modo que $f(x) < g(x)$ en cada entorno perforado de a. Sin embargo, por Corolario 4.1,

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

y

$$\lim_{x \to a} (x - a)^2 = 0.$$

Por tanto, $L = M = 0$.

La desigualdad era estricta punto a punto, pero la separación entre ambas funciones, $(x - a)^2$, se hacía arbitrariamente pequeña al aproximarnos a a.

Así,

$f(x) < g(x)$ localmente $\quad\not\Longrightarrow\quad$ $L < M$.

Lo que sí basta para conservar una desigualdad estricta es una **separación positiva uniforme**. Si existe algún número $\eta > 0$ y algún radio $r > 0$ tales que $f(x) + \eta \leq g(x)$ siempre que $0 < |x - a| < r$, entonces Proposición 4.4 aplicado a $f + \eta$ y g produce $L + \eta \leq M$.

Como $\eta > 0$, $L < M$.

La diferencia entre las dos situaciones es cuantitativa: una desigualdad estricta puede estrecharse hasta desaparecer; una brecha fija η no.

Cuando una función queda atrapada entre dos aproximaciones. La preservación del orden parte de la existencia de los dos límites y compara sus valores. El problema siguiente es distinto: queremos usar desigualdades para **demostrar que un límite existe**. Supongamos que tres funciones $f, g, h \colon A \to \mathbb{R}$ satisfacen, suficientemente cerca de a, $f(x) \leq g(x) \leq h(x)$, y que las funciones exteriores se aproximan al mismo número:

$$f(x) \to L, \qquad h(x) \to L.$$

Intuitivamente, cuando x entra en un entorno suficientemente pequeño de a, los valores de $f(x)$ y $h(x)$ quedan dentro de cualquier banda vertical alrededor de L. Si $g(x)$ debe permanecer entre ellos, tampoco puede escapar de esa banda. Esta idea produce el teorema del sándwich funcional.

TEOREMA 4.3. **Teorema del sándwich funcional.** Sean

$$f, g, h \colon A \to \mathbb{R},$$

y sea a un punto de acumulación de A. Supongamos que existe $r > 0$ tal que

$$f(x) \leq g(x) \leq h(x)$$

para todo $x \in A$ con

$$0 < |x - a| < r.$$

Si

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

y

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

entonces

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

Demostración. Sea $\varepsilon > 0$.
Como

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

existe $\delta_f > 0$ tal que

$$0 < |x - a| < \delta_f, \qquad x \in A,$$

implica

$$|f(x) - L| < \varepsilon.$$

Equivalentemente,

$$L - \varepsilon < f(x) < L + \varepsilon.$$

Del mismo modo, como

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

existe $\delta_h > 0$ tal que

$$0 < |x - a| < \delta_h, \qquad x \in A,$$

implica

$$|h(x) - L| < \varepsilon,$$

y por tanto

$$L - \varepsilon < h(x) < L + \varepsilon.$$

Debemos hacer compatibles tres controles locales:

1. el radio r donde sabemos que $f \le g \le h$;
2. el radio δ_f donde f está a menos de ε de L;
3. el radio δ_h donde h está a menos de ε de L.

Tomamos

$$\delta = \min\{r, \delta_f, \delta_h\}.$$

Sea ahora $x \in A$ tal que

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

Entonces se cumplen simultáneamente

$$f(x) \le g(x) \le h(x),$$

$$L - \varepsilon < f(x),$$

y

$$h(x) < L + \varepsilon.$$

Encadenando las desigualdades obtenemos

$$L - \varepsilon < f(x) \le g(x) \le h(x) < L + \varepsilon.$$

En particular,

$$L - \varepsilon < g(x) < L + \varepsilon,$$

lo cual equivale a

$$|g(x) - L| < \varepsilon.$$

Como $\varepsilon > 0$ era arbitrario, concluimos que

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

Aquí hemos elegido deliberadamente una prueba directa ε–δ. El resultado también podría transportarse desde el teorema del sándwich secuencial mediante Teorema 4.2, exactamente como hicimos con Proposición 4.4. Pero la prueba directa revela con especial claridad la geometría local:

las dos paredes exteriores entran en la banda $(L-\varepsilon, L+\varepsilon)$, y la función intermedia queda obligada a entrar con ellas.

No podemos presuponer el límite de la función intermedia. Hay una forma tentadora, pero circular, de intentar justificar el teorema. Podríamos escribir informalmente $f(x) \le g(x) \le h(x)$ y después decir

$$L \le \lim_{x\to a} g(x) \le L.$$

El problema es que esa línea ya contiene la expresión

$$\lim_{x\to a} g(x),$$

como si supiéramos de antemano que ese límite existe. Pero precisamente eso es lo que el teorema debe demostrar. El sándwich realiza dos tareas a la vez:

1. prueba la **existencia** del límite de g;
2. identifica su valor como L.

La demostración ε–δ anterior no presupone ninguna de las dos.

Las dos funciones exteriores deben cerrar sobre el mismo número. La coincidencia de los límites exteriores es esencial. Que una función quede localmente entre otras dos funciones con límites distintos no obliga a que tenga límite. Por ejemplo, cerca de 0 podemos considerar

$$f(x) = 0, \qquad h(x) = 1,$$

y una función g que tome únicamente los valores 0 y 1 de manera oscilante. La desigualdad $0 \le g(x) \le 1$ no contiene información suficiente para determinar un único valor límite, porque las paredes exteriores no se aproximan entre sí. Las cotas deben converger **al mismo número**: esa coincidencia hace que el intervalo entre ellas termine contenido en cualquier entorno del candidato. Dos cotas con límites distintos solo encerrarían los valores en una franja de anchura positiva.

Geométricamente,

las paredes no solo encierran a g; además se cierran sobre L.

La forma absoluta del sándwich. En muchos problemas resulta artificial buscar dos funciones exteriores por separado. Lo natural es estimar directamente el error respecto del candidato L. Supongamos que existe una función $r\colon A \to \mathbb{R}$ con

$$\lim_{x\to a} r(x) = 0,$$

y que, en algún entorno perforado de a, $|g(x) - L| \le r(x)$. Como el lado izquierdo es no negativo, esa desigualdad fuerza localmente $r(x) \ge 0$. Además, $-r(x) \le g(x) - L \le r(x)$. Por Proposición 4.3, $-r(x) \to 0$, y por hipótesis $r(x) \to 0$. El teorema del sándwich produce $g(x) - L \to 0$.

Finalmente, usando de nuevo el álgebra de límites, $g(x) \to L$. Podemos registrar esta forma operativa como

$$\boxed{|g(x) - L| \le r(x) \text{ localmente}, \qquad r(x) \to 0 \quad \Longrightarrow \quad g(x) \to L.}$$

Esta formulación será particularmente útil cuando la función tenga oscilaciones, cambios de signo o factores cuyo límite individual resulte difícil de analizar. En lugar de calcular exactamente $g(x)$, buscamos una cota sencilla para su **error**.

Una oscilación amortiguada. El ejemplo siguiente muestra la fuerza de esa estrategia.

Ejemplo 4.6. **Una oscilación que desaparece por estar dominada.** Definamos

$$u\colon \mathbb{R} \setminus \{0\} \to \mathbb{R}, \qquad u(x) = x^2 \sigma\left(\frac{1}{x}\right).$$

El punto 0 es punto de acumulación de $\mathbb{R} \setminus \{0\}$, de modo que el límite está legítimamente planteado. Queremos estudiar

$$\lim_{x \to 0} x^2 \sigma\left(\frac{1}{x}\right).$$

El factor

$$\sigma\left(\frac{1}{x}\right)$$

oscila cuando x se aproxima a 0. No necesitamos determinar un límite para ese factor. Lo único que utilizaremos es la desigualdad elemental $|\sigma t| \le 1$ para todo $t \in \mathbb{R}$.

Por tanto, para todo $x \neq 0$,

$$\left| x^2 \sigma\left(\frac{1}{x}\right) \right| = x^2 \left| \sigma\left(\frac{1}{x}\right) \right| \le x^2.$$

Por Corolario 4.1,

$$\lim_{x \to 0} x^2 = 0.$$

Aplicando la forma absoluta del teorema del sándwich con $r(x) = x^2$, obtenemos

$$\boxed{\lim_{x \to 0} x^2 \sigma\left(\frac{1}{x}\right) = 0.}$$

El ejemplo merece una lectura detenida. La ley del producto de §4.5 no es aquí el camino adecuado: para usarla necesitaríamos disponer separadamente del límite de

$$\sigma\left(\frac{1}{x}\right)$$

cuando $x \to 0$.

El sándwich evita esa exigencia. No intenta comprender toda la información del factor oscilante. Solo utiliza el dato cuantitativo

$$\left| \sigma\left(\frac{1}{x}\right) \right| \le 1.$$

El factor x^2 hace entonces que el tamaño completo del producto quede atrapado por una función que tiende a cero.

El patrón es

$$\boxed{\text{factor uniformemente acotado} \times \text{factor cuyo tamaño tiende a } 0 \longrightarrow \text{producto que tiende a } 0,}$$

siempre que la estimación correspondiente haya sido demostrada.

El límite puede depender de una estimación menos precisa que la fórmula. El ejemplo anterior ilustra un principio metodológico importante. Para demostrar un límite no siempre necesitamos la descripción exacta de la función. La igualdad $u(x) = x^2\sigma(1/x)$ contiene mucha información: signo, oscilación, frecuencia de cambio y valores concretos. Para el límite en 0, casi toda esa información resulta innecesaria. La desigualdad más pobre $|u(x)| \leq x^2$ contiene menos datos, pero contiene exactamente los datos que necesitamos. Esto cambia el tipo de pregunta que conviene formular durante una prueba. En vez de preguntar siempre

¿puedo calcular exactamente esta expresión?,

conviene preguntar también

¿puedo controlar su tamaño mediante algo cuyo límite ya conozco?

En cálculo, una estimación adecuada puede ser más útil que una fórmula exacta.

Cómo fabricar una cota útil. Supongamos que queremos demostrar

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

Una estrategia frecuente consiste en comenzar por el error $|g(x) - L|$. Después intentamos transformarlo o estimarlo hasta obtener $|g(x) - L| \leq r(x)$, donde $r(x) \to 0$. La forma absoluta del sándwich termina entonces la prueba. El procedimiento puede resumirse así:

$$\boxed{\text{candidato } L \to |g(x) - L| \to \text{cota } r(x) \to r(x) \to 0 \to g(x) \to L.}$$

La cota no necesita ser óptima. Solo debe ser **suficientemente fuerte** para tender a cero. Por ejemplo, si cerca de a obtenemos $|g(x) - L| \leq 7|x - a|$, entonces no necesitamos mejorar el factor 7. Como

$$\lim_{x\to a} |x - a| = 0,$$

el lado derecho tiende a cero y el límite queda demostrado. Del mismo modo, una estimación $|g(x) - L| \leq |x - a| + |x - a|^2$ es suficiente porque ambas partes del lado derecho tienden a cero por el álgebra ya establecida.

Sincronizar desigualdades locales. Las pruebas de esta sección vuelven a utilizar una operación que ya apareció repetidamente: $\delta = \min\{\delta_1, \delta_2, \ldots, \delta_k\}$. Cada radio puede controlar una propiedad distinta:

- una desigualdad $f \leq g$;
- la proximidad de $f(x)$ a su límite;
- la proximidad de $h(x)$ al suyo;
- una estimación auxiliar necesaria para acotar el error.

Tomar el mínimo produce un único entorno perforado donde **todas** esas condiciones son válidas simultáneamente. Esto es el equivalente funcional de tomar un máximo de índices en una prueba con sucesiones:

$$\boxed{\begin{array}{c}\text{sucesiones: } N = \max\{N_1, \dots, N_k\},\\ \text{funciones: } \delta = \min\{\delta_1, \dots, \delta_k\}.\end{array}}$$

En ambos casos buscamos una región común donde todos los controles previamente obtenidos puedan utilizarse a la vez.

Tres errores que conviene evitar. 1. Una sola cota no determina un límite.

De $f(x) \le g(x)$ localmente y $f(x) \to L$ no se sigue que $g(x) \to L$. La función superior podría permanecer lejos de L.

2. Estar entre dos funciones no basta si sus límites son distintos.

El teorema del sándwich necesita que las dos funciones exteriores converjan al **mismo** número.

3. Una desigualdad estricta puede perder su estrictez en el límite.

De $f(x) < g(x)$ localmente solo podemos concluir, en general, $L \le M$. Para forzar $L < M$ necesitamos información cuantitativa adicional, como una separación positiva fija. Los tres errores tienen una raíz común: una desigualdad debe leerse junto con la información sobre **cómo se comporta la separación entre sus términos al acercarnos al punto**.

Qué hemos ganado. A la teoría algebraica de §4.5 añadimos ahora una teoría de orden local.

Sabemos que

$$f(x) \le g(x) \text{ localmente}, \qquad f(x) \to L, \qquad g(x) \to M \quad \Longrightarrow \quad L \le M,$$

y que

$$f(x) \le g(x) \le h(x) \text{ localmente}, \qquad f(x), h(x) \to L \quad \Longrightarrow \quad g(x) \to L.$$

Disponemos además de la forma operativa

$$|g(x) - L| \le r(x) \text{ localmente}, \qquad r(x) \to 0 \quad \Longrightarrow \quad g(x) \to L.$$

Con ella podemos controlar funciones oscilantes sin comprender por separado todos sus factores. La teoría desarrollada hasta aquí sigue estudiando aproximaciones a un punto sin distinguir desde qué lado llegan los puntos del dominio. Pero esa distinción puede ser decisiva. En §4.7 introduciremos límites laterales y veremos que una función puede aproximarse a un valor desde la izquierda y a otro distinto desde la derecha, impidiendo la existencia del límite bilateral.

4.7. Acercarse desde un lado: límites laterales

Hasta ahora hemos escrito

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

sin distinguir por qué lado de a llegan los puntos del dominio. La definición de Definición 4.2 observa todos los puntos de A que entran en un entorno perforado de a: $V_\delta^*(a) \cap A$. Pero un entorno perforado se descompone naturalmente en dos regiones: $V_\delta^*(a) = (a - \delta, a) \cup (a, a + \delta)$. La primera contiene puntos situados a la izquierda de a; la segunda, puntos situados a la derecha.

En muchos problemas ambas regiones contribuyen al acercamiento. En otros, solo una de ellas contiene puntos del dominio. Y también puede ocurrir que la función tenga comportamientos distintos en cada lado. Por eso necesitamos separar dos preguntas:

$$\boxed{\text{¿qué hace } f(x) \text{ cuando } x \text{ se acerca a } a \text{ desde la izquierda?}}$$

y

$$\boxed{\text{¿qué hace } f(x) \text{ cuando } x \text{ se acerca a } a \text{ desde la derecha?}}$$

La teoría no cambia de naturaleza. Seguiremos usando la misma lógica ε–δ, pero restringiremos los puntos admitidos por una desigualdad de orden.

Acumulación por izquierda y por derecha. Antes de hablar de límites laterales debemos repetir la precaución lógica de §4.1. Un límite lateral solo debe plantearse si realmente existen puntos del dominio arbitrariamente próximos a a por el lado correspondiente.

Definición 4.3. **Punto de acumulación lateral y límite lateral.** Sean $A \subseteq \mathbb{R}$, $a \in \mathbb{R}$ y $f\colon A \to \mathbb{R}$.

Diremos que a es **punto de acumulación de A por la izquierda** si

$$\forall \delta > 0\ \exists x \in A: \qquad a - \delta < x < a.$$

Equivalentemente,

$$(a-\delta, a) \cap A \neq \varnothing \qquad \text{para todo } \delta > 0.$$

Diremos que a es **punto de acumulación de A por la derecha** si

$$\forall \delta > 0\ \exists x \in A: \qquad a < x < a + \delta,$$

es decir, si

$$(a, a+\delta) \cap A \neq \varnothing \qquad \text{para todo } \delta > 0.$$

Si a es punto de acumulación de A por la izquierda, diremos que

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

si

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in A: \qquad a - \delta < x < a \Longrightarrow |f(x) - L| < \varepsilon.$$

Si a es punto de acumulación de A por la derecha, diremos que

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

si

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in A: \qquad a < x < a + \delta \Longrightarrow |f(x) - L| < \varepsilon.$$

Los signos − y + escritos sobre el punto de aproximación no son signos algebraicos aplicados a a. Indican **dirección de acercamiento**:

$$x \to a^- \quad \text{significa acercarse con } x < a,$$

mientras que

$$x \to a^+ \quad \text{significa acercarse con } x > a.$$

El punto a continúa excluido en ambos casos. En la definición izquierda, $x < a$ garantiza automáticamente $x \neq a$; en la derecha, $x > a$ hace lo mismo.

Ser punto de acumulación no significa disponer de ambos lados. Si a es punto de acumulación de A por la izquierda o por la derecha, entonces es punto de acumulación de A en el sentido de Definición 4.1. Pero la conversa no es cierta.

Consideremos

$$A = [0, \infty), \qquad a = 0.$$

Ya vimos en §4.1 que 0 es punto de acumulación de A: hay puntos positivos del dominio arbitrariamente próximos a 0. Más precisamente, 0 es punto de acumulación **por la derecha**. Sin embargo, no es punto de acumulación por la izquierda, porque $(-\delta, 0) \cap [0, \infty) = \varnothing$ para todo $\delta > 0$. Por tanto, bajo nuestra convención, tiene sentido preguntar por

$$\lim_{x \to 0^+} f(x),$$

pero no por un límite izquierdo de f en 0 si el dominio sigue siendo $[0, \infty)$. Esto evita una frase imprecisa que aparece con frecuencia:

> «Para que exista un límite deben existir los dos límites laterales».

No siempre. Esa afirmación necesita una hipótesis sobre el dominio. Si cerca de a el dominio solo contiene puntos por un lado, el límite ordinario de Definición 4.2 examina únicamente esos puntos disponibles.

Por ejemplo, sea

$$f\colon [0, \infty) \to \mathbb{R}, \qquad f(x) = x.$$

Dado $\varepsilon > 0$, tomando $\delta = \varepsilon$ obtenemos, para todo $x \in [0, \infty)$,

$$0 < |x| < \delta \Longrightarrow |f(x) - 0| = x < \varepsilon.$$

Así,

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

La misma elección demuestra

$$\lim_{x \to 0^+} f(x) = 0.$$

No hay un límite izquierdo que debamos comparar, porque 0 no es punto de acumulación del dominio por la izquierda.

Los límites laterales son límites ordinarios sobre dominios restringidos. La nueva notación puede parecer una teoría independiente, pero en realidad es una especialización de lo que ya construimos.

Definamos

$$A_- := A \cap (-\infty, a), \qquad A_+ := A \cap (a, \infty).$$

Entonces a es punto de acumulación de A por la izquierda exactamente cuando es punto de acumulación de A_-. Del mismo modo, es punto de acumulación por la derecha exactamente cuando es punto de acumulación de A_+.

Además,

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

significa exactamente

$$\lim_{x \to a} f|_{A_-}(x) = L,$$

mientras que

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

significa

$$\lim_{x \to a} f|_{A_+}(x) = L.$$

Esta observación es importante porque nos entrega inmediatamente toda la teoría ya demostrada para límites ordinarios, aplicada a las restricciones correspondientes. En particular:

- cada límite lateral, cuando existe, es único por Teorema 4.1;
- las leyes algebraicas de Proposición 4.3 se aplican lateralmente;
- el orden y el sándwich de §4.6 se aplican lateralmente;
- el criterio secuencial de Teorema 4.2 se aplica a sucesiones que permanezcan en el lado adecuado.

Por ejemplo,

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

si y solo si para toda sucesión (x_n) de puntos de A con

$$x_n < a, \qquad x_n \to a,$$

se cumple $f(x_n) \to L$.

La versión derecha se obtiene sustituyendo $x_n < a$ por $x_n > a$. No hemos creado una segunda teoría de límites. Hemos aprendido a **restringir la dirección desde la cual se permite aproximarse**.

Cuando hay puntos por ambos lados. Si a es punto de acumulación de A tanto por la izquierda como por la derecha, cada entorno perforado contiene puntos admisibles de ambos lados. Tendremos que controlar las dos formas de aproximación. En esa situación sí aparece el criterio clásico que conecta los tres límites.

TEOREMA 4.4. **Criterio bilateral mediante límites laterales.** Sea $f\colon A \to \mathbb{R}$ y supongamos que a es punto de acumulación de A por la izquierda y por la derecha. Entonces, para $L \in \mathbb{R}$,

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

si y solo si

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

Demostración. Supongamos primero que

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

Sea $\varepsilon > 0$. Por la definición bilateral existe $\delta > 0$ tal que, para todo $x \in A$,

$$0 < |x - a| < \delta \Longrightarrow |f(x) - L| < \varepsilon.$$

Si

$$a - \delta < x < a,$$

entonces automáticamente

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

y por tanto

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

Así, el mismo radio δ demuestra el límite izquierdo. El argumento para el límite derecho es idéntico: de

$$a < x < a + \delta$$

se deduce

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

Por consiguiente,

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

Recíprocamente, supongamos que ambos límites laterales existen y valen L. Sea $\varepsilon > 0$. Del límite izquierdo obtenemos un radio $\delta_- > 0$ tal que

$$a - \delta_- < x < a, \qquad x \in A,$$

implica

$$|f(x) - L| < \varepsilon.$$

Del límite derecho obtenemos $\delta_+ > 0$ tal que

$$a < x < a + \delta_+, \qquad x \in A,$$

implica la misma estimación.
Tomemos

$$\delta := \min\{\delta_-, \delta_+\}.$$

Sea ahora $x \in A$ tal que

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

Como $x \neq a$, por tricotomía solo hay dos posibilidades.
Si $x < a$, entonces

$$a - \delta < x < a,$$

y como $\delta \leq \delta_-$, obtenemos

$$|f(x) - L| < \varepsilon.$$

Si $x > a$, entonces

$$a < x < a + \delta,$$

y como $\delta \leq \delta_+$, obtenemos nuevamente

$$|f(x) - L| < \varepsilon.$$

En ambos casos se cumple la condición bilateral. Por tanto,

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

La demostración muestra con precisión qué hace el mínimo de radios:

$$\boxed{\text{control izquierdo} + \text{control derecho} \xrightarrow{\min\{\delta_-, \delta_+\}} \text{control bilateral.}}$$

También muestra que no basta con que los dos límites laterales **existan**. Deben coincidir.

Dos valores laterales distintos destruyen el límite bilateral.

Ejemplo 4.7. **La función signo: dos acercamientos, dos valores.** Consideremos

$$f\colon \mathbb{R} \setminus \{0\} \to \mathbb{R}, \qquad f(x) = \frac{|x|}{x}.$$

Si $x < 0$, entonces $|x| = -x$, de modo que $f(x) = \frac{-x}{x} = -1$. Por tanto, en todo el lado izquierdo de 0 la función es exactamente constante e igual a -1. De aquí se sigue inmediatamente que

$$\lim_{x\to 0^-} \frac{|x|}{x} = -1.$$

Si $x > 0$, entonces $|x| = x$, y por ello $f(x) = \frac{x}{x} = 1$.
Así,

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

Los dos límites laterales existen, pero $-1 \neq 1$. Como 0 es punto de acumulación del dominio por ambos lados, Teorema 4.4 implica que el límite bilateral

$$\lim_{x\to 0} \frac{|x|}{x}$$

no existe como número real.

El ejemplo permite ver el fenómeno sin esconderlo detrás de cálculo algebraico. A la izquierda, todos los valores próximos son -1; a la derecha, todos son 1. Ningún número real puede describir simultáneamente ambos comportamientos. También podríamos extender la función a 0 asignando cualquier valor, por ejemplo $f(0) = 0$, sin alterar ninguno de los límites laterales ni reparar el límite bilateral. El punto exacto continúa siendo irrelevante para el límite.

El criterio secuencial permite diagnosticar la discrepancia lateral. El ejemplo anterior también puede leerse con Teorema 4.2. Tomemos

$$x_n = -\frac{1}{n+1}, \qquad y_n = \frac{1}{n+1}.$$

Ambas sucesiones convergen a 0 y evitan el punto 0. Pero $f(x_n) = -1$ para todo n, mientras que $f(y_n) = 1$ para todo n. Por tanto,

$$f(x_n) \to -1, \qquad f(y_n) \to 1.$$

Una sola sucesión que llega desde la izquierda detecta el límite izquierdo; otra que llega desde la derecha detecta el derecho. Juntas muestran por qué ningún único límite bilateral puede sobrevivir. Este patrón será útil al refutar límites:

dos aproximaciones admisibles al mismo punto + dos límites de imagen distintos $\Longrightarrow$ no existe límite bilateral.

La diferencia respecto de §4.4 es que ahora la elección de las sucesiones está guiada por la geometría lateral del dominio.

Qué debe auditarse antes de usar límites laterales. Antes de escribir un límite lateral conviene comprobar tres cosas. Primero, **el lado debe contener puntos del dominio arbitrariamente próximos al centro**. Si no los contiene, la condición sería vacua y no adoptaremos esa expresión como límite lateral. Segundo, la notación a^- o a^+ no dice nada sobre el signo de a. Por ejemplo, $x \to (-3)^+$ significa acercarse a -3 mediante valores mayores que -3. Tercero, cuando ambos lados están disponibles, hay que distinguir tres preguntas:

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

Las dos primeras estudian restricciones distintas del dominio. La tercera exige compatibilidad entre todo punto disponible suficientemente próximo a a. En resumen, si ambos lados existen geométricamente,

$$\boxed{\text{límite bilateral} \iff \text{límite izquierdo} = \text{límite derecho.}}$$

Pero si el dominio solo se aproxima por un lado, el límite ordinario sigue siendo relativo a ese dominio y no exige inventar un acercamiento inexistente. La siguiente ampliación será de otra naturaleza. Hasta aquí todos los límites han tenido valores reales finitos y el punto de aproximación ha sido un número real. En §4.8 permitiremos que la salida escape sin cota y también estudiaremos qué significa que la **entrada** se aleje hacia $+\infty$ o $-\infty$.

4.8. Cuando la salida o la entrada escapan: límites infinitos y límites en infinito

Hasta ahora hemos estudiado límites funcionales de la forma

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

donde tanto el punto de aproximación a como el valor límite L son números reales. Hay, sin embargo, dos ampliaciones distintas que aparecen de manera natural. En la primera, la **entrada sigue acercándose a un punto real** a, pero los valores de la función crecen o decrecen sin quedar finalmente cerca de ningún número real. Por ejemplo, cuando x se aproxima a 0 con $x \neq 0$, $\frac{1}{x^2}$ supera cualquier barrera positiva si $|x|$ es suficientemente pequeño. En la segunda, ya no hacemos que la entrada se acerque a un número real. Preguntamos qué ocurre con $f(x)$ cuando x se hace arbitrariamente grande o arbitrariamente negativo. Por ejemplo, $\frac{2x+1}{x-3}$ se acerca a 2 cuando x avanza hacia valores positivos cada vez mayores.

Conviene separar cuidadosamente las dos ideas:

$$\boxed{\begin{array}{rl} x \to a, \quad f(x) \to \pm\infty & \text{salida que escapa,} \\ x \to \pm\infty, \quad f(x) \to L \text{ o } \pm\infty & \text{entrada que escapa.} \end{array}}$$

En ambos casos, el símbolo ∞ abrevia una condición cuantificada. No designa un número real.

Una salida que termina más allá de toda barrera. En `T1-C04` definimos $a_n \to +\infty$ diciendo que, dada cualquier barrera real M, toda la cola de la sucesión termina por encima de M. Ahora sustituimos la cola por un entorno perforado.

Definición 4.4. **Límites funcionales infinitos en un punto.** Sea

$$f\colon A \to \mathbb{R},$$

y sea a un punto de acumulación de A.

Diremos que

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

si para todo número real M existe $\delta > 0$ tal que, para todo $x \in A$,

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

En símbolos,

$$\forall M \in \mathbb{R}\ \exists \delta > 0\ \forall x \in A: \quad 0 < |x - a| < \delta \Longrightarrow f(x) > M.$$

Diremos que

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

si para todo número real M existe $\delta > 0$ tal que, para todo $x \in A$,

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

En símbolos,

$$\forall M \in \mathbb{R}\ \exists \delta > 0\ \forall x \in A: \quad 0 < |x - a| < \delta \Longrightarrow f(x) < M.$$

La estructura lógica es paralela a la de un límite real, pero el desafío ha cambiado. Para

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

alguien fija una tolerancia $\varepsilon > 0$ alrededor de L y nosotros buscamos un radio que fuerce $|f(x) - L| < \varepsilon$.

Para

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

alguien fija una barrera real M y debemos encontrar un radio que fuerce $f(x) > M$.

Por tanto,

barrera arbitraria $\longrightarrow$ radio adecuado $\longrightarrow$ todos los valores próximos quedan más allá de la barrera.

La palabra **todos** vuelve a ser decisiva. No basta con encontrar valores de f arbitrariamente grandes cerca de a. El crecimiento hacia $+\infty$ exige que, dentro de algún entorno perforado suficientemente pequeño, **ningún valor vuelva a caer por debajo de la barrera fijada**.

$+\infty$ y $-\infty$ no son valores de la función. La escritura

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

no afirma que exista un número real llamado $+\infty$ al que se aproximan los valores. Tampoco afirma que la función tome el valor $+\infty$. Seguimos trabajando con $f(x) \in \mathbb{R}$ para cada x de su dominio. El símbolo describe un patrón de escape: $\forall M \in \mathbb{R}$, los valores de f terminan localmente por encima de M. De manera análoga, $f(x) \to -\infty$ significa que los valores terminan por debajo de **toda** barrera real.

Así,

$$+\infty \notin \mathbb{R}, \qquad -\infty \notin \mathbb{R},$$

y una función que tiene límite infinito en a no tiene allí un límite real finito en el sentido de Definición 4.2. La incompatibilidad puede verse directamente. Si, por ejemplo,

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

entonces, dada cualquier cota real B, existe un entorno perforado en el cual $f(x) > B$. Por tanto, la función no está localmente acotada superiormente alrededor de a. Esto contrasta con Proposición 4.2, donde demostramos que un límite real finito fuerza acotación local.

Un ejemplo rector: $1/x^2$ cerca de cero.

Ejemplo 4.8. **Un límite infinito en un punto.** Consideremos

$$f\colon \mathbb{R} \setminus \{0\} \to \mathbb{R}, \qquad f(x) = \frac{1}{x^2}.$$

Entonces

$$\lim_{x\to 0} \frac{1}{x^2} = +\infty.$$

Demostración. Sea $M \in \mathbb{R}$ una barrera arbitraria. Definamos $K := |M| + 1$.

Entonces

$$K > 0 \qquad \text{y} \qquad K > M.$$

Tomemos $\delta := \frac{1}{\sqrt{K}} > 0$.

Si $x \neq 0$ y $0 < |x| < \delta$, entonces $x^2 < \delta^2 = \frac{1}{K}$.

Como $x^2 > 0$ y $1/K > 0$, al tomar recíprocos se invierte el orden: $\frac{1}{x^2} > K > M$. Así, para toda barrera real M hemos construido un radio δ tal que

$$0 < |x| < \delta \quad \Longrightarrow \quad \frac{1}{x^2} > M.$$

Por Definición 4.4,

$$\lim_{x\to 0} \frac{1}{x^2} = +\infty.$$

■

La prueba contiene el mismo movimiento que una demostración ε–δ: partimos de la condición que deseamos en la **salida**, $\frac{1}{x^2} > M$, y la convertimos en una condición suficiente sobre la **entrada**, $|x| < \frac{1}{\sqrt{K}}$. La diferencia es que ya no intentamos entrar en una banda horizontal alrededor de un número real. Intentamos quedar por encima de una altura arbitraria.

Los laterales siguen importando. La definición anterior se combina con §4.7 simplemente restringiendo el dominio al lado correspondiente.

Por ejemplo,

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

La primera afirmación dice que, dada cualquier barrera M, los valores $1/x$ terminan por encima de M cuando $x > 0$ se acerca suficientemente a 0. La segunda dice que, dada cualquier barrera M, los valores $1/x$ terminan por debajo de M cuando $x < 0$ se acerca suficientemente a 0. Por tanto, no sería correcto escribir

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

ni

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

El comportamiento bilateral no escapa en una única dirección. La información lateral sigue siendo necesaria para diagnosticarlo.

Ahora hacemos escapar la entrada. La segunda ampliación modifica el significado de la expresión situada debajo de la flecha. Cuando escribimos $x \to +\infty$, no afirmamos que x se acerque a un punto llamado $+\infty$. Queremos decir que x termina superando cualquier umbral real que fijemos. Pero aquí aparece el mismo problema de vacuidad que encontramos en §4.1. Supongamos que el dominio A estuviera acotado superiormente. Entonces existiría una cota B tal que $x \leq B$ para todo $x \in A$. Si eligiéramos un umbral $R > B$, no existiría ningún $x \in A$ con $x > R$. Cualquier implicación cuyo antecedente fuera $x > R$ quedaría automáticamente verdadera. De ese modo podríamos fabricar falsos límites en $+\infty$ exactamente como podíamos fabricar falsos límites en puntos aislados. Por eso exigiremos que el dominio tenga puntos arbitrariamente grandes.

Definición 4.5. **Límites cuando la variable tiende a $\pm\infty$.** Sea

$$f\colon A \to \mathbb{R}.$$

Para hablar de $x \to +\infty$ exigiremos que A sea **no acotado superiormente**; es decir,

$$\forall R \in \mathbb{R}\ \exists x \in A: \quad x > R.$$

Para hablar de $x \to -\infty$ exigiremos que A sea **no acotado inferiormente**; es decir,

$$\forall R \in \mathbb{R}\ \exists x \in A: \quad x < R.$$

Sea $L \in \mathbb{R}$.

Diremos que

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

si para todo $\varepsilon > 0$ existe $R \in \mathbb{R}$ tal que, para todo $x \in A$,

$$x > R \quad \Longrightarrow \quad |f(x) - L| < \varepsilon.$$

Diremos que

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

si para todo $\varepsilon > 0$ existe $R \in \mathbb{R}$ tal que, para todo $x \in A$,

$$x < R \quad \Longrightarrow \quad |f(x) - L| < \varepsilon.$$

Si además queremos describir una salida infinita, reemplazamos la banda $|f(x) - L| < \varepsilon$ por una condición de barrera. Por ejemplo,

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

significa

$$\forall M \in \mathbb{R}\ \exists R \in \mathbb{R}\ \forall x \in A: \quad x > R \Longrightarrow f(x) > M.$$

Las otras tres combinaciones quedan definidas de manera explícita.
Escribiremos

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

si

$$\forall M \in \mathbb{R}\ \exists R \in \mathbb{R}\ \forall x \in A: \quad x > R \Longrightarrow f(x) < M.$$

Escribiremos

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

si

$$\forall M \in \mathbb{R}\ \exists R \in \mathbb{R}\ \forall x \in A: \quad x < R \Longrightarrow f(x) > M.$$

Finalmente,

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

si

$$\forall M \in \mathbb{R}\ \exists R \in \mathbb{R}\ \forall x \in A: \quad x < R \Longrightarrow f(x) < M.$$

En las cuatro combinaciones, la dirección de la entrada determina si controlamos $x > R$ o $x < R$, y la dirección de la salida determina si exigimos $f(x) > M$ o $f(x) < M$.

Esta definición contiene dos mecanismos independientes. La condición sobre la **entrada** decide qué región del dominio debemos controlar:

$$x > R \qquad \text{o} \qquad x < R.$$

La condición sobre la **salida** decide qué comportamiento exigimos allí:

$$|f(x) - L| < \varepsilon, \qquad f(x) > M, \qquad \text{o} \qquad f(x) < M.$$

Por eso una expresión como

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

contiene dos símbolos $+\infty$ con funciones lógicas distintas. El primero describe que la **entrada** termina más allá de todo umbral; el segundo describe que la **salida** termina más allá de toda barrera.

El dominio vuelve a formar parte del problema. La exigencia de no acotación no es un detalle técnico. Compara exactamente con la hipótesis de punto de acumulación usada para límites en puntos finitos:

$x \to a$	puntos del dominio arbitrariamente próximos a a
$x \to +\infty$	puntos del dominio arbitrariamente grandes
$x \to -\infty$	puntos del dominio arbitrariamente negativos

Por ejemplo, el dominio $A = \mathbb{N}$ es no acotado superiormente, de modo que tiene sentido estudiar límites de funciones sobre A cuando $x \to +\infty$. En cambio, no es no acotado inferiormente, por lo que no consideraremos $x \to -\infty$ dentro de ese dominio. Del mismo modo, $A = (-\infty, 5]$ permite estudiar $x \to -\infty$, pero no $x \to +\infty$. Una vez más, la fórmula aislada no decide qué aproximaciones están disponibles: lo decide el dominio completo.

Un límite finito cuando $x \to +\infty$.

Ejemplo 4.9. **Una función racional que se estabiliza en infinito.** Consideremos

$$f\colon \mathbb{R} \setminus \{3\} \to \mathbb{R}, \qquad f(x) = \frac{2x+1}{x-3}.$$

Entonces

$$\lim_{x \to +\infty} \frac{2x+1}{x-3} = 2.$$

Demostración. Sea $\varepsilon > 0$. Queremos encontrar un umbral R tal que $x > R$ fuerce

$$\left| \frac{2x+1}{x-3} - 2 \right| < \varepsilon.$$

Simplifiquemos primero el error: $\frac{2x+1}{x-3} - 2 = \frac{2x+1-2x+6}{x-3} = \frac{7}{x-3}$.
Si imponemos $x > 3 + \frac{7}{\varepsilon}$, entonces $x - 3 > \frac{7}{\varepsilon} > 0$.
Por tanto,

$$\left| \frac{2x+1}{x-3} - 2 \right| = \frac{7}{x-3} < \varepsilon.$$

Basta elegir $R := 3 + \frac{7}{\varepsilon}$.
Así, para todo x del dominio con $x > R$, $|f(x) - 2| < \varepsilon$, y concluimos que

$$\lim_{x \to +\infty} \frac{2x+1}{x-3} = 2.$$

■

La prueba vuelve a mostrar una arquitectura conocida:

tolerancia en la salida
$\longrightarrow$ umbral en la entrada
$\longrightarrow$ control de todos los puntos posteriores del dominio.

El papel de R es análogo al de N para sucesiones. La diferencia es que ahora la variable recorre puntos reales del dominio y no índices naturales.

También podemos tener salida infinita cuando la entrada escapa. Consideremos $f(x) = x^2$.

Cuando $x \to +\infty$, también $x^2 \to +\infty$.

En efecto, dada una barrera $M \in \mathbb{R}$, tomemos $K := |M| + 1$ y $R := \sqrt{K}$.

Si $x > R$, entonces $x > 0$ y

$$x^2 > R^2 = K > M.$$

Así,

$$\lim_{x \to +\infty} x^2 = +\infty.$$

El mismo ejemplo satisface

$$\lim_{x \to -\infty} x^2 = +\infty,$$

pero ahora la región de entrada que debe controlarse es una semirrecta de la forma $x < R$ con R suficientemente negativo. Esto muestra que el signo del infinito de la **entrada** no determina el signo del infinito de la **salida**.

La lectura secuencial sigue disponible. Las definiciones de esta sección son la traducción funcional de los patrones que ya construimos para sucesiones en `T1-C04`.

Por ejemplo, si

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

y (x_n) es cualquier sucesión de puntos del dominio con

$$x_n \neq a, \qquad x_n \to a,$$

entonces $f(x_n) \to +\infty$ en el sentido de Definición 3.5. La razón es directa. Fijada una barrera M, Definición 4.4 proporciona un radio $\delta > 0$. Como $x_n \to a$, existe N tal que $0 < |x_n - a| < \delta$ para todo $n \geq N$. Entonces $f(x_n) > M$ para toda esa cola.

Análogamente, si

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

y una sucesión (x_n) de puntos de A satisface $x_n \to +\infty$, entonces $f(x_n) \to L$. Fijado $\varepsilon > 0$, elegimos el umbral R dado por Definición 4.5. Como $x_n \to +\infty$, toda la cola termina satisfaciendo $x_n > R$, y por tanto $|f(x_n) - L| < \varepsilon$. No necesitamos todavía formular como teorema todas las conversas secuenciales posibles. Lo importante aquí es reconocer que la misma lógica de **umbral + cola** vuelve a aparecer dentro de los límites funcionales en infinito.

No acotación tampoco basta para una función. La advertencia de `T1-C04` se conserva. Que una función tome valores arbitrariamente grandes cerca de a no implica que

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

Por ejemplo, cerca de 0 la función $f(x) = \frac{1}{x}$ toma valores positivos arbitrariamente grandes y valores negativos arbitrariamente pequeños. No existe un entorno perforado bilateral en el que todos los valores queden por encima de una barrera positiva, ni uno en el que todos queden por

debajo de una barrera negativa. Del mismo modo, una función puede ser no acotada cuando $x \to +\infty$ sin tender a $+\infty$ ni a $-\infty$. El patrón relevante no es

aparecen valores muy grandes,

sino

después de algún umbral, todos los valores quedan del lado exigido de la barrera.

No haremos aritmética ficticia con ∞. Las leyes algebraicas de Proposición 4.3 fueron demostradas para límites reales finitos. Las definiciones nuevas no autorizan a tratar $+\infty$ o $-\infty$ como si fueran números reales adicionales.

Por ejemplo, cuando $x \to +\infty$,

$$x \to +\infty \qquad \text{y} \qquad -x \to -\infty.$$

Pero distintas combinaciones con esos mismos tipos de comportamiento pueden producir resultados completamente diferentes. No existe por tanto una regla numérica obtenida escribiendo simplemente $+\infty + (-\infty)$. Del mismo modo, expresiones informales como

$$\frac{\infty}{\infty}, \qquad 0 \cdot \infty, \qquad \infty - \infty$$

no son operaciones realizadas dentro de $\mathbb{R}$. En este capítulo no desarrollaremos una teoría general para tales patrones. Cuando necesitemos calcular un límite concreto, volveremos a definiciones, estimaciones, álgebra legítima sobre cantidades reales y transformaciones justificadas.

Un mapa de las cuatro direcciones. Conviene cerrar la sección separando las dos preguntas que pueden variar independientemente. La **entrada** puede dirigirse hacia

$$a, \qquad +\infty, \qquad -\infty.$$

La **salida** puede aproximarse a un número real L o escapar hacia

$$+\infty \qquad \text{o} \qquad -\infty.$$

Cada combinación se interpreta mediante cuantificadores, no mediante una imagen informal de «llegar al infinito».

El principio unificador es:

fijar una exigencia sobre la salida
$\longrightarrow$ encontrar una región suficientemente avanzada de la entrada
$\longrightarrow$ controlar todos los puntos del dominio en esa región.

Cuando $x \to a$, esa región es un entorno perforado.

Cuando $x \to +\infty$, es una semirrecta $x > R$.

Cuando $x \to -\infty$, es una semirrecta $x < R$.

Con esto ya disponemos de casi todo el vocabulario de límites necesario para el capítulo. En §4.9 volveremos a límites reales finitos para abordar una operación especialmente delicada: **componer límites** sin ocultar qué ocurre cuando la función interior toma exactamente el punto que el límite exterior excluye.

4.9. Componer límites sin esconder la hipótesis delicada

Hasta ahora hemos aprendido a calcular y comparar límites de una misma variable mediante álgebra, orden, sándwich, aproximaciones laterales y distintos tipos de infinito. Hay, sin embargo, una operación que merece una auditoría propia: **la composición**.

Supongamos que

$$g(x) \to b \qquad (x \to a)$$

y que, por otra parte,

$$f(y) \to L \qquad (y \to b).$$

La tentación natural es encadenar ambas afirmaciones y concluir inmediatamente

$$f(g(x)) \to L \qquad (x \to a).$$

La idea es correcta en muchas situaciones, pero existe un pequeño agujero lógico que no debemos ocultar.

El límite exterior no controla lo que ocurre exactamente en b. Recordemos qué afirma

$$\lim_{y \to b} f(y) = L.$$

Para cada $\varepsilon > 0$ existe un radio $\eta > 0$ tal que

$$0 < |y - b| < \eta \quad \Longrightarrow \quad |f(y) - L| < \varepsilon.$$

La condición contiene dos desigualdades:

$$0 < |y - b| \qquad \text{y} \qquad |y - b| < \eta.$$

La segunda dice que y está suficientemente cerca de b. La primera dice que y **no es exactamente** b. Ahora bien, del límite interior

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

podemos conseguir, para x suficientemente próximo a a, $|g(x) - b| < \eta$. Pero esta desigualdad no garantiza por sí sola $0 < |g(x) - b|$. Podría ocurrir que $g(x) = b$ para algunos puntos, o incluso para todos los puntos, arbitrariamente próximos a a. Ese es el lugar exacto donde una demostración ingenua de la composición puede romperse. El diagnóstico puede resumirse así:

el límite interior entrega $\lvert g(x) - b \rvert < \eta$,
el límite exterior exige $0 < \lvert g(x) - b \rvert < \eta$.

Falta controlar la desigualdad estricta de la izquierda.

La condición de no impacto local. Una manera limpia de cerrar ese hueco consiste en exigir que, suficientemente cerca de a y sin usar el propio punto a, la función interior no tome exactamente el valor b. Es decir, pediremos que exista algún radio $r > 0$ tal que

$$x \in A, \qquad 0 < |x - a| < r \quad \Longrightarrow \quad g(x) \neq b.$$

Esta condición es enteramente local. No exige que $g(x) \neq b$ en todo el dominio, sino únicamente en algún entorno perforado de a. Además, si b ni siquiera pertenece al codominio efectivo B de g, la condición es automática: como $g(x) \in B$, nunca puede ocurrir $g(x) = b$. El caso delicado aparece cuando $b \in B$ y la función interior puede golpear ese punto.

Teorema 4.5. **Composición de límites bajo condición de no impacto local.** Sean

$$g\colon A \to B, \qquad f\colon B \to \mathbb{R},$$

con $A, B \subseteq \mathbb{R}$. Supongamos que a es punto de acumulación de A, que b es punto de acumulación de B y que

$$\lim_{x \to a} g(x) = b, \qquad \lim_{y \to b} f(y) = L.$$

Supongamos además que existe $r > 0$ tal que, para todo $x \in A$,

$$0 < |x - a| < r \quad \Longrightarrow \quad g(x) \neq b.$$

Entonces

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

Demostración. Sea $\varepsilon > 0$.

Como

$$\lim_{y \to b} f(y) = L,$$

existe $\eta > 0$ tal que, para todo $y \in B$,

$$0 < |y - b| < \eta \quad \Longrightarrow \quad |f(y) - L| < \varepsilon.$$

Ahora usamos el límite interior. Como

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

existe $\delta_g > 0$ tal que, para todo $x \in A$,

$$0 < |x - a| < \delta_g \quad \Longrightarrow \quad |g(x) - b| < \eta.$$

Por hipótesis existe además $r > 0$ tal que

$$0 < |x - a| < r \quad \Longrightarrow \quad g(x) \neq b.$$

Sincronizamos ambos controles tomando $\delta := \min\{\delta_g, r\}$.

Sea ahora $x \in A$ y supongamos $0 < |x - a| < \delta$.

Como $\delta \leq \delta_g$, $|g(x) - b| < \eta$.

Como $\delta \le r$, $g(x) \neq b$, y por tanto $0 < |g(x) - b|$.

Juntando las dos desigualdades obtenemos $0 < |g(x) - b| < \eta$. Además, $g(x) \in B$ porque $g\colon A \to B$. Podemos entonces aplicar el control proporcionado por el límite exterior y concluir $|f(g(x)) - L| < \varepsilon$. Hemos probado que para todo $\varepsilon > 0$ existe $\delta > 0$ tal que

$$x \in A, \qquad 0 < |x - a| < \delta \quad \Longrightarrow \quad |f(g(x)) - L| < \varepsilon.$$

Por definición,

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

■

La arquitectura de la prueba merece quedar visible:

$$\begin{gathered} \varepsilon \xrightarrow{f(y)\to L} \eta \\ \eta \xrightarrow{g(x)\to b} \delta_g \\ \text{no impacto local} \to r \\ \delta = \min\{\delta_g, r\} \to 0 < |g(x) - b| < \eta \to |f(g(x)) - L| < \varepsilon. \end{gathered}$$

El mínimo cumple aquí la misma función que en otras pruebas locales del capítulo: sincroniza condiciones que deben ser verdaderas simultáneamente.

Un caso donde la condición es automática. Consideremos

$$A = \mathbb{R} \setminus \{0\}, \qquad B = \mathbb{R} \setminus \{1\},$$

y definamos

$$g\colon A \to B, \qquad g(x) = x + 1,$$

junto con

$$f\colon B \to \mathbb{R}, \qquad f(y) = \frac{y^2 - 1}{y - 1}.$$

Cuando $x \to 0$ dentro de A, $g(x) = x + 1 \to 1$.

Por otra parte, el límite ya estudiado de la función racional con agujero da

$$\lim_{y \to 1} \frac{y^2 - 1}{y - 1} = 2.$$

Como $1 \notin B$, ningún valor de g puede ser igual a 1. La condición de no impacto es automática, y Teorema 4.5 permite concluir

$$\lim_{x \to 0} f(g(x)) = 2.$$

De hecho, para $x \neq 0$,

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

de modo que el resultado también puede comprobarse directamente. Este ejemplo muestra que la condición adicional no es una complicación artificial. Simplemente verifica que los valores producidos por la función interior pertenecen a la región perforada sobre la cual el límite exterior ejerce control.

Qué ocurre si omitimos la condición. Ahora construiremos el fallo más sencillo posible.

Ejemplo 4.10. **La función interior golpea exactamente el punto que el límite exterior ignora.** Definamos

$$g\colon \mathbb{R} \to \mathbb{R}, \qquad g(x) = 0$$

para todo $x \in \mathbb{R}$, y

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(y) = \begin{cases} 1, & y = 0, \\ 0, & y \neq 0. \end{cases}$$

Entonces

$$\lim_{x \to 0} g(x) = 0$$

y

$$\lim_{y \to 0} f(y) = 0,$$

pero $f(g(x)) = 1$ para todo x, de modo que

$$\lim_{x \to 0} f(g(x)) = 1 \neq 0.$$

Verifiquemos cada afirmación.

La primera es inmediata: para cualquier $\varepsilon > 0$ y cualquier x, $|g(x) - 0| = 0 < \varepsilon$.

Por tanto,

$$g(x) \to 0 \qquad (x \to 0).$$

Para la segunda, fijemos $\varepsilon > 0$. Si $0 < |y| < 1$, entonces necesariamente $y \neq 0$, y por definición $f(y) = 0$.

Así, $|f(y) - 0| = 0 < \varepsilon$.

Por consiguiente,

$$\lim_{y \to 0} f(y) = 0.$$

Sin embargo, la composición evalúa siempre a f **exactamente** en el punto que el límite exterior excluye:

$$g(x) = 0 \quad \Longrightarrow \quad f(g(x)) = f(0) = 1.$$

La condición de no impacto no falla ocasionalmente: falla para todo punto del dominio. El contraejemplo localiza con precisión el problema. El límite

$$\lim_{y \to 0} f(y) = 0$$

no contiene ninguna información sobre $f(0)$. Por eso no puede controlar una composición que insiste en evaluar justamente $f(0)$.

La hipótesis es suficiente, no necesaria. Debemos interpretar correctamente Teorema 4.5. La condición $g(x) \neq b$ cerca de a es una condición **suficiente** para garantizar la composición a partir de los dos límites dados. No afirmamos que sea necesaria en cada ejemplo concreto. Puede ocurrir que $g(x) = b$ en puntos arbitrariamente próximos a a y que, aun así, la composición tenga el límite esperado por alguna razón adicional. Por ejemplo, si el valor efectivo de f en b coincide con el comportamiento límite exterior, los impactos en b pueden dejar de causar problemas. Pero esa información no está contenida en la afirmación

$$\lim_{y \to b} f(y) = L.$$

Por eso no podemos eliminar sin más la hipótesis de no impacto de un teorema general basado únicamente en los dos límites perforados. En el capítulo siguiente introduciremos una noción que incorpora explícitamente el valor de la función en el punto. Allí la composición adquirirá una formulación más familiar. Aquí no utilizaremos esa teoría antes de haberla definido.

Lectura secuencial de la misma prueba. El criterio secuencial permite comprobar la misma arquitectura desde otro ángulo. Sea (x_n) una sucesión de puntos de A tal que

$$x_n \neq a, \qquad x_n \to a.$$

Como $g(x) \to b$, el criterio secuencial da $g(x_n) \to b$. La condición de no impacto local asegura además que, a partir de algún índice, $g(x_n) \neq b$. Por tanto, la cola de $(g(x_n))$ es una sucesión admisible para aplicar el criterio secuencial al límite exterior. De

$$f(y) \to L \qquad (y \to b)$$

obtenemos entonces $f(g(x_n)) \to L$.

Como esto sucede para toda sucesión admisible (x_n), Teorema 4.2 recupera

$$f(g(x)) \to L \qquad (x \to a).$$

Esta segunda lectura no sustituye la demostración directa: la confirma y muestra nuevamente cómo `T1-C04` continúa operando dentro de la teoría funcional.

Protocolo de auditoría para una composición. Antes de escribir automáticamente

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

conviene detener esa notación, porque todavía no hemos demostrado una regla general que permita evaluar f en un límite como si fuera una operación algebraica. En este capítulo el procedimiento seguro es:

1. identificar
$$\lim_{x \to a} g(x) = b;$$
2. verificar que el límite exterior
$$\lim_{y \to b} f(y) = L$$
está legítimamente planteado;

3. comprobar el tipado de la composición: $g(x)$ debe pertenecer al dominio de f;
4. auditar el punto excluido: comprobar que $g(x) \neq b$ en algún entorno perforado de a, salvo que se disponga de otra información que trate explícitamente los impactos en b;
5. solo entonces encadenar los controles de los dos límites.

El núcleo lógico puede condensarse en una pregunta:

¿la salida de la función interior entra realmente en la región donde el límite exterior controla a f?

Esta pregunta será una de las estaciones del laboratorio de la sección siguiente. Con §4.9 queda cerrada la teoría nueva de límites finitos prevista para el capítulo. En §4.10 no añadiremos definiciones ni teoremas: organizaremos todas las herramientas construidas hasta aquí en un **laboratorio de límite funcional** antes de pasar al banco de ejercicios y soluciones.

4.10. Laboratorio de límite funcional

Hasta aquí hemos construido las herramientas del capítulo una por una. Sabemos decidir si un punto admite aproximaciones dentro de un dominio; formular y negar un límite mediante ε–δ; traducir la aproximación funcional a sucesiones; usar álgebra, orden y sándwich; distinguir límites laterales, infinitos y en infinito; y componer límites sin esconder el caso en que la función interior golpea exactamente el punto excluido por el límite exterior.

Ahora cambia la tarea.

En las seis estaciones que siguen **no introduciremos teoría nueva**. El problema será diagnosticar primero qué tipo de límite tenemos delante, seleccionar la herramienta adecuada y auditar las hipótesis que una solución demasiado rápida podría pasar por alto. Conviene adoptar este protocolo general:

dominio y tipo de aproximación → objetivo de salida → herramienta → prueba → auditoría local

La pregunta ya no será solamente

> ¿cuánto vale este límite?,

sino también

> **¿está legítimamente planteado, qué resultado permite resolverlo y qué hipótesis local hace funcionar ese resultado?**

Estación 1 — La fórmula no decide si el límite está planteado. Consideremos la misma regla $x \longmapsto x^2$ bajo dos dominios distintos:

$$f\colon [0,\infty) \to \mathbb{R}, \qquad f(x) = x^2,$$

y

$$g\colon (\{0\} \cup [1,\infty)) \to \mathbb{R}, \qquad g(x) = x^2.$$

Problema. Decidir si, dentro de la convención de este capítulo, tiene sentido estudiar

$$\lim_{x\to 0} f(x)$$

y

$$\lim_{x\to 0} g(x).$$

No calculemos todavía ningún valor. Primero auditemos el dominio.

Diagnóstico para f. El dominio es $A = [0, \infty)$.

Debemos preguntar si 0 es punto de acumulación de A. Sea $\delta > 0$. Tomemos

$$h := \min\left\{\frac{\delta}{2}, \frac{1}{2}\right\} > 0.$$

Entonces $h \in [0, \infty)$ y

$$0 < |h - 0| = h < \delta.$$

Por tanto, todo entorno perforado de 0 contiene puntos del dominio. Así,

$$0 \text{ es punto de acumulación de } [0, \infty).$$

El límite

$$\lim_{x\to 0} f(x)$$

está legítimamente planteado, aunque la aproximación disponible ocurra solamente desde la derecha.

Diagnóstico para g. Ahora el dominio es $B = \{0\} \cup [1, \infty)$.

Si elegimos $\delta = \frac{1}{2}$, entonces $V^*_{1/2}(0) \cap B = \varnothing$.

No existen puntos del dominio distintos de 0 a distancia menor que $1/2$ del centro. Por tanto, 0 es aislado en B y **no** es punto de acumulación. En este libro no formulamos entonces

$$\lim_{x\to 0} g(x).$$

Lectura de la estación. Las dos funciones utilizan la misma expresión algebraica, pero tienen geometrías locales distintas:

$$\boxed{\text{misma fórmula} \not\Longrightarrow \text{mismos problemas de límite disponibles.}}$$

Antes de transformar una expresión debemos mirar el dominio. En un límite finito, la primera pregunta es si el punto de aproximación es punto de acumulación del dominio. En un límite cuando $x \to \pm\infty$, la pregunta análoga será si el dominio es no acotado en la dirección correspondiente.

Estación 2 — Descubrir $\delta(\varepsilon)$ antes de escribir la prueba. Queremos demostrar directamente que

$$\lim_{x\to 2} x^2 = 4.$$

Problema. Construir una elección explícita de δ a partir de ε y explicar por qué aparece una restricción local adicional.

Diagnóstico. La definición nos pide controlar $|x^2 - 4|$.

Factorizamos: $|x^2 - 4| = |x - 2|\,|x + 2|$.

El primer factor es exactamente el error de entrada. El segundo no es pequeño cuando $x \to 2$; necesitamos únicamente impedir que crezca sin control. Por eso hacemos primero una elección auxiliar. Si imponemos $|x - 2| < 1$, entonces $1 < x < 3$, y de ahí $3 < x + 2 < 5$.

En particular, $|x + 2| < 5$.

Así obtenemos la estimación local $|x^2 - 4| < 5|x - 2|$. Ahora ya podemos diseñar el radio.

Demostración. Sea $\varepsilon > 0$. Elegimos $\delta := \min\left\{1, \frac{\varepsilon}{5}\right\}$.

Sea $x \in \mathbb{R}$ tal que $0 < |x - 2| < \delta$.

Como $\delta \leq 1$, tenemos $|x - 2| < 1$, y por tanto $|x + 2| < 5$.

Además, como $\delta \leq \frac{\varepsilon}{5}$, se cumple $|x - 2| < \frac{\varepsilon}{5}$.

Por consiguiente,

$$|x^2 - 4| = |x - 2|\,|x + 2| < \frac{\varepsilon}{5} \cdot 5 = \varepsilon.$$

Hemos demostrado que

$$\boxed{\lim_{x \to 2} x^2 = 4.}$$

Lectura de la estación. La elección $\delta = \min\left\{1, \frac{\varepsilon}{5}\right\}$ combina dos trabajos diferentes:

- el radio 1 mantiene bajo control el factor $|x + 2|$;
- el radio $\varepsilon/5$ hace suficientemente pequeño el factor $|x - 2|$.

Este patrón reaparece en muchas pruebas:

$$\boxed{\text{restricción local auxiliar} + \text{restricción dependiente de } \varepsilon \to \delta = \min\{\cdots\}.}$$

La búsqueda algebraica ocurre **antes** de la demostración formal. Su función es descubrir qué testigo $\delta(\varepsilon)$ debe elegirse.

Estación 3 — Una sola sucesión puede destruir un límite funcional. Consideremos

$$h\colon \mathbb{R} \setminus \{0\} \to \mathbb{R}, \qquad h(x) = \sigma\left(\frac{1}{x}\right).$$

Problema. Demostrar que

$$\lim_{x \to 0} h(x)$$

no existe como límite real finito, utilizando el criterio secuencial.

Diagnóstico. Una prueba directa a partir de la negación ε–δ sería posible, pero la oscilación sugiere una herramienta más económica. Por Teorema 4.2, si el límite funcional existiera, entonces **toda** sucesión de puntos del dominio que evitara 0 y convergiera a 0 produciría una sucesión de imágenes convergente al mismo número. Por tanto, basta encontrar una sola sucesión admisible cuyas imágenes no converjan.

Construcción de la sucesión testigo. Definamos, para $n \geq 0$, $x_n = \frac{2}{\omega(2n+1)}$.

Todos los términos son positivos y distintos de 0. Además, $x_n = \frac{2}{\omega}\frac{1}{2n+1} \to 0$, porque $1/(2n+1) \to 0$.

Sin embargo, $\frac{1}{x_n} = \frac{\omega(2n+1)}{2} = \frac{\omega}{2} + n\omega$.

Por tanto, $h(x_n) = \sigma\left(\frac{\omega}{2} + n\omega\right) = (-1)^n$.

La sucesión $((-1)^n)$ no converge en $\mathbb{R}$, como ya sabemos desde `T1-C04`. Si $h(x)$ tuviera un límite real cuando $x \to 0$, Teorema 4.2 obligaría a que $(h(x_n))$ convergiera. Como no lo hace, concluimos

$$\boxed{\lim_{x\to 0} \sigma\left(\frac{1}{x}\right) \text{ no existe en } \mathbb{R}.}$$

Lectura de la estación. No hemos necesitado describir todos los caminos posibles hacia 0. Para **probar** un límite, debemos controlar todos los acercamientos; para **refutarlo**, basta encontrar un acercamiento permitido que viole la conducta exigida. El criterio secuencial convierte esa asimetría en una estrategia:

un límite funcional afirma universalidad sobre sucesiones admisibles; una sola sucesión testigo puede refutarlo.

No hemos introducido teoría de subsucesiones: construimos directamente una sucesión concreta de puntos del dominio.

Estación 4 — El álgebra comprime controles locales que no deben olvidarse. Consideremos

$$F(x) = \frac{(x+1)(2x-1)}{x-1}, \qquad x \neq 1.$$

Problema. Calcular

$$\lim_{x\to 2} F(x)$$

mediante el álgebra de límites y, después, identificar los controles locales que justifican el producto y el cociente.

Cálculo comprimido. Por los límites elementales y Proposición 4.3, $x + 1 \to 3$,

$$2x - 1 \to 3,$$

y $x - 1 \to 1$ cuando $x \to 2$.

El límite del denominador es no nulo. Por tanto, la ley del cociente es aplicable y obtenemos

$$\boxed{\lim_{x\to 2} \frac{(x+1)(2x-1)}{x-1} = 9.}$$

La cuenta es breve. Pero el teorema está comprimiendo información local real.

Qué se esconde detrás del producto. Los factores

$$x+1 \qquad \text{y} \qquad 2x-1$$

tienen límites reales finitos. Por Proposición 4.2, ambos quedan acotados en algún entorno perforado de 2. En este ejemplo podemos verlo directamente. Si $|x-2|<\frac{1}{2}$, entonces $\frac{3}{2}<x<\frac{5}{2}$.

Así, $\frac{5}{2}<x+1<\frac{7}{2}$ y $2<2x-1<4$.

Ninguno de los factores puede crecer arbitrariamente dentro de ese entorno. Ese es el tipo de control que permite convertir errores pequeños de los factores en un error pequeño del producto.

Qué se esconde detrás del cociente. La cuestión más delicada es el denominador. Del mismo control $|x-2|<\frac{1}{2}$ obtenemos $\frac{1}{2}<x-1<\frac{3}{2}$.

En particular, $|x-1|>\frac{1}{2}$.

Por tanto, en ese entorno el denominador no solo es distinto de cero: está **separado de cero por una distancia positiva**. Esto es la versión concreta de la separación local de Proposición 4.2 que la demostración general del cociente necesita.

Lectura de la estación. Cuando usamos Proposición 4.3 podemos escribir una prueba muy corta, pero no debemos olvidar qué condiciones hacen posible esa compresión:

producto: límites finitos $\Rightarrow$ acotación local;
cociente: límite no nulo del denominador $\Rightarrow$ separación local de cero.

Aplicar un teorema no significa borrar sus hipótesis; significa que ya hemos demostrado una vez el mecanismo que las convierte en conclusión.

Estación 5 — Clasificar el límite antes de intentar calcularlo. Sea

$$p\colon \mathbb{R}\setminus\{1\}\to\mathbb{R}, \qquad p(x)=\frac{1}{x-1}.$$

Problema. Analizar el comportamiento de p en tres situaciones:

$$x\to 1^+, \qquad x\to 1^-, \qquad x\to+\infty.$$

La dificultad principal no es algebraica. Es reconocer que las tres preguntas pertenecen a tipos diferentes de límite.

Desde la derecha de 1. Afirmamos que

$$\lim_{x\to 1^+}\frac{1}{x-1}=+\infty.$$

Sea $M\in\mathbb{R}$. Definamos $K:=|M|+1>0$ y tomemos $\delta:=\frac{1}{K}$.

Si $1<x<1+\delta$, entonces $0<x-1<\frac{1}{K}$.

Al tomar recíprocos positivos, $\frac{1}{x-1}>K>M$.

Esto prueba el límite lateral infinito positivo.

Desde la izquierda de 1. Afirmamos ahora que

$$\lim_{x\to 1^-}\frac{1}{x-1}=-\infty.$$

Sea nuevamente $M\in\mathbb{R}$, tomemos $K:=|M|+1$ y $\delta:=\frac{1}{K}$.

Si $1-\delta<x<1$, entonces $-\frac{1}{K}<x-1<0$.

Como el denominador es negativo y su magnitud es menor que $1/K$, $\frac{1}{x-1} < -K < M$. Por tanto, el límite lateral izquierdo es $-\infty$. Los dos lados no describen la misma salida. En consecuencia, no existe un límite bilateral real ni un límite bilateral $+\infty$ o $-\infty$ en 1.

Cuando la entrada tiende a $+\infty$. Ahora cambia por completo la geometría de la entrada. El dominio es no acotado superiormente, así que la pregunta está legítimamente planteada.

Afirmamos que

$$\lim_{x\to+\infty} \frac{1}{x-1} = 0.$$

Sea $\varepsilon > 0$. Elegimos $R := 1 + \frac{1}{\varepsilon}$.

Si $x > R$, entonces $x - 1 > \frac{1}{\varepsilon} > 0$, y por tanto

$$\left|\frac{1}{x-1}\right| = \frac{1}{x-1} < \varepsilon.$$

Así queda probado el límite finito en infinito.

Lectura de la estación. Una misma fórmula ha producido tres comportamientos distintos:

$$\boxed{\begin{aligned} \lim_{x\to1^+} \frac{1}{x-1} &= +\infty, \\ \lim_{x\to1^-} \frac{1}{x-1} &= -\infty, \\ \lim_{x\to+\infty} \frac{1}{x-1} &= 0. \end{aligned}}$$

Antes de manipular una expresión debemos identificar qué cuantificadores corresponden al problema:

- entorno perforado bilateral;
- entorno perforado lateral;
- barreras de salida;
- o umbrales de entrada en $\pm\infty$.

Clasificar correctamente el límite es parte de resolverlo.

Estación 6 — Auditar una composición antes de sustituir límites. Definamos

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(y) = \begin{cases} 7, & y = 1, \\ 2, & y \neq 1, \end{cases}$$

y consideremos dos funciones interiores:

$$g\colon \mathbb{R} \to \mathbb{R}, \qquad g(x) = 1 + x^2,$$

y

$$h\colon \mathbb{R} \to \mathbb{R}, \qquad h(x) = 1.$$

Cuando $x \to 0$,

$$g(x) \to 1 \qquad \text{y} \qquad h(x) \to 1.$$

Además,

$$\lim_{y\to 1} f(y) = 2.$$

Problema. Decidir qué ocurre con $f(g(x))$ y $f(h(x))$ cuando $x \to 0$, sin aplicar una regla de sustitución automática.

Primera composición: $f \circ g$. Para todo $x \neq 0$, $x^2 > 0$, y por tanto $g(x) = 1 + x^2 > 1$.

Así, $g(x) \neq 1$ en todo entorno perforado de 0. La hipótesis delicada de Teorema 4.5 se cumple. Como $g(x) \to 1$ y

$$f(y) \to 2 \qquad (y \to 1),$$

concluimos

$$\boxed{\lim_{x\to 0} f(g(x)) = 2.}$$

De hecho, para $x \neq 0$ tenemos directamente $f(g(x)) = 2$.

Obsérvese que $g(0) = 1$ y, en consecuencia, $f(g(0)) = 7$.

Eso no afecta al límite: la condición de no impacto se exige en el **entorno perforado**, no necesariamente en el centro.

Segunda composición: $f \circ h$. Ahora $h(x) = 1$ para todos los x. Por tanto, aunque $h(x) \to 1$, la función interior golpea exactamente el punto problemático en cada punto del entorno: $h(x) = 1$.

Así, $f(h(x)) = f(1) = 7$ para todo x, y entonces

$$\boxed{\lim_{x\to 0} f(h(x)) = 7.}$$

No obtenemos 2.

Lectura de la estación. Los datos

$$g(x) \to 1, \qquad h(x) \to 1, \qquad f(y) \to 2 \quad (y \to 1)$$

no bastan por sí solos para tratar ambas composiciones de la misma manera.

La diferencia es local:

$$\boxed{\begin{array}{c} g(x) \neq 1 \text{ para } x \neq 0 \text{ próximo a } 0, \\ h(x) = 1 \text{ para todo } x. \end{array}}$$

La pregunta correcta antes de componer es la que cerró §4.9:

$$\boxed{\text{¿la salida de la función interior entra en la región perforada donde el límite exterior controla a } f?}$$

Un mapa de decisiones para límites de funciones. Las seis estaciones muestran que «calcular un límite» no designa una única técnica. Antes de hacer álgebra conviene diagnosticar el problema. El mapa mínimo del capítulo es:

¿el acercamiento está legítimamente disponible en el dominio?
↓
¿qué tipo de salida buscamos: real, $+\infty$ o $-\infty$?
↓
¿prueba directa, criterio secuencial, álgebra, orden, sándwich, laterales o infinito?
↓
¿hay una hipótesis local oculta que auditar?

Podemos convertirlo en preguntas operativas.

Si $x \to a$ con a finito, preguntamos primero si a es punto de acumulación del dominio. Si solo hay puntos disponibles por un lado, debemos decidir si corresponde un límite ordinario relativo al dominio o un límite lateral explícito. Si a es aislado, la formulación queda fuera de la convención de este capítulo.

Si $x \to +\infty$ o $x \to -\infty$, comprobamos antes que el dominio sea no acotado en esa dirección. Después identificamos si la salida debe aproximarse a un real o superar/descender toda barrera.

Si nos dan un candidato real L, una prueba directa puede comenzar simplificando $|f(x)-L|$ hasta descubrir qué condición sobre $|x-a|$ basta para hacerlo menor que ε. Las restricciones auxiliares se sincronizan normalmente mediante un mínimo de radios.

Si sospechamos que el límite no existe, el criterio secuencial puede convertir el problema en la construcción de una sucesión admisible que produzca imágenes incompatibles con cualquier límite real. Para refutar un candidato concreto también podemos usar directamente la negación ε–δ.

Si la función está construida a partir de otras con límites conocidos, el álgebra de Proposición 4.3 comprime la prueba. Pero debemos auditar qué queda detrás de esa compresión: acotación local para productos y separación respecto de cero para cocientes.

Si una parte de la expresión oscila pero permanece acotada, conviene buscar una dominación en valor absoluto y utilizar Teorema 4.3 en vez de intentar asignar un límite a cada factor por separado.

Si aparecen lados distintos, calculamos cada comportamiento sobre el dominio restringido correspondiente. Solo cuando el punto acumula por ambos lados podemos aplicar Teorema 4.4 para reconstruir un límite bilateral a partir de dos laterales iguales.

Si queremos componer límites, no sustituimos mecánicamente. Primero verificamos el tipado de la composición y después auditamos si la función interior puede tomar exactamente el punto que el límite exterior excluye. Teorema 4.5 resuelve el caso en que ese impacto queda evitado localmente.

Qué debe poder hacer ahora el lector. Para justificar un límite hay que identificar qué puntos admite el dominio y qué condición debe mantenerse en todos los argumentos suficientemente próximos. Esa decisión determina la definición o el teorema que podemos utilizar:

- cuál es el dominio efectivo y qué puntos de aproximación están disponibles;
- si el problema exige ε–δ, barreras, umbrales en infinito o información lateral;
- si estamos usando una definición o un resultado ya demostrado;
- cuáles son las hipótesis locales delicadas de ese resultado;
- si conviene probar directamente o refutar mediante una sucesión;
- dónde aparece acotación local o separación respecto de cero;
- si una composición transfiere realmente puntos al entorno perforado correcto;

- y qué herramientas posteriores —en particular continuidad— estamos evitando para mantener la cadena demostrativa del capítulo.

El capítulo comenzó traduciendo la cola de una sucesión a un entorno perforado de una función. Después de estas seis estaciones, esa traducción se ha convertido en un repertorio coordinado:

ε–δ,
criterio secuencial,
álgebra y control local,
orden y sándwich,
laterales,
barreras y límites en infinito,
composición auditada.

Los ejercicios combinan estas decisiones. La información del enunciado tendrá que orientar la elección de una prueba directa, una comparación, un argumento secuencial o un análisis de los límites laterales, con las hipótesis que corresponden a cada caso.

4.11. Ejercicios y soluciones

Llegamos al banco final de `T1-C05`. Hasta aquí las técnicas aparecieron primero por separado y después, en §4.10, como estaciones guiadas. Ahora desaparece esa señalización: antes de calcular habrá que decidir si el problema exige analizar el dominio, construir una prueba ε–δ, usar el criterio secuencial, aplicar álgebra u orden, separar límites laterales, trabajar con barreras o auditar una composición. El banco contiene exactamente cuarenta ejercicios, distribuidos en siete niveles:

$$7A + 7B + 7C + 6D + 5E + 5F + 3G = 40.$$

Los niveles no miden solo dificultad algebraica. Los niveles D y E obligan a delimitar afirmaciones falsas mediante diagnósticos y contraejemplos; F reconstruye resultados del capítulo desde sus ideas básicas; G mezcla varias capas y exige decidir primero qué límite está realmente planteado. Todos los problemas pueden resolverse con los capítulos publicados desde los números reales hasta el presente. No se necesitan continuidad, IVT/EVT, compactitud, derivadas, regla de l'Hôpital, series, Bolzano–Weierstrass, criterio de Cauchy ni topología general.

Nivel A — Reconocimiento y comprensión. Ejercicio A1. Acumulación antes de límite. Para cada par (A, a) decide si a es punto de acumulación de A, punto aislado de A o ninguna de las dos cosas:

1. $A = [0, 2]$, $a = 0$;
2. $A = \mathbb{Z}$, $a = 3$;
3. $A = \{1/n : n \geq 1\}$, $a = 0$;
4. $A = \{1/n : n \geq 1\}$, $a = 1/4$;
5. $A = (0, 1)$, $a = 2$.

Indica en cuáles de los cinco casos nuestra convención permite plantear un límite $\lim_{x \to a} f(x)$ para una función con dominio A.

Ejercicio A2. Leer una definición sin invertir cuantificadores. Escribe en palabras la afirmación

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in A: \quad 0 < |x - a| < \delta \Longrightarrow |f(x) - L| < \varepsilon.$$

Después explica por qué cada una de las siguientes lecturas es incorrecta:

1. «existe un $\delta > 0$ que funciona para todos los $\varepsilon > 0$»;
2. «para cada x próximo a a puedo elegir después un δ que sirva para ese x».

Ejercicio A3. Límite y valor puntual. Sea

$$f(x) = \begin{cases} 3x - 1, & x \neq 2, \\ 100, & x = 2. \end{cases}$$

Determina $f(2)$ y $\lim_{x \to 2} f(x)$. Explica por qué no hay contradicción si ambos números son distintos.

Ejercicio A4. Qué sucesiones son admisibles. Para estudiar $\lim_{x \to 0} f(x)$ mediante el criterio secuencial, decide cuáles de las siguientes sucesiones pueden utilizarse directamente como sucesiones de prueba y justifica tu respuesta:

1. $x_n = 1/(n+1)$;
2. $x_n = (-1)^n/(n+1)$;
3. $x_n = 0$ para todo n;
4. $x_n = 1 + 1/(n+1)$;
5. $x_n = 0$ si n es par y $x_n = 1/(n+1)$ si n es impar.

Ejercicio A5. Dos consecuencias locales. Supón que

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

Indica qué puede afirmarse en cada caso y qué hipótesis adicional se necesita, si alguna:

1. f está acotada en algún entorno perforado de a;
2. si $L \neq 0$, entonces $f(x)$ permanece separado de 0 en algún entorno perforado de a;
3. si $L = 0$, entonces $f(x)$ permanece separado de 0 en algún entorno perforado de a.

Ejercicio A6. Clasificar antes de calcular. Para

$$f(x) = \frac{1}{x - 2}, \qquad x \neq 2,$$

clasifica —sin intentar tratarlos como un único límite— los comportamientos cuando

$$x \to 2^-, \qquad x \to 2^+, \qquad x \to +\infty, \qquad x \to -\infty.$$

Indica en cada caso si la salida es real, $+\infty$ o $-\infty$.

Ejercicio A7. El dominio también importa en infinito. Sean

$$A = (-\infty, 5], \qquad B = [5, \infty).$$

Para una función $f : A \to \mathbb{R}$, ¿tiene sentido bajo la convención de este capítulo estudiar $x \to +\infty$? ¿Y $x \to -\infty$? Responde las mismas preguntas para una función $g : B \to \mathbb{R}$.

Nivel B — Aplicación directa. Ejercicio B1. Álgebra de límites. Calcula

$$\lim_{x \to 2} \frac{x^2 + 3x - 1}{2x + 5}$$

utilizando únicamente el álgebra de límites funcionales. Audita explícitamente la hipótesis del denominador.

Ejercicio B2. Oscilación amortiguada. Demuestra que

$$\lim_{x \to 0} x^3 \kappa\left(\frac{1}{x}\right) = 0.$$

No intentes asignar un límite separado a $\kappa(1/x)$.

Ejercicio B3. Una función por tramos. Sea

$$f(x) = \begin{cases} 2x + 1, & x < 1, \\ x^2 + 2, & x > 1. \end{cases}$$

Calcula los límites laterales en $x = 1$ y decide si existe el límite bilateral.

Ejercicio B4. Una barrera arbitraria. Demuestra desde la definición de límite infinito que

$$\frac{1}{(x-3)^2} \to +\infty \qquad (x \to 3).$$

Da una elección explícita de δ en función de una barrera real M.

Ejercicio B5. Cociente al infinito. Demuestra que

$$\lim_{x \to +\infty} \frac{5x - 4}{2x + 1} = \frac{5}{2}.$$

Tu argumento debe producir un umbral $R = R(\varepsilon)$.

Ejercicio B6. Composición segura. Sea $g(x) = 1 + x^2$ y sea $f : (1, \infty) \to \mathbb{R}$ dada por

$$f(y) = \frac{y^2 - 1}{y - 1}.$$

Calcula

$$\lim_{x \to 0} f(g(x))$$

auditando las hipótesis de la composición.

Ejercicio B7. Un límite que dos sucesiones destruyen. Para

$$f(x) = \frac{|x|}{x}, \qquad x \neq 0,$$

usa dos sucesiones explícitas para demostrar que $\lim_{x\to 0} f(x)$ no existe como límite real.

Nivel C — Combinación estructural. Ejercicio C1. Descubrir un radio. Demuestra directamente desde la definición ε–δ que

$$\lim_{x\to 3} x^2 = 9.$$

La prueba debe explicar de dónde sale una elección del tipo $\delta = \min\{1, \frac{\varepsilon}{7}\}$.

Ejercicio C2. Cociente con control local. Demuestra directamente que

$$\lim_{x\to 1} \frac{x+1}{x+2} = \frac{2}{3}.$$

No basta citar la ley del cociente: encuentra una cota local para el denominador y produce un $\delta(\varepsilon)$ explícito.

Ejercicio C3. Qué se esconde en la ley del producto. Supón

$$\lim_{x\to a} f(x) = 2, \qquad \lim_{x\to a} g(x) = -3.$$

Demuestra directamente que

$$\lim_{x\to a} f(x)g(x) = -6$$

partiendo de la identidad

$$f(x)g(x) + 6 = f(x)(g(x) + 3) - 3(f(x) - 2).$$

Tu prueba debe indicar dónde se usa la acotación local de f.

Ejercicio C4. Criterio secuencial más sándwich. Sea $f(x) = \frac{|x|}{1+x^2}$.
Usa el criterio secuencial para demostrar que

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

En la prueba secuencial utiliza una estimación sencilla en lugar de calcular un cociente de límites.

Ejercicio C5. Composición con valor puntual adverso. Define $f : \mathbb{R} \setminus \{-1\} \to \mathbb{R}$ por

$$f(y) = \begin{cases} 4, & y = 0, \\ \dfrac{1-y}{1+y}, & y \neq 0, \end{cases}$$

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

1. Demuestra que $\lim_{y\to 0} f(y) = 1$.

2. Verifica la hipótesis de no impacto para g cuando $x \to 0$.
3. Deduce $\lim_{x\to 0} f(g(x))$.
4. Compara el resultado con $f(g(0))$.

Ejercicio C6. Un mismo cociente, cuatro regímenes. Para

$$f(x) = \frac{x+1}{x-1}, \qquad x \neq 1,$$

determina

$$\lim_{x\to 1^-} f(x), \qquad \lim_{x\to 1^+} f(x), \qquad \lim_{x\to +\infty} f(x), \qquad \lim_{x\to -\infty} f(x).$$

Explica por qué no existe un límite bilateral en $x = 1$.

Ejercicio C7. Un dominio discreto con límite infinito. Sea

$$A = \left\{ \frac{1}{n} : n \in \mathbb{N}_{>0} \right\}$$

y define

$$f : A \to \mathbb{R}, \qquad f(x) = \frac{1}{x}.$$

Demuestra que 0 es punto de acumulación de A y que

$$f(x) \to +\infty \qquad (x \to 0,\ x \in A).$$

La prueba del límite debe formularse con una barrera real M y un radio δ.

Nivel D — Inversión y diagnóstico. Ejercicio D1. Evaluar no es tomar límite. Decide si es verdadera o falsa la afirmación:

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

Si es falsa, da un contraejemplo donde a sea punto de acumulación del dominio y explica exactamente qué información falta.

Ejercicio D2. El límite no exige valor puntual. Decide si es verdadera o falsa:

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

Si es falsa, construye un contraejemplo cuya función ni siquiera esté definida en a.

Ejercicio D3. Acotación no implica convergencia local. Decide si es verdadera o falsa:

Si f está acotada en algún entorno perforado de a, entonces $\lim_{x\to a} f(x)$ existe.

Conserva la acotación en tu contraejemplo.

Ejercicio D4. Una tabla no es una prueba. Un estudiante calcula

$$f(0.9),\ f(0.99),\ f(0.999),\ f(1.001),\ f(1.01),\ f(1.1)$$

y observa que todos esos valores están muy cerca de 5. Concluye que

$$\lim_{x\to 1} f(x) = 5.$$

Explica por qué la inferencia es inválida y construye una función que coincida con el valor 5 en todos esos puntos, pero cuyo límite en 1 no sea 5.

Ejercicio D5. Componer no es sustituir automáticamente. Decide si es verdadera o falsa:

> Si $g(x) \to b$ cuando $x \to a$ y $f(y) \to L$ cuando $y \to b$, entonces siempre $f(g(x)) \to L$.

Si es falsa, conserva ambos límites de las hipótesis y haz fallar solo la conclusión.

Ejercicio D6. Fallar bilateralmente no destruye los laterales. Decide si es verdadera o falsa:

> Si $\lim_{x\to a} f(x)$ no existe, entonces ninguno de los dos límites laterales existe.

Da un contraejemplo en el que ambos límites laterales existan pero sean distintos.

Nivel E — Contraejemplos. Ejercicio E1. Tres valores puntuales, un mismo límite. Construye tres funciones $f, g, h : \mathbb{R} \to \mathbb{R}$ que coincidan para todo $x \neq 0$, tengan valores distintos en 0 y satisfagan

$$\lim_{x\to 0} f(x) = \lim_{x\to 0} g(x) = \lim_{x\to 0} h(x).$$

Verifica explícitamente el límite común.

Ejercicio E2. Oscilación sin límite. Construye una función acotada en $\mathbb{R} \setminus \{0\}$ que oscile infinitamente al acercarse a 0 y no tenga límite real allí. Demuestra la inexistencia mediante dos sucesiones.

Ejercicio E3. Dos laterales incompatibles. Construye una función $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ tal que

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

Verifica ambos límites directamente y explica por qué no existe el bilateral.

Ejercicio E4. Acotada en cada escala, sin límite. Construye una función $f : \mathbb{R} \to \mathbb{R}$ que satisfaga $0 \leq f(x) \leq 1$ para todo x, pero no tenga límite en ningún punto real. Puedes utilizar la densidad de $\mathbb{Q}$ y de $\mathbb{R} \setminus \mathbb{Q}$ demostrada en `T1-C02`.

Ejercicio E5. Impacto exacto que rompe una composición. Construye funciones $g : \mathbb{R} \to \mathbb{R}$ y $f : \mathbb{R} \to \mathbb{R}$ y números a, b, L tales que

$$g(x) \to b \quad (x \to a), \qquad f(y) \to L \quad (y \to b),$$

pero

$$\lim_{x \to a} f(g(x)) \neq L.$$

Tu ejemplo debe fallar precisamente porque $g(x) = b$ para puntos arbitrariamente próximos a a.

Nivel F — Descubrimiento guiado. Ejercicio F1. Reconstruir la negación ε–δ. Parte de

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x \in A : \quad 0 < |x - a| < \delta \Longrightarrow |f(x) - L| < \varepsilon.$$

Niega la afirmación paso a paso hasta obtener una formulación equivalente que comience con $\exists \varepsilon_0 > 0$. Explica por qué el último cuantificador debe producir un punto testigo x que depende del radio elegido.

Ejercicio F2. De una negación a una sucesión testigo. Supón que a es punto de acumulación de A y que

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

1. Usa la negación ε–δ para fijar un $\varepsilon_0 > 0$.
2. Para cada n, toma $\delta_n = 1/(n+1)$ y elige un punto $x_n \in A$ que viole la condición de límite.
3. Demuestra que $x_n \to a$, que $x_n \neq a$ y que $f(x_n) \not\to L$.

Así reconstruirás el mecanismo central de la dirección contrapositiva del criterio secuencial.

Ejercicio F3. Descubrir la acotación local. Supón

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

1. Aplica la definición con $\varepsilon = 1$.
2. Demuestra que existe $\delta > 0$ tal que $|f(x)| < |L| + 1$ siempre que $0 < |x - a| < \delta$.
3. Explica por qué esto prueba acotación local perforada, pero no necesariamente acotación global.

Ejercicio F4. Reconstruir la ley del producto. Sean

$$\lim_{x \to a} f(x) = L, \qquad \lim_{x \to a} g(x) = M.$$

1. Reescribe $f(x)g(x) - LM$ como suma de dos términos donde aparezcan $g(x) - M$ y $f(x) - L$.
2. Usa la acotación local de una de las funciones para controlar el término mixto.
3. Sincroniza los radios con un mínimo.

4. Concluye directamente desde ε–δ que $f(x)g(x) \to LM$.

Ejercicio F5. Reconstruir el criterio bilateral. Supón que a es punto de acumulación del dominio tanto por la izquierda como por la derecha.

1. Demuestra que si $\lim_{x\to a} f(x) = L$, entonces ambos límites laterales valen L.
2. Supón ahora que ambos límites laterales valen L. Dados los radios laterales δ_- y δ_+ asociados a una misma tolerancia ε, elige un radio bilateral que funcione.
3. Concluye la equivalencia completa.

Nivel G — Desafíos. Ejercicio G1. Un límite con radical y agujero. Define, para $x > -3$ y $x \neq 1$, $f(x) = \frac{\sqrt{x+3}-2}{x-1}$.

Demuestra directamente desde ε–δ que

$$\lim_{x\to 1} f(x) = \frac{1}{4}.$$

No invoques continuidad de la raíz. La racionalización debe convertirse en una estimación cuantitativa que produzca un radio explícito.

Ejercicio G2. La oscilación sobrevive, el producto no. Define

$$f(x) = x\sigma\left(\frac{1}{x^2}\right), \qquad x \neq 0.$$

1. Demuestra mediante el criterio secuencial que $f(x) \to 0$ cuando $x \to 0$.
2. Construye dos sucesiones admisibles que demuestren que $\sigma(1/x^2)$, considerada por sí sola, no tiene límite cuando $x \to 0$.
3. Explica por qué no existe contradicción entre 1 y 2.

Ejercicio G3. Dominio irregular, laterales y composición. Sea

$$A = (-1, 0) \cup \left\{ \frac{1}{n} : n \in \mathbb{N}_{>0} \right\}.$$

Define $h : A \to [0, \infty)$ por

$$h(x) = \begin{cases} x^2, & -1 < x < 0, \\ 0, & x = 1/n \text{ para algún } n \geq 1, \end{cases}$$

y define $f : [0, \infty) \to \mathbb{R}$ por

$$f(y) = \begin{cases} 7, & y = 0, \\ \dfrac{1}{1+y}, & y > 0. \end{cases}$$

1. Demuestra que 0 es punto de acumulación de A por ambos lados.
2. Demuestra que $h(x) \to 0$ cuando $x \to 0$ dentro de A.
3. Demuestra que $f(y) \to 1$ cuando $y \to 0$ dentro de $[0, \infty)$.
4. Calcula los límites laterales de $f(h(x))$ cuando $x \to 0$ dentro de A.

5. Decide si existe el límite bilateral y explica exactamente qué hipótesis de Teorema 4.5 falla.

Soluciones. Las soluciones siguen exactamente el orden de los ejercicios. En A–C se explicita la herramienta utilizada; en D–E se verifica que los contraejemplos conserven las hipótesis pertinentes; en F–G se hace visible la arquitectura completa de la prueba y la frontera de no circularidad.

Soluciones del nivel A. **Solución A1.**

1. Para $A = [0, 2]$ y $a = 0$, todo entorno perforado de 0 contiene puntos de A a la derecha. Por ejemplo, dado $\delta > 0$ podemos tomar
$$x = \min\left\{\frac{\delta}{2}, 1\right\} > 0.$$
Entonces $x \in [0, 2]$ y $0 < |x| < \delta$. Por tanto, 0 es punto de acumulación.
2. Para $A = \mathbb{Z}$ y $a = 3$, el radio $\delta = 1/2$ satisface $V_{1/2}(3) \cap \mathbb{Z} = \{3\}$. Así, 3 es punto aislado.
3. Para $A = \{1/n : n \geq 1\}$ y $a = 0$, dado $\delta > 0$ la propiedad arquimediana permite elegir $n > 1/\delta$. Entonces $0 < \frac{1}{n} < \delta$. Por tanto, 0 es punto de acumulación aunque $0 \notin A$.
4. El punto $1/4$ pertenece a A y es aislado. Sus vecinos en la lista son $1/3$ y $1/5$; basta tomar un radio menor que la mitad de la menor de las distancias
$$\frac{1}{3} - \frac{1}{4}, \qquad \frac{1}{4} - \frac{1}{5}.$$
5. Para $A = (0, 1)$ y $a = 2$, el punto no pertenece a A y tampoco es de acumulación. Por ejemplo, $V_{1/2}(2) = (3/2, 5/2)$ no contiene puntos de $(0, 1)$.

Nuestra convención permite plantear $\lim_{x\to a} f(x)$ precisamente en los casos 1 y 3, porque allí a es punto de acumulación del dominio.

Solución A2.

La afirmación dice:

> dada cualquier tolerancia positiva ε alrededor de L, existe un radio positivo δ alrededor de a tal que todo punto x del dominio, distinto de a y a distancia menor que δ de a, produce un valor $f(x)$ a distancia menor que ε de L.

El orden lógico es

$$\varepsilon \longrightarrow \delta(\varepsilon) \longrightarrow \text{todos los } x \text{ del entorno perforado.}$$

La lectura 1 intercambia los dos primeros cuantificadores. Exigiría $\exists \delta > 0\ \forall \varepsilon > 0$, que es mucho más fuerte: un mismo radio tendría que producir una precisión de salida arbitrariamente pequeña. La lectura 2 altera el último tramo. Si observáramos primero un punto x y eligiéramos después un δ adaptado a él, estaríamos usando una dependencia de la forma $\delta = \delta(\varepsilon, x)$, pero la definición exige que δ quede fijado antes de cuantificar universalmente sobre x.

Solución A3.
Por definición puntual, $f(2) = 100$.
Sin embargo, para $x \neq 2$, $f(x) = 3x - 1$.
Por el límite de una función lineal obtenido desde el álgebra de límites,

$$\lim_{x \to 2} f(x) = 3 \cdot 2 - 1 = 5.$$

No hay contradicción porque la definición del límite usa $0 < |x - 2| < \delta$, y excluye exactamente el punto $x = 2$. El valor puntual puede modificarse sin alterar el comportamiento en el entorno perforado.

Solución A4.
El criterio secuencial para $x \to 0$ exige dos condiciones:

$$x_n \to 0 \qquad \text{y} \qquad x_n \neq 0 \text{ para todo } n.$$

1. $x_n = 1/(n+1)$: **admisible**. Converge a 0 y nunca vale 0.
2. $x_n = (-1)^n/(n+1)$: **admisible**. Su valor absoluto es $1/(n+1) \to 0$ y ningún término es 0.
3. $x_n = 0$: **no admisible**. Converge a 0, pero toca el centro en todos los índices.
4. $x_n = 1 + 1/(n+1)$: **no admisible** para este límite. Converge a 1, no a 0.
5. La sucesión que vale 0 en índices pares y $1/(n+1)$ en los impares converge a 0, pero **no es admisible directamente** porque contiene infinitos términos iguales al centro. El criterio exige que la sucesión usada evite 0.

Solución A5.

1. Es verdadero sin hipótesis adicional: todo límite real finito implica acotación local en algún entorno perforado. Aplicando la definición con $\varepsilon = 1$ obtenemos $|f(x) - L| < 1$, y entonces
$$|f(x)| \leq |f(x) - L| + |L| < |L| + 1.$$
2. También es verdadero, pero requiere $L \neq 0$. Tomando $\varepsilon = \frac{|L|}{2}$ obtenemos localmente $|f(x) - L| < \frac{|L|}{2}$. Por la desigualdad triangular inversa,
$$|f(x)| \geq |L| - |f(x) - L| > \frac{|L|}{2} > 0.$$
3. Es falso en general. Si $L = 0$, la función puede acercarse a 0 y tomar valores arbitrariamente pequeños. Por ejemplo, $f(x) = x$ satisface $f(x) \to 0$ cuando $x \to 0$, pero no puede permanecer separada de 0 por una cota positiva fija en ningún entorno perforado de 0.

Solución A6.
Si $x < 2$, entonces $x - 2 < 0$ y su magnitud tiende a 0 cuando $x \to 2^-$. Por tanto,

$$\frac{1}{x - 2} \to -\infty \qquad (x \to 2^-).$$

Si $x > 2$, el denominador es positivo y tiende a 0, así que

$$\frac{1}{x-2} \to +\infty \qquad (x \to 2^+).$$

En cambio, cuando $|x|$ crece, $\frac{1}{x-2}$ tiende al número real 0. Más precisamente,

$$\lim_{x\to+\infty} \frac{1}{x-2} = 0, \qquad \lim_{x\to-\infty} \frac{1}{x-2} = 0.$$

El primer símbolo de infinito describe la dirección de la **entrada**; $\pm\infty$ en la salida describe escape de los **valores**. Son papeles distintos.

Solución A7.

El conjunto $A = (-\infty, 5]$ es no acotado hacia $-\infty$, pero sí está acotado superiormente. Por tanto, para $f : A \to \mathbb{R}$ tiene sentido estudiar $x \to -\infty$, pero no $x \to +\infty$ bajo la convención del capítulo. Análogamente, $B = [5, \infty)$ es no acotado hacia $+\infty$ y acotado inferiormente. Para $g : B \to \mathbb{R}$ tiene sentido estudiar $x \to +\infty$, pero no $x \to -\infty$. La condición de no acotación en la dirección pertinente cumple aquí el mismo papel lógico que la acumulación en un punto finito: garantiza que haya entradas disponibles arbitrariamente lejos.

Soluciones del nivel B. **Solución B1.**

Por el álgebra de límites, $x^2+3x-1 \to 2^2+3\cdot 2-1 = 9$ cuando $x \to 2$, y $2x+5 \to 2\cdot 2+5 = 9$. El límite del denominador es distinto de cero. Por Proposición 4.2, el denominador queda separado de 0 en algún entorno perforado de 2, de modo que la ley del cociente de Proposición 4.3 es aplicable. Entonces

$$\boxed{\lim_{x\to 2} \frac{x^2+3x-1}{2x+5} = 1.}$$

La auditoría del denominador no consiste en observar solamente que $2x + 5 \neq 0$ en $x = 2$; lo relevante es que su **límite** es $9 \neq 0$, lo cual garantiza control local.

Solución B2.

Para todo $x \neq 0$,

$$\left| x^3 \kappa\left(\frac{1}{x}\right) \right| \leq |x|^3,$$

porque $|\kappa t| \leq 1$.

Además,

$$|x|^3 \to 0 \qquad (x \to 0).$$

Por la forma absoluta del teorema del sándwich,

$$\boxed{\lim_{x\to 0} x^3 \kappa\left(\frac{1}{x}\right) = 0.}$$

No necesitamos que $\kappa(1/x)$ tenga límite. Basta que permanezca acotado mientras el factor x^3 aplasta la oscilación.

Solución B3.
Por la izquierda de 1 se usa la primera rama:

$$\lim_{x\to 1^-} f(x) = \lim_{x\to 1^-} (2x+1) = 3.$$

Por la derecha se usa la segunda:

$$\lim_{x\to 1^+} f(x) = \lim_{x\to 1^+} (x^2+2) = 3.$$

Los dos límites laterales existen y coinciden. Como el dominio acumula por ambos lados de 1, Teorema 4.4 da

$$\boxed{\lim_{x\to 1} f(x) = 3.}$$

La función ni siquiera necesita estar definida en $x = 1$ para que esta conclusión sea válida.

Solución B4.
Debemos probar que para toda barrera real M existe $\delta > 0$ tal que

$$0 < |x-3| < \delta \quad \Longrightarrow \quad \frac{1}{(x-3)^2} > M.$$

Tomemos $K = |M| + 1 > 0$ y elijamos $\delta = \frac{1}{\sqrt{K}}$.
Si $0 < |x-3| < \delta$, entonces $0 < (x-3)^2 < \frac{1}{K}$.
Al tomar recíprocos positivos,

$$\frac{1}{(x-3)^2} > K = |M| + 1 > M.$$

Como M era arbitrario,

$$\boxed{\frac{1}{(x-3)^2} \to +\infty \quad (x \to 3).}$$

Solución B5.
Restamos el candidato:

$$\frac{5x-4}{2x+1} - \frac{5}{2} = \frac{2(5x-4) - 5(2x+1)}{2(2x+1)} = -\frac{13}{2(2x+1)}.$$

Así,

$$\left|\frac{5x-4}{2x+1} - \frac{5}{2}\right| = \frac{13}{2|2x+1|}.$$

Para $x \geq 0$ tenemos $2x+1 > 2x$, y por tanto $\frac{13}{2(2x+1)} < \frac{13}{4x}$ si $x > 0$.
Dado $\varepsilon > 0$, elegimos

$$R = \text{máx}\left\{1, \frac{13}{4\varepsilon}\right\}.$$

Si $x > R$, entonces

$$\left|\frac{5x-4}{2x+1}-\frac{5}{2}\right| < \frac{13}{4x} < \varepsilon.$$

Por definición,

$$\boxed{\lim_{x\to+\infty}\frac{5x-4}{2x+1}=\frac{5}{2}.}$$

Solución B6.
Primero,

$$g(x) = 1 + x^2 \to 1 \qquad (x \to 0).$$

El dominio exterior es $(1,\infty)$. Para $y > 1$,

$$f(y) = \frac{y^2-1}{y-1} = y + 1,$$

de modo que

$$\lim_{y\to 1} f(y) = 2$$

dentro del dominio $(1,\infty)$.

La condición delicada es automática: si $x \neq 0$, entonces $g(x) = 1 + x^2 > 1$, y por tanto $g(x) \neq 1$. Además $g(x) \in (1,\infty)$ para $x \neq 0$, así que la composición está correctamente tipada en un entorno perforado de 0. Aplicando Teorema 4.5,

$$\boxed{\lim_{x\to 0} f(g(x)) = 2.}$$

De hecho, para $x \neq 0$, $f(g(x)) = g(x) + 1 = 2 + x^2$, lo que confirma el resultado directamente.

Solución B7.
Tomemos

$$x_n = \frac{1}{n+1}, \qquad y_n = -\frac{1}{n+1}.$$

Ambas sucesiones evitan 0 y satisfacen

$$x_n \to 0, \qquad y_n \to 0.$$

Sin embargo, $f(x_n) = 1$ para todo n, mientras que $f(y_n) = -1$ para todo n. Si existiera un límite real L de $f(x)$ cuando $x \to 0$, el criterio secuencial obligaría simultáneamente a

$$1 \to L \qquad \text{y} \qquad -1 \to L,$$

lo que contradice la unicidad del límite secuencial. Por tanto,

$$\boxed{\lim_{x\to 0}\frac{|x|}{x} \text{ no existe.}}$$

Soluciones del nivel C. **Solución C1.**

Queremos hacer pequeño $|x^2 - 9|$.

Factorizamos: $|x^2 - 9| = |x - 3|\,|x + 3|$.

El factor $|x - 3|$ es precisamente el que controla la definición. Debemos acotar localmente $|x + 3|$. Si imponemos primero $|x - 3| < 1$, entonces $2 < x < 4$, y por tanto $|x + 3| < 7$. Así, bajo esa restricción auxiliar, $|x^2 - 9| < 7|x - 3|$.

Dado $\varepsilon > 0$, elegimos $\delta = \min\left\{1, \frac{\varepsilon}{7}\right\}$.

Si $0 < |x - 3| < \delta$, entonces $|x - 3| < 1$ y

$$|x^2 - 9| < 7|x - 3| < 7\delta \leq \varepsilon.$$

Por tanto,

$$\boxed{\lim_{x\to 3} x^2 = 9.}$$

El mínimo sincroniza dos necesidades distintas: mantener x en una región donde $|x + 3|$ esté controlado y alcanzar la precisión ε.

Solución C2.

Restamos el candidato: $\frac{x+1}{x+2} - \frac{2}{3} = \frac{3x+3-2x-4}{3(x+2)} = \frac{x-1}{3(x+2)}$.

Por tanto,

$$\left|\frac{x+1}{x+2} - \frac{2}{3}\right| = \frac{|x-1|}{3|x+2|}.$$

Necesitamos impedir que el denominador sea pequeño. Si $|x - 1| < 1$, entonces $0 < x < 2$, y en particular $x + 2 > 2$.

Así, $3|x + 2| > 6$ y obtenemos

$$\left|\frac{x+1}{x+2} - \frac{2}{3}\right| < \frac{|x-1|}{6}.$$

Dado $\varepsilon > 0$, tomemos $\delta = \min\{1, 6\varepsilon\}$.

Entonces $0 < |x - 1| < \delta$ implica

$$\left|\frac{x+1}{x+2} - \frac{2}{3}\right| < \frac{|x-1|}{6} < \frac{\delta}{6} \leq \varepsilon.$$

Luego

$$\boxed{\lim_{x\to 1} \frac{x+1}{x+2} = \frac{2}{3}.}$$

Solución C3.

Sea $\varepsilon > 0$. Como $f(x) \to 2$, la acotación local garantiza que existe un radio $r > 0$ y una constante, por ejemplo $B > 0$, tales que

$$0 < |x - a| < r \quad \Longrightarrow \quad |f(x)| \leq B.$$

Podemos obtener concretamente $B = 3$ aplicando la definición con tolerancia 1: localmente

$$|f(x)-2|<1 \quad \Longrightarrow \quad |f(x)|<3.$$

Usamos ahora la identidad dada:

$$f(x)g(x)+6 = f(x)(g(x)+3)-3(f(x)-2).$$

Por desigualdad triangular,

$$|f(x)g(x)+6| \le |f(x)|\,|g(x)+3| + 3|f(x)-2|.$$

Buscamos que cada término sea menor que $\varepsilon/2$. Del límite $g(x)\to -3$ obtenemos un radio $\delta_g>0$ tal que

$$0<|x-a|<\delta_g \quad \Longrightarrow \quad |g(x)+3|<\frac{\varepsilon}{2B}.$$

Del límite $f(x)\to 2$ obtenemos un radio $\delta_f>0$ tal que

$$0<|x-a|<\delta_f \quad \Longrightarrow \quad |f(x)-2|<\frac{\varepsilon}{6}.$$

Tomamos $\delta = \min\{r,\delta_g,\delta_f\}$.
Entonces $|f(x)g(x)+6| < B\frac{\varepsilon}{2B} + 3\frac{\varepsilon}{6} = \varepsilon$.
Por tanto,

$$\boxed{f(x)g(x)\to -6.}$$

El punto delicado es la presencia de $|f(x)|$: antes de poder controlar el producto debemos saber que ese factor no crece sin control cerca de a.

Solución C4.
Sea (x_n) cualquier sucesión admisible para el criterio secuencial:

$$x_n\neq 0, \qquad x_n\to 0.$$

Entonces

$$0\le \frac{|x_n|}{1+x_n^2} \le |x_n|,$$

porque $1+x_n^2\ge 1$.
Como $|x_n|\to 0$, el teorema del sándwich para sucesiones da

$$\frac{|x_n|}{1+x_n^2}\to 0.$$

Esto vale para toda sucesión admisible. Por Teorema 4.2,

$$\boxed{\lim_{x\to 0}\frac{|x|}{1+x^2}=0.}$$

La estimación evita cualquier necesidad de tratar el cociente mediante una ley más fuerte de lo necesario.

Solución C5.
Para $y \neq 0$, $f(y) = \frac{1-y}{1+y}$.
Por el álgebra de límites,

$$\lim_{y\to 0} f(y) = \frac{1-0}{1+0} = 1,$$

pues el límite del denominador es $1 \neq 0$.
Ahora

$$g(x) = x^2 \to 0 \qquad (x \to 0).$$

Además, si $x \neq 0$, $g(x) = x^2 > 0$, de modo que $g(x) \neq 0$ en todo entorno perforado de 0. Se cumple la condición de no impacto de Teorema 4.5. Por consiguiente,

$$\boxed{\lim_{x\to 0} f(g(x)) = 1.}$$

En cambio, $f(g(0)) = f(0) = 4$.
La discrepancia es perfectamente compatible con la teoría: el límite exterior controla valores de $f(y)$ para $y \neq 0$ próximos a 0, y la función interior evita 0 cuando $x \neq 0$.

Solución C6.
Cerca de $x = 1$ el numerador satisface $x + 1 \to 2 > 0$. Por la izquierda, $x - 1 \to 0^-$, así que

$$\boxed{\frac{x+1}{x-1} \to -\infty \qquad (x \to 1^-).}$$

Por la derecha, $x - 1 \to 0^+$, y por tanto

$$\boxed{\frac{x+1}{x-1} \to +\infty \qquad (x \to 1^+).}$$

Como los comportamientos laterales son incompatibles, no existe límite bilateral real ni un único límite bilateral $+\infty$ o $-\infty$ en $x = 1$. Para $x \to \pm\infty$, dividimos numerador y denominador por x:

$$\frac{x+1}{x-1} = \frac{1+1/x}{1-1/x}.$$

Como $1/x \to 0$ tanto para $x \to +\infty$ como para $x \to -\infty$,

$$\boxed{\lim_{x\to+\infty} \frac{x+1}{x-1} = 1, \qquad \lim_{x\to-\infty} \frac{x+1}{x-1} = 1.}$$

Solución C7.
Primero probamos acumulación. Dado $\delta > 0$, por la propiedad arquimediana existe n tal que $n > \frac{1}{\delta}$. Entonces $0 < \frac{1}{n} < \delta$, y $1/n \in A$. Por tanto, 0 es punto de acumulación de A. Ahora queremos probar que $f(x) = \frac{1}{x} \to +\infty$ cuando $x \to 0$ dentro de A. Sea $M \in \mathbb{R}$ una barrera arbitraria. Tomemos $K = |M| + 1 > 0$ y elijamos $\delta = \frac{1}{K}$.
Si $x \in A$ y $0 < |x| < \delta$, entonces $x > 0$ y $0 < x < \frac{1}{K}$.

Al tomar recíprocos positivos,

$$f(x) = \frac{1}{x} > K > |M| \geq M.$$

Así,

$$\boxed{f(x) \to +\infty \qquad (x \to 0,\ x \in A).}$$

La discreción del dominio no impide el límite: lo esencial es que el dominio tenga puntos arbitrariamente próximos a 0.

Soluciones del nivel D. **Solución D1.**

La afirmación es falsa. Definamos

$$f(x) = \begin{cases} 0, & x = 0, \\ 1, & x \neq 0. \end{cases}$$

Tomamos $a = 0$ y $L = 0$. Entonces $f(0) = 0 = L$.

Sin embargo, para todo $x \neq 0$, $f(x) = 1$, y por tanto

$$\lim_{x \to 0} f(x) = 1 \neq 0.$$

El punto 0 es punto de acumulación de $\mathbb{R}$, así que el límite está legítimamente planteado. Lo que falta en la afirmación es información sobre los valores de $f(x)$ en un **entorno perforado** de a. El dato $f(a) = L$ controla un solo punto; el límite controla todos los puntos suficientemente próximos distintos del centro.

Solución D2.

La afirmación es falsa. Consideremos

$$f : \mathbb{R} \setminus \{0\} \to \mathbb{R}, \qquad f(x) = x.$$

El punto 0 es punto de acumulación del dominio y

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

De hecho, $|f(x) - 0| = |x|$, de modo que basta elegir $\delta = \varepsilon$.

Pero $f(0)$ no existe porque $0 \notin \mathrm{Dom}(f)$.

La definición de límite no exige que el centro pertenezca al dominio; exige que sea punto de acumulación de él.

Solución D3.

La afirmación es falsa. Tomemos

$$f(x) = \frac{|x|}{x}, \qquad x \neq 0.$$

La función está acotada globalmente, pues $|f(x)| = 1$ para todo $x \neq 0$. En particular, está acotada en cualquier entorno perforado de 0.

Sin embargo,

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

Los laterales son distintos, así que el límite bilateral no existe. La teoría demuestra la implicación

$$\text{límite real finito} \Longrightarrow \text{acotación local},$$

pero no su recíproca.

Solución D4.

Una cantidad finita de observaciones no controla todos los puntos de ningún entorno perforado. Construyamos una función que engañe exactamente a esa tabla.

Sea

$$S = \{0.9, 0.99, 0.999, 1.001, 1.01, 1.1\}$$

y definamos

$$f(x) = \begin{cases} 5, & x \in S, \\ 0, & x \notin S. \end{cases}$$

Los seis valores observados por el estudiante son exactamente 5. No obstante, el punto de S más cercano a 1 está a distancia 0.001. Si elegimos, por ejemplo, $\delta = \frac{1}{2000} = 0.0005$, entonces $0 < |x-1| < \delta$ implica $x \notin S$, y por tanto $f(x) = 0$.

Así,

$$\boxed{\lim_{x\to 1} f(x) = 0,}$$

no 5.

Una tabla puede sugerir una conjetura, pero una prueba de límite necesita controlar **todos** los puntos del dominio suficientemente próximos al centro.

Solución D5.

La afirmación es falsa sin una hipótesis adicional. Definamos $g(x) = 0$ para todo x y

$$f(y) = \begin{cases} 1, & y = 0, \\ 0, & y \neq 0. \end{cases}$$

Tomemos $a = 0$, $b = 0$ y $L = 0$. Entonces

$$g(x) \to 0 \qquad (x \to 0),$$

y

$$f(y) \to 0 \qquad (y \to 0),$$

porque el límite exterior ignora el valor puntual $f(0) = 1$. Pero $f(g(x)) = f(0) = 1$ para todo x. Por tanto,

$$\lim_{x\to 0} f(g(x)) = 1 \neq 0.$$

La hipótesis ausente es el **no impacto local**: $g(x) \neq b$ en algún entorno perforado de a, o alguna información alternativa que controle explícitamente qué ocurre cuando la función interior toma el valor b.

Solución D6.
La afirmación es falsa. Tomemos

$$f(x) = \begin{cases} 2, & x < 0, \\ -3, & x > 0. \end{cases}$$

Entonces

$$\lim_{x \to 0^-} f(x) = 2$$

y

$$\lim_{x \to 0^+} f(x) = -3.$$

Ambos límites laterales existen, pero son distintos. Por Teorema 4.4, el límite bilateral no existe. Así, el fallo bilateral puede deberse precisamente a una incompatibilidad entre dos comportamientos laterales perfectamente bien definidos.

Soluciones del nivel E. **Solución E1.**
Una construcción es $f(x) = x^2$ para todo x, y

$$g(x) = \begin{cases} 7, & x = 0, \\ x^2, & x \neq 0, \end{cases} \qquad h(x) = \begin{cases} -4, & x = 0, \\ x^2, & x \neq 0. \end{cases}$$

Las tres funciones coinciden en todo entorno perforado de 0:

$$f(x) = g(x) = h(x) = x^2 \qquad (x \neq 0).$$

Como

$$\lim_{x \to 0} x^2 = 0,$$

la estabilidad bajo coincidencia local de Proposición 4.1 da

$$\boxed{\lim_{x \to 0} f(x) = \lim_{x \to 0} g(x) = \lim_{x \to 0} h(x) = 0.}$$

Sin embargo,

$$f(0) = 0, \qquad g(0) = 7, \qquad h(0) = -4.$$

Esto materializa la irrelevancia del valor puntual para un límite perforado.

Solución E2.
Tomemos

$$f(x) = \sigma\left(\frac{1}{x}\right), \qquad x \neq 0.$$

La función está acotada porque $|f(x)| \leq 1$.
Definamos $x_n = \frac{1}{\frac{\omega}{2}+2\omega n}$ y

$$y_n = \frac{1}{\frac{3\omega}{2} + 2\omega n}.$$

Ambas sucesiones son positivas, evitan 0 y convergen a 0. Pero $f(x_n) = \sigma\left(\frac{\omega}{2} + 2\omega n\right) = 1$, mientras que

$$f(y_n) = \sigma\left(\frac{3\omega}{2} + 2\omega n\right) = -1.$$

Si existiera un límite real en 0, el criterio secuencial obligaría a que ambas sucesiones de imágenes convergieran al mismo número. No ocurre. Por tanto,

$$\boxed{\lim_{x\to 0} \sigma(1/x) \text{ no existe.}}$$

Solución E3.
Podemos definir

$$f(x) = \begin{cases} 2, & x < 0, \\ -3, & x > 0. \end{cases}$$

Para probar el límite izquierdo, sea $\varepsilon > 0$. Cualquier $\delta > 0$ sirve: si $-\delta < x < 0$, entonces $|f(x) - 2| = 0 < \varepsilon$.
Por tanto,

$$\lim_{x\to 0^-} f(x) = 2.$$

Análogamente, para $0 < x < \delta$, $|f(x) - (-3)| = 0 < \varepsilon$, y así

$$\lim_{x\to 0^+} f(x) = -3.$$

Como los valores laterales son diferentes,

$$\boxed{\lim_{x\to 0} f(x) \text{ no existe.}}$$

Solución E4.
Consideremos la función de Dirichlet

$$f(x) = \begin{cases} 1, & x \in \mathbb{Q}, \\ 0, & x \notin \mathbb{Q}. \end{cases}$$

Claramente, $0 \leq f(x) \leq 1$ para todo $x \in \mathbb{R}$.

Fijemos un punto arbitrario $a \in \mathbb{R}$ y supongamos, buscando contradicción, que existe

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

Tomemos $\varepsilon = \frac{1}{3}$.

Entonces debería existir $\delta > 0$ tal que todo x con $0 < |x - a| < \delta$ satisficiera $|f(x) - L| < \frac{1}{3}$. Por la densidad de $\mathbb{Q}$ y de los irracionales, en ese mismo entorno perforado existen un racional q y un irracional r. Por tanto, $|1 - L| < \frac{1}{3}$ y $|L| < \frac{1}{3}$. Pero entonces, por desigualdad triangular,

$$1 = |1 - 0| \leq |1 - L| + |L| < \frac{2}{3},$$

contradicción.

Como a era arbitrario, f no tiene límite en ningún punto real.

Solución E5.

Tomemos

$$a = 0, \qquad b = 0, \qquad L = 0,$$

y definamos $g(x) = 0$ para todo x, junto con

$$f(y) = \begin{cases} 1, & y = 0, \\ 0, & y \neq 0. \end{cases}$$

Entonces

$$g(x) \to 0 \qquad (x \to 0)$$

y

$$f(y) \to 0 \qquad (y \to 0).$$

Sin embargo, $g(x) = b$ para **todos** los puntos próximos a a, no solo para algunos. En consecuencia, $f(g(x)) = f(0) = 1$, y

$$\boxed{\lim_{x \to 0} f(g(x)) = 1 \neq L.}$$

El ejemplo conserva exactamente los dos límites de partida y hace fallar la conclusión únicamente porque la función interior cae sobre el punto que el límite exterior excluye.

Soluciones del nivel F. **Solución F1.**

La afirmación original es

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x \in A : \quad P(\varepsilon, \delta, x),$$

con

$$P(\varepsilon, \delta, x) : \quad 0 < |x - a| < \delta \Longrightarrow |f(x) - L| < \varepsilon.$$

Negar un cuantificador universal lo convierte en existencial y negar uno existencial lo convierte en universal. Así,

$$\neg\Big[\forall\varepsilon > 0\ \exists\delta > 0\ \forall x \in A : P\Big]$$

se transforma en

$$\exists\varepsilon_0 > 0\ \forall\delta > 0\ \exists x \in A : \neg P(\varepsilon_0, \delta, x).$$

Ahora negamos la implicación. La negación de $A \Longrightarrow B$ es A y $\neg B$.

Por tanto,

$$\boxed{\exists\varepsilon_0 > 0\ \forall\delta > 0\ \exists x \in A : \quad 0 < |x - a| < \delta \quad \text{y} \quad |f(x) - L| \geq \varepsilon_0.}$$

El punto testigo aparece **después** del cuantificador universal sobre δ. Por eso puede depender del radio: cada escala puede requerir un punto distinto que haga visible el fallo.

Solución F2.

Como el límite candidato L falla, por la negación de la definición existe una tolerancia fija $\varepsilon_0 > 0$ tal que para todo $\delta > 0$ existe $x \in A$ con $0 < |x - a| < \delta$ y $|f(x) - L| \geq \varepsilon_0$.

Para cada $n \in \mathbb{N}$ elegimos $\delta_n = \frac{1}{n+1}$.

La negación suministra un punto $x_n \in A$ tal que $0 < |x_n - a| < \frac{1}{n+1}$ y $|f(x_n) - L| \geq \varepsilon_0$. La primera desigualdad garantiza inmediatamente $x_n \neq a$. Además, $0 \leq |x_n - a| < \frac{1}{n+1} \to 0$, de modo que, por sándwich, $x_n \to a$. Sin embargo, la distancia de $f(x_n)$ a L nunca baja de la cantidad positiva fija ε_0. Por tanto, $f(x_n) \not\to L$. Hemos construido una sucesión admisible que destruye el candidato L. Este es exactamente el mecanismo usado en la dirección contrapositiva de Teorema 4.2.

Solución F3.

Aplicamos la definición del límite con la tolerancia particular $\varepsilon = 1$.

Existe entonces $\delta > 0$ tal que

$$0 < |x - a| < \delta \quad \Longrightarrow \quad |f(x) - L| < 1.$$

Por desigualdad triangular,

$$|f(x)| \leq |f(x) - L| + |L| < 1 + |L|.$$

Así,

$$\boxed{0 < |x - a| < \delta \Longrightarrow |f(x)| < |L| + 1.}$$

Esto es acotación **local perforada**. No dice nada sobre valores de f lejos de a y tampoco controla necesariamente $f(a)$ si el punto pertenece al dominio. Una función puede tener límite finito en a y ser no acotada en regiones distantes.

Solución F4.

Usamos la descomposición

$$f(x)g(x) - LM = f(x)(g(x) - M) + M(f(x) - L).$$

Como $f(x) \to L$, por acotación local existe $r > 0$ tal que

$$0 < |x-a| < r \quad \Longrightarrow \quad |f(x)| < B,$$

donde podemos tomar $B = |L| + 1 > 0$.

Sea $\varepsilon > 0$. Del límite $g(x) \to M$ obtenemos $\delta_g > 0$ tal que

$$0 < |x-a| < \delta_g \quad \Longrightarrow \quad |g(x) - M| < \frac{\varepsilon}{2B}.$$

Del límite $f(x) \to L$ obtenemos $\delta_f > 0$ tal que

$$0 < |x-a| < \delta_f \quad \Longrightarrow \quad |f(x) - L| < \frac{\varepsilon}{2(|M|+1)}.$$

Tomamos $\delta = \text{mín}\{r, \delta_g, \delta_f\}$.

Si $0 < |x-a| < \delta$, entonces

$$\begin{aligned} |f(x)g(x) - LM| &\leq |f(x)|\,|g(x) - M| + |M|\,|f(x) - L| \\ &< B\frac{\varepsilon}{2B} + |M|\frac{\varepsilon}{2(|M|+1)} \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Por tanto,

$$\boxed{f(x)g(x) \to LM.}$$

El uso de $|M| + 1$ evita tener que separar artificialmente el caso $M = 0$.

Solución F5.

Supongamos primero

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

Dado $\varepsilon > 0$, existe $\delta > 0$ tal que para todo punto del dominio con $0 < |x-a| < \delta$ se cumple $|f(x) - L| < \varepsilon$. En particular, la misma desigualdad vale al restringir a puntos con $x < a$ y al restringir a puntos con $x > a$. Por tanto,

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

Recíprocamente, supongamos que ambos límites laterales valen L. Dado $\varepsilon > 0$, existen radios $\delta_- > 0$ y $\delta_+ > 0$ tales que

$$a - \delta_- < x < a \Longrightarrow |f(x) - L| < \varepsilon$$

y

$$a < x < a + \delta_+ \Longrightarrow |f(x) - L| < \varepsilon,$$

para los puntos pertinentes del dominio.

Tomemos $\delta = \text{mín}\{\delta_-, \delta_+\}$.

Si x pertenece al dominio y $0 < |x-a| < \delta$, entonces necesariamente ocurre uno de dos casos: $x < a$ o $x > a$. En el primero se aplica el control izquierdo; en el segundo, el derecho. En ambos casos, $|f(x) - L| < \varepsilon$.

Por tanto,

$$\boxed{\lim_{x\to a} f(x) = L \iff \lim_{x\to a^-} f(x) = L \text{ y } \lim_{x\to a^+} f(x) = L.}$$

La hipótesis de acumulación por ambos lados garantiza que las dos afirmaciones laterales tengan contenido.

Soluciones del nivel G. **Solución G1.**

Para $x > -3$ y $x \neq 1$, racionalizamos:

$$\frac{\sqrt{x+3}-2}{x-1}\cdot\frac{\sqrt{x+3}+2}{\sqrt{x+3}+2} = \frac{x+3-4}{(x-1)(\sqrt{x+3}+2)}.$$

Como $x \neq 1$, $f(x) = \frac{1}{\sqrt{x+3}+2}$.

Ahora estudiamos directamente la distancia al candidato $1/4$:

$$\begin{aligned}\left|\frac{1}{\sqrt{x+3}+2} - \frac{1}{4}\right| &= \frac{|2-\sqrt{x+3}|}{4(\sqrt{x+3}+2)} \\ &= \frac{|1-x|}{4(\sqrt{x+3}+2)^2}.\end{aligned}$$

Como $\sqrt{x+3} \geq 0$ en el dominio, $\sqrt{x+3}+2 \geq 2$, y por tanto $4(\sqrt{x+3}+2)^2 \geq 16$. De aquí obtenemos la estimación decisiva:

$$\left|f(x) - \frac{1}{4}\right| \leq \frac{|x-1|}{16}.$$

Sea $\varepsilon > 0$. Tomemos $\delta = \min\{1, 16\varepsilon\}$.

Si x pertenece al dominio y $0 < |x-1| < \delta$, entonces

$$\left|f(x) - \frac{1}{4}\right| \leq \frac{|x-1|}{16} < \frac{\delta}{16} \leq \varepsilon.$$

Por tanto,

$$\boxed{\lim_{x\to 1} \frac{\sqrt{x+3}-2}{x-1} = \frac{1}{4}.}$$

No se utilizó continuidad de la raíz. La racionalización convirtió el problema en una desigualdad elemental controlada por $|x-1|$.

Solución G2.

1. El producto amortiguado. Sea (x_n) cualquier sucesión admisible:

$$x_n \neq 0, \qquad x_n \to 0.$$

Entonces

$$\left|x_n \sigma\left(\frac{1}{x_n^2}\right)\right| \leq |x_n|.$$

Como $|x_n| \to 0$, el sándwich secuencial implica

$$x_n\sigma\left(\frac{1}{x_n^2}\right) \to 0.$$

Esto vale para toda sucesión admisible, así que por Teorema 4.2

$$\boxed{\lim_{x\to 0} x\sigma(1/x^2) = 0.}$$

2. La oscilación interior no tiene límite. Definamos

$$u_n = \frac{1}{\sqrt{\frac{\omega}{2} + 2\omega n}}$$

y

$$v_n = \frac{1}{\sqrt{\frac{3\omega}{2} + 2\omega n}}.$$

Ambas sucesiones son positivas, evitan 0 y convergen a 0. Sin embargo,

$$\sigma\left(\frac{1}{u_n^2}\right) = \sigma\left(\frac{\omega}{2} + 2\omega n\right) = 1,$$

mientras que

$$\sigma\left(\frac{1}{v_n^2}\right) = \sigma\left(\frac{3\omega}{2} + 2\omega n\right) = -1.$$

Por el criterio secuencial, $\sigma(1/x^2)$ no tiene límite cuando $x \to 0$.

3. No hay contradicción. El producto no se ha calculado como producto de dos límites. Se ha probado mediante la estimación $|x\sigma(1/x^2)| \leq |x|$.

La oscilación puede persistir indefinidamente y, aun así, quedar comprimida por un factor cuya magnitud tiende a 0.

Solución G3.

1. Acumulación por ambos lados. Por la izquierda, dado $\delta > 0$ podemos tomar

$$x = -\min\left\{\frac{\delta}{2}, \frac{1}{2}\right\}.$$

Entonces $x \in (-1, 0)$ y $0 < |x| < \delta$.

Por la derecha, la propiedad arquimediana permite elegir n con $\frac{1}{n} < \delta$. Entonces $1/n \in A$, $1/n > 0$ y $|1/n| < \delta$. Por tanto, 0 es punto de acumulación de A desde ambos lados.

2. Límite de la función interior. Sea $\varepsilon > 0$ y elijamos $\delta = \min\{1, \sqrt{\varepsilon}\}$.

Sea $x \in A$ con $0 < |x| < \delta$.

Si $x < 0$, entonces

$$|h(x) - 0| = x^2 < \delta^2 \leq \varepsilon.$$

Si $x > 0$, por la forma de A tenemos $x = 1/n$ para algún n y, por definición, $h(x) = 0$, de modo que $|h(x)| = 0 < \varepsilon$.

Así,

$$\boxed{h(x) \to 0 \qquad (x \to 0,\ x \in A).}$$

3. Límite exterior. El dominio de f es $[0, \infty)$, de modo que 0 es un punto de acumulación disponible solo desde la derecha, pero el límite ordinario relativo al dominio está bien definido.

Para $y > 0$, $f(y) = \frac{1}{1+y}$.

Entonces

$$|f(y) - 1| = \left| \frac{1}{1+y} - 1 \right| = \frac{y}{1+y} \leq y.$$

Dado $\varepsilon > 0$, basta elegir $\eta = \varepsilon$. Si $0 < y < \eta$, entonces $|f(y) - 1| \leq y < \varepsilon$.

Por tanto,

$$\boxed{f(y) \to 1 \qquad (y \to 0,\ y \in [0, \infty)).}$$

El valor $f(0) = 7$ no interviene en ese límite perforado.

4. Composición por cada lado. Si $x \to 0^-$ dentro de A, entonces $x \in (-1, 0)$ y $h(x) = x^2 > 0$.

Por tanto, $f(h(x)) = \frac{1}{1+x^2} \to 1$.

Así,

$$\boxed{\lim_{x \to 0^-} f(h(x)) = 1.}$$

En cambio, los puntos de A a la derecha de 0 son exactamente $x = 1/n$. Para todos ellos, $h(1/n) = 0$, y en consecuencia $f(h(1/n)) = f(0) = 7$.

Por tanto,

$$\boxed{\lim_{x \to 0^+} f(h(x)) = 7}$$

relativamente al dominio A.

5. Diagnóstico final. Los laterales existen pero son distintos. Luego, por Teorema 4.4,

$$\boxed{\lim_{x \to 0} f(h(x)) \text{ no existe.}}$$

No hay contradicción con Teorema 4.5. La hipótesis de no impacto local falla: por pequeño que sea $r > 0$, existen puntos derechos $x = \frac{1}{n}$ con $0 < x < r$ y $h(x) = 0$. La función interior golpea exactamente el punto $b = 0$ que el límite exterior excluye, y lo hace arbitrariamente cerca del punto de aproximación. El lado izquierdo evita ese impacto y hereda el límite exterior 1; el lado derecho lo golpea siempre y hereda el valor puntual 7.

Cierre del banco. Los cuarenta ejercicios han recorrido la arquitectura completa del capítulo sin introducir herramientas posteriores. El hilo que conviene conservar es:

$$\boxed{\text{dominio y acumulación} \to \text{tipo de límite} \to \text{método de prueba} \to \text{auditoría de hipótesis locales.}}$$

En particular, una solución rigurosa ya debe distinguir entre el valor puntual y el comportamiento perforado, saber cuándo una sucesión refuta un candidato, reconocer cuándo un producto necesita acotación local, cuándo un cociente necesita separación respecto de cero, cuándo dos

laterales controlan el bilateral y cuándo una composición requiere evitar el punto ignorado por el límite exterior.

Capítulo 5

Continuidad en la recta: intervalos, compacidad y teoremas fundamentales

En los ejemplos periódicos usaremos las funciones triangulares σ y κ definidas en C03, con $\omega = 2$. Sus cotas $|\sigma|, |\kappa| \leq 1$ y los valores $\sigma(\omega/2 + n\omega) = (-1)^n$ proceden directamente de aquella definición por tramos; no se presupone trigonometría.

Las funciones σ y κ son continuas: en cada tramo son afines y en los puntos de unión coinciden ambos valores laterales con el valor de la función. Cada punto tiene un entorno que corta solo finitos tramos, así que el criterio de empalme se aplica también a su prolongación periódica. El límite funcional describe lo que ocurre con las imágenes de puntos próximos a un centro, sin imponer una condición al valor tomado en ese centro. Su definición utiliza un entorno perforado: $0 < |x - a| < \delta$. Ese pequeño agujero en el centro era deliberado. Permitía que

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

tuviera sentido aunque $f(a)$ no estuviera definida, y permitía también que el límite existiera aunque el valor puntual fuese distinto de L.

Ahora cambia la pregunta.

La continuidad exige relacionar ese comportamiento próximo con el propio valor $f(a)$. Para formularla tendremos que incluir el centro en el control y comprobar primero que pertenece al dominio. Ese cambio parece pequeño, pero abre una teoría nueva. Consideremos, por ejemplo, las funciones $f(x) = x + 1$ y

$$g(x) = \begin{cases} x + 1, & x \neq 1, \\ 7, & x = 1. \end{cases}$$

Ambas tienen exactamente el mismo comportamiento en todo entorno perforado de 1. Por tanto,

$$\lim_{x \to 1} f(x) = 2 \qquad \text{y} \qquad \lim_{x \to 1} g(x) = 2.$$

Pero en el centro ocurre algo diferente:

$$f(1) = 2, \qquad g(1) = 7.$$

El límite no distingue estas dos situaciones porque, por construcción, ignora el valor central. La continuidad sí deberá distinguirlas. La idea que gobernará este capítulo puede resumirse así:

el límite controla los alrededores;	la continuidad exige que el centro encaje con ellos.

Al principio esta será todavía una propiedad local. Sin embargo, cuando la combinemos con la completitud de $\mathbb{R}$ y con la estructura de los intervalos, aparecerán consecuencias globales mucho

más fuertes: valores intermedios, existencia de extremos, compactitud de intervalos cerrados y continuidad uniforme. La pregunta central del capítulo será, por tanto,

¿cómo puede una condición local, punto por punto, producir información sobre todo un intervalo?

Antes de responderla globalmente debemos fijar con precisión qué significa ser continuo **en un solo punto**.

5.1. Continuidad en un punto: controlar también el centro

Del entorno perforado al entorno completo. Recordemos la forma del límite finito introducido en Definición 4.2. Bajo la hipótesis de que a sea punto de acumulación del dominio A, escribir

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

significa que

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in A: \qquad 0 < |x-a| < \delta \Longrightarrow |f(x) - L| < \varepsilon.$$

Hay dos rasgos que debemos mirar con atención. Primero, la entrada está **perforada**: $0 < |x-a|$.

Así que $x = a$ queda excluido.

Segundo, el valor objetivo L es un número que se determina por el comportamiento alrededor de a. No tiene por qué coincidir con $f(a)$ y, de hecho, $f(a)$ podría ni siquiera existir. Para hablar de continuidad cambiaremos exactamente esos dos aspectos. Queremos que el punto a pertenezca al dominio y queremos que, cuando x esté cerca de a, el valor $f(x)$ esté cerca del valor **efectivamente tomado por la función** en a. Por tanto, el centro deja de estar perforado y el objetivo de salida deja de ser un número independiente: será $f(a)$.

Definición 5.1. **Continuidad en un punto y en un conjunto.** Sea

$$f\colon A \to \mathbb{R}, \qquad A \subseteq \mathbb{R},$$

y sea $a \in A$.

Diremos que f es **continua en** a —relativamente a su dominio A— si

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in A: \qquad |x-a| < \delta \Longrightarrow |f(x) - f(a)| < \varepsilon.$$

Diremos que f es **continua en** A si es continua en cada punto de A. Si $E \subseteq A$, diremos que f es **continua sobre** E cuando la restricción

$$f|_E\colon E \to \mathbb{R}$$

es continua en cada punto de E.

La última frase hace explícito algo que ya aprendimos en `T1-C03`: el dominio forma parte de la función. Cuando restringimos una función, también restringimos las maneras en que la variable puede aproximarse a un punto. No debemos leer la definición como una fórmula para memorizar, sino como un mecanismo de control. El orden de los cuantificadores es

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall x \in A.$$

La tolerancia de salida ε se fija primero. Después debemos encontrar un radio de entrada δ, que puede depender de ε y del punto a, pero no del punto particular x que aparezca más tarde. Una vez elegido ese radio, **todo** x del dominio suficientemente cercano a a debe satisfacer $|f(x) - f(a)| < \varepsilon$.

En palabras:

> por pequeña que sea la tolerancia que se nos exija alrededor de $f(a)$, podemos encerrar a a en un entorno suficientemente pequeño para que todos los valores de la función correspondientes a puntos del dominio dentro de ese entorno queden dentro de la tolerancia requerida.

Dos cambios mínimos, una diferencia conceptual decisiva. Comparemos lado a lado las dos condiciones.

Para el límite:

$$0 < |x - a| < \delta \Longrightarrow |f(x) - L| < \varepsilon.$$

Para la continuidad: $|x - a| < \delta \Longrightarrow |f(x) - f(a)| < \varepsilon$.

Han desaparecido dos grados de libertad. Ya no podemos ignorar $x = a$, y ya no podemos escoger como objetivo un número L distinto del valor puntual. La función debe ser compatible consigo misma en el centro. Podemos visualizarlo mediante la cadena

$$\boxed{\text{entrada próxima a } a \Longrightarrow \text{salida próxima a } f(a).}$$

La palabra **próxima** debe entenderse cuantitativamente: para toda precisión de salida existe una precisión de entrada que funciona para todos los puntos permitidos por el dominio.

El centro ya participa
En la definición de límite, $x = a$ estaba excluido porque queríamos estudiar el comportamiento alrededor del punto independientemente del valor puntual. En continuidad ocurre lo contrario. Si $x = a$, entonces $|f(x)-f(a)| = 0$, y el centro satisface automáticamente cualquier tolerancia positiva. La continuidad no ignora el punto: lo incorpora de una manera compatible con sus alrededores.

Cómo puede fallar la continuidad. La negación de la definición es también informativa. f **no** es continua en a si existe alguna tolerancia positiva que no puede ser garantizada por ningún radio de entrada. En símbolos:

$$\exists \varepsilon_0 > 0 \; \forall \delta > 0 \; \exists x \in A : \qquad |x - a| < \delta \quad \text{y} \quad |f(x) - f(a)| \geq \varepsilon_0.$$

Obsérvese que ahora no aparece la condición $0 < |x - a|$. No la hemos olvidado. Simplemente estamos negando la definición de continuidad, cuya condición de entrada incluye el centro. Más adelante utilizaremos esta forma para diagnosticar discontinuidades. Por ahora lo importante es leerla correctamente:

> hay una distancia vertical fija respecto de $f(a)$ que sigue siendo violada por algún punto del dominio, por mucho que reduzcamos el entorno alrededor de a.

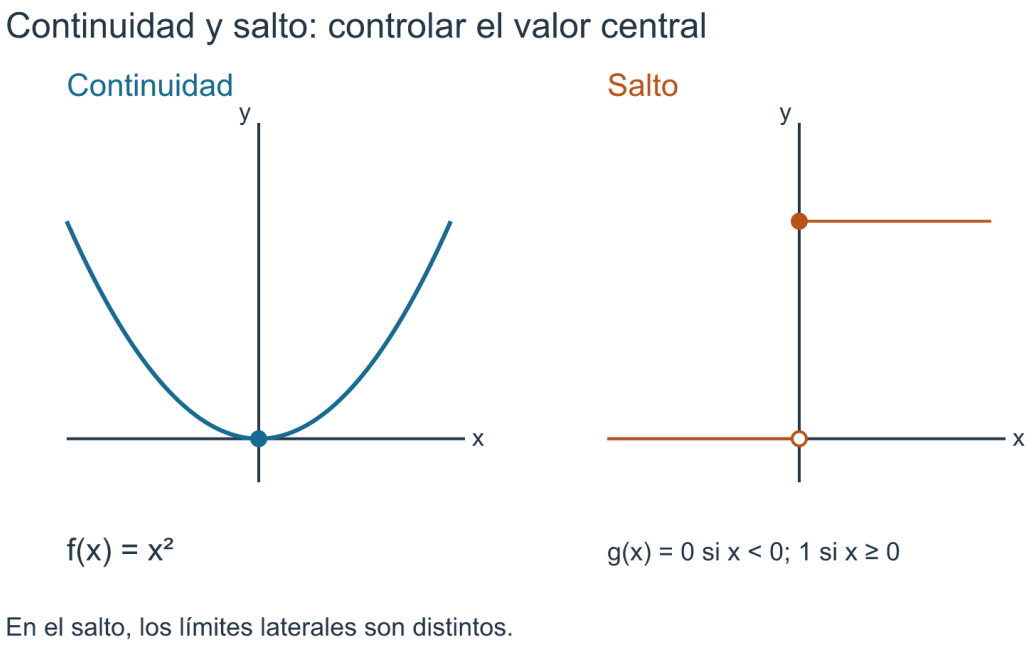


FIGURA 5.1. El cuadrado es continuo en 0. En cambio, la función que vale 0 a la izquierda y 1 desde 0 presenta un salto: sus límites laterales no coinciden.

Una primera prueba directa: el cuadrado es continuo. Antes de relacionar continuidad y límite, conviene demostrar una continuidad directamente desde la definición.

Consideremos

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = x^2,$$

y fijemos un punto arbitrario $a \in \mathbb{R}$.

Queremos demostrar que f es continua en a. Dado $\varepsilon > 0$, debemos controlar

$$|f(x) - f(a)| = |x^2 - a^2| = |x - a|\,|x + a|.$$

Aparece exactamente la misma dificultad que ya vimos en las pruebas ε–δ de límites: no basta controlar $|x-a|$; también necesitamos controlar localmente el segundo factor. Imponemos primero la condición auxiliar $|x - a| < 1$. Por la desigualdad triangular, $|x| \le |x - a| + |a| < 1 + |a|$, y por tanto $|x + a| \le |x| + |a| < 2|a| + 1$.

Ahora basta elegir

$$\delta = \min\left\{1, \frac{\varepsilon}{2|a| + 1}\right\}.$$

Si $|x - a| < \delta$, entonces

$$\begin{aligned} |f(x) - f(a)| &= |x - a|\,|x + a| \\ &< \delta(2|a| + 1) \\ &\le \varepsilon. \end{aligned}$$

Hemos demostrado que x^2 es continua en el punto arbitrario a, y por consiguiente es continua en $\mathbb{R}$. La arquitectura de la prueba merece quedar visible:

factorizar el error $\to$ acotar localmente un factor $\to$ hacer pequeño el otro $\to$ $\delta = \text{mín}\{\text{controles}\}$.

No hemos necesitado una nueva técnica de estimación. La continuidad reutiliza la ingeniería ε–δ desarrollada en `T1-C05`; lo nuevo es **qué estamos obligados a aproximar**.

El dominio decide desde dónde podemos acercarnos. La definición exige $x \in A$.

Toda continuidad es relativa al dominio de la función. Los puntos que intervienen en la estimación deben pertenecer a él, aunque un entorno real del centro contenga también puntos ajenos al dominio.

Supongamos que $f\colon [0,1] \to \mathbb{R}$.

Para estudiar la continuidad en 0 no preguntamos qué ocurre con $f(x)$ para números negativos cercanos a 0, porque esos números no pertenecen al dominio.

La condición es simplemente

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in [0,1] : |x| < \delta \Longrightarrow |f(x) - f(0)| < \varepsilon.$$

Todos los puntos relevantes se encuentran a la derecha de 0 porque el dominio no ofrece otros. Del mismo modo, en el extremo 1 solo intervienen puntos del intervalo situados a la izquierda. No necesitamos añadir a la definición general una cláusula especial para extremos. La restricción del dominio ya contiene toda la geometría necesaria. Cuando $a < b$ y $f\colon [a,b] \to \mathbb{R}$, la continuidad en el extremo izquierdo equivale a

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

y la continuidad en el extremo derecho equivale a

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

Estas son simplemente las formulaciones laterales del límite relativo al dominio desarrolladas en §4.7.

Cambiar el dominio puede cambiar la continuidad
Si $E \subseteq A$, la restricción $f|_E$ puede tener mejores propiedades de continuidad que f sobre A, porque algunos modos de aproximación dejan de estar disponibles. No estamos alterando los valores de la función en E; estamos alterando qué puntos pueden acercarse a cada centro.

El punto aislado: continuidad sin límite perforado. Aquí aparece una diferencia fundamental con el capítulo anterior. En `T1-C05` decidimos no formular

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

cuando a no es punto de acumulación del dominio. La razón era que un entorno perforado vacío volvería vacuamente verdadera la condición de límite para cualquier candidato L, destruyendo la unicidad. La continuidad no tiene ese problema. Si a es aislado, existe un entorno cuyo único punto del dominio es a. Como la definición de continuidad incluye el centro, para ese único argumento tenemos $|f(a) - f(a)| = 0$. Por tanto, toda función es continua en sus puntos aislados.

Esta afirmación, junto con la relación exacta entre continuidad y límite en puntos de acumulación, merece quedar registrada formalmente.

Proposición 5.1. **Continuidad, límite y puntos aislados.** Sea

$$f\colon A \to \mathbb{R}, \qquad a \in A.$$

1. Si a es un punto aislado de A, entonces f es continua en a.
2. Si a es un punto de acumulación de A, entonces

$$f \text{ es continua en } a \iff \lim_{x\to a} f(x) = f(a).$$

Idea de la prueba
En un punto aislado basta escoger un radio que no contenga ningún otro punto del dominio: la única entrada posible es $x = a$, y el error de salida es exactamente 0. En un punto de acumulación, continuidad y límite usan prácticamente la misma desigualdad. La diferencia es que el límite trabaja con $x \neq a$. Para pasar del límite a la continuidad basta separar dos casos: $x = a$, donde el error es 0, y $x \neq a$, donde actúa la hipótesis de límite.

Demostración. Supongamos primero que a es aislado en A. Por definición, existe $r > 0$ tal que $V_r(a) \cap A = \{a\}$.

Sea $\varepsilon > 0$ arbitrario. Elegimos $\delta = r$.

Si $x \in A$ y $|x - a| < \delta$, entonces $x \in V_r(a) \cap A$, de modo que necesariamente $x = a$.

Por consiguiente,

$$|f(x) - f(a)| = |f(a) - f(a)| = 0 < \varepsilon.$$

Así, f es continua en a.

Supongamos ahora que a es punto de acumulación de A.

($\Rightarrow$) Supongamos que f es continua en a. Sea $\varepsilon > 0$. Por continuidad, existe $\delta > 0$ tal que para todo $x \in A$, $|x - a| < \delta \Longrightarrow |f(x) - f(a)| < \varepsilon$.

En particular, la misma conclusión vale para los puntos $x \in A$ que satisfacen la condición más fuerte $0 < |x - a| < \delta$. Por la definición de límite,

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

($\Leftarrow$) Supongamos ahora que

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

Sea $\varepsilon > 0$. Por la definición de límite, existe $\delta > 0$ tal que

$$x \in A, \qquad 0 < |x - a| < \delta \Longrightarrow |f(x) - f(a)| < \varepsilon.$$

Tomemos cualquier $x \in A$ con $|x - a| < \delta$.

Hay dos casos.

Si $x = a$, entonces $|f(x) - f(a)| = 0 < \varepsilon$.

Si $x \neq a$, entonces $0 < |x - a| < \delta$, y la hipótesis de límite proporciona $|f(x) - f(a)| < \varepsilon$. En ambos casos se cumple la condición de continuidad. Por tanto, f es continua en a. ■

No hay contradicción entre los dos capítulos
En un punto aislado, nuestra convención de `T1-C05` **no formula** un límite perforado porque no hay puntos del dominio que se aproximen al centro evitando el centro. Sin embargo, la continuidad sí está perfectamente definida porque requiere $a \in A$ y permite $x = a$. Por eso pueden coexistir sin conflicto estas dos afirmaciones: «el límite en a no se formula» y

«la función es continua en a».

Un dominio con dos geometrías locales distintas. La mejor manera de fijar esta diferencia es trabajar con una sola función cuyo dominio contenga, al mismo tiempo, un punto aislado y un punto de acumulación.

Ejemplo 5.1. **Continuidad en un punto aislado y en un punto de acumulación.** Sea $A = \{0\} \cup [1, \infty)$ y definamos $f\colon A \to \mathbb{R}$ por

$$f(0) = 100, \qquad f(x) = x^2 \quad \text{si } x \geq 1.$$

Entonces f es continua tanto en 0 como en 1, pero por razones completamente distintas.

En 0. El punto 0 es aislado en A. Por ejemplo, $V_{1/2}(0) \cap A = \{0\}$.

Por Proposición 5.1, f es continua en 0 cualquiera sea el valor que hayamos asignado a $f(0)$. El número 100 no necesita parecerse en absoluto a los valores x^2 de la otra componente del dominio. Esto no es una anomalía. Localmente, alrededor de 0, no existen otros puntos del dominio con los cuales comparar $f(0)$. Además, bajo la convención de `T1-C05`, no formulamos

$$\lim_{x \to 0} f(x),$$

porque 0 no es punto de acumulación de A.

En 1. La situación cambia. El punto 1 sí es punto de acumulación de A; de hecho, hay puntos de $[1, \infty)$ arbitrariamente próximos a 1 por la derecha.

Aquí la continuidad exige un control real de los valores vecinos.

Sea $\varepsilon > 0$. Elegimos $\delta = \min\left\{1, \frac{\varepsilon}{3}\right\}$.

Si $x \in A$ y $|x - 1| < \delta$, entonces $\delta \leq 1$ excluye el punto aislado 0, y necesariamente $1 \leq x < 2$.

Por tanto, $|x + 1| < 3$.

Como $f(1) = 1$, obtenemos

$$\begin{aligned} |f(x) - f(1)| &= |x^2 - 1| \\ &= |x - 1|\,|x + 1| \\ &< 3|x - 1| \\ &< 3\delta \\ &\leq \varepsilon. \end{aligned}$$

Así, f es continua en 1.

El contraste es exacto:

- en 0, la geometría del dominio hace innecesario controlar otros puntos;
- en 1, existen puntos del dominio arbitrariamente cercanos y la continuidad exige una estimación ε–δ genuina.

La fórmula de la función no basta para decidir qué prueba corresponde. Primero debemos mirar **cómo se acerca el dominio al punto**.

Un agujero reparado y un valor incompatible. Volvamos ahora al ejemplo que abrió el capítulo.

Definamos $g\colon \mathbb{R} \to \mathbb{R}$ por

$$g(x) = \begin{cases} x+1, & x \neq 1, \\ 7, & x = 1. \end{cases}$$

En `T1-C05` ya disponemos de las herramientas para afirmar que

$$\lim_{x\to 1} g(x) = 2,$$

porque en un entorno perforado de 1 la función coincide con $x+1$.

Pero $g(1) = 7$.

Como 1 es punto de acumulación de $\mathbb{R}$, Proposición 5.1 nos dice inmediatamente que g no es continua en 1:

$$\lim_{x\to 1} g(x) = 2 \neq 7 = g(1).$$

Si en cambio definimos

$$\widetilde{g}(x) = \begin{cases} x+1, & x \neq 1, \\ 2, & x = 1, \end{cases}$$

entonces

$$\lim_{x\to 1} \widetilde{g}(x) = 2 = \widetilde{g}(1),$$

y por tanto $\widetilde{g}$ sí es continua en 1. La discontinuidad procede aquí de la incompatibilidad entre el valor central y el comportamiento de los puntos próximos:

comportamiento perforado + valor central compatible = continuidad.

Qué significa «continua en un intervalo». La expresión

f es continua en $[a,b]$

no introduce una nueva definición. Significa que la función $f\colon [a,b] \to \mathbb{R}$ es continua en **cada** punto de su dominio.

Por tanto:

- en cada punto interior $c \in (a,b)$, el dominio permite aproximaciones desde ambos lados;
- en a, solo importan aproximaciones desde la derecha;
- en b, solo importan aproximaciones desde la izquierda.

Esta forma relativa al dominio será crucial cuando lleguemos a los teoremas globales del capítulo. El teorema del valor intermedio y el teorema de los valores extremos hablarán de funciones continuas sobre intervalos cerrados sin exigir que la función esté definida fuera del intervalo. No añadiremos valores artificiales a la izquierda de a ni a la derecha de b. La continuidad se juzga en el espacio donde la función realmente vive.

Continuidad no es una propiedad de la fórmula aislada. A esta altura podemos reunir varias lecciones anteriores. En `T1-C03` aprendimos que una fórmula no determina por sí sola una función: importan dominio y codominio. En `T1-C05` aprendimos que un límite depende de cómo el dominio se aproxima al punto. Ahora vemos que la continuidad hereda ambas dependencias. La expresión $x \mapsto x^2$ puede aparecer en funciones con dominios distintos, y esos dominios determinan qué puntos deben compararse localmente. La pregunta correcta no es simplemente

«¿la fórmula x^2 es continua?»

sino

«¿es continua esta función, con este dominio, en este punto?»

Después de demostrar teoremas generales podremos abreviar el lenguaje y decir cómodamente que los polinomios son continuos. Pero esa abreviatura estará respaldada por una estructura precisa.

Cuatro diferencias que conviene conservar. Podemos resumir la transición desde límite a continuidad en la siguiente tabla.

Pregunta	Límite funcional en a	Continuidad en a
¿Debe a pertenecer al dominio?	No necesariamente	Sí
¿Se admite $x = a$ en la condición?	No: entorno perforado	Sí: entorno completo relativo al dominio
¿Cuál es el objetivo de salida?	Un candidato L	El valor forzado $f(a)$
¿Qué ocurre si a es aislado?	No formulamos el límite bajo nuestra convención	Toda función es continua en a

En un punto de acumulación, Proposición 5.1 une las dos columnas:

$$\boxed{f \text{ continua en } a \iff \lim_{x \to a} f(x) = f(a).}$$

Pero esta equivalencia no debe hacernos olvidar que la definición directa de continuidad es más amplia: también cubre correctamente puntos aislados y extremos de dominios.

Qué hemos ganado. La continuidad en un punto ha reorganizado tres piezas que ya conocíamos:

$$\boxed{\text{dominio} + \text{valor puntual} + \text{control } \varepsilon\text{–}\delta \longrightarrow \text{continuidad local.}}$$

Ahora sabemos que:

1. la continuidad se define relativamente al dominio;
2. el centro $x = a$ forma parte de la condición;
3. en puntos de acumulación, continuidad equivale a que el límite exista y coincida con $f(a)$;
4. en puntos aislados, toda función es automáticamente continua;
5. en extremos de intervalos, la propia geometría del dominio convierte la continuidad en una condición unilateral;
6. cambiar el valor puntual puede destruir o reparar continuidad sin alterar el límite perforado.

Todavía hemos hablado de un punto a la vez. En la sección siguiente construiremos herramientas que permitan reconocer continuidad sin repetir desde cero una prueba ε–δ en cada ejemplo. El primer paso será traducir la definición al lenguaje de sucesiones. Después demostraremos que las operaciones algebraicas y la composición preservan continuidad bajo las hipótesis adecuadas. Allí veremos, además, por qué la composición de funciones continuas ya no necesita la condición de «no impacto» que fue indispensable para la composición de límites en §4.9.

5.2. Criterio secuencial, álgebra y composición de funciones continuas

La definición de continuidad de §5.1 es directa y completa, pero no queremos reconstruir desde cero una prueba ε–δ cada vez que aparezca una suma, un producto, un cociente o una composición. Disponemos ya de una teoría extensa de límites. La pregunta natural es:

¿podemos transportar esa teoría a la continuidad sin perder de vista el valor central?

La respuesta es afirmativa, pero antes debemos ajustar con precisión el criterio secuencial. En `T1-C05` demostramos que el límite funcional

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

puede caracterizarse mediante sucesiones $x_n \to a$ cuyos términos **evitan** el centro: $x_n \neq a$. Esa exclusión era necesaria porque el límite perforado no controla $f(a)$. Para continuidad la situación cambia. La propia definición incluye el centro y exige compatibilidad con $f(a)$. Por tanto, una sucesión utilizada para probar continuidad puede tocar a, puede hacerlo infinitas veces e incluso puede ser la sucesión constante $x_n = a$. Esta diferencia aparentemente pequeña es exactamente la que permitirá formular un criterio secuencial válido también en puntos aislados.

El criterio secuencial de continuidad.

Teorema 5.1. **Criterio secuencial de continuidad.** Sea

$$f\colon A \to \mathbb{R}, \qquad A \subseteq \mathbb{R},$$

y sea $a \in A$. Entonces son equivalentes:

1. f es continua en a.
2. Para toda sucesión (x_n) de puntos de A tal que
$$x_n \to a,$$
se cumple
$$f(x_n) \to f(a).$$

No se exige que $x_n \neq a$.

Demostración. Demostraremos ambas implicaciones.

(1) $\Rightarrow$ (2)**.** Supongamos que f es continua en a y sea (x_n) una sucesión cualquiera de puntos de A con

$$x_n \to a.$$

Debemos demostrar que

$$f(x_n) \to f(a).$$

Sea $\varepsilon > 0$. Por continuidad de f en a, existe $\delta > 0$ tal que, para todo $x \in A$,

$$|x - a| < \delta \Longrightarrow |f(x) - f(a)| < \varepsilon.$$

Como $x_n \to a$, existe N tal que, para todo $n \geq N$,

$$|x_n - a| < \delta.$$

Cada x_n pertenece a A, de modo que para $n \geq N$ podemos aplicar el control de continuidad y obtenemos

$$|f(x_n) - f(a)| < \varepsilon.$$

Esto demuestra que

$$f(x_n) \to f(a).$$

Obsérvese que no necesitamos demostrar

$$0 < |x_n - a|.$$

Si algún término satisface $x_n = a$, entonces simplemente

$$|f(x_n) - f(a)| = 0,$$

y ese término ya está perfectamente controlado.

(2) $\Rightarrow$ (1)**.** Probaremos la contraposición. Supongamos que f **no** es continua en a.

Por la negación de la definición de continuidad existe un número

$$\varepsilon_0 > 0$$

tal que, para todo $\delta > 0$, existe $x \in A$ con

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

y

$$|f(x) - f(a)| \geq \varepsilon_0.$$

Para cada $n \in \mathbb{N}$, aplicamos esta afirmación con

$$\delta_n = \frac{1}{n+1}.$$

Podemos elegir $x_n \in A$ de manera que

$$|x_n - a| < \frac{1}{n+1}$$

y

$$|f(x_n) - f(a)| \geq \varepsilon_0.$$

La primera desigualdad y el teorema del sándwich para sucesiones implican

$$x_n \to a.$$

Pero la segunda desigualdad vale para **todo** n, así que la sucesión $(f(x_n))$ nunca entra en la banda de radio ε_0 alrededor de $f(a)$. Por tanto,

$$f(x_n) \not\to f(a).$$

Hemos construido una sucesión de puntos de A que converge a a y viola la condición (2). Esta es la contraposición de (2) $\Rightarrow$ (1).

Concluimos que ambas condiciones son equivalentes.

Hay un detalle lógico interesante en la segunda mitad de la prueba. Aunque no impusimos explícitamente $x_n \neq a$, los testigos fabricados por la discontinuidad no pueden ser iguales a a. En efecto, si $x_n = a$, entonces $|f(x_n) - f(a)| = 0$, lo que contradice $|f(x_n) - f(a)| \geq \varepsilon_0 > 0$. La propia violación de continuidad expulsa automáticamente al centro de la sucesión testigo.

La diferencia exacta con el criterio de límite funcional. Conviene colocar los dos criterios uno junto al otro. Para el límite funcional de Teorema 4.2:

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

si y solo si

$$x_n \in A, \qquad x_n \neq a, \qquad x_n \to a \Longrightarrow f(x_n) \to L.$$

Para la continuidad de Teorema 5.1: f continua en a si y solo si

$$x_n \in A, \qquad x_n \to a \Longrightarrow f(x_n) \to f(a).$$

La diferencia es estructural:

límite: el centro está perforado y las sucesiones deben evitarlo, continuidad: el centro está controlado y las sucesiones pueden tocarlo.

No estamos añadiendo una excepción al criterio secuencial. Estamos reflejando exactamente la diferencia entre las dos definiciones.

Los puntos aislados vistos mediante sucesiones. El criterio secuencial ofrece una segunda explicación de la continuidad automática en puntos aislados demostrada en Proposición 5.1. Sea a un punto aislado de A. Existe entonces $r > 0$ tal que $A \cap V_r(a) = \{a\}$. Si una sucesión (x_n) de puntos de A satisface $x_n \to a$, entonces, eventualmente, $|x_n - a| < r$. Pero dentro de ese entorno el único punto permitido del dominio es a. Por tanto existe N tal que

$$x_n = a \qquad (n \geq N).$$

Así, $f(x_n) = f(a)$ eventualmente, y en consecuencia $f(x_n) \to f(a)$.

Por Teorema 5.1, f es continua en a. La continuidad en un punto aislado puede leerse, por tanto, de dos maneras equivalentes:

visión ε–δ : un entorno pequeño solo contiene a, visión secuencial: toda sucesión que converge a a es finalmente constante en a.

De límites de sucesiones a álgebra de funciones continuas. El criterio secuencial nos permite ahora reutilizar casi toda la maquinaria de `T1-C04`. Supongamos que f y g son continuas en a. Si $x_n \to a$, entonces

$$f(x_n) \to f(a), \qquad g(x_n) \to g(a).$$

En ese momento ya no estamos ante un problema nuevo de continuidad: estamos ante dos sucesiones reales convergentes. Podemos aplicar su álgebra y después regresar al lenguaje funcional mediante Teorema 5.1. Esta idea concentra varias reglas en una sola proposición.

Proposición 5.2. **Álgebra y composición de funciones continuas.** Sea $A \subseteq \mathbb{R}$, sea $a \in A$ y sean

$$f, g\colon A \to \mathbb{R}$$

funciones continuas en a. Entonces:

1. para todo $c \in \mathbb{R}$, la función cf es continua en a;
2. $f + g$ y $f - g$ son continuas en a;
3. fg es continua en a;
4. $|f|$ es continua en a;
5. si $g(a) \neq 0$ y
$$C := \{x \in A : g(x) \neq 0\},$$
entonces $a \in C$ y el cociente
$$\frac{f}{g}\colon C \to \mathbb{R}$$
es continuo en a relativamente a su dominio C.

Además, sean $B \subseteq \mathbb{R}$,

$$g\colon A \to B$$

y

$$F\colon B \to \mathbb{R}.$$

Si g es continua en a y F es continua en $g(a)$, entonces la composición

$$F \circ g\colon A \to \mathbb{R}$$

es continua en a.

Las funciones constantes y la identidad son continuas en todos los puntos de sus dominios naturales.

Demostración. Comenzamos con las operaciones algebraicas.

Sea (x_n) una sucesión arbitraria de puntos de A tal que

$$x_n \to a.$$

Como f y g son continuas en a, Teorema 5.1 da

$$f(x_n) \to f(a), \qquad g(x_n) \to g(a).$$

Aplicando el álgebra de límites secuenciales de `T1-C04`, obtenemos

$$cf(x_n) \to cf(a),$$

$$f(x_n) + g(x_n) \to f(a) + g(a),$$

$$f(x_n) - g(x_n) \to f(a) - g(a),$$

$$f(x_n)g(x_n) \to f(a)g(a),$$

y

$$|f(x_n)| \to |f(a)|.$$

Como estas conclusiones valen para toda sucesión (x_n) del dominio que converge a a, Teorema 5.1 implica que cf, $f+g$, $f-g$, fg y $|f|$ son continuas en a.

Para el cociente supongamos

$$g(a) \neq 0.$$

Entonces $a \in C$. Sea (x_n) una sucesión arbitraria de puntos de C tal que

$$x_n \to a.$$

Como $C \subseteq A$, la continuidad de f y g en a nos da nuevamente

$$f(x_n) \to f(a), \qquad g(x_n) \to g(a) \neq 0.$$

Cada $g(x_n)$ es no nulo porque $x_n \in C$. La ley secuencial del cociente produce

$$\frac{f(x_n)}{g(x_n)} \longrightarrow \frac{f(a)}{g(a)}.$$

El lado derecho es precisamente el valor del cociente en a. Por el criterio secuencial, f/g es continua en a relativamente a C.

Pasemos a la composición. Sea $\varepsilon > 0$.

Como F es continua en $g(a)$, existe $\eta > 0$ tal que, para todo $y \in B$,

$$|y - g(a)| < \eta \Longrightarrow |F(y) - F(g(a))| < \varepsilon.$$

Como g es continua en a, existe $\delta > 0$ tal que, para todo $x \in A$,

$$|x - a| < \delta \Longrightarrow |g(x) - g(a)| < \eta.$$

Por tanto, si $x \in A$ y $|x-a| < \delta$, entonces $g(x) \in B$ y

$$|g(x) - g(a)| < \eta.$$

Aplicando el control exterior obtenemos

$$|F(g(x)) - F(g(a))| < \varepsilon.$$

Pero

$$(F \circ g)(a) = F(g(a)),$$

de modo que

$$|(F \circ g)(x) - (F \circ g)(a)| < \varepsilon.$$

Esto demuestra que $F \circ g$ es continua en a.

Finalmente, si $k(x) = c$ es constante, entonces

$$|k(x) - k(a)| = 0$$

para todos los puntos del dominio, por lo que k es continua. Para la identidad $i(x) = x$ basta elegir $\delta = \varepsilon$, pues

$$|i(x) - i(a)| = |x - a|.$$

Queda demostrada la proposición.

La prueba anterior utiliza dos estilos deliberadamente diferentes. Para el álgebra, el criterio secuencial comprime argumentos que ya fueron desarrollados en `T1-C04` y `T1-C05`. No tiene sentido volver a reconstruir cada producto desde una estimación ε–δ si la teoría secuencial ya encapsula ese trabajo. Para la composición, en cambio, hemos usado una prueba directa. La razón es pedagógica: esa prueba deja visible la diferencia exacta con el teorema de composición de límites de §4.9.

Por qué desaparece la hipótesis de no impacto. Recordemos la dificultad de Teorema 4.5.

Si solo sabemos que $g(x) \to b$ y

$$F(y) \to L \qquad (y \to b),$$

el límite exterior controla puntos que satisfacen $0 < |y - b| < \eta$. Después de sustituir $y = g(x)$ necesitamos, por tanto, $0 < |g(x) - b| < \eta$. El límite interior proporciona la desigualdad superior, pero no garantiza $g(x) \neq b$. Por eso §4.9 necesitó una condición adicional de no impacto. En continuidad, la condición exterior es distinta. Si F es continua en b, tenemos $|y - b| < \eta \Longrightarrow |F(y) - F(b)| < \varepsilon$.

No aparece $0 < |y - b|$.

El propio punto $y = b$ está permitido y, de hecho, está controlado exactamente porque $|F(b) - F(b)| = 0$. Por eso, después de poner $y = g(x)$, solo necesitamos demostrar $|g(x) - b| < \eta$. Si ocurre que $g(x) = b$, no aparece ningún problema: el error exterior es simplemente cero. La diferencia puede condensarse así:

límite exterior	:	$0 < \lvert y - b \rvert < \eta$,
continuidad exterior	:	$\lvert y - b \rvert < \eta$.

La desaparición de la hipótesis de no impacto no es una regla nueva añadida por conveniencia. Es una consecuencia directa de que la continuidad **controla el centro**.

Un golpe exacto al punto ya no causa problemas.

Ejemplo 5.2. **Composición continua aunque la función interior golpee siempre el punto exterior.** Definamos

$$g\colon \mathbb{R} \to \mathbb{R}, \qquad g(x) = 1$$

para todo x, y

$$F\colon \mathbb{R} \to \mathbb{R}, \qquad F(y) = (y-1)^2 + 7.$$

Estudiemos la composición en $a = 0$.

La función g es continua en 0 porque $|g(x) - g(0)| = |1 - 1| = 0$ para todo x. La función F es continua en 1. Podemos verlo directamente: si $|y - 1| < \delta$, entonces

$$|F(y) - F(1)| = |(y-1)^2| = |y-1|^2.$$

Dado $\varepsilon > 0$, basta tomar, por ejemplo, $\delta = \min\{1, \varepsilon\}$.

Si $|y - 1| < \delta$, entonces $|y - 1| < 1$ y

$$|F(y) - F(1)| = |y-1|^2 < |y-1| < \delta \leq \varepsilon.$$

Por Proposición 5.2, $F \circ g$ es continua en 0. En realidad, $(F \circ g)(x) = F(1) = 7$ para todo x, de modo que la composición es constante. Lo importante es que la función interior golpea exactamente el punto exterior $g(x) = 1$ no solo en $x = 0$, sino para **todo** x. La condición de no impacto de §4.9 falla de la forma más fuerte posible. Sin embargo, para continuidad no hace falta. El contraste con `#exm-t1-0051` es preciso. Allí una función exterior podía tener

$$\lim_{y \to 1} H(y) = 2$$

pero un valor excepcional $H(1) = 7$.

La composición con la función interior constante $g(x) = 1$ veía siempre el valor excepcional y quedaba igual a 7, no al límite exterior 2. Aquí no existe esa discordancia porque la continuidad de F en 1 exige que el comportamiento próximo sea compatible con el valor $F(1) = 7$. Por eso golpear exactamente el centro deja de ser una obstrucción.

El ejemplo es deliberadamente extremo. Si incluso una función interior que permanece **siempre** en el punto exterior puede componerse sin problemas con una función continua, entonces queda claro que el obstáculo de §4.9 pertenecía a la naturaleza perforada del límite, no a la composición en sí misma.

Polinomios y funciones racionales como primeras familias continuas. En §4.5 ya demostramos, sin utilizar continuidad, que para todo polinomio p,

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

y que, si $q(a) \neq 0$,

$$\lim_{x \to a} \frac{p(x)}{q(x)} = \frac{p(a)}{q(a)}.$$

Ahora podemos encapsular esos hechos en el nuevo lenguaje.

Corolario 5.1. **Continuidad de polinomios, funciones racionales y composiciones elementales autorizadas.** Todo polinomio real

$$p\colon \mathbb{R} \to \mathbb{R}$$

es continuo en $\mathbb{R}$.

Si p y q son polinomios y

$$D := \{x \in \mathbb{R} : q(x) \neq 0\},$$

entonces la función racional

$$r\colon D \to \mathbb{R}, \qquad r(x) = \frac{p(x)}{q(x)}$$

es continua en todo punto de D. En consecuencia, cualquier expresión obtenida mediante un número finito de sumas, diferencias, productos, valores absolutos, cocientes en puntos donde el denominador no se anula y composiciones de funciones cuya continuidad ya haya sido establecida es continua en los puntos de su dominio natural donde todas esas operaciones están definidas.

Demostración. Las funciones constantes y la identidad son continuas por Proposición 5.2. Aplicando repetidamente las leyes de suma, producto y multiplicación por constantes, obtenemos la continuidad de cada monomio

$$x \mapsto c_k x^k$$

y después de toda suma finita de monomios. Por tanto, todo polinomio es continuo.

Sea ahora $a \in D$. Entonces

$$q(a) \neq 0.$$

Los polinomios p y q son continuos en a, así que la ley del cociente de Proposición 5.2 implica que p/q es continua en a relativamente a su dominio natural D. Como a era arbitrario, la función racional es continua en todo D. La última afirmación se obtiene aplicando finitamente las cláusulas de Proposición 5.2, auditando en cada cociente que el denominador sea no nulo y en cada composición que la imagen de la función interior pertenezca al dominio de la exterior.

Este corolario cambia de manera importante nuestra forma de trabajar. Por ejemplo, para demostrar que $R(x) = \frac{x^2+3x+1}{x^2+1}$ es continua en un punto $a \in \mathbb{R}$, ya no necesitamos abrir la definición ε–δ ni calcular nuevamente un límite. Basta observar que el numerador y el denominador son polinomios y que $a^2 + 1 > 0$. Por tanto, el denominador no se anula y Corolario 5.1 autoriza la continuidad. Las estimaciones de `T1-C05` justificaron las leyes de límite. El criterio secuencial y Proposición 5.2 permiten aplicarlas a la continuidad: una vez verificadas sus hipótesis, podemos construir funciones continuas a partir de otras sin repetir las estimaciones de cada operación.

La continuidad se puede demostrar por construcción. A partir de esta sección aparecen dos estrategias principales. La primera es volver a la definición cuando necesitamos comprender una estimación concreta:

$$\varepsilon \longrightarrow \delta \longrightarrow |f(x) - f(a)| < \varepsilon.$$

La segunda es reconocer que una función se ha construido a partir de piezas cuya continuidad ya conocemos:

$$\boxed{\text{piezas continuas} + \text{operaciones que preservan continuidad} \longrightarrow \text{función continua.}}$$

La segunda estrategia será cada vez más frecuente, pero siempre estará subordinada a una auditoría de hipótesis. Para una suma no aparece una condición adicional. Para un cociente debemos verificar que el denominador no se anule en el punto. Para una composición debemos verificar el tipado $g(A) \subseteq B$ o, equivalentemente, que la salida de la función interior sea una entrada legítima para la exterior. La continuidad simplifica muchas pruebas; no elimina la necesidad de leer dominios e hipótesis.

Qué hemos ganado. En esta sección hemos convertido la continuidad en una propiedad que puede estudiarse con dos lenguajes intercambiables:

$$\boxed{\varepsilon\text{–}\delta \quad \longleftrightarrow \quad \text{sucesiones.}}$$

El criterio secuencial difiere del criterio para límites funcionales en un punto esencial: las sucesiones pueden tocar el centro porque la continuidad también lo controla. A partir de allí obtuvimos una máquina de construcción:

$$\boxed{\text{constantes/identidad} \to \text{álgebra} \to \text{cocientes auditados} \to \text{composición} \to \text{familias continuas.}}$$

Y la comparación con §4.9 mostró exactamente por qué la composición continua es más limpia que la composición de límites perforados:

$$\boxed{\text{continuidad exterior controla también el punto de impacto.}}$$

La siguiente pregunta será negativa y, por ello, igualmente importante: **¿de qué maneras puede fallar la continuidad?**

En §5.3 estudiaremos discontinuidades removibles, saltos y oscilaciones persistentes. No construiremos una clasificación exhaustiva; utilizaremos esos tres mecanismos como laboratorio para aprender a diagnosticar qué parte de la definición se ha roto y, cuando sea posible, cómo repararla.

5.3. Cómo falla la continuidad: reparación, salto y oscilación

Para diagnosticar una discontinuidad podemos volver a la misma exigencia que usamos al demostrar continuidad: las entradas suficientemente próximas deben producir salidas próximas al valor central. Los ejemplos siguientes permitirán identificar de qué manera puede fallar ese control. La definición de continuidad en un punto $a \in A$ exige que, al imponer entradas del dominio suficientemente próximas a a, todas las salidas queden próximas a $f(a)$. Por tanto, hay tres preguntas que conviene separar:

1. ¿pertenece el punto a al dominio?;
2. ¿existe un comportamiento perforado bien definido cuando $x \to a$?;
3. si ese comportamiento existe, ¿coincide con el valor que la función toma en el centro?

Esta separación evita confundir defectos muy diferentes. A veces toda la dificultad está concentrada en un único valor mal asignado. Otras veces los dos lados del dominio quieren acercarse a valores distintos. Y en otros casos ni siquiera existe un valor de salida hacia el cual se estabilice el comportamiento perforado. La idea de esta sección será:

diagnosticar primero el comportamiento alrededor del punto y solo después mirar el valor central.

Discontinuidad y defecto removible. En §5.1 ya escribimos la negación cuantificada de continuidad. Ahora fijaremos el vocabulario necesario para distinguir un defecto que puede repararse cambiando un solo valor de otro que no puede repararse de ese modo.

Definición 5.2. **Discontinuidad y reparación removible.** Sea

$$f\colon A \to \mathbb{R}, \qquad A \subseteq \mathbb{R}.$$

Sea $a \in A$. Diremos que f es **discontinua en** a si no es continua en a, es decir, si

$$\exists \varepsilon_0 > 0 \; \forall \delta > 0 \; \exists x \in A : \qquad |x - a| < \delta \quad \text{y} \quad |f(x) - f(a)| \geq \varepsilon_0.$$

Supongamos además que a es punto de acumulación de A. Diremos que la discontinuidad en a es **removible** si existe un límite finito

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

Como f es discontinua en a, necesariamente

$$L \neq f(a).$$

En ese caso, la función reparada

$$\widetilde{f}(x) = \begin{cases} f(x), & x \neq a, \\ L, & x = a, \end{cases}$$

es continua en a.

Si $a \notin A$, no diremos que f es discontinua **en** a, porque $f(a)$ no está definido. Si, sin embargo, a es punto de acumulación de A y existe un límite finito

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

diremos que a es un **agujero removible** para f: la extensión

$$\widehat{f}\colon A \cup \{a\} \to \mathbb{R}, \qquad \widehat{f}(x) = \begin{cases} f(x), & x \in A, \\ L, & x = a, \end{cases}$$

es continua en a.

La distinción del último párrafo es lógica, no terminológica solamente. Para preguntar si f es continua en a necesitamos, por definición, que a pertenezca al dominio. Si el punto falta, no tenemos todavía una función cuyo valor central pueda ser comparado con los valores vecinos. Lo que sí podemos preguntar es si el comportamiento perforado determina **qué valor habría que añadir** para obtener una extensión continua.

La reparación, cuando existe, es única. Supongamos primero que $a \in A$ y que la discontinuidad es removible. Tenemos

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

Cambiar el valor de f en el único punto a no altera este límite perforado. Por tanto,

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

Como además $\widetilde{f}(a) = L$,

Proposición 5.1 implica que $\widetilde{f}$ es continua en a.

Pero hay algo más: **ningún otro valor podría funcionar**. Si definiéramos otra función g que coincidiera con f fuera de a y asignáramos $g(a) = M$, entonces el comportamiento perforado seguiría dando

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

Si g fuese continua en a, Proposición 5.1 exigiría

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

Por unicidad del límite, $M = L$.

El límite perforado no solo nos dice que el defecto puede repararse: **fuerza el único valor que lo repara**. El mismo argumento vale para un agujero fuera del dominio. Si existe una extensión continua al punto faltante, el valor añadido debe ser exactamente el límite perforado.

Removible no significa que el límite falle
En una discontinuidad removible ocurre justamente lo contrario: el límite perforado **sí existe** y es finito. Lo que falla es la compatibilidad entre ese comportamiento y el valor central:

$$\lim_{x \to a} f(x) = L \qquad \text{pero} \qquad f(a) \neq L.$$

Por eso basta corregir un único valor.

Un defecto concentrado en un solo punto.

Ejemplo 5.3. **Reparar una discontinuidad removible y rellenar un agujero.** Definamos $f\colon \mathbb{R} \to \mathbb{R}$ por

$$f(x) = \begin{cases} \dfrac{x^2 - 1}{x - 1}, & x \neq 1, \\ 100, & x = 1. \end{cases}$$

Entonces f es discontinua en 1, pero su discontinuidad es removible. La única reparación continua consiste en sustituir $f(1) = 100$ por $f(1) = 2$.

Para $x \neq 1$ podemos factorizar $x^2 - 1 = (x - 1)(x + 1)$, y por tanto $f(x) = x + 1$. El comportamiento perforado es así completamente regular:

$$\lim_{x \to 1} f(x) = \lim_{x \to 1} (x + 1) = 2.$$

Sin embargo, $f(1) = 100$.

Como 1 es punto de acumulación de $\mathbb{R}$, Proposición 5.1 da inmediatamente

$$2 = \lim_{x \to 1} f(x) \neq f(1),$$

de modo que f no es continua en 1.

Definamos ahora

$$\widetilde{f}(x) = \begin{cases} \dfrac{x^2 - 1}{x - 1}, & x \neq 1, \\ 2, & x = 1. \end{cases}$$

Entonces

$$\lim_{x \to 1} \widetilde{f}(x) = 2 = \widetilde{f}(1),$$

y por tanto $\widetilde{f}$ es continua en 1. Podemos mirar el mismo fenómeno desde otra función. Sea

$$u \colon \mathbb{R} \setminus \{1\} \to \mathbb{R}, \qquad u(x) = \frac{x^2 - 1}{x - 1}.$$

No es correcto decir que u es discontinua en 1, porque $1 \notin \mathrm{Dom}(u)$. Sí es correcto afirmar que 1 es un agujero removible de u, porque

$$\lim_{x \to 1} u(x) = 2,$$

y la extensión que añade el valor 2 en el punto faltante es continua. El diagnóstico completo separa, por tanto, tres objetos que gráficamente podrían parecer similares:

- una función definida en el centro con el valor equivocado;
- una función cuyo dominio tiene un agujero;
- la extensión continua obtenida al introducir el valor forzado por el límite.

Cuando cambiar el centro no basta. El mecanismo anterior funciona porque **todo el comportamiento alrededor del punto ya converge hacia un único número**. Si ese número no existe, no hay valor central capaz de fabricar continuidad. Dos fallos canónicos ilustran esta situación. El primero es un **salto**: existen límites laterales finitos, pero quieren llegar a números diferentes. El segundo es una **oscilación persistente**: al acercarnos al punto podemos encontrar secuencias cuyos valores de salida se mantienen separados, de manera que ni siquiera hay un único candidato a límite perforado. No elevaremos estas palabras a una clasificación exhaustiva de todas las discontinuidades. Las usaremos como mecanismos diagnósticos concretos.

El valor central no puede arreglar el comportamiento perforado
Cambiar $f(a)$ afecta únicamente el punto a.
No cambia:

- los límites laterales;
- el límite perforado, si existe;
- las imágenes de sucesiones que se acercan a a evitando el centro.

Por eso una modificación puntual puede reparar una discontinuidad removible, pero no puede reconciliar dos límites laterales distintos ni detener una oscilación que ocurre en todos los entornos perforados.

Salto y oscilación bajo el mismo microscopio.

EJEMPLO 5.4. **Dos discontinuidades que no se reparan cambiando el valor central.** Consideremos, para un número arbitrario $c \in \mathbb{R}$, la función

$$J_c\colon \mathbb{R} \to \mathbb{R}, \qquad J_c(x) = \begin{cases} -1, & x < 0, \\ c, & x = 0, \\ 1, & x > 0, \end{cases}$$

y la función

$$O\colon \mathbb{R} \to \mathbb{R}, \qquad O(x) = \begin{cases} \sigma(1/x), & x \neq 0, \\ 0, & x = 0. \end{cases}$$

J_c tiene una discontinuidad de salto en 0 para **todo** valor de c. La función O tiene una discontinuidad por oscilación persistente en 0. Ninguna de las dos puede repararse modificando solamente el valor central.

El salto. Para J_c, los límites laterales son inmediatos:

$$\lim_{x\to 0^-} J_c(x) = -1, \qquad \lim_{x\to 0^+} J_c(x) = 1.$$

Ambos existen y son finitos, pero son distintos. Por Teorema 4.4, el límite bilateral

$$\lim_{x\to 0} J_c(x)$$

no existe.

Obsérvese que el número c no ha participado en absoluto en este cálculo. Podemos escoger

$$c = -1, \qquad c = 0, \qquad c = 1,$$

o cualquier otro real: los valores laterales siguen siendo -1 y 1. Como 0 es punto de acumulación de $\mathbb{R}$, si J_c fuese continua en 0, Proposición 5.1 forzaría la existencia del límite bilateral y su igualdad con $J_c(0)$. Pero ese límite no existe. Por tanto, J_c es discontinua en 0 para todo c. El defecto no está concentrado en el centro. Está en la incompatibilidad entre **dos comportamientos laterales**.

La oscilación. Para O, no tenemos dos valores laterales constantes entre los cuales escoger. La salida sigue alternando entre valores separados por una distancia fija arbitrariamente cerca de 0.

Consideremos las sucesiones

$$x_n = \frac{2}{(4n+1)\omega}, \qquad y_n = \frac{2}{(4n+3)\omega}, \qquad n \geq 0.$$

Ambas son positivas y satisfacen

$$x_n \to 0, \qquad y_n \to 0.$$

Además, $\frac{1}{x_n} = \frac{(4n+1)\omega}{2} = \frac{\omega}{2} + 2n\omega$, de modo que $O(x_n) = 1$ para todo n. Del mismo modo, $\frac{1}{y_n} = \frac{(4n+3)\omega}{2} = \frac{3\omega}{2} + 2n\omega$, y por tanto $O(y_n) = -1$ para todo n.

Así,

$$O(x_n) \to 1, \qquad O(y_n) \to -1.$$

Si O fuese continua en 0, el criterio secuencial Teorema 5.1 obligaría a que **toda** sucesión del dominio convergente a 0 tuviera imágenes convergentes a $O(0) = 0$. Las dos sucesiones construidas lo contradicen. Por tanto, O no es continua en 0. Más aún, cambiar $O(0)$ por cualquier número c tampoco ayudaría. Las sucesiones x_n e y_n evitan el centro y seguirían produciendo respectivamente los valores constantes 1 y -1. No existe un único número que pueda ser límite de ambas imágenes. El comportamiento perforado no se estabiliza en un valor. Por ello el defecto no puede describirse mediante dos límites laterales finitos diferentes.

Tres diagnósticos, tres estructuras distintas. Podemos reunir los tres mecanismos estudiados.

Mecanismo	Comportamiento perforado en a	Papel de $f(a)$	¿Se repara cambiando solo $f(a)$?
Removible	existe un único límite finito L	está ausente o no coincide con L	Sí; el único valor posible es L
Salto	los laterales finitos existen pero son distintos	no puede reconciliar los dos lados	No
Oscilación persistente	no existe un único límite; distintas sucesiones producen comportamientos incompatibles	no controla la oscilación perforada	No

La tabla sugiere un protocolo de diagnóstico que será útil en ejercicios y pruebas posteriores.

Primero: auditar el dominio. Si $a \notin A$, no hablamos todavía de continuidad de f en a; preguntamos por una posible extensión.

Segundo: si a es punto de acumulación, estudiar el comportamiento perforado. Puede bastar un límite bilateral, límites laterales o sucesiones testigo.

Tercero: solo si existe un límite finito único, compararlo con el valor central.

La cadena lógica es

dominio → comportamiento perforado → valor central → diagnóstico de continuidad.

Esta cadena impide un error frecuente: mirar primero $f(a)$ y tratar de adivinar continuidad a partir de ese único número. La continuidad es local, pero nunca es una propiedad de un punto aislado de la gráfica cuando alrededor del punto existen entradas del dominio.

Una discontinuidad puede tener un valor perfectamente definido. Nada en la definición de discontinuidad exige que la función tenga un «agujero» literal. Las funciones J_c y O están definidas en 0. También podemos escoger sus valores centrales de manera aparentemente razonable. Sin embargo, eso no altera el comportamiento de sus alrededores.

Por otra parte, una función con un punto ausente de su dominio puede poseer un comportamiento perforado tan regular que exista una extensión continua única. Así que debemos mantener separadas las afirmaciones

$$\boxed{\text{«la función está definida en } a\text{»}}$$

y

$$\boxed{\text{«la función es continua en } a\text{».}}$$

La primera es una condición de dominio. La segunda impone compatibilidad entre el centro y todos los modos permitidos de aproximación.

Por qué esta sección precede al teorema del valor intermedio. Los ejemplos anteriores muestran exactamente qué puede ocurrir cuando la continuidad falla. Una función con salto puede pasar de valores negativos a positivos sin tomar necesariamente los valores que quedan entre ambos lados. Una función oscilante puede visitar repetidamente regiones distintas sin estabilizarse. Una discontinuidad removible, en cambio, conserva un comportamiento perforado único y solo falla en el punto central. La pregunta natural es ahora la inversa:

$$\boxed{\text{si una función es continua en todo un intervalo,¿puede todavía saltarse un valor intermedio?}}$$

Responder esa pregunta ya no será un problema puramente local. Tendremos que combinar continuidad con la estructura de orden de $\mathbb{R}$ y, por primera vez en este capítulo, hacer entrar materialmente la **completitud** mediante un supremo. Ese será el contenido de §5.4: el teorema del valor intermedio.

5.4. De lo local a atravesar valores: teorema del valor intermedio

Por qué aparece este teorema. En §5.3 vimos tres maneras en que la continuidad puede fallar. Una discontinuidad removible deja intacto un único comportamiento perforado, pero un salto puede pasar de un nivel a otro sin recorrer los valores intermedios. Eso sugiere una pregunta que ya no es puramente local:

$$\boxed{\begin{array}{c}\text{si una función es continua en todos los puntos de un}\\ \text{intervalo,}\\ \text{¿puede saltarse un valor situado entre dos valores que}\\ \text{sí toma?}\end{array}}$$

La intuición gráfica dice que no. Si una curva pasa de una altura por debajo de y a otra por encima de y, parece inevitable que cruce la altura y. Pero esa imagen no es una demostración. Una gráfica solo representa el fenómeno; no explica qué propiedad de $\mathbb{R}$ obliga a que exista el punto de cruce. Aquí aparecerá por primera vez en este capítulo una interacción decisiva entre dos ideas construidas por separado:

$$\boxed{\text{continuidad local} + \text{completitud de } \mathbb{R} \Longrightarrow \text{existencia global de un valor intermedio.}}$$

El objetivo de esta sección no será únicamente conocer el teorema. Queremos entender **por qué la prueba necesita un supremo** y cómo la continuidad convierte ese supremo en un punto donde la función toma exactamente el valor buscado.

Exploración: ¿dónde debería estar el cruce? Supongamos provisionalmente que $f\colon [a,b] \to \mathbb{R}$ es continua y que $f(a) < y < f(b)$. Buscamos un punto $c \in (a,b)$ tal que $f(c) = y$. Una primera tentación sería buscar «el último punto antes de cruzar y». Pero esa frase es demasiado fuerte: no sabemos que f sea creciente, ni siquiera que los puntos donde $f(x) \leq y$ formen un solo intervalo. Podrían aparecer varias regiones del dominio en las que la función estuviera por debajo del nivel y. Lo que sí podemos hacer es reunir **todos** esos puntos: $S = \{x \in [a,b] : f(x) \leq y\}$.

¿Qué sabemos inmediatamente?

Como $f(a) < y$, tenemos $a \in S$.

Así que S no es vacío. Además, $S \subseteq [a,b]$, de modo que b es una cota superior de S. En este momento aparece una pregunta genuina:

¿necesitamos que S tenga un máximo?

No. Y de hecho todavía no tenemos derecho a afirmar que lo tenga. Lo que la completitud de $\mathbb{R}$ sí garantiza es algo más flexible: todo conjunto real no vacío y acotado superiormente posee un supremo. Por tanto existe $c = \sup S$. Este punto c es nuestro candidato al cruce. La elección no es arbitraria. El supremo captura una **frontera de orden**: ningún punto de S puede estar a la derecha de c, pero hay puntos de S tan próximos a c por la izquierda como queramos. Todavía falta el paso decisivo. Saber que c es una frontera del conjunto S no implica automáticamente que $f(c) = y$. La continuidad tendrá que convertir la frontera del dominio en una igualdad de valores.

Predecir la prueba antes de verla. Antes del enunciado formal, intentemos anticipar qué debería ocurrir.

Hay tres posibilidades:

$$f(c) < y, \qquad f(c) = y, \qquad f(c) > y.$$

Si $f(c) < y$, la continuidad debería obligar a que puntos un poco a la **derecha** de c sigan teniendo valores menores que y. Esos puntos pertenecerían a S y quedarían por encima de su supremo: imposible. Si $f(c) > y$, la continuidad debería obligar a que puntos suficientemente próximos a c tengan valores mayores que y. Pero la propiedad aproximativa del supremo exige puntos de S arbitrariamente próximos a c por la izquierda. Esos mismos puntos deberían satisfacer a la vez $f(x) \leq y$ y $f(x) > y$, otra imposibilidad.

Solo puede quedar $f(c) = y$.

Para completar la prueba basta coordinar esos controles en las desigualdades siguientes:

$$\boxed{S \xrightarrow{\text{completitud}} c = \sup S \xrightarrow{\text{continuidad}} f(c) = y.}$$

El enunciado formal.

TEOREMA 5.2. **Teorema del valor intermedio (Bolzano–Cauchy).** Sea

$$f\colon [a,b] \to \mathbb{R}$$

una función continua, con $a < b$. Si $y \in \mathbb{R}$ satisface

$$\min\{f(a), f(b)\} \leq y \leq \max\{f(a), f(b)\},$$

entonces existe $c \in [a,b]$ tal que

$$f(c) = y.$$

Si y está estrictamente entre $f(a)$ y $f(b)$, entonces puede elegirse

$$c \in (a, b).$$

Anatomía del teorema. Conviene separar con precisión sus componentes.

Premisa 1: el dominio es un intervalo cerrado. No se trata de dos puntos aislados. Todos los reales situados entre a y b pertenecen al dominio.

Premisa 2: f es continua en todo $[a, b]$. La condición local debe estar disponible precisamente en el punto frontera que construiremos.

Premisa 3: y está entre los valores extremos. El nivel buscado queda encerrado verticalmente entre $f(a)$ y $f(b)$.

Conclusión: existe al menos un punto del intervalo cuya imagen es y.

El tipo lógico es una afirmación de **existencia**. El teorema no afirma:

- que el punto c sea único;
- que podamos escribir una fórmula explícita para c;
- que f sea monótona;
- que podamos localizar inmediatamente c con una precisión prescrita.

Estas fronteras importan. El valor intermedio garantiza que el nivel se alcanza; no resuelve por sí solo el problema de encontrar dónde ni cuántas veces.

Estrategia de la demostración. La prueba formal tendrá cinco bloques.

1. Reducir al caso verdaderamente no trivial $f(a) < y < f(b)$.
2. Construir $S = \{x \in [a, b] : f(x) \leq y\}$.
3. Usar completitud para definir $c = \sup S$.
4. Descartar $f(c) < y$ extendiendo S hacia la derecha mediante continuidad.
5. Descartar $f(c) > y$ combinando continuidad con puntos de S arbitrariamente próximos al supremo.

Lo esencial es distinguir la **estrategia global** de las estimaciones locales. La estrategia es fabricar una frontera mediante orden y completitud. Las desigualdades ε–δ solo ejecutarán el mecanismo que impide que la imagen de esa frontera quede por encima o por debajo del nivel objetivo.

Demostración. Demostración. Si $y = f(a)$ o $y = f(b)$, la conclusión es inmediata tomando respectivamente $c = a$ o $c = b$.

Supongamos entonces que y está estrictamente entre los valores extremos.

Comenzaremos por el caso $f(a) < y < f(b)$.

Definimos $S = \{x \in [a, b] : f(x) \leq y\}$.

Como $f(a) < y$, tenemos $a \in S$, así que S es no vacío. Además, $S \subseteq [a, b]$, por lo que está acotado superiormente por b. Por la completitud de $\mathbb{R}$, existe $c = \sup S$.

Demostraremos que $f(c) = y$.

Primer descarte: $f(c) < y$ es imposible. Supongamos, para obtener una contradicción, que $f(c) < y$.

Tomemos la tolerancia positiva $\varepsilon = \frac{y - f(c)}{2} > 0$.

Como f es continua en c, existe $\delta_0 > 0$ tal que para todo $x \in [a,b]$, $|x-c| < \delta_0 \Longrightarrow |f(x)-f(c)| < \varepsilon$. Además, $c \neq b$, porque de ser $c = b$ tendríamos $f(c) = f(b) > y$, contrariamente a la suposición $f(c) < y$.

Por tanto, $b - c > 0$.

Definamos

$$h = \text{mín}\left\{\frac{\delta_0}{2}, \frac{b-c}{2}\right\} > 0$$

y consideremos $x = c + h$.

Entonces $c < x < b$ y $|x - c| = h < \delta_0$.

Por continuidad, $|f(x) - f(c)| < \varepsilon$, de donde

$$f(x) < f(c) + \varepsilon = f(c) + \frac{y - f(c)}{2} = \frac{f(c) + y}{2} < y.$$

Así, $x \in S$.

Pero $x > c$, lo que contradice que c sea una cota superior de S.

Por tanto, $f(c) \not< y$.

Segundo descarte: $f(c) > y$ es imposible. Supongamos ahora que $f(c) > y$.

Tomemos $\varepsilon = \frac{f(c)-y}{2} > 0$.

Por continuidad en c, existe $\delta_1 > 0$ tal que

$$x \in [a,b], \qquad |x - c| < \delta_1 \Longrightarrow |f(x) - f(c)| < \varepsilon.$$

Como $c = \sup S$, la caracterización aproximativa del supremo garantiza que existe $s \in S$ tal que $c - \delta_1 < s \leq c$. Pero bajo la suposición $f(c) > y$ tenemos $c \notin S$. Por consiguiente, $s < c$, y en particular $|s - c| < \delta_1$.

La continuidad da entonces $|f(s) - f(c)| < \varepsilon$.

Por la desigualdad anterior,

$$f(s) > f(c) - \varepsilon = f(c) - \frac{f(c) - y}{2} = \frac{f(c) + y}{2} > y.$$

Sin embargo, $s \in S$ significa precisamente $f(s) \leq y$.

Hemos obtenido una contradicción. Por tanto, $f(c) \not> y$. Al no poder ocurrir ni $f(c) < y$ ni $f(c) > y$, la tricotomía del orden real obliga a que $f(c) = y$.

Esto demuestra el caso $f(a) < y < f(b)$.

Si en cambio $f(a) > y > f(b)$, consideramos la función $-f$, que es continua por Proposición 5.2. Entonces $-f(a) < -y < -f(b)$, y el caso ya demostrado proporciona $c \in (a,b)$ tal que $-f(c) = -y$.

Por tanto, $f(c) = y$.

Queda demostrado el teorema. ■

Lectura de la demostración: qué hizo realmente cada pieza. La prueba contiene varias líneas, pero su motor es mucho más compacto.

1. ¿Por qué se introdujo S? Porque queríamos transformar una pregunta sobre valores de la función en una pregunta de orden sobre puntos del dominio. El conjunto $S = \{x : f(x) \leq y\}$ codifica todos los lugares donde todavía estamos a la altura y o por debajo de ella.

2. ¿Por qué un supremo y no un máximo? Porque nada garantiza de antemano que exista un «último» punto de S. El supremo existe por completitud aunque no sepamos todavía si pertenece a S. Precisamente la continuidad terminará mostrando que su imagen es y.

3. ¿Dónde entra exactamente la completitud? En una sola transición crucial:

$$S \neq \varnothing, \qquad S \text{ acotado superiormente} \Longrightarrow \exists \sup S \in \mathbb{R}.$$

Todo lo que sigue utiliza orden, la propiedad aproximativa del supremo y continuidad.

4. ¿Qué hace la continuidad? Convierte una desigualdad estricta en el valor central en una desigualdad del mismo signo en todo un pequeño entorno. Si $f(c) < y$, crea puntos de S a la derecha de c. Si $f(c) > y$, crea una zona alrededor de c donde ningún punto puede pertenecer a S. Ambas posibilidades contradicen el significado de $c = \sup S$. La idea puede resumirse así:

$$\boxed{\text{supremo} = \text{frontera de orden}, \qquad \text{continuidad} = \text{compatibilidad local},}$$

$$\boxed{\text{frontera de orden} + \text{compatibilidad local} \Longrightarrow \text{nivel alcanzado}.}$$

Prueba de estrés de las hipótesis. Un teorema central se comprende mejor cuando sabemos qué se rompe al retirar sus hipótesis.

Sin continuidad: un salto puede evitar el valor intermedio. Consideremos en $[-1, 1]$ la función

$$J(x) = \begin{cases} -1, & x < 0, \\ 1, & x \geq 0. \end{cases}$$

Tenemos

$$J(-1) = -1, \qquad J(1) = 1,$$

de modo que 0 está entre los valores de los extremos.

Sin embargo, $J(x) \neq 0$ para todo $x \in [-1, 1]$.

El fallo no es misterioso: J no es continua en 0. Este es exactamente el tipo de salto estudiado en §5.3.

Sin estructura de intervalo: continuidad relativa no basta. Tomemos el dominio $A = \{-1, 1\}$ y la función

$$f \colon A \to \mathbb{R}, \qquad f(x) = x.$$

Los dos puntos del dominio son aislados, así que f es continua en ambos por Proposición 5.1.

Además,

$$f(-1) = -1, \qquad f(1) = 1.$$

Pero no existe $c \in A$ con $f(c) = 0$.

La continuidad puntual no crea puntos que faltan en el dominio. La geometría de intervalo es una parte real del teorema.

Si reemplazamos $\mathbb{R}$ por $\mathbb{Q}$: el análogo falla sin completitud. Intentemos repetir el fenómeno en el cuerpo ordenado $\mathbb{Q}$. Allí el análogo del intervalo cerrado $[1,2]$ es $\mathbb{Q} \cap [1,2]$. La función $q \longmapsto q^2 - 2$ es continua relativamente a ese dominio racional. Además,

$$1^2 - 2 = -1 < 0, \qquad 2^2 - 2 = 2 > 0.$$

Pero no existe ningún racional q tal que $q^2 - 2 = 0$, como demostramos al comienzo de `T1-C02`. El nivel cero está «entre» los valores extremos, pero el punto de cruce sería $\sqrt{2}$, que no pertenece a $\mathbb{Q}$. Este ejemplo revela por qué la completitud no es una decoración técnica de la prueba. El teorema es una manifestación concreta de que la recta real **no tiene ese tipo de huecos**.

Primer corolario: cambio de signo implica raíz. La forma más utilizada del teorema toma como nivel objetivo $y = 0$.

Corolario 5.2. **Teorema de Bolzano para cambio de signo.** Sea

$$f\colon [a,b] \to \mathbb{R}$$

continua. Si

$$f(a)f(b) < 0,$$

entonces existe $c \in (a,b)$ tal que

$$f(c) = 0.$$

Demostración. La desigualdad $f(a)f(b) < 0$ significa que $f(a)$ y $f(b)$ tienen signos opuestos. Por tanto 0 está estrictamente entre ambos valores. Aplicando Teorema 5.2 con $y = 0$, obtenemos un punto $c \in (a,b)$ tal que $f(c) = 0$.

■

Variante con extremos
Si solo sabemos $f(a)f(b) \leq 0$, la misma idea garantiza una raíz en $[a,b]$. Puede ocurrir que la raíz sea uno de los extremos.

Segundo corolario: la imagen continua de un intervalo es un intervalo. El teorema puede reformularse sin mencionar un nivel particular.

Corolario 5.3. **La imagen continua de un intervalo es un intervalo.** Sea $I \subseteq \mathbb{R}$ un intervalo y sea

$$f\colon I \to \mathbb{R}$$

continua. Entonces

$$f(I)$$

es un intervalo.

Demostración. Tomemos dos valores $u, v \in f(I)$ con $u < v$.

Queremos demostrar que todo número $y \in [\,u, v\,]$ pertenece también a $f(I)$. Por definición de imagen, existen $x_u, x_v \in I$ tales que

$$f(x_u) = u, \qquad f(x_v) = v.$$

Como I es un intervalo, contiene todos los puntos situados entre x_u y x_v. Si definimos

$$\alpha = \min\{x_u, x_v\}, \qquad \beta = \max\{x_u, x_v\},$$

entonces $[\alpha, \beta] \subseteq I$.

La restricción de f a $[\alpha, \beta]$ sigue siendo continua. El número y está entre los valores de f en los extremos x_u y x_v, aunque estos puntos puedan aparecer en cualquiera de los dos órdenes dentro del dominio. Por Teorema 5.2 existe $c \in [\alpha, \beta]$ tal que $f(c) = y$.

Como $c \in I$, concluimos que $y \in f(I)$.

Por tanto, $f(I)$ contiene junto con cualesquiera dos de sus puntos todos los reales situados entre ellos. Esa es precisamente la propiedad de ser un intervalo. ■

Este corolario transforma una afirmación de existencia puntual en una afirmación estructural:

$$\boxed{I \text{ intervalo} + f \text{ continua} \Longrightarrow f(I) \text{ intervalo.}}$$

No afirma todavía que la imagen sea cerrada ni acotada. Esas propiedades requerirán hipótesis adicionales y aparecerán más adelante mediante compactitud y el teorema de los valores extremos.

Ejemplo: demostrar existencia sin resolver la ecuación.

Ejemplo 5.5. **Una raíz garantizada sin fórmula explícita.** Consideremos la ecuación $x^5 + x - 1 = 0$.

La función $p(x) = x^5 + x - 1$ es un polinomio y, por Corolario 5.1, es continua en $\mathbb{R}$.

Además, $p(0) = -1 < 0$ y $p(1) = 1 > 0$.

Por Corolario 5.2 existe al menos un número $c \in (0, 1)$ tal que $c^5 + c - 1 = 0$.

El resultado es genuinamente informativo aunque no hayamos calculado c. Podemos incluso mejorar inmediatamente el encierro con una sola evaluación:

$$p\left(\frac{1}{2}\right) = \frac{1}{32} + \frac{1}{2} - 1 = -\frac{15}{32} < 0,$$

mientras $p(1) > 0$.

Por Bolzano existe una raíz en

$$\left(\frac{1}{2}, 1\right).$$

Pero debemos distinguir cuidadosamente tres preguntas:

1. **¿existe una raíz?** — el teorema del valor intermedio responde que sí;
2. **¿es única?** — el IVT por sí solo no lo decide;
3. **¿cómo aproximarla con error controlado?** — esa será la tarea de la bisección en §5.5.

Variación: convertir un nivel cualquiera en un problema de raíces. Buscar un punto donde $f(x) = y$ es equivalente a buscar una raíz de $g(x) = f(x) - y$. Si f es continua, Proposición 5.2 implica que g también lo es. Además, $f(a) < y < f(b)$ equivale a $g(a) < 0 < g(b)$. Así, el teorema del valor intermedio y el teorema de Bolzano son dos lecturas del mismo mecanismo:

$$\boxed{\text{alcanzar un nivel } y \longleftrightarrow \text{hacer cero la función } f - y.}$$

Esta traducción será útil repetidamente: muchos problemas de existencia pueden convertirse en problemas de encontrar ceros de una función continua.

Reconstrucción de la prueba. Sin releer la demostración completa, conviene poder responder estas preguntas.

1. ¿Por qué el conjunto $S = \{x : f(x) \leq y\}$ es no vacío y está acotado?
2. ¿Qué información da $c = \sup S$ que no necesitaría dar un máximo?
3. Si $f(c) < y$, ¿qué punto nuevo construye la continuidad y por qué contradice que c sea cota superior?
4. Si $f(c) > y$, ¿qué punto proporciona la caracterización aproximativa del supremo y por qué entra en conflicto con la continuidad?
5. ¿En qué línea exacta dejaría de funcionar la prueba si trabajáramos solamente en $\mathbb{Q}$?

La pregunta central de autoexplicación es:

> ¿qué tenía que ocurrir para que la frontera de orden $c = \sup S$ se convirtiera en un punto de cruce?

La respuesta es: la continuidad debía impedir que la función permaneciera estrictamente a un lado del nivel y en un entorno de esa frontera.

Qué hemos conectado. Esta sección realiza el primer gran paso local $\to$ global del capítulo. La cadena completa es

$$\boxed{\begin{array}{c}\text{orden en } \mathbb{R} + \text{completitud} \longrightarrow \text{frontera } c = \sup S,\\ \text{continuidad en } c \longrightarrow \text{la frontera no puede quedar por encima ni por debajo de } y,\\ \therefore\ f(c) = y.\end{array}}$$

De aquí obtenemos tres formas equivalentes de pensar el fenómeno:

- **nivel:** una función continua atraviesa todos los valores entre dos valores que ya toma;
- **raíz:** una función continua que cambia de signo debe anularse;
- **imagen:** la imagen continua de un intervalo vuelve a ser un intervalo.

El teorema todavía es existencial. Sabe que el punto está allí, pero no nos entrega un procedimiento cuantitativo para encerrarlo con precisión creciente. La siguiente sección añadirá precisamente esa dimensión constructiva: conservaremos un cambio de signo y construiremos una cadena de intervalos cerrados cada vez más pequeños. Así, el IVT dejará de ser solo una garantía de existencia y se convertirá en el punto de partida del **método de bisección**.

5.5. Bisección: convertir existencia en localización

Por qué aparece: existir no es todavía localizar. El teorema de Bolzano nos permite pasar de un cambio de signo a una afirmación de existencia. Si una función continua satisface $f(a) < 0 < f(b)$, sabemos que existe al menos un punto $c \in (a, b)$ con $f(c) = 0$. Pero el enunciado

no nos dice dónde está ese punto dentro del intervalo. Ésta es una diferencia conceptual importante. Podemos distinguir tres preguntas:

existencia	:	¿hay una raíz?
localización	:	¿en qué intervalo sabemos que está?
precisión	:	¿qué tan pequeño es ese intervalo?

§5.4 respondió la primera. Ahora queremos responder simultáneamente las otras dos.

La idea inicial es elemental: si una raíz está garantizada entre dos puntos cuyos valores tienen signos opuestos, cortemos el intervalo por la mitad y conservemos una mitad donde siga habiendo signos opuestos.

Después repetimos.

La pregunta que debe guiar la sección es:

¿qué información debemos conservar en cada corte para no perder la raíz?

La respuesta será un **invariante**.

Un corte antes de formalizar el proceso. Supongamos $f(a_0) < 0 < f(b_0)$ y definamos el punto medio $m_0 = \frac{a_0+b_0}{2}$. Por tricotomía, exactamente una de las siguientes posibilidades ocurre:

1. $f(m_0) = 0$;
2. $f(m_0) < 0$;
3. $f(m_0) > 0$.

En el primer caso hemos encontrado una raíz exacta. Si $f(m_0) < 0$, entonces $f(m_0) < 0 < f(b_0)$, y podemos descartar la mitad izquierda y conservar $[m_0, b_0]$. Si, en cambio, $f(m_0) > 0$, entonces $f(a_0) < 0 < f(m_0)$, y conservamos $[a_0, m_0]$. En cualquiera de los dos casos no terminales han ocurrido dos cosas a la vez:

el intervalo se ha reducido a la mitad + el cambio de signo se ha conservado.

La segunda propiedad es la que mantiene viva la información de existencia; la primera es la que aporta precisión cuantitativa.

El invariante que sostiene todo el método. Si el procedimiento no se ha detenido, queremos construir intervalos $I_n = [a_n, b_n]$ tales que, para todo n, $f(a_n) < 0 < f(b_n)$. Éste será el invariante de signos.

Al mismo tiempo exigiremos $I_{n+1} \subseteq I_n$ y $b_n - a_n = \frac{b_0-a_0}{2^n}$.

La segunda igualdad muestra que el tamaño del intervalo no depende de la forma concreta de la gráfica. Cada paso divide la incertidumbre horizontal exactamente por 2. Conviene anticipar ya una consecuencia: $\frac{b_0-a_0}{2^n} \longrightarrow 0$. Por tanto, si conseguimos mantener el proceso indefinidamente, los intervalos encajados no podrán conservar un segmento de longitud positiva en común. Deberán concentrarse en un único punto. Pero todavía queda una cuestión esencial:

¿por qué ese único punto tendrá que ser una raíz de f?

La respuesta no vendrá de la geometría de los intervalos por sí sola. Allí entrará la continuidad.

Anatomía del teorema. Antes de demostrarlo, separemos con precisión sus piezas.

Hipótesis. Tenemos una función continua

$$f\colon [a_0, b_0] \to \mathbb{R}, \qquad a_0 < b_0,$$

con $f(a_0)f(b_0) \leq 0$.

Si uno de los valores extremos es cero, ya tenemos una raíz. El caso genuino de bisección es, por tanto, $f(a_0)f(b_0) < 0$. Multiplicando f por -1 si fuese necesario, podemos orientar los signos como $f(a_0) < 0 < f(b_0)$.

Conclusión. O bien algún punto medio es ya una raíz y el procedimiento termina en un número finito de pasos, o bien obtenemos una cadena infinita de intervalos encajados que determina un único punto c, y ese punto satisface $f(c) = 0$.

Además, la longitud de los intervalos proporciona una cota explícita del error.

Tipo lógico. No es un teorema de unicidad de raíces. La función puede tener varias raíces en $[a_0, b_0]$. Lo que será único es el punto común de **la cadena concreta de intervalos que construyamos** cuando sus longitudes tienden a cero.

Este matiz evita una confusión frecuente:

unicidad del punto localizado por una cadena $\neq$ unicidad global de la raíz de f.

Teorema 5.3. **Teorema de bisección.** Sea

$$f\colon [a_0, b_0] \to \mathbb{R}$$

continua, con $a_0 < b_0$ y

$$f(a_0)f(b_0) \leq 0.$$

Entonces existe $c \in [a_0, b_0]$ tal que

$$f(c) = 0.$$

En el caso no trivial $f(a_0)f(b_0) < 0$, orientemos los signos —reemplazando f por $-f$ si es necesario— de modo que

$$f(a_0) < 0 < f(b_0).$$

Definamos recursivamente, mientras no aparezca una raíz exacta,

$$m_n = \frac{a_n + b_n}{2}.$$

- Si $f(m_n) = 0$, el proceso termina y $c = m_n$.
- Si $f(m_n) < 0$, tomamos
$$[a_{n+1}, b_{n+1}] = [m_n, b_n].$$
- Si $f(m_n) > 0$, tomamos
$$[a_{n+1}, b_{n+1}] = [a_n, m_n].$$

Si el proceso no termina en un número finito de pasos, los intervalos

$$I_n = [a_n, b_n]$$

satisfacen, para todo n,

$$I_{n+1} \subseteq I_n,$$

$$f(a_n) < 0 < f(b_n),$$

y

$$b_n - a_n = \frac{b_0 - a_0}{2^n}.$$

Existe entonces un único punto

$$\{c\} = \bigcap_{n=0}^{\infty} I_n,$$

y se cumple

$$a_n \to c, \qquad b_n \to c, \qquad f(c) = 0.$$

Además, para el punto medio m_n de I_n,

$$|m_n - c| \leq \frac{b_n - a_n}{2} = \frac{b_0 - a_0}{2^{n+1}}.$$

Demostración. Si $f(a_0)f(b_0) = 0$, uno de los extremos es raíz y no hay nada que demostrar. Supongamos, pues, $f(a_0)f(b_0) < 0$. Tras reemplazar f por $-f$ si es necesario, podemos suponer $f(a_0) < 0 < f(b_0)$. Construimos los intervalos según la regla del enunciado. Si para algún n ocurre $f(m_n) = 0$, hemos encontrado una raíz y la prueba termina. Supongamos desde ahora que esto nunca ocurre.

1. El invariante de signos se conserva.

Fijemos n y supongamos $f(a_n) < 0 < f(b_n)$.

Como $f(m_n) \neq 0$, solo hay dos posibilidades.

Si $f(m_n) < 0$, definimos

$$a_{n+1} = m_n, \qquad b_{n+1} = b_n,$$

y entonces $f(a_{n+1}) < 0 < f(b_{n+1})$.

Si $f(m_n) > 0$, definimos

$$a_{n+1} = a_n, \qquad b_{n+1} = m_n,$$

y nuevamente $f(a_{n+1}) < 0 < f(b_{n+1})$.

Por inducción, el cambio de signo queda preservado en todos los intervalos.

2. Los intervalos están encajados y su longitud se divide por dos.

Cada I_{n+1} es una de las dos mitades cerradas de I_n. Por tanto, $I_{n+1} \subseteq I_n$ y $b_{n+1} - a_{n+1} = \frac{b_n - a_n}{2}$.

Iterando, $b_n - a_n = \frac{b_0 - a_0}{2^n}$.

Como $2^{-n} \to 0$, obtenemos $b_n - a_n \to 0$.

3. Los intervalos determinan un único punto.

Los I_n son cerrados, no vacíos y encajados. Por el principio de intervalos cerrados encajados demostrado en `T1-C02`, existe al menos un punto $c \in \bigcap_{n=0}^{\infty} I_n$. La contracción de las longitudes fuerza la unicidad. En efecto, si c y d pertenecieran a todos los intervalos, entonces para todo n, $|c-d| \le b_n - a_n$. El lado derecho tiende a 0, de modo que necesariamente $|c-d| = 0$, y por tanto $c = d$.

Así, $\bigcap_{n=0}^{\infty} I_n = \{c\}$.

4. Los extremos convergen al punto común.

Como $c \in [a_n, b_n]$, $0 \le c - a_n \le b_n - a_n$ y $0 \le b_n - c \le b_n - a_n$.

El término de la derecha tiende a 0. Por el teorema del sándwich, $a_n \to c$ y $b_n \to c$.

5. La continuidad convierte el punto localizado en una raíz.

Hasta aquí solo hemos usado orden, bisección e intervalos encajados. Todavía no hemos demostrado que $f(c) = 0$.

Como f es continua en c y

$$a_n \to c, \qquad b_n \to c,$$

el criterio secuencial de continuidad Teorema 5.1 da $f(a_n) \to f(c)$ y $f(b_n) \to f(c)$. Pero el invariante nos asegura para todo n que $f(a_n) < 0 < f(b_n)$. La preservación débil del orden al pasar al límite implica $f(c) \le 0$ y $f(c) \ge 0$.

Por tanto, $f(c) = 0$.

6. Cota del error del punto medio.

Como $c \in [a_n, b_n]$ y $m_n = \frac{a_n + b_n}{2}$, la distancia máxima desde el punto medio a cualquier punto del intervalo es la mitad de su longitud. Luego $|m_n - c| \le \frac{b_n - a_n}{2} = \frac{b_0 - a_0}{2^{n+1}}$.

Queda demostrado el teorema. ■

Leer la prueba por capas. La demostración anterior contiene varias ideas distintas. Conviene no reducirla a «repetir puntos medios». La **estrategia** es conservar una propiedad que garantiza la posibilidad de una raíz mientras reducimos geométricamente la región donde puede estar.

La **arquitectura** es

cambio de signo
$\to$ intervalos encajados
$\to$ diámetro $\to 0$
$\to$ punto único c
$\to$ continuidad
$\to f(c) = 0.$

El **mecanismo de existencia del punto** c es el principio de intervalos cerrados encajados, que descansa finalmente en la completitud de $\mathbb{R}$. El **mecanismo que convierte** c **en raíz** es distinto: continuidad + convergencia de los extremos + preservación del orden en límites. La **técnica cuantitativa** es simplemente $|I_n| = 2^{-n}|I_0|$. Y el **cierre** de la prueba no es que los intervalos sean pequeños. Es la igualdad $f(c) = 0$.

Dónde se usa realmente la continuidad
La construcción de los intervalos puede ejecutarse usando únicamente comparaciones de signos. La continuidad no es necesaria para partir cada intervalo por la mitad. Es necesaria al final, cuando queremos concluir que el punto común de los intervalos no puede conservar un valor estrictamente negativo ni estrictamente positivo. Esta separación será útil para entender el contraejemplo siguiente.

Prueba de estrés: sin continuidad podemos localizar un salto, no una raíz. Consideremos $h\colon [0,1] \to \mathbb{R}$,

$$h(x) = \begin{cases} -1, & x < \frac{1}{3}, \\ 1, & x \geq \frac{1}{3}. \end{cases}$$

Tenemos

$$h(0) = -1 < 0 < 1 = h(1).$$

Podemos aplicar mecánicamente la regla de partir el intervalo y conservar una mitad cuyos extremos tengan signos opuestos. Los intervalos se harán cada vez más pequeños y se concentrarán alrededor de $c = \frac{1}{3}$.

Sin embargo,

$$h\left(\frac{1}{3}\right) = 1,$$

y la función no tiene ninguna raíz.

¿Qué falló?

No falló la bisección geométrica. Tampoco falló el encajamiento. Falló exactamente el paso

$$a_n, b_n \to c \quad \Longrightarrow \quad h(a_n), h(b_n) \to h(c),$$

porque h no es continua en c. Este ejemplo muestra que la hipótesis de continuidad no es ornamental: es la pieza que impide que el proceso converja hacia un **salto** en lugar de hacia un cero.

Un ejemplo completo: localizar una raíz de $x^3 + x - 1$.

Ejemplo 5.6. **Bisección de una raíz con error controlado.** Consideremos $p(x) = x^3 + x - 1$.

Como p es un polinomio, es continua en $\mathbb{R}$. Además, $p(0) = -1 < 0$ y $p(1) = 1 > 0$. Por Bolzano existe al menos una raíz en $(0,1)$. Aplicaremos bisección para localizar una de ellas.

Partimos de $I_0 = [0,1]$.

Los primeros cortes son:

n	$I_n = [a_n, b_n]$	m_n	signo de $p(m_n)$	intervalo conservado
0	$[0,1]$	$\frac{1}{2}$	$p(\frac{1}{2}) = -\frac{3}{8} < 0$	$[\frac{1}{2}, 1]$
1	$[\frac{1}{2}, 1]$	$\frac{3}{4}$	$p(\frac{3}{4}) = \frac{11}{64} > 0$	$[\frac{1}{2}, \frac{3}{4}]$
2	$[\frac{1}{2}, \frac{3}{4}]$	$\frac{5}{8}$	$p(\frac{5}{8}) = -\frac{67}{512} < 0$	$[\frac{5}{8}, \frac{3}{4}]$

n	$I_n = [a_n, b_n]$	m_n	signo de $p(m_n)$	intervalo conservado
3	$[\frac{5}{8}, \frac{3}{4}]$	$\frac{11}{16}$	$p(\frac{11}{16}) = \frac{51}{4096} > 0$	$[\frac{5}{8}, \frac{11}{16}]$
4	$[\frac{5}{8}, \frac{11}{16}]$	$\frac{21}{32}$	$p(\frac{21}{32}) = -\frac{2003}{32768} < 0$	$[\frac{21}{32}, \frac{11}{16}]$
5	$[\frac{21}{32}, \frac{11}{16}]$	$\frac{43}{64}$	$p(\frac{43}{64}) = -\frac{6509}{262144} < 0$	$[\frac{43}{64}, \frac{11}{16}]$
6	$[\frac{43}{64}, \frac{11}{16}]$	$\frac{87}{128}$	$p(\frac{87}{128}) = -\frac{13241}{2097152} < 0$	$[\frac{87}{128}, \frac{11}{16}]$
7	$[\frac{87}{128}, \frac{11}{16}]$	$\frac{175}{256}$	$p(\frac{175}{256}) = \frac{50959}{16777216} > 0$	$[\frac{87}{128}, \frac{175}{256}]$

Después de ocho decisiones sabemos, por tanto, que la raíz localizada c satisface $\frac{87}{128} \le c \le \frac{175}{256}$. Como $\frac{87}{128} = \frac{174}{256}$, el intervalo final tiene longitud $\frac{175}{256} - \frac{174}{256} = \frac{1}{256}$. Si usamos su punto medio $q = \frac{349}{512}$, entonces $|q - c| \le \frac{1}{512}$. Numéricamente, $q = 0.681640625$, por lo que ya tenemos una aproximación con error absoluto garantizado inferior a 0.001954. Lo importante no es el decimal concreto. Es que la garantía de error proviene **antes** de conocer la raíz:

$$\boxed{\text{número de bisecciones} \Longrightarrow \text{longitud del intervalo} \Longrightarrow \text{cota del error.}}$$

Cuántos pasos hacen falta para una precisión dada. Si el intervalo inicial tiene longitud $L = b_0 - a_0$, tras n bisecciones el intervalo tiene longitud $\frac{L}{2^n}$. Usando su punto medio como aproximación, el error satisface $|m_n - c| \le \frac{L}{2^{n+1}}$. Por tanto, para garantizar un error menor que una tolerancia $\varepsilon > 0$, basta elegir n de modo que $\frac{L}{2^{n+1}} < \varepsilon$.

Equivalentemente, $2^{n+1} > \frac{L}{\varepsilon}$.

No necesitamos saber de antemano dónde está la raíz. Solo necesitamos decidir cuántas veces reducir la incertidumbre a la mitad. Ésta es una forma especialmente limpia de **velocidad de localización**: el ancho del intervalo decrece geométricamente con razón $1/2$.

Existencia, localización y cálculo efectivo no son la misma afirmación. Hay una sutileza conceptual que conviene conservar. El teorema de Bolzano afirma clásicamente $\exists c \in [a, b] : f(c) = 0$. El teorema de bisección añade una construcción recursiva de intervalos, siempre que en cada paso podamos decidir cuál de las alternativas

$$f(m_n) < 0, \qquad f(m_n) = 0, \qquad f(m_n) > 0$$

se verifica.

En matemática clásica, la tricotomía del orden real permite formular esa selección sin dificultad. Sin embargo, si queremos convertir el esquema en un **algoritmo numérico implementable** para una función dada por algún procedimiento de cálculo, necesitamos además poder evaluar el signo de $f(m_n)$ con suficiente fiabilidad. El teorema abstracto de continuidad no suministra por sí solo un método computacional para hacerlo.

Así conviene distinguir:

existencia matemática ≠ procedimiento de localización ≠ implementación computacional concreta.

La bisección acerca mucho más las tres ideas que el IVT puro, pero no las identifica lógicamente.

Otras hipótesis bajo presión. ¿Es necesario que haya cambio de signo?

Es una condición suficiente para esta versión del método, no una condición necesaria para que exista una raíz. Por ejemplo, $f(x) = x^2$ tiene una raíz en 0, pero en $[-1, 1]$ los extremos tienen el mismo signo.

Por otra parte, tener el mismo signo en los extremos tampoco garantiza que haya una raíz: para $g(x) = x^2 + 1$ no existe ninguna. Por tanto, $f(a)f(b) > 0$ no permite decidir por sí solo si hay ceros dentro del intervalo.

¿Basta con que los intervalos estén encajados?

No para obtener un punto único. La cadena constante $[0, 1] \supseteq [0, 1] \supseteq [0, 1] \supseteq \cdots$ tiene una intersección completa igual a $[0, 1]$. La propiedad decisiva de la bisección es que $|I_n| \to 0$. El encajamiento produce compatibilidad entre las localizaciones; la contracción produce precisión y unicidad del punto común.

Reconstrucción de la demostración. Sin releer el teorema, el lector debería poder reconstruir la prueba respondiendo estas preguntas:

1. ¿qué propiedad debe mantenerse después de cada corte para conservar la posibilidad de una raíz?
2. ¿por qué cada paso reduce exactamente a la mitad la longitud?
3. ¿qué resultado anterior garantiza que todos los intervalos encajados tienen un punto común?
4. ¿por qué la condición $|I_n| \to 0$ hace único ese punto?
5. ¿cómo se demuestra que $a_n \to c$ y $b_n \to c$ sin invocar Bolzano–Weierstrass?
6. ¿en qué paso exacto entra la continuidad?
7. ¿por qué el contraejemplo escalonado conserva la geometría de la bisección pero destruye la conclusión $f(c) = 0$?
8. ¿de dónde sale la cota $|m_n - c| \leq \frac{b_0 - a_0}{2^{n+1}}$?

La pregunta de autoexplicación central es:

> ¿qué tenía que ocurrir para transformar una garantía cualitativa de existencia en una localización cuantitativa de la raíz?

La respuesta tiene dos partes:

preservar un invariante + hacer tender a cero la escala de incertidumbre.

Qué hemos ganado. El IVT y la bisección cumplen funciones distintas y complementarias. El IVT dice:

«una raíz debe existir».

La bisección añade:

«podemos encerrarla en intervalos cuya longitud conocemos exactamente».

La cadena conceptual queda ahora

$$\boxed{\begin{array}{c}\text{continuidad + cambio de signo}\\ \to \text{existencia}\\ \to \text{invariante de signos}\\ \to \text{intervalos encajados}\\ \to \text{contracción geométrica}\\ \to \text{raíz localizada con error controlado.}\end{array}}$$

Hasta aquí hemos transformado información local y completitud en existencia y localización. La siguiente sección dará otro salto: ya no seguiremos una sola cadena de intervalos, sino una familia posiblemente infinita de entornos. El problema será entender cuándo podemos reemplazar una cobertura infinita por **finitos** controles. Eso nos llevará a recubrimientos abiertos y a la compactitud de $[a, b]$.

5.6. Recubrimientos abiertos y compactitud de un intervalo cerrado

De muchos controles locales a un control finito. La continuidad es una propiedad local. Si f es continua en cada punto de $[a, b]$, entonces para cada centro x y para cada tolerancia adecuada podemos encontrar algún entorno alrededor de x en el que la función queda controlada. Pero hay una dificultad inmediata. El intervalo contiene infinitos puntos. Por tanto, una afirmación de la forma

$$\text{«para cada } x \in [a, b] \text{ existe un entorno } V_{r_x}(x)\text{»}$$

puede producir una familia infinita de radios

$$r_x > 0, \qquad x \in [a, b].$$

Nada nos autoriza todavía a tomar simplemente $\min_{x\in[a,b]} r_x$, porque ese mínimo podría no existir. Incluso si todos los radios son positivos, su ínfimo podría ser 0. Por ejemplo, la familia numérica $r_n = \frac{1}{n}$ está formada exclusivamente por números positivos, pero $\inf_{n\geq 1} r_n = 0$.

Así que el paso

$$\text{control local en cada punto} \longrightarrow \text{un único control global}$$

no es automático.

Necesitamos otra idea.

En lugar de intentar controlar simultáneamente una familia infinita de radios, preguntaremos si, de todos los entornos locales disponibles, podemos conservar **solo un número finito** y seguir cubriendo todo el intervalo. Esta es la función de la compactitud. La pregunta guía de la sección es

$$\boxed{\text{¿cuándo puede una cobertura potencialmente infinita reducirse a finitos controles sin dejar huecos?}}$$

En la recta real, el intervalo cerrado y acotado $[a, b]$ posee exactamente la propiedad que necesitamos. Antes de demostrarla debemos precisar cuatro conceptos: conjunto abierto, recubrimiento abierto, subrecubrimiento finito y compactitud.

Abiertos: disponer de margen alrededor de cada punto. En §4.1 introdujimos el entorno $V_r(x) = \{y \in \mathbb{R} : |y - x| < r\}$.

Ahora lo utilizaremos para describir conjuntos que no terminan abruptamente en ninguno de sus propios puntos.

DEFINICIÓN 5.3. **Conjunto abierto, recubrimiento abierto y compactitud en $\mathbb{R}$.**

1. Un conjunto $U \subseteq \mathbb{R}$ es **abierto** si para todo $x \in U$ existe $r > 0$ tal que
$$V_r(x) \subseteq U.$$
2. Sea $K \subseteq \mathbb{R}$. Una familia $\mathcal{U}$ de subconjuntos abiertos de $\mathbb{R}$ es un **recubrimiento abierto** de K si
$$K \subseteq \bigcup_{U \in \mathcal{U}} U.$$
Equivalentemente, todo punto de K pertenece al menos a un miembro de $\mathcal{U}$.
3. Un **subrecubrimiento finito** de K extraído de $\mathcal{U}$ es una familia finita
$$U_1, \ldots, U_m \in \mathcal{U}$$
tal que
$$K \subseteq U_1 \cup \cdots \cup U_m.$$
4. Un conjunto $K \subseteq \mathbb{R}$ es **compacto** si todo recubrimiento abierto de K admite un subrecubrimiento finito.

Hay varios cuantificadores escondidos en la última frase. Conviene hacerlos visibles. Decir que K es compacto significa:

para toda familia abierta $\mathcal{U}$ que cubra K, existe una subfamilia finita de $\mathcal{U}$ que todavía cubre K.

No estamos afirmando que los recubrimientos de K sean finitos. Pueden ser infinitos, no numerables e incluso muy redundantes.

La afirmación es otra:

> por complicada que sea una cobertura abierta de K, siempre podemos **extraer** de ella una cantidad finita de miembros que ya realizan todo el trabajo de cobertura.

Esta palabra, *extraer*, es importante. No se nos permite inventar conjuntos nuevos después de ver la cobertura. El subrecubrimiento debe estar formado por miembros de la familia original.

Leer correctamente la definición de abierto. La condición $x \in U \Longrightarrow \exists r > 0 : \ V_r(x) \subseteq U$ dice que cada punto de U dispone de un pequeño margen alrededor suyo que permanece completamente dentro de U. El radio puede depender del punto. Por ejemplo, $U = (0, 1)$ es abierto. Si $x \in (0, 1)$, entonces tanto x como $1 - x$ son positivos. Podemos tomar, por ejemplo, $r = \frac{1}{2} \min\{x, 1 - x\} > 0$.

Entonces $V_r(x) \subseteq (0, 1)$.

En cambio, $[0, 1]$ no es abierto en $\mathbb{R}$: el punto 0 pertenece al conjunto, pero ningún entorno completo $V_r(0)$ queda contenido en $[0, 1]$, porque todo entorno de 0 contiene números negativos. Esto no constituye ningún defecto del intervalo cerrado. Ser abierto y ser compacto son propiedades diferentes.

No confundamos el conjunto cubierto con los conjuntos que lo cubren
En un recubrimiento abierto de K, los miembros de la familia $\mathcal{U}$ deben ser abiertos en $\mathbb{R}$. El propio conjunto K **no necesita ser abierto**. De hecho, nuestro teorema central dirá precisamente que el intervalo $[a, b]$, que en general no es abierto, es compacto.

Dos comprobaciones elementales. La recta completa $\mathbb{R}$ es abierta: para cualquier $x \in \mathbb{R}$ podemos tomar cualquier radio positivo y el entorno correspondiente sigue contenido en $\mathbb{R}$. El conjunto vacío también es abierto. La condición $\forall x \in \varnothing\ \exists r > 0 :\ V_r(x) \subseteq \varnothing$ es verdadera porque no existe ningún punto $x \in \varnothing$ que pueda violarla. No necesitaremos construir ahora una teoría general de operaciones entre abiertos, interior, clausura o frontera. Para este capítulo basta una propiedad funcional de la apertura:

si $c \in U$ y U es abierto, entonces U contiene un intervalo real alrededor de c.

Ese pequeño margen será la pieza decisiva de la demostración de compactitud.

Una primera compactitud sin dificultad: los conjuntos finitos. Antes de enfrentarnos con $[a, b]$, examinemos un caso donde la extracción finita es inmediata. Sea $K = \{x_1, \dots, x_n\}$ un conjunto finito y sea $\mathcal{U}$ un recubrimiento abierto de K. Como x_1 está cubierto, existe algún $U_1 \in \mathcal{U}$ con $x_1 \in U_1$. Del mismo modo, para cada j podemos escoger un miembro $U_j \in \mathcal{U}$ que contenga a x_j.

Entonces $K \subseteq U_1 \cup \cdots \cup U_n$.

Así, todo conjunto finito es compacto. Obsérvese que aquí la apertura de los U_j ni siquiera fue necesaria: la finitud de K resolvió el problema por sí sola. El intervalo $[a, b]$ es radicalmente diferente. Contiene infinitos puntos, así que no podemos elegir ingenuamente un conjunto de la cobertura por cada punto y esperar terminar después de un número finito de pasos. Necesitaremos usar la estructura de los números reales.

Un intervalo abierto puede fallar: la cobertura que se escapa hacia 0. Antes de demostrar el resultado positivo conviene ver una cobertura donde ninguna selección finita funciona.

Ejemplo 5.7. **Un recubrimiento abierto de $(0, 1)$ sin subrecubrimiento finito.** Para cada entero $n \geq 2$, definamos

$$U_n = \left(\frac{1}{n}, 1\right).$$

La familia $\mathcal{U} = \{U_n : n \geq 2\}$ es un recubrimiento abierto de $(0, 1)$, pero ningún número finito de sus miembros cubre todo $(0, 1)$.

Primero comprobemos que realmente cubre el intervalo. Sea $x \in (0, 1)$. Como $x > 0$, la propiedad arquimediana permite escoger n suficientemente grande para que $\frac{1}{n} < x$.

Entonces

$$x \in \left(\frac{1}{n}, 1\right) = U_n.$$

Por tanto, $(0, 1) \subseteq \bigcup_{n \geq 2} U_n$.

Ahora supongamos que escogemos solo un número finito de miembros, $U_{n_1}, \dots, U_{n_m}$.

Sea $N = \max\{n_1, \dots, n_m\}$.

Como los conjuntos crecen con el índice, $U_{n_j} \subseteq U_N$ para todo j, y por consiguiente

$$U_{n_1} \cup \cdots \cup U_{n_m} = U_N = \left(\frac{1}{N}, 1\right).$$

Pero el punto $x = \frac{1}{N+1}$ pertenece a $(0,1)$ y no pertenece a U_N, porque $\frac{1}{N+1} < \frac{1}{N}$. Así, ninguna subfamilia finita cubre $(0,1)$.

En consecuencia,

$$\boxed{(0,1) \text{ no es compacto.}}$$

El mecanismo del fallo es muy concreto. Cada vez que conservamos solo finitos intervalos, queda un pequeño tramo próximo a 0 que escapa de la cobertura finita.

Lo que este ejemplo sí demuestra
El intervalo $(0,1)$ es acotado, pero no es compacto. Por tanto, **acotación por sí sola no implica compactitud**. No vamos a deducir todavía una caracterización general de los subconjuntos compactos de $\mathbb{R}$. Nuestro objetivo inmediato es más concreto: demostrar directamente que $[a,b]$ sí posee la propiedad de extracción finita.

Predecir la demostración: ¿hasta dónde hemos cubierto finitamente? Sea ahora $\mathcal{U}$ un recubrimiento abierto arbitrario de $[a,b]$. La dificultad es que no sabemos cuántos miembros contiene $\mathcal{U}$, ni cómo se superponen, ni qué tamaños tienen. Una estrategia razonable es no intentar elegir inmediatamente un subrecubrimiento de todo el intervalo. Preguntemos primero:

> ¿hasta qué punto x podemos garantizar que el segmento inicial $[a,x]$ ya se cubre con finitos miembros de $\mathcal{U}$?

Esta pregunta produce el conjunto

$$S = \{x \in [a,b] : [a,x] \text{ admite un subrecubrimiento finito de } \mathcal{U}\}.$$

La idea es casi idéntica en espíritu a la prueba del IVT de §5.4. Allí construimos un conjunto de puntos situados antes de un nivel de cruce y tomamos su supremo. Aquí construimos un conjunto de puntos **hasta los cuales hemos conseguido cobertura finita** y volveremos a tomar un supremo. La completitud servirá para fabricar una frontera $c = \sup S$. Después la apertura de la cobertura deberá impedir que esa frontera se detenga antes de b.

La arquitectura prevista es

cobertura abierta de $[a, b]$
$\downarrow$
$S = \{x : [a, x] \text{ tiene cobertura finita}\}$
$\downarrow$ completitud
$c = \sup S$
$\downarrow$ apertura alrededor de c
la cobertura finita puede avanzar más allá de c
$\downarrow$
$c = b$
$\downarrow$
$[a, b]$ tiene un subrecubrimiento finito.

Antes de leer la prueba formal, conviene localizar las dos preguntas delicadas:

1. ¿por qué un abierto que contiene a c permite **extender** una cobertura finita que llega suficientemente cerca de c?;
2. aun después de demostrar $c = b$, ¿por qué sabemos que b mismo pertenece a S y no es solamente el supremo de S?

La segunda pregunta es esencial. Saber $\sup S = b$ no significa automáticamente $b \in S$. Ya aprendimos en §1.3 que un supremo no tiene por qué ser un máximo.

El teorema de compactitud del intervalo cerrado.

TEOREMA 5.4. **Compactitud de un intervalo cerrado.** Sean $a, b \in \mathbb{R}$ con $a \leq b$. Entonces el intervalo

$$[a, b]$$

es compacto. Es decir, todo recubrimiento abierto de $[a, b]$ contiene un subrecubrimiento finito.

Anatomía del teorema. Dato geométrico. El conjunto es un intervalo de la forma $[a, b]$: contiene sus dos extremos y está acotado entre ellos.

Dato de cobertura. Se nos entrega una familia arbitraria $\mathcal{U}$ de abiertos de $\mathbb{R}$ cuya unión contiene todo el intervalo.

Conclusión. Debemos encontrar miembros $U_1, \ldots, U_m \in \mathcal{U}$ en número finito tales que $[a, b] \subseteq U_1 \cup \cdots \cup U_m$.

El teorema no dice cuántos miembros necesitaremos, ni proporciona una cota universal para m. La cantidad depende de la cobertura concreta. Tampoco afirma que $[a, b]$ pueda cubrirse por finitos abiertos: eso sería trivial, porque el propio $\mathbb{R}$ lo cubre con un solo abierto. La fuerza del resultado está en el orden de cuantificadores:

cualquiera sea el recubrimiento abierto dado, podemos extraer de él una subfamilia finita.

Demostración: estrategia global. Si $a = b$, el intervalo es el singleton $\{a\}$ y ya sabemos que todo conjunto finito es compacto.

Supongamos desde ahora $a < b$.

Sea $\mathcal{U}$ un recubrimiento abierto arbitrario de $[a, b]$.

Definimos

$$S = \{x \in [a,b] : [a,x] \text{ puede cubrirse con finitos miembros de } \mathcal{U}\}.$$

La prueba tendrá cuatro etapas:

1. demostrar que S es no vacío y está acotado superiormente;
2. usar completitud para definir $c = \sup S$;
3. demostrar que $c = b$ usando la apertura de un miembro de la cobertura que contiene a c;
4. demostrar finalmente que $b \in S$.

La cuarta etapa no puede omitirse.

Etapa 1 — El conjunto S está listo para aplicar completitud. Como $\mathcal{U}$ cubre $[a,b]$, existe algún $U_a \in \mathcal{U}$ tal que $a \in U_a$. El conjunto $[a,a] = \{a\}$ queda cubierto por el único miembro U_a. Por tanto, $a \in S$.

Así, S es no vacío.

Además, por definición, $S \subseteq [a,b]$, así que b es una cota superior de S. Hemos obtenido exactamente las dos hipótesis necesarias para aplicar la completitud de $\mathbb{R}$.

Etapa 2 — Aquí entra la completitud. Como S es no vacío y está acotado superiormente, existe $c = \sup S$.

Además, $a \le c \le b$.

Este es el único paso de la prueba en que invocamos directamente el axioma de completitud.

Localicemos la completitud

La cobertura puede ser completamente irregular. No obtenemos c de una fórmula ni de una sucesión convergente. Lo obtenemos porque el conjunto de puntos hasta los cuales existe una cobertura finita es no vacío y acotado superiormente:

$$S \neq \varnothing, \qquad S \le b \quad \Longrightarrow \quad c = \sup S \in \mathbb{R}.$$

Sin la propiedad del supremo, esta arquitectura de prueba se detendría aquí.

Etapa 3 — La frontera no puede detenerse antes de b. Queremos demostrar $c = b$.

Supongamos, buscando una contradicción, que $c < b$. Como $c \in [a,b]$ y $\mathcal{U}$ cubre todo $[a,b]$, existe algún $U \in \mathcal{U}$ con $c \in U$. El conjunto U es abierto. Por tanto, existe $r > 0$ tal que $V_r(c) \subseteq U$. Esta es la ventaja concreta de la apertura: U no contiene solamente el punto c; contiene un **margen completo** alrededor de él.

Definamos

$$\eta = \min\left\{\frac{r}{2}, \frac{b-c}{2}\right\} > 0$$

y $d = c + \eta$.

Entonces $c < d < b$ y, como $\eta \le r/2$, $d - c < r$.

Necesitamos ahora una cobertura finita que llegue suficientemente cerca de c desde la izquierda.

Si $c = a$, podemos tomar simplemente $s = a \in S$.

Si $c > a$, aplicamos la caracterización aproximativa del supremo con tolerancia η y obtenemos algún $s \in S$ tal que $c - \eta < s \le c$. En ambos casos podemos disponer de un punto $s \in S$ que satisface $c - \eta < s \le c$. Como $s \in S$, existen finitos miembros $U_1, \ldots, U_m \in \mathcal{U}$ que cubren $[a,s]$.

Ahora observemos qué ocurre en el tramo restante $[s, d]$. Si $x \in [s, d]$, entonces $s \leq x \leq d$, y por las desigualdades

$$c - \eta < s \qquad \text{y} \qquad d = c + \eta$$

obtenemos

$$|x - c| \leq \eta \leq \frac{r}{2} < r.$$

Por tanto, $x \in V_r(c) \subseteq U$.

Así, $[s, d] \subseteq U$.

La familia finita $U_1, \ldots, U_m, U$ cubre entonces todo $[a, d]$.

Por definición de S, $d \in S$.

Pero $d > c$, lo cual contradice que c sea una cota superior de S. La suposición $c < b$ es imposible. Como ya sabíamos que $c \leq b$, concluimos

$$\boxed{c = b.}$$

Qué acaba de ocurrir. La parte central de la prueba puede comprimirse en una imagen conceptual:

$$\boxed{\begin{array}{c}\text{cobertura finita hasta } s + \text{abierto con margen alrededor de } c \\ \Longrightarrow \text{cobertura finita un poco más allá de } c.\end{array}}$$

El supremo pretendía ser la frontera más lejana alcanzable por coberturas finitas. La apertura impide que una frontera interior del intervalo sea realmente final. Por eso la única frontera posible es el extremo b.

Etapa 4 — Supremo igual a b todavía no significa $b \in S$. Hasta ahora sabemos $\sup S = b$.

Necesitamos demostrar $b \in S$.

Como $\mathcal{U}$ cubre $[a, b]$, existe algún $U_b \in \mathcal{U}$ tal que $b \in U_b$. Por apertura de U_b, existe $r > 0$ tal que $V_r(b) \subseteq U_b$. Como $b = \sup S$, la caracterización aproximativa del supremo proporciona algún $s \in S$ con $b - \frac{r}{2} < s \leq b$. Si $s = b$, ya tenemos $b \in S$ y hemos terminado. Si $s < b$, como $s \in S$, existe una familia finita $U_1, \ldots, U_m \in \mathcal{U}$ que cubre $[a, s]$.

Para todo $x \in [s, b]$ se cumple

$$0 \leq b - x \leq b - s < \frac{r}{2} < r,$$

de modo que $x \in V_r(b) \subseteq U_b$.

Por tanto, $[a, b] \subseteq U_1 \cup \cdots \cup U_m \cup U_b$.

Hemos extraído un subrecubrimiento finito de la cobertura original $\mathcal{U}$. Como $\mathcal{U}$ era un recubrimiento abierto arbitrario, concluimos que $[a, b]$ es compacto. ■

Anatomía de la demostración. La demostración contiene dos mecanismos distintos que conviene no mezclar.

Mecanismo 1 — completitud.

$$S \neq \varnothing, \quad S \text{ acotado superiormente} \Longrightarrow c = \sup S.$$

Este mecanismo produce una frontera real.

Mecanismo 2 — apertura.

$$c \in U, \quad U \text{ abierto} \Longrightarrow \exists r > 0 : \ V_r(c) \subseteq U.$$

Este mecanismo convierte la pertenencia de un solo punto en un margen que permite prolongar la cobertura.

Ambos se combinan así:

completitud fabrica la frontera; apertura demuestra que una frontera interior no puede detenernos.

Prueba de estrés I — ¿por qué la cobertura debe ser abierta? Podríamos preguntarnos si el teorema seguiría siendo cierto para coberturas por conjuntos arbitrarios.

La respuesta es no.

Considérese la familia $\mathcal{S} = \{\{x\} : x \in [a, b]\}$.

Su unión es exactamente $\bigcup_{x\in[a,b]}\{x\} = [a, b]$.

Así que $\mathcal{S}$ cubre el intervalo. Sin embargo, si $a < b$, cualquier subfamilia finita $\{x_1\}, \ldots, \{x_m\}$ solo cubre finitos puntos y deja sin cubrir todos los demás.

No existe subrecubrimiento finito.

Esto no contradice Teorema 5.4 porque los singletons $\{x\}$ no son abiertos en $\mathbb{R}$. Y ahora podemos identificar exactamente dónde fallaría la demostración: de $c \in \{c\}$ no podemos deducir ningún radio $r > 0$ tal que $V_r(c) \subseteq \{c\}$. El mecanismo que permitía avanzar más allá del supremo desaparece.

Prueba de estrés II — ¿por qué no basta con una cobertura finita inventada después? Supongamos que una familia infinita $\mathcal{U}$ cubre K. La compactitud no pregunta si podemos encontrar **alguna otra** familia finita de abiertos que cubra K. Para un conjunto acotado eso podría ser muy fácil y no conservaría la información local original. La exigencia es mucho más precisa:

$U_1, \ldots, U_m$ deben pertenecer a la cobertura original $\mathcal{U}$.

Esta condición es exactamente la que hará útil la compactitud en §5.7 y §5.8. Allí cada abierto llevará asociado un dato local —una cota o un radio de continuidad— y necesitaremos conservar finitos de **esos mismos datos**, no reemplazarlos por abiertos arbitrarios.

Prueba de estrés III — ¿dónde importa que el extremo derecho pertenezca al intervalo? El ejemplo Ejemplo 5.7 ya mostró que $(0, 1)$ no es compacto. Allí la cobertura

$$\left(\frac{1}{n}, 1\right), \qquad n \geq 2,$$

se aproxima indefinidamente al extremo ausente 0. Después de cualquier selección finita sigue quedando un punto del dominio más cerca de 0 que todos los umbrales $1/n$ seleccionados. En $[a, b]$, en cambio, ambos extremos pertenecen al conjunto. En la demostración, esto se ve al comienzo y al final:

- $a \in [a, b]$ nos permite iniciar S con $a \in S$;
- $b \in [a, b]$ garantiza que algún abierto de la cobertura contiene al extremo y permite cerrar la última brecha.

No estamos demostrando todavía el teorema general «compacto si y solo si cerrado y acotado» para subconjuntos de $\mathbb{R}$. Pero ya vemos por qué la presencia efectiva de los extremos resulta estructural en este caso.

Prueba de estrés IV — ¿dónde importa la completitud de $\mathbb{R}$? La apertura por sí sola no crea la frontera c. La prueba requiere pasar de

$$S \neq \varnothing \quad \text{y} \quad S \text{ acotado superiormente}$$

a $c = \sup S \in \mathbb{R}$.

Este paso es exactamente el axioma de completitud usado en §1.5. La lógica es paralela a la del teorema del valor intermedio:

IVT: $S = \{x : f(x) \leq y\} \xrightarrow{\sup}$ frontera de cruce,
compactitud: $S = \{x : [a, x]$ tiene cobertura finita$\} \xrightarrow{\sup}$ frontera de cobertura.

En ambos casos, una propiedad local impide que la frontera obtenida por completitud quede en una posición incorrecta.

- En IVT, la continuidad impide que $f(c)$ quede estrictamente por encima o por debajo del nivel buscado.
- En compactitud, la apertura impide que la cobertura finita se detenga en un punto interior.

En ambos argumentos, la completitud proporciona un supremo y el control local permite avanzar hasta él o más allá. La misma combinación resuelve dos preguntas globales diferentes.

Compactitud no significa «pequeño». La palabra puede sugerir informalmente que el conjunto es corto, estrecho o de pequeña longitud. Esa intuición es engañosa. El intervalo $[-10^{100}, 10^{100}]$ es compacto, aunque sea enorme en longitud.

En cambio, $(0, 1)$ no es compacto, aunque esté contenido en un intervalo de longitud 1. La compactitud no mide tamaño métrico bruto. Expresa una propiedad de **extracción finita frente a coberturas abiertas**.

Compactitud tampoco significa «tener pocos puntos». El intervalo $[a, b]$ con $a < b$ contiene infinitos puntos y, sin embargo, es compacto. Por tanto, la compactitud no es finitud del conjunto. La relación correcta es otra:

infinitos puntos pueden estar organizados de tal manera
que toda cobertura abierta se reduzca a finitos testigos.

Ese es precisamente el poder que necesitaremos para transformar información local en información global.

Por qué este teorema pertenece al capítulo de continuidad. A primera vista, Teorema 5.4 no menciona funciones. No hay ninguna f en su enunciado. Sin embargo, es una de las herramientas centrales del capítulo porque las propiedades de continuidad producen información **punto por punto**. Más adelante tendremos situaciones del tipo

$$\forall x \in [a,b] \quad \exists V_{r_x}(x) \quad \text{con cierta propiedad local } P_x.$$

Los entornos $V_{r_x}(x)$ formarán un recubrimiento abierto de $[a,b]$. La compactitud permitirá sustituir toda esa familia por $V_{r_{x_1}}(x_1), \ldots, V_{r_{x_m}}(x_m)$. Entonces una operación finita —por ejemplo tomar máx$\{M_1, \ldots, M_m\}$ o mín$\{r_1, \ldots, r_m\}$

— convertirá los controles locales seleccionados en un control global.

Esta es la máquina conceptual:

control local en cada punto
$\to$ recubrimiento abierto
$\to$ subrecubrimiento finito
$\to$ operación finita
$\to$ control global.

No ejecutaremos todavía esas aplicaciones. Serán el corazón del teorema de los valores extremos y de Heine–Cantor.

Lo que no estamos introduciendo todavía. La palabra *compacto* pertenece a una teoría mucho más amplia. Más adelante podrá estudiarse en espacios métricos y topológicos, junto con otras caracterizaciones y equivalencias. Aquí deliberadamente mantenemos una jurisdicción mínima.

Solo hemos definido:

- abiertos de $\mathbb{R}$ mediante entornos;
- recubrimientos abiertos;
- subrecubrimientos finitos;
- compactitud por la propiedad de extracción finita.

No necesitamos todavía desarrollar sistemáticamente:

- interior;
- adherencia;
- frontera;
- bases topológicas;
- compactitud secuencial general;
- equivalencias con Bolzano–Weierstrass;
- espacios métricos o topológicos abstractos.

Frontera de no circularidad
La demostración de Teorema 5.4 **no usa Bolzano–Weierstrass** ni extrae subsucesiones convergentes.
Su cadena demostrativa es

$$\text{completitud} \to \sup S \to \text{apertura} \to \text{extensión finita} \to \text{compactitud de } [a,b].$$

Bolzano–Weierstrass permanecerá reservado para el Tomo II.

Reconstrucción de la prueba. Sin releer la demostración completa, el lector debería poder responder:

1. ¿qué propiedad representa la pertenencia $x \in S$?;
2. ¿por qué $a \in S$?;
3. ¿por qué S está acotado superiormente?;
4. ¿en qué línea exacta usamos la completitud de $\mathbb{R}$?;
5. si $c = \sup S < b$, ¿qué información adicional proporciona que un abierto U contenga a c?;
6. ¿por qué necesitamos un punto $s \in S$ muy próximo a c por la izquierda?;
7. ¿cómo se combinan la cobertura finita de $[a,s]$ y el abierto U para cubrir $[a,d]$ con $d > c$?;
8. ¿por qué eso contradice que c sea cota superior?;
9. después de demostrar $c = b$, ¿por qué todavía hace falta un último argumento para obtener $b \in S$?;
10. ¿qué paso desaparecería si los miembros de la cobertura no fueran abiertos?

La pregunta central de autoexplicación es:

> ¿qué tenía que ocurrir para convertir una cantidad potencialmente infinita de controles locales en una cantidad finita sin perder ningún punto del intervalo?

La respuesta es:

la completitud debe producir una frontera de cobertura +la apertura debe impedir que esa frontera se detenga antes del extremo.

Qué hemos ganado. Hasta §5.5, la completitud se había usado para construir o localizar puntos:

- supremos;
- raíces;
- fronteras de cruce;
- puntos comunes de intervalos encajados.

Ahora aparece una capacidad nueva. La completitud, combinada con apertura, permite obtener **finitud efectiva dentro de una estructura infinita**:

$[a,b]$ contiene infinitos puntos,	pero toda cobertura abierta se reduce a una cobertura finita.

Esta finitud no es un detalle técnico. En la siguiente sección la utilizaremos para demostrar que una función continua en $[a,b]$ no solo está localmente controlada: está **globalmente acotada** y

además alcanza sus valores extremos. Ese será el contenido de §5.7: el teorema de los valores extremos.

5.7. Los extremos sí se alcanzan: teorema de los valores extremos

La sección anterior construyó una herramienta que, a primera vista, no hablaba de funciones: $[a, b]$ es compacto. Ahora veremos por qué esa propiedad era exactamente la pieza que faltaba. La continuidad es, en principio, una condición local. En cada punto c podemos controlar $f(x)$ cuando x permanece suficientemente cerca de c. Pero una función definida en todo un intervalo contiene potencialmente infinitos puntos, y los radios de control pueden variar de un punto a otro.

La pregunta es entonces:

¿cómo pasamos de una cota local distinta en cada punto a una cota válida para todo $[a, b]$?

La respuesta será la compactitud. Pero el teorema que queremos demostrar dice algo todavía más fuerte. No afirma solamente que la función quede encerrada entre dos números. Afirma que existen puntos del dominio donde los valores extremos **se alcanzan**. Conviene separar desde el principio esas dos afirmaciones. Una función puede estar acotada superiormente sin poseer máximo. Por ejemplo,

$$f(x) = x, \qquad x \in (0, 1),$$

satisface $f(x) < 1$ para todo $x \in (0, 1)$, de modo que 1 es cota superior y de hecho $\sup f((0, 1)) = 1$. Sin embargo, no existe $x \in (0, 1)$ tal que $f(x) = 1$. Así que debemos distinguir cuidadosamente:

tener supremo $\neq$ alcanzar el supremo.

El teorema de esta sección demostrará que, para una función continua sobre un intervalo cerrado, esa diferencia desaparece.

El enunciado.

TEOREMA 5.5. **Teorema de los valores extremos.** Sean $a, b \in \mathbb{R}$ con $a \leq b$ y sea

$$f\colon [a, b] \to \mathbb{R}$$

continua.

Entonces existen puntos

$$x_m, x_M \in [a, b]$$

tales que

$$f(x_m) \leq f(x) \leq f(x_M) \qquad \text{para todo } x \in [a, b].$$

Equivalentemente, f alcanza en $[a, b]$ un mínimo y un máximo global.

Si $a = b$, la afirmación es inmediata: el dominio contiene un solo punto y podemos tomar $x_m = x_M = a$. Por tanto, en la demostración interesante supondremos $a < b$.

Antes de demostrarlo: dos problemas distintos. El teorema contiene en realidad dos tareas. La primera es demostrar que la imagen $f([a,b])$ está acotada. Solo después de saber eso podremos escribir legítimamente $M = \sup f([a,b])$ y $m = \inf f([a,b])$. La segunda tarea es demostrar que esas fronteras pertenecen realmente a la imagen:

$$M \in f([a,b]), \qquad m \in f([a,b]).$$

Es decir, necesitamos puntos x_M, x_m con

$$f(x_M) = M, \qquad f(x_m) = m.$$

La arquitectura completa será

continuidad local
$\to$ acotación local
$\to$ recubrimiento abierto
$\to$ compactitud
$\to$ acotación global
$\to$ sup / ínf
$\to$ intervalos encajados
$\to$ alcanzamiento.

La prueba combina, por tanto, dos mecanismos ya construidos en el libro:

1. **compactitud**, para transformar infinitos controles locales en finitos controles globales;
2. **bisección e intervalos cerrados encajados**, para localizar un punto donde el supremo de la imagen se materializa.

No utilizaremos Bolzano–Weierstrass ni extraeremos subsucesiones.

Primera etapa: la continuidad produce acotación local. Fijemos un punto arbitrario $c \in [a,b]$.

Como f es continua en c, podemos aplicar la definición con una tolerancia muy sencilla: $\varepsilon = 1$. Existe entonces un radio $r_c > 0$ tal que, para todo $x \in [a,b]$, $|x-c| < r_c \Longrightarrow |f(x) - f(c)| < 1$.

Por la desigualdad triangular,

$$|f(x)| \le |f(x) - f(c)| + |f(c)| < 1 + |f(c)|.$$

Así, dentro del entorno $V_{r_c}(c)$, la función está acotada por el número $B_c = |f(c)| + 1$.

Obsérvese la estructura:

$$\forall c \in [a,b] \quad \exists r_c > 0 \quad \exists B_c > 0$$

con la propiedad $x \in [a,b] \cap V_{r_c}(c) \Longrightarrow |f(x)| < B_c$.

Esto todavía **no** es una cota global. El número B_c puede depender del centro c, y también el radio r_c puede deteriorarse al movernos por el intervalo. La continuidad punto a punto nos ha dado una familia potencialmente infinita de controles.

Segunda etapa: compactitud convierte infinitos controles en finitos. Los entornos $\mathcal{V} = \{V_{r_c}(c) : c \in [a,b]\}$ forman un recubrimiento abierto de $[a,b]$. En efecto, cada $c \in [a,b]$ pertenece a su propio entorno $c \in V_{r_c}(c)$. Por Teorema 5.4, el intervalo $[a,b]$ es compacto. Por tanto, existe una subfamilia finita $V_{r_{c_1}}(c_1), \ldots, V_{r_{c_N}}(c_N)$ que todavía cubre todo $[a,b]$. Cada uno de esos entornos viene acompañado de una cota local $B_{c_i} = |f(c_i)| + 1$. Como ahora solo tenemos **finitos** números, podemos tomar su máximo:

$$B = \max_{1 \le i \le N} B_{c_i}.$$

Sea $x \in [a,b]$ arbitrario. Como los entornos seleccionados cubren el intervalo, existe algún índice i tal que $x \in V_{r_{c_i}}(c_i)$.

Entonces $|f(x)| < B_{c_i} \le B$.

Por tanto,

$$\boxed{|f(x)| \le B \qquad \text{para todo } x \in [a,b].}$$

Hemos probado que f está globalmente acotada.

Dónde entra exactamente la compactitud
La continuidad por sí sola produjo una cota **distinta para cada entorno local**. La compactitud no creó nuevas cotas. Su función fue permitir seleccionar solo un número finito de ellas: $B_{c_1}, \ldots, B_{c_N}$.
Solo después pudimos formar $\max\{B_{c_1}, \ldots, B_{c_N}\}$.
El salto local $\to$ global ocurre precisamente en la extracción de una subcubierta finita.

Ya podemos hablar del supremo de la imagen. Como $[a,b]$ no es vacío, tampoco lo es $f([a,b])$. Y acabamos de demostrar que este conjunto está acotado superior e inferiormente.

Por completitud de $\mathbb{R}$, existen $M = \sup f([a,b])$ y $m = \inf f([a,b])$.

Pero todavía no sabemos que M o m sean valores efectivamente tomados por f. El siguiente paso será demostrar $\exists x_M \in [a,b] : f(x_M) = M$.

Tercera etapa: conservar el supremo mientras bisecamos el dominio. Comenzamos con $I_0 = [a,b]$.

Por definición de M, $\sup f(I_0) = M$.

Supongamos que ya hemos construido un intervalo cerrado $I_n = [a_n, b_n]$ tal que $\sup f(I_n) = M$. Tomemos su punto medio $q_n = \frac{a_n + b_n}{2}$ y dividamos I_n en las dos mitades cerradas

$$L_n = [a_n, q_n], \qquad R_n = [q_n, b_n].$$

Como $I_n = L_n \cup R_n$, tenemos $f(I_n) = f(L_n) \cup f(R_n)$.

Los dos conjuntos de valores son no vacíos y están acotados superiormente, porque están contenidos en $f([a,b])$.

Definamos

$$M_L = \sup f(L_n), \qquad M_R = \sup f(R_n).$$

Entonces $\sup f(I_n) = \max\{M_L, M_R\}$.

¿Por qué?

Primero, todo valor de $f(I_n)$ pertenece a una de las dos imágenes, de modo que $\max\{M_L, M_R\}$ es cota superior de $f(I_n)$. Segundo, cualquier cota superior de $f(I_n)$ es también cota superior de $f(L_n)$ y de $f(R_n)$, y por tanto debe ser mayor o igual que ambos supremos. Así, debe ser mayor o igual que su máximo. Por la unicidad del supremo, $M = \sup f(I_n) = \max\{M_L, M_R\}$. Por consiguiente, al menos una de las dos mitades satisface $\sup f(\text{mitad}) = M$. Elegimos una de ellas y la llamamos I_{n+1}. De este modo construimos inductivamente una cadena $I_0 \supseteq I_1 \supseteq I_2 \supseteq \cdots$ de intervalos cerrados encajados que satisfacen simultáneamente $\sup f(I_n) = M$ para todo n, y $|I_n| = \frac{b-a}{2^n}$.

La longitud tiende a cero: $|I_n| \to 0$.

El invariante de la bisección
En §5.5 la bisección conservaba el **cambio de signo**. Aquí conservamos una propiedad diferente:

$$\boxed{\sup f(I_n) = M.}$$

La geometría del procedimiento es la misma —intervalos cerrados encajados de longitud decreciente—, pero el invariante que guía qué mitad conservar ha cambiado.

Cuarta etapa: los intervalos encajados determinan un punto. Por el principio de intervalos cerrados encajados desarrollado en `T1-C02`, existe un único punto $c \in \bigcap_{n=0}^{\infty} I_n$.

En particular, $c \in I_n$ para todo n.

Queremos demostrar que $f(c) = M$.

Sabemos ya que $f(c) \le M$, porque M es cota superior de toda la imagen $f([a,b])$. La única posibilidad que debemos descartar es $f(c) < M$. Supongamos, buscando una contradicción, que eso ocurre.

Definamos $\varepsilon = \frac{M - f(c)}{2} > 0$.

Por continuidad de f en c, existe $\delta > 0$ tal que, para todo $x \in [a,b]$, $|x - c| < \delta \implies |f(x) - f(c)| < \varepsilon$.

En particular,

$$f(x) < f(c) + \varepsilon = f(c) + \frac{M - f(c)}{2} = \frac{M + f(c)}{2}.$$

Pongamos $K = \frac{M + f(c)}{2}$.

Como $f(c) < M$, $K < M$.

Ahora utilizamos que las longitudes de los intervalos tienden a cero. Existe n tal que $|I_n| < \delta$.

Como $c \in I_n$, todo $x \in I_n$ satisface $|x - c| \le |I_n| < \delta$.

Por tanto, $f(x) < K$ para todo $x \in I_n$.

Así, K es una cota superior de $f(I_n)$.

Pero $K < M$ y, por construcción, $\sup f(I_n) = M$.

Eso es imposible: el supremo no puede ser mayor que una cota superior del mismo conjunto.

La contradicción muestra que $f(c) \not< M$.

Como ya sabíamos que $f(c) \le M$, concluimos

$$\boxed{f(c) = M.}$$

Tomamos entonces $x_M = c$.

El máximo global está alcanzado.

El mínimo se obtiene sin repetir toda la prueba. Consideremos la función $g = -f$.

Por el álgebra de funciones continuas de Proposición 5.2, g es continua en $[a, b]$. Acabamos de demostrar que toda función continua en $[a, b]$ alcanza su máximo. Por tanto, existe $x_m \in [a, b]$ tal que $g(x) \leq g(x_m)$ para todo $x \in [a, b]$.

Es decir, $-f(x) \leq -f(x_m)$.

Multiplicando por -1 e invirtiendo el sentido de la desigualdad, $f(x_m) \leq f(x)$ para todo $x \in [a, b]$. Así, f alcanza también su mínimo global. Hemos demostrado Teorema 5.5. ■

Anatomía de la demostración. La prueba tiene dos columnas que conviene no mezclar.

Columna A: demostrar acotación.

$$\boxed{\text{continuidad en cada } c \to \text{cota local } B_c \to \{V_{r_c}(c)\}_{c\in[a,b]} \to \text{subcubierta finita} \to B = \max B_{c_i}.}$$

Aquí la compactitud es la herramienta decisiva.

Columna B: demostrar alcanzamiento.

$$\boxed{f([a,b]) \text{ acotada} \to M = \sup f([a,b]) \to \sup f(I_n) = M \to I_n \downarrow \{c\} \to \text{continuidad en } c \to f(c) = M.}$$

Aquí la herramienta decisiva es la combinación

$$\text{supremo} + \text{intervalos encajados} + \text{continuidad}.$$

La primera columna explica por qué el supremo existe. La segunda explica por qué deja de ser una frontera abstracta y se convierte en un valor de la función.

Por qué no hemos usado Bolzano–Weierstrass. Una demostración muy común del teorema de los valores extremos procede de otra manera: toma una sucesión (x_n) tal que $f(x_n) \to M$, extrae de (x_n) una subsucesión convergente y usa continuidad. Ese argumento es correcto cuando ya se dispone del teorema de Bolzano–Weierstrass. Pero en nuestra arquitectura todavía no lo hemos introducido. La prueba presente reemplaza esa extracción por una construcción explícita de intervalos cerrados encajados: $I_0 \supseteq I_1 \supseteq I_2 \supseteq \cdots$ con $\sup f(I_n) = M$ y $|I_n| \to 0$. El punto c aparece como intersección de esos intervalos, no como límite de una subsucesión seleccionada.

Frontera de no circularidad

En esta demostración no se usa:

- Bolzano–Weierstrass;
- compactitud secuencial;
- teoría sistemática de subsucesiones;
- derivadas;
- Rolle;
- teorema del valor medio;
- integración.

La dependencia efectiva es

$$\boxed{\text{continuidad} + \text{compactitud de } [a,b] + \text{supremo} + \text{intervalos encajados} \Longrightarrow \text{valores extremos.}}$$

El dominio abierto muestra qué puede fallar.

EJEMPLO 5.8. **Una función continua y acotada que no alcanza sus extremos.** Consideremos

$$f\colon (0,1) \to \mathbb{R}, \qquad f(x) = x.$$

La función es continua y está acotada, pero no alcanza ni su supremo ni su ínfimo.

En efecto, para todo $x \in (0,1)$, $0 < x < 1$, de modo que $0 < f(x) < 1$. Por tanto, la imagen es $f((0,1)) = (0,1)$.

Así, $\inf f((0,1)) = 0$ y $\sup f((0,1)) = 1$.

Pero $0 \notin f((0,1))$ y $1 \notin f((0,1))$.

No existe ningún $x_m \in (0,1)$ con $f(x_m) = 0$, ni ningún $x_M \in (0,1)$ con $f(x_M) = 1$. Este ejemplo es especialmente instructivo porque no falla la continuidad y tampoco falla la acotación. Lo que falla es la estructura del dominio necesaria para forzar el alcanzamiento.

La comparación es exacta:

Dominio	Función	Continua	Acotada	Alcanza máximo y mínimo
$[0,1]$	$f(x) = x$	sí	sí	sí
$(0,1)$	$f(x) = x$	sí	sí	no

En el intervalo cerrado,

$$f(0) = 0, \qquad f(1) = 1.$$

En el intervalo abierto, precisamente los puntos que realizarían los extremos han sido excluidos.

Continuidad sola ni siquiera garantiza acotación global. El ejemplo anterior muestra que continua + acotada no garantiza alcanzamiento sobre un dominio no compacto. Pero podemos decir todavía más: la continuidad por sí sola tampoco garantiza acotación global.

Considérese

$$h\colon (0,1) \to \mathbb{R}, \qquad h(x) = \frac{1}{x}.$$

La función es continua en cada punto de $(0,1)$, pero no está acotada superiormente. Dado cualquier $B > 0$, podemos escoger $x \in (0,1)$ suficientemente pequeño para que $\frac{1}{x} > B$.

Por ejemplo, si tomamos

$$x < \min\left\{1, \frac{1}{B}\right\},$$

obtenemos $h(x) > B$.

Así quedan separadas tres propiedades distintas:

$$\boxed{\text{continuidad} \not\Rightarrow \text{acotación global} \not\Rightarrow \text{alcanzamiento de extremos.}}$$

Sobre $[a, b]$, la compactitud es la pieza adicional que permite cerrar ambas brechas.

¿Es realmente necesario que el intervalo sea cerrado? El teorema no afirma que una función continua sobre un intervalo abierto **nunca** alcance extremos. Por ejemplo, $f(x) = -x^2$ sobre $(-1, 1)$ alcanza un máximo en $x = 0$. Fuera de un dominio compacto, una función continua puede alcanzar sus extremos, pero la continuidad por sí sola ya no lo garantiza. Habrá que justificar su existencia en cada caso. Esto es una distinción lógica importante:

$$\boxed{\begin{array}{l}\text{hipótesis suficiente para garantizar una conclusión}\\ \neq \text{condición necesaria en cada ejemplo particular.}\end{array}}$$

El intervalo cerrado y acotado proporciona una estructura uniforme bajo la cual **toda** función continua satisface la conclusión del teorema.

¿Es necesaria la continuidad? También aquí debemos distinguir garantía general de ejemplos particulares. Una función discontinua puede perfectamente alcanzar máximo y mínimo. Por ejemplo, cualquier función definida sobre un conjunto finito los alcanza. Pero si eliminamos continuidad no podemos conservar el teorema para todas las funciones sobre $[a, b]$.

Considérese $g\colon [0, 1] \to \mathbb{R}$ definida por

$$g(x) = \begin{cases} x, & 0 \le x < 1, \\ 0, & x = 1. \end{cases}$$

La función está acotada y el dominio es compacto, pero $\sup g([0, 1]) = 1$ sin que exista $x \in [0, 1]$ tal que $g(x) = 1$. La discontinuidad en 1 permite que el valor extremo quede como una frontera no alcanzada. Por tanto, las dos piezas del teorema cumplen funciones diferentes:

$$\boxed{\text{compactitud del dominio} + \text{continuidad de la función} \Longrightarrow \text{alcanzamiento global.}}$$

Una consecuencia geométrica combinando IVT y EVT. Teorema 5.5 nos da puntos $x_m, x_M \in [a, b]$ con

$$m = f(x_m), \qquad M = f(x_M),$$

y $m \le f(x) \le M$ para todo $x \in [a, b]$.

Por Corolario 5.3, la imagen continua de un intervalo es un intervalo. Como la imagen contiene sus extremos m y M y no puede contener valores fuera de ellos, necesariamente

$$\boxed{f([a, b]) = [m, M].}$$

Esta fórmula reúne dos teoremas globales diferentes:

- EVT proporciona los extremos y garantiza que pertenecen a la imagen;
- IVT garantiza que **todos los valores entre ellos** también pertenecen a la imagen.

Así, para una función continua sobre un intervalo cerrado,

$$\boxed{\text{la imagen vuelve a ser un intervalo cerrado y acotado.}}$$

No registramos aquí un nuevo corolario formal porque esta consecuencia será reutilizada como problema de síntesis en el banco del capítulo.

El papel de cada hipótesis. Conviene auditar el teorema completo.

Hipótesis 1: f está definida en todo $[a, b]$.

Sin los extremos del dominio disponibles, valores límite de la imagen pueden quedar sin realizar, como ocurre con x en $(0, 1)$.

Hipótesis 2: f es continua.

La continuidad produce las cotas locales y, al final, impide que el supremo persista en intervalos arbitrariamente pequeños alrededor de un punto cuyo valor estuviera estrictamente por debajo de él.

Hipótesis 3: el dominio es compacto.

La compactitud transforma las cotas locales en una cota global finita.

Completitud de $\mathbb{R}$.

Aparece detrás de dos mecanismos ya establecidos:

1. la compactitud de $[a, b]$ fue demostrada mediante un supremo;
2. el principio de intervalos cerrados encajados también se apoya en la completitud.

Por tanto, aunque el enunciado del EVT no mencione la palabra *completitud*, su prueba forma parte de la cadena

$$\boxed{\text{completitud de } \mathbb{R} \to \text{compactitud / intervalos encajados} \to \text{EVT.}}$$

Reconstrucción de la demostración. Sin releer la prueba, el lector debería poder responder estas preguntas:

1. ¿por qué la continuidad en un punto produce una cota local si tomamos $\varepsilon = 1$?
2. ¿por qué los entornos asociados a todos los puntos forman un recubrimiento abierto de $[a, b]$?
3. ¿qué parte exacta de la compactitud permite pasar de infinitas cotas locales a una sola cota global?
4. ¿por qué necesitamos primero demostrar acotación antes de definir $M = \sup f([a, b])$?
5. al bisecar I_n, ¿por qué al menos una mitad debe conservar supremo M?
6. ¿cuál es el invariante que sustituye al cambio de signo de §5.5?
7. ¿qué resultado garantiza un único punto c en todos los intervalos I_n?
8. si $f(c) < M$, ¿por qué la continuidad produce una cota $K < M$ válida en un entorno de c?
9. ¿por qué, para n grande, todo I_n queda dentro de ese entorno?
10. ¿por qué eso contradice $\sup f(I_n) = M$?
11. ¿cómo se obtiene el mínimo sin repetir toda la construcción?
12. ¿en qué parte de la prueba sería ilegítimo invocar Bolzano–Weierstrass según la arquitectura actual del libro?

La pregunta central de autoexplicación es:

> ¿por qué una función continua no puede mantener su supremo eternamente «a punto de alcanzarse» dentro de intervalos cerrados cada vez más pequeños sin terminar alcanzándolo en el único punto que sobrevive?

La respuesta puede comprimirse en

$$\boxed{\sup f(I_n) = M + |I_n| \to 0 + \bigcap I_n = \{c\} + \text{continuidad en } c \Longrightarrow f(c) = M.}$$

Qué hemos ganado. El teorema de los valores extremos completa un segundo gran paso local → global. La continuidad nos daba, alrededor de cada punto, información que podía depender del punto. La compactitud permitió reducir esos controles a un conjunto finito y fabricar una cota global. Después, completitud e intervalos encajados transformaron una frontera abstracta de la imagen en un valor efectivamente alcanzado. El resultado final es mucho más fuerte que simple acotación:

$$\boxed{\exists x_m, x_M \in [a,b] \quad \forall x \in [a,b]: \quad f(x_m) \le f(x) \le f(x_M).}$$

Hasta aquí, la compactitud nos ha permitido pasar de lo local a una **cota global**. La siguiente sección exigirá algo todavía más uniforme. En continuidad ordinaria, el radio δ puede depender del punto. Preguntaremos si, sobre $[a,b]$, podemos escoger **un solo radio** que funcione simultáneamente para todos los puntos del intervalo. Eso nos llevará a la continuidad uniforme y al teorema de Heine–Cantor.

5.8. Una sola δ para todo el intervalo: continuidad uniforme

Hasta ahora la continuidad se ha formulado **punto por punto**. Si fijamos $a \in A$ y exigimos que f sea continua en a, entonces, dado $\varepsilon > 0$, podemos escoger un radio $\delta = \delta(a,\varepsilon)$ que dependa del punto donde estamos trabajando. Ese detalle de dependencia no era un problema. Para demostrar continuidad en un punto concreto solo necesitábamos un radio que funcionara alrededor de **ese** punto. Pero el teorema de los valores extremos y la compactitud de $[a,b]$ nos han enseñado a mirar el intervalo completo como una sola unidad.

La nueva pregunta es:

$$\boxed{\text{¿puede existir un único radio de entrada que funcione simultáneamente en todo el dominio?}}$$

Esta pregunta conduce a una noción más fuerte que la continuidad ordinaria.

Dos órdenes de cuantificadores. Conviene comenzar comparando las dos estructuras lógicas. Decir que $f\colon A \to \mathbb{R}$ es continua en todo A significa:

$$\forall a \in A\ \forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in A: \qquad |x-a| < \delta \Longrightarrow |f(x)-f(a)| < \varepsilon.$$

El punto a se fija **antes** de escoger δ. Por tanto, el radio puede depender de a: $\delta = \delta(a,\varepsilon)$. La continuidad uniforme exigirá algo más fuerte: dado ε, debemos escoger primero una única δ y esa misma δ deberá funcionar después para **todos los pares de puntos** del dominio.

Definición 5.4. **Continuidad uniforme.** Sea

$$f\colon A \to \mathbb{R}, \qquad A \subseteq \mathbb{R}.$$

Diremos que f es **uniformemente continua en** A si

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x, y \in A: \qquad |x-y| < \delta \Longrightarrow |f(x)-f(y)| < \varepsilon.$$

La diferencia debe leerse en el orden de los cuantificadores.

Para continuidad ordinaria:

$$\boxed{\forall a \in A\ \forall \varepsilon > 0\ \exists \delta(a,\varepsilon) > 0.}$$

Para continuidad uniforme:

$$\boxed{\forall \varepsilon > 0\ \exists \delta(\varepsilon) > 0\ \forall x, y \in A.}$$

En la segunda fórmula, cuando llega el turno de elegir x e y, la δ ya ha sido fijada. No podemos cambiarla según la región del dominio donde aparezcan los puntos.

Qué significa geométricamente la palabra «uniforme». En continuidad puntual, alrededor de cada centro a podemos imaginar un radio propio: $a \longmapsto \delta_a$. En unas zonas del dominio ese radio puede ser grande y en otras muy pequeño. La continuidad uniforme pregunta si podemos encontrar, para cada tolerancia vertical ε, una sola escala horizontal que controle **todos** los pares de puntos suficientemente próximos.

Es decir,

$$\boxed{\text{puntos de entrada próximos entre sí} \Longrightarrow \text{salidas próximas entre sí,}}$$

con una proximidad de entrada que no depende de dónde están situados los puntos. Esta formulación explica también por qué la definición ya no usa un centro distinguido a y un punto variable x. Ambos puntos aparecen simétricamente: $x, y \in A$.

Continuidad uniforme implica continuidad. La nueva condición es realmente más fuerte. Supongamos que f es uniformemente continua en A y fijemos un punto arbitrario $a \in A$. Sea $\varepsilon > 0$. La continuidad uniforme proporciona una $\delta > 0$ tal que para todos $x, y \in A$, $|x - y| < \delta \Longrightarrow |f(x) - f(y)| < \varepsilon$.

Ahora ponemos $y = a$.

Entonces, para todo $x \in A$, $|x - a| < \delta \Longrightarrow |f(x) - f(a)| < \varepsilon$.

Eso es exactamente continuidad de f en a.

Como a era arbitrario,

$$\boxed{\text{uniformemente continua en } A \Longrightarrow \text{continua en } A.}$$

La recíproca será falsa en general. Antes de construir el contraejemplo, conviene notar otra consecuencia inmediata. Si $E \subseteq A$ y f es uniformemente continua en A, entonces la restricción $f|_E$ es uniformemente continua en E. La razón es simple: la misma δ que controlaba todos los pares de puntos de A controla, en particular, todos los pares de puntos del subconjunto E.

Cómo se niega la continuidad uniforme. Para demostrar que una función **no** es uniformemente continua debemos negar cuidadosamente la definición.

La afirmación

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x, y \in A: \quad |x - y| < \delta \Longrightarrow |f(x) - f(y)| < \varepsilon$$

falla exactamente cuando existe una tolerancia vertical fija $\varepsilon_0 > 0$ tal que ninguna escala horizontal funciona globalmente:

$$\boxed{\exists \varepsilon_0 > 0\ \forall \delta > 0\ \exists x, y \in A: \quad |x - y| < \delta \quad \text{y} \quad |f(x) - f(y)| \geq \varepsilon_0.}$$

Esta es la herramienta correcta para un contraejemplo. No basta decir que «la gráfica se vuelve muy empinada» o que «la función crece mucho». Debemos producir, para **cada** $\delta > 0$, dos puntos separados por menos de δ cuyas imágenes mantengan una separación vertical fija.

El cuadrado: continuo en todas partes, pero no uniformemente continuo en $\mathbb{R}$.

Ejemplo 5.9. **Continuidad sin continuidad uniforme.** La función

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = x^2,$$

es continua en $\mathbb{R}$, pero no es uniformemente continua en $\mathbb{R}$.

Ya sabemos que x^2 es continua en cada punto real. Recordemos cómo aparecía la dependencia del punto en una prueba directa. Al estudiar continuidad en un centro a, una elección posible era

$$\delta = \min\left\{1, \frac{\varepsilon}{2|a|+1}\right\}.$$

Cuando $|a|$ crece, el denominador $2|a|+1$ también crece, de modo que el radio disponible puede hacerse arbitrariamente pequeño. Eso sugiere que no existirá una única δ que funcione en toda la recta. Pero una sugerencia no es una prueba. Demostremos la negación formal de continuidad uniforme.

Tomemos $\varepsilon_0 = 1$.

Sea $\delta > 0$ arbitrario. Definimos

$$x = \frac{1}{\delta}, \qquad y = \frac{1}{\delta} + \frac{\delta}{2}.$$

Entonces $|x - y| = \frac{\delta}{2} < \delta$.

Sin embargo,

$$\begin{aligned} |f(y) - f(x)| &= \left| \left(\frac{1}{\delta} + \frac{\delta}{2}\right)^2 - \left(\frac{1}{\delta}\right)^2 \right| \\ &= \left| 1 + \frac{\delta^2}{4} \right| \\ &= 1 + \frac{\delta^2}{4} \\ &> 1. \end{aligned}$$

Por tanto, para **todo** $\delta > 0$ hemos encontrado dos puntos $x, y \in \mathbb{R}$ con $|x - y| < \delta$ pero $|f(x) - f(y)| \geq 1$. Esto demuestra que x^2 no es uniformemente continua en $\mathbb{R}$. La obstrucción no consiste simplemente en que x^2 sea no acotada. Lo decisivo es que, lejos del origen, puntos horizontalmente muy próximos pueden producir variaciones verticales de tamaño fijo.

Un contraste: el valor absoluto sí tiene un control global. La función $f(x) = |x|$ es uniformemente continua en $\mathbb{R}$. La desigualdad triangular inversa nos da $\big||x| - |y|\big| \leq |x - y|$.

Dado $\varepsilon > 0$, basta escoger $\delta = \varepsilon$.

Si $|x - y| < \delta$, entonces $\big||x| - |y|\big| < \varepsilon$.

Aquí el control no contiene ninguna referencia a la posición de x o y. La misma estimación funciona cerca del origen, cerca de 10^6 y cerca de -10^{100}. Esta es una señal típica de uniformidad:

$$\boxed{|f(x) - f(y)| \leq C|x - y|}$$

con una constante C independiente de los puntos produce inmediatamente continuidad uniforme, tomando $\delta = \frac{\varepsilon}{C}$ cuando $C > 0$. No desarrollaremos todavía una teoría general de este

tipo de estimaciones. Solo registramos el patrón porque ayuda a reconocer cuándo un control es verdaderamente global.

El problema que resolverá la compactitud. El ejemplo x^2 en $\mathbb{R}$ demuestra que

$$\text{continuidad} \not\Rightarrow \text{continuidad uniforme}$$

sobre dominios arbitrarios.

Pero nuestra situación central no es un dominio arbitrario. Estamos trabajando con un intervalo cerrado $[a, b]$. En §5.6 demostramos que este intervalo es compacto. La pregunta es ahora si esa propiedad global puede impedir que los radios locales de continuidad se deterioren indefinidamente. Para cada punto $x \in [a, b]$, la continuidad nos proporciona algún radio $r_x > 0$. Podemos imaginar una familia completa $\{r_x : x \in [a, b]\}$. Una tentación sería escribir $\delta = \text{ínf}_{x\in[a,b]} r_x$ o incluso hablar informalmente del «mínimo de todos los radios».

Eso no está justificado.

Aunque cada r_x sea positivo, el ínfimo de una familia infinita de números positivos puede ser cero. Por ejemplo, $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$ son todos positivos, pero

$$\text{ínf}\left\{1, \frac{1}{2}, \frac{1}{3}, \ldots\right\} = 0.$$

Así que la continuidad puntual, por sí sola, no permite tomar un mínimo global de radios. La compactitud debe intervenir **antes**. Su función será reducir la familia potencialmente infinita de controles locales a una familia finita. Solo entonces podremos tomar un mínimo positivo.

El truco del radio mitad. Hay todavía una segunda dificultad. Supongamos que alrededor de cada punto x tenemos un radio r_x tal que

$$|z - x| < r_x \Longrightarrow |f(z) - f(x)| < \frac{\varepsilon}{2}.$$

Si una subcubierta finita nos entrega centros $x_1, \ldots, x_m$, no basta con decir:

tomemos el mínimo de $r_{x_1}, \ldots, r_{x_m}$.

Debemos garantizar que dos puntos u y v próximos entre sí puedan compararse **a través de un mismo centro**. Para conseguirlo cubriremos el intervalo no con los entornos completos $V_{r_x}(x)$, sino con los entornos reducidos $V_{r_x/2}(x)$. Entonces, si u cae dentro de uno de esos entornos reducidos y v está suficientemente cerca de u, la desigualdad triangular permitirá colocar también a v dentro del entorno completo de radio r_x. Esta pequeña reducción de radio es el engranaje técnico que sincroniza dos puntos.

Heine–Cantor en un intervalo cerrado.

TEOREMA 5.6. **Teorema de Heine–Cantor en** $[a, b]$**.** Sea

$$f\colon [a, b] \to \mathbb{R}$$

continua. Entonces f es uniformemente continua en $[a, b]$.

Arquitectura de la prueba
La cadena que debemos construir es

continuidad punto a punto
$\to$ un radio r_x para cada x
$\to$ recubrimiento por radios mitad
$\to$ subrecubrimiento finito
$\to \min r_{x_i}$
$\to$ una δ global.

La compactitud no crea los radios locales: esos vienen de la continuidad. La compactitud permite conservar solo finitísimos de ellos. El mínimo finito convierte entonces información local en una escala global.

Demostración. Sea $\varepsilon > 0$ arbitrario.

Para cada punto $x \in [a, b]$, la continuidad de f en x proporciona un radio $r_x > 0$ tal que para todo $z \in [a, b]$,

$$|z - x| < r_x \Longrightarrow |f(z) - f(x)| < \frac{\varepsilon}{2}.$$

Consideremos ahora los entornos abiertos $V_{r_x/2}(x)$. Cuando x recorre $[a, b]$, estos entornos forman un recubrimiento abierto de $[a, b]$, porque cada punto x pertenece a su propio entorno: $x \in V_{r_x/2}(x)$. Por la compactitud de $[a, b]$ demostrada en Teorema 5.4, existe una subcubierta finita:

$$[a, b] \subseteq V_{r_{x_1}/2}(x_1) \cup \cdots \cup V_{r_{x_m}/2}(x_m).$$

Los números $r_{x_1}, \ldots, r_{x_m}$ son positivos y son solo finitos.

Por tanto,

$$r_* = \min_{1 \le i \le m} r_{x_i}$$

existe y satisface $r_* > 0$.

Definimos $\delta = \frac{r_*}{2}$.

Queremos demostrar que esta única δ funciona para todos los pares de puntos del intervalo.

Sean $u, v \in [a, b]$ tales que $|u - v| < \delta$.

Como los entornos reducidos cubren $[a, b]$, existe algún índice $i \in \{1, \ldots, m\}$ tal que $u \in V_{r_{x_i}/2}(x_i)$.

Es decir, $|u - x_i| < \frac{r_{x_i}}{2}$.

Además, $\delta = \frac{r_*}{2} \le \frac{r_{x_i}}{2}$.

Por la desigualdad triangular,

$$\begin{aligned} |v - x_i| &\leq |v - u| + |u - x_i| \\ &< \delta + \frac{r_{x_i}}{2} \\ &\leq \frac{r_{x_i}}{2} + \frac{r_{x_i}}{2} \\ &= r_{x_i}. \end{aligned}$$

Así, tanto u como v pertenecen al entorno **completo** de radio r_{x_i} alrededor del mismo centro x_i.

Por la elección de r_{x_i}, $|f(u) - f(x_i)| < \frac{\varepsilon}{2}$ y $|f(v) - f(x_i)| < \frac{\varepsilon}{2}$.

Finalmente,

$$\begin{aligned} |f(u) - f(v)| &\leq |f(u) - f(x_i)| + |f(x_i) - f(v)| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon. \end{aligned}$$

Hemos encontrado una $\delta > 0$ que depende de ε, pero no de u ni de v, y que satisface $|u - v| < \delta \Longrightarrow |f(u) - f(v)| < \varepsilon$ para todos $u, v \in [a, b]$. Por tanto, f es uniformemente continua en $[a, b]$. ■

Dónde entra exactamente la compactitud. Es importante no resumir la demostración como

> «una función continua en un intervalo cerrado es uniforme porque el intervalo es compacto».

Eso nombra el teorema, pero oculta el mecanismo. Antes de usar compactitud tenemos infinitos datos locales: $\forall x \in [a, b]\ \exists r_x > 0$. Después de usar compactitud obtenemos finitísimos centros: $x_1, \ldots, x_m$. Solo en ese momento podemos formar

$$\min_{1 \leq i \leq m} r_{x_i} > 0.$$

Por tanto, la función lógica de la compactitud es:

$$\boxed{\text{familia infinita de controles locales} \xrightarrow{\text{compactitud}} \text{familia finita} \xrightarrow{\text{mín}} \text{control uniforme.}}$$

La operación «tomar el mínimo» no es la parte profunda. La parte profunda es justificar que solo necesitamos mirar finitísimos radios.

Por qué usamos radios mitad. También conviene aislar la función del factor $1/2$. La subcubierta nos garantiza que u pertenece a algún entorno reducido: $|u - x_i| < \frac{r_{x_i}}{2}$. Luego queremos admitir que v se mueva una distancia adicional menor que δ.

Si hemos asegurado $\delta \leq \frac{r_{x_i}}{2}$, entonces $|v - x_i| < \frac{r_{x_i}}{2} + \frac{r_{x_i}}{2} = r_{x_i}$.

Ese es exactamente el radio donde sabemos controlar la función. Sin la reserva proporcionada por el radio mitad, el hecho de que u estuviera dentro de $V_{r_{x_i}}(x_i)$ no dejaría margen suficiente para garantizar que un punto v cercano a u continuara dentro del mismo entorno de control.

La técnica es:

$$\boxed{\text{cubrir con la mitad} \quad \Longrightarrow \quad \text{controlar con el radio completo.}}$$

El mismo cuadrado cambia de comportamiento al cambiar el dominio. En Ejemplo 5.9 demostramos que x^2 no es uniformemente continua en $\mathbb{R}$. Pero para cualquier intervalo cerrado $[a, b]$, la restricción $x^2|_{[a,b]}$ sí es uniformemente continua por Teorema 5.6. La fórmula no ha cambiado. Lo que ha cambiado es el dominio. Esto retoma una lección que atraviesa todo el libro:

las propiedades de una función dependen de la función completa, no de una fórmula aislada.

En $\mathbb{R}$, podemos desplazarnos hacia regiones donde la variación de x^2 frente a pequeños incrementos se vuelve arbitrariamente grande. En un intervalo cerrado $[a, b]$, la compactitud impide escapar indefinidamente hacia esas regiones.

Una prueba directa del cuadrado sobre un intervalo acotado. Heine–Cantor proporciona una razón estructural. También podemos ver la uniformidad directamente. Supongamos, por ejemplo, que $x, y \in [-R, R]$ con $R > 0$.

Entonces $|x^2 - y^2| = |x - y|\,|x + y|$.

Como $|x + y| \le |x| + |y| \le 2R$, obtenemos $|x^2 - y^2| \le 2R|x - y|$.

Dado $\varepsilon > 0$, basta tomar $\delta = \frac{\varepsilon}{2R}$ cuando $R > 0$.

Así, $|x - y| < \delta \Longrightarrow |x^2 - y^2| < \varepsilon$.

Obsérvese la diferencia con la prueba de continuidad puntual. Allí aparecía una cota que dependía del centro a. Aquí la acotación del dominio proporciona una cota global: $|x + y| \le 2R$. Este ejemplo permite ver dos vías distintas hacia la uniformidad:

1. una vía **algebraica**, cuando disponemos de una estimación global explícita;
2. una vía **estructural**, dada por Heine–Cantor para cualquier función continua sobre $[a, b]$.

La segunda es mucho más poderosa porque no necesita conocer una fórmula útil para f.

La continuidad uniforme no afirma que la función sea acotada en cualquier dominio. Debemos evitar una inferencia apresurada. La función identidad $f(x) = x$ es uniformemente continua en $\mathbb{R}$, pues $|f(x) - f(y)| = |x - y|$. Sin embargo, no está acotada.

Por tanto,

continuidad uniforme $\not\Rightarrow$ acotación global

sobre dominios arbitrarios.

La acotación global de una función continua en $[a, b]$ proviene del teorema de los valores extremos, no de la definición de continuidad uniforme por sí sola. Las propiedades responden preguntas diferentes:

- continuidad uniforme controla **variaciones entre puntos próximos**;
- acotación controla **el tamaño absoluto de los valores**.

Una diferencia que prepara la integración. La continuidad puntual afirma, en esencia:

> fijado el punto donde estoy, puedo adaptar mi escala de entrada a ese punto.

La continuidad uniforme afirma:

> una vez fijada la precisión de salida, puedo escoger una escala de entrada válida en todo el dominio.

Ese cambio será fundamental más adelante. Cuando dividamos un intervalo en muchas partes y queramos controlar simultáneamente la oscilación de una función en todas ellas, no podremos permitir que cada punto requiera una escala incompatible con las demás. Heine–Cantor garantiza precisamente que, sobre un intervalo cerrado,

$$\boxed{\text{continuidad} + \text{compactitud} \Longrightarrow \text{una escala común de control.}}$$

La teoría formal de integrabilidad de Riemann se desarrollará más adelante. Aquí solo registramos el puente conceptual: la continuidad uniforme es la forma de continuidad que permite sincronizar estimaciones sobre todo un intervalo.

Reconstrucción de la prueba de Heine–Cantor. Sin releer la demostración, el lector debería poder responder:

1. ¿qué diferencia de cuantificadores separa continuidad de continuidad uniforme?
2. ¿por qué uniformidad implica continuidad ordinaria?
3. ¿cómo se niega correctamente la continuidad uniforme?
4. ¿cómo demuestra x^2 en $\mathbb{R}$ que la recíproca «continua $\Rightarrow$ uniformemente continua» es falsa?
5. para un $\varepsilon > 0$ fijo, ¿qué proporciona la continuidad en cada punto $x \in [a, b]$?
6. ¿por qué los entornos $V_{r_x/2}(x)$ forman un recubrimiento abierto?
7. ¿qué extrae exactamente la compactitud de ese recubrimiento?
8. ¿por qué el mínimo de los radios seleccionados es estrictamente positivo?
9. ¿por qué no podríamos tomar sin más el mínimo o el ínfimo de todos los radios antes de usar compactitud?
10. ¿qué función cumple el factor $1/2$ en los entornos de la cubierta?
11. si u pertenece a $V_{r_{x_i}/2}(x_i)$ y $|u - v| < \delta$, ¿cómo demuestra la desigualdad triangular que $v \in V_{r_{x_i}}(x_i)$?
12. ¿por qué comparar simultáneamente $f(u)$ y $f(v)$ con $f(x_i)$ produce la estimación final?
13. ¿en qué punto exacto de la prueba se transforma una familia infinita de decisiones locales en una decisión global?

La respuesta estructural es

$$\boxed{\forall x\ \exists r_x \xrightarrow{\text{compactitud}} r_{x_1}, \dots, r_{x_m} \xrightarrow{\text{mín}} \delta > 0 \xrightarrow{\text{radio mitad}} \forall u, v.}$$

Qué hemos ganado. La continuidad uniforme añade una nueva capa a la cadena local $\to$ global del capítulo. Primero teníamos continuidad en cada punto:

control local dependiente del centro.

Después la compactitud nos permitió sincronizar finitísimos controles:

subcubierta finita.

Finalmente el mínimo de esos radios produjo

un único control global.

Así obtenemos

$$\boxed{f \text{ continua en } [a,b] \Longrightarrow f \text{ uniformemente continua en } [a,b].}$$

La próxima sección utilizará otra vez el teorema del valor intermedio, pero con un objetivo distinto. Preguntaremos qué fuerza adicional aparece cuando una función continua sobre un intervalo es además **inyectiva**. Veremos que la continuidad impide que una función inyectiva cambie arbitrariamente de orientación: deberá ser estrictamente monótona. Después usaremos esa estructura para demostrar que su inversa sobre la imagen también es continua.

5.9. Continuidad + inyectividad en un intervalo: monotonía e inversa continua

Hasta ahora la continuidad ha producido dos tipos de información global sobre intervalos. Por un lado, el teorema del valor intermedio nos dijo que una función continua no puede **saltar valores**. Por otro, la compactitud de un intervalo cerrado nos permitió transformar controles locales en conclusiones globales de acotación, alcanzamiento y uniformidad. Ahora añadiremos una hipótesis distinta: inyectividad.

La pregunta es:

$$\boxed{\text{¿qué aspecto puede tener una función continua que nunca repite un valor?}}$$

Una función arbitraria puede subir, bajar, volver a subir y cruzar varias veces una misma altura. Pero si además es inyectiva, repetir una altura está prohibido. La continuidad convierte esa prohibición en una rigidez mucho más fuerte: sobre un intervalo, la función debe elegir una sola orientación y conservarla.

La propiedad intermedia que impide picos y valles. Supongamos que $x_1 < x_2 < x_3$ son tres puntos de un intervalo I y que $f\colon I \to \mathbb{R}$ es continua e inyectiva. Como f es inyectiva, los tres números

$$f(x_1), \qquad f(x_2), \qquad f(x_3)$$

son distintos.

¿Qué ocurriría si $f(x_2)$ quedara por encima de ambos valores extremos?

Supongamos, por ejemplo, que $f(x_2) > \max\{f(x_1), f(x_3)\}$.

Podemos elegir un número y tal que $\max\{f(x_1), f(x_3)\} < y < f(x_2)$.

Como I es un intervalo,

$$[x_1, x_2] \subseteq I, \qquad [x_2, x_3] \subseteq I.$$

Por el teorema del valor intermedio aplicado al primer subintervalo, existe $u \in (x_1, x_2)$ tal que $f(u) = y$. Aplicando el mismo teorema al segundo subintervalo, existe $v \in (x_2, x_3)$ tal que $f(v) = y$. Pero $u \neq v$, de modo que $f(u) = f(v)$ contradice la inyectividad. El caso en que $f(x_2)$ quede por debajo de ambos extremos es completamente análogo. Por consiguiente, el valor del punto intermedio debe quedar **estrictamente entre** los valores de los extremos:

$$\boxed{\min\{f(x_1), f(x_3)\} < f(x_2) < \max\{f(x_1), f(x_3)\}.}$$

Esta propiedad suele describirse diciendo que la función preserva el *betweenness*: un punto situado entre otros dos no puede convertirse en un pico o un valle nuevo sin obligar a repetir algún valor por continuidad.

De no tener picos ni valles a elegir una orientación. La propiedad anterior todavía no dice explícitamente «creciente» o «decreciente». Debemos demostrar que la orientación no puede cambiar de una región del intervalo a otra. Tomemos dos puntos $p < q$ de I. Por inyectividad, $f(p) \neq f(q)$.

Hay exactamente dos posibilidades: $f(p) < f(q)$ o $f(p) > f(q)$.

Supongamos primero que $f(p) < f(q)$.

Queremos demostrar que entonces $x < y \Longrightarrow f(x) < f(y)$ para **todos** $x, y \in I$. Tomemos $x < y$ arbitrarios. Ordenemos los puntos distintos que aparecen entre $p,\ q,\ x,\ y$ y llamémoslos

$$z_1 < z_2 < \cdots < z_m, \qquad 2 \leq m \leq 4.$$

Para cada tres puntos consecutivos $z_i < z_{i+1} < z_{i+2}$, la propiedad anterior obliga a que $f(z_{i+1})$ quede estrictamente entre $f(z_i)$ y $f(z_{i+2})$. Eso tiene una consecuencia muy concreta: los dos incrementos consecutivos $f(z_{i+1}) - f(z_i)$ y $f(z_{i+2}) - f(z_{i+1})$ deben tener el **mismo signo**. Por propagación a lo largo de la cadena finita, todos los incrementos adyacentes $f(z_{j+1}) - f(z_j)$ tienen el mismo signo. Como $p < q$ y sabemos que $f(p) < f(q)$, ese signo común no puede ser negativo. Por tanto es positivo. En particular, como $x < y$ aparecen en la misma cadena ordenada, $f(x) < f(y)$.

Así, f es estrictamente creciente.

Si en cambio $f(p) > f(q)$, el mismo argumento muestra que todos los incrementos tienen signo negativo y, por tanto, f es estrictamente decreciente. La continuidad ha eliminado la posibilidad de cambiar de orientación sin repetir una altura.

El teorema completo.

Teorema 5.7. **Continuidad, inyectividad, monotonía e inversa.** Sea $I \subseteq \mathbb{R}$ un intervalo y sea

$$f\colon I \to \mathbb{R}$$

continua e inyectiva.

Si I contiene al menos dos puntos, entonces f es exactamente de uno de estos dos tipos:

1. estrictamente creciente en I;
2. estrictamente decreciente en I.

Además, la imagen

$$J := f(I)$$

es un intervalo, la función inversa

$$f^{-1}\colon J \to I$$

está bien definida y es continua en todo punto de J, relativamente a ese dominio. Si I consta de un solo punto, la afirmación sobre la inversa sigue siendo trivial y la cuestión de orientación no añade contenido.

Demostración. La monotonía estricta ya ha sido establecida por el argumento anterior: el teorema del valor intermedio impide picos y valles interiores, y la inyectividad fuerza a que la orientación elegida entre dos puntos se propague a todo el intervalo.

Por Corolario 5.3, como I es un intervalo y f es continua, $J = f(I)$ es también un intervalo. Como f es inyectiva, y es sobreyectiva sobre su propia imagen J, resulta biyectiva como aplicación $f\colon I \to J$. Por la teoría de inversas de `T1-C03`, existe una única función $f^{-1}\colon J \to I$. Resta demostrar que esta inversa es continua. Supongamos primero que f es estrictamente creciente.

Fijemos $y_0 \in J$ y escribamos $x_0 = f^{-1}(y_0)$.

Sea $\varepsilon > 0$.

Nuestro objetivo es encontrar $\delta > 0$ tal que, para todo $y \in J$, $|y - y_0| < \delta \Longrightarrow |f^{-1}(y) - x_0| < \varepsilon$. La dificultad es que x_0 podría ser un punto interior de I, un extremo perteneciente a I, o estar arbitrariamente próximo a un extremo que no pertenece al intervalo. Por eso trabajaremos **relativamente al dominio**, igual que en §5.1. Consideremos primero el lado izquierdo. Si $I \cap (-\infty, x_0 - \varepsilon] = \varnothing$, entonces ningún punto del dominio puede quedar a distancia al menos ε de x_0 por la izquierda. En ese lado no necesitamos imponer ninguna condición adicional. Si, por el contrario, $I \cap (-\infty, x_0 - \varepsilon] \neq \varnothing$, existe algún $z \in I$ con $z \leq x_0 - \varepsilon$. Como I es un intervalo y $z < x_0$, el punto $u = x_0 - \frac{\varepsilon}{2}$ pertenece a I.

La monotonía estricta da $f(u) < f(x_0) = y_0$.

Definimos entonces la separación positiva $d_- = y_0 - f(u) > 0$. De modo análogo, para el lado derecho hay dos posibilidades. Si $I \cap [x_0 + \varepsilon, \infty) = \varnothing$, no existe ningún punto del dominio demasiado lejano por la derecha. Si ese conjunto no es vacío, entonces $v = x_0 + \frac{\varepsilon}{2}$ pertenece a I, y la monotonía estricta produce $f(v) > f(x_0) = y_0$.

Definimos $d_+ = f(v) - y_0 > 0$.

Elegimos ahora $\delta > 0$ tomando el mínimo de las separaciones d_- y d_+ correspondientes a los lados que realmente estén activos. Si solo hay un lado activo, tomamos la separación de ese lado; si ninguno lo está, podemos tomar, por ejemplo, $\delta = 1$. Sea ahora $y \in J$ tal que $|y - y_0| < \delta$, y pongamos $x = f^{-1}(y)$.

Queremos probar $x_0 - \varepsilon < x < x_0 + \varepsilon$.

En el lado izquierdo, si no había puntos de I en $(-\infty, x_0 - \varepsilon]$, la desigualdad $x > x_0 - \varepsilon$ es automática. Si el lado izquierdo estaba activo, entonces $\delta \leq d_-$ y

$$y > y_0 - \delta \geq y_0 - d_- = f(u).$$

Como f es estrictamente creciente,

$$f(x) > f(u) \Longrightarrow x > u = x_0 - \frac{\varepsilon}{2} > x_0 - \varepsilon.$$

El lado derecho es simétrico. Si estaba inactivo, la cota $x < x_0 + \varepsilon$ es automática. Si estaba activo, entonces $\delta \leq d_+$ y

$$y < y_0 + \delta \leq y_0 + d_+ = f(v),$$

de donde, por monotonía estricta,

$$x < v = x_0 + \frac{\varepsilon}{2} < x_0 + \varepsilon.$$

Por tanto, $|x - x_0| < \varepsilon$.

Como $x = f^{-1}(y)$, hemos demostrado que

$$|y - y_0| < \delta \Longrightarrow |f^{-1}(y) - f^{-1}(y_0)| < \varepsilon.$$

Así, f^{-1} es continua en y_0.

Como y_0 era arbitrario, f^{-1} es continua en todo J cuando f es estrictamente creciente. Si f es estrictamente decreciente, consideramos $F = -f$. Por Proposición 5.2, F es continua; además es inyectiva y estrictamente creciente. Por el caso ya demostrado, la inversa de $F\colon I \to -J$ es continua.

Como $F^{-1}(z) = f^{-1}(-z)$, obtenemos $f^{-1}(y) = F^{-1}(-y)$.

La función $y \mapsto -y$ es continua, y la composición de funciones continuas lo es por Proposición 5.2. Por tanto f^{-1} es continua también en el caso decreciente. ■

Dónde entró realmente cada hipótesis. La demostración contiene tres mecanismos diferentes y conviene no mezclarlos. La **continuidad** entra primero mediante IVT:

si aparece un pico o un valle interior, algún valor intermedio debe repetirse.

La **inyectividad** convierte esa repetición forzada en una contradicción. El hecho de que el dominio sea un **intervalo** garantiza dos cosas:

1. que $[x_1, x_2]$ y $[x_2, x_3]$ permanezcan dentro del dominio al aplicar IVT;
2. que los puntos de barrera $x_0 \pm \varepsilon/2$ pertenezcan al dominio cuando existe realmente un tramo suficientemente largo hacia ese lado.

Finalmente, una vez obtenida la monotonía estricta, la continuidad de la inversa se demuestra por **orden**: valores de salida suficientemente próximos a y_0 quedan atrapados entre imágenes de puntos situados a ambos lados de x_0.

La arquitectura completa es

$$\boxed{\text{continuidad} + \text{inyectividad} + \text{intervalo} \xrightarrow{\text{IVT}} \text{monotonía estricta} \xrightarrow{\text{orden}} \text{inversa continua.}}$$

No aparece ninguna derivada.

Por qué la continuidad no puede eliminarse. En `T1-C03` ya vimos que una función estrictamente monótona no tiene por qué ser continua.

Por ejemplo,

$$f(x) = \begin{cases} x, & x < 0, \\ x + 1, & x \geq 0 \end{cases}$$

es estrictamente creciente en $\mathbb{R}$, pero tiene un salto en 0. Su imagen es $(-\infty, 0) \cup [1, \infty)$, que **no** es un intervalo. Esto muestra exactamente lo que aporta la continuidad al teorema: impide huecos en la imagen y, mediante IVT, transforma la mera monotonía en una estructura sin saltos.

La conclusión

$$f(I) \text{ es un intervalo}$$

no proviene de la monotonía estricta por sí sola; proviene de continuidad + IVT.

Una consecuencia fundamental: continuidad de la raíz cuadrada.

Ejemplo 5.10. **El cuadrado restringido y la continuidad de $\sqrt{\ }$.** Consideremos

$$f\colon [0, \infty) \to [0, \infty), \qquad f(x) = x^2.$$

Entonces f es continua, estrictamente creciente y biyectiva. Su inversa es $f^{-1}(y) = \sqrt{y}$. Por Teorema 5.7, $\sqrt{}: [0,\infty) \to [0,\infty)$ es continua en todo su dominio.

La continuidad de f proviene de Corolario 5.1, porque x^2 es un polinomio. La monotonía estricta sobre $[0,\infty)$ puede verificarse directamente: si $0 \le x < y$, entonces $y^2 - x^2 = (y-x)(x+y) > 0$, por lo que $x^2 < y^2$. La existencia de la raíz no negativa de todo $y \ge 0$ fue establecida mediante completitud en `T1-C02`; por tanto la imagen del cuadrado restringido es exactamente $[0,\infty)$. Así, un resultado que antes conocíamos como existencia algebraico-ordenada, $y \mapsto \sqrt{y}$, adquiere ahora una nueva propiedad analítica:

$$\boxed{\sqrt{} \text{ es continua en } [0,\infty).}$$

En el extremo 0 esta continuidad debe leerse relativamente al dominio. De hecho, podemos verificarla directamente.

Sea $\varepsilon > 0$ y tomemos $\delta = \varepsilon^2$.

Si $y \in [0,\infty)$ y $|y - 0| < \delta$, entonces $0 \le y < \varepsilon^2$, y por la monotonía del cuadrado en los no negativos, $0 \le \sqrt{y} < \varepsilon$.

Por tanto, $|\sqrt{y} - \sqrt{0}| < \varepsilon$.

No necesitamos extender la raíz cuadrada a números negativos para hablar de continuidad en el extremo del dominio.

La inversa refleja el orden. Cuando f es estrictamente creciente, Corolario 2.2 ya nos dice que f^{-1} también es estrictamente creciente sobre $f(I)$. Cuando f es estrictamente decreciente, su inversa es también estrictamente decreciente. El teorema de esta sección añade algo que `T1-C03` todavía no podía afirmar:

si la función original es además continua sobre un intervalo, entonces la inversa no solo conserva el orden: también es continua.

Así quedan conectadas tres capas construidas en capítulos distintos:

$$\text{estructura funcional} \to \text{orden} \to \text{continuidad}.$$

Reconstrucción de la demostración. Sin releer la prueba, conviene poder responder:

1. si $x_1 < x_2 < x_3$, ¿por qué $f(x_2)$ no puede quedar por encima de $f(x_1)$ y $f(x_3)$?
2. ¿dónde aparece exactamente el teorema del valor intermedio en esa contradicción?
3. ¿por qué la inyectividad convierte dos cruces de la misma altura en una imposibilidad?
4. ¿qué significa que $f(x_2)$ quede estrictamente entre $f(x_1)$ y $f(x_3)$?
5. ¿por qué esa propiedad fuerza a todos los incrementos sucesivos a tener el mismo signo?
6. ¿cómo basta comparar inicialmente un solo par $p < q$ para fijar la orientación global?
7. ¿por qué $f(I)$ es un intervalo?
8. ¿qué resultado de `T1-C03` permite definir $f^{-1}: f(I) \to I$?
9. para demostrar continuidad de la inversa, ¿qué significan los lados «activos» alrededor de x_0?
10. ¿por qué, si existe un punto de I a distancia al menos ε por la izquierda, también pertenece a I el punto $x_0 - \varepsilon/2$?

11. ¿cómo transforma la monotonía una separación horizontal alrededor de x_0 en una separación vertical alrededor de $y_0 = f(x_0)$?
12. ¿por qué el argumento funciona también en extremos del intervalo sin añadir valores fuera del dominio?
13. ¿cómo se reduce el caso decreciente al creciente mediante $-f$?
14. ¿qué hipótesis impide el salto del ejemplo estrictamente creciente de `T1-C03`?

El núcleo lógico puede resumirse así:

$$\boxed{\text{IVT impide saltarse alturas} + \text{inyectividad impide repetirlas} \Longrightarrow \text{una sola orientación.}}$$

Y una vez fijada esa orientación,

$$\boxed{\text{orden de } f + \text{barreras laterales} \Longrightarrow \text{continuidad de } f^{-1}.}$$

Qué hemos ganado. Con esta sección se cierra la última pieza teórica nueva del capítulo. Partimos de continuidad como una condición local en un punto y hemos llegado a una familia de resultados globales:

$$\boxed{\begin{array}{c}\text{continuidad local}\\ \Downarrow\\ \text{IVT / bisección / compactitud / EVT / Heine–Cantor}\\ \Downarrow\\ \text{monotonía e inversas continuas bajo inyectividad.}\end{array}}$$

La siguiente sección no introducirá nuevas definiciones ni nuevos teoremas. Funcionará como **laboratorio de continuidad**: reunirá las herramientas del capítulo y obligará a decidir, en cada situación, qué hipótesis está produciendo realmente cada conclusión.

5.10. Laboratorio de continuidad: del control local al global

Hasta aquí hemos construido la teoría de continuidad por capas. Primero aprendimos a controlar una función en un punto; después transportamos ese control mediante sucesiones, álgebra y composición; distinguimos discontinuidades; atravesamos valores mediante IVT; localizamos raíces por bisección; convertimos controles locales en información global mediante compactitud y EVT; uniformizamos radios con Heine–Cantor; y finalmente demostramos que continuidad + inyectividad sobre un intervalo fuerzan monotonía estricta e inversa continua.

Ahora cambia la tarea.

En las seis estaciones que siguen **no aparecerá teoría nueva**. Tampoco introduciremos nuevos objetos formales. La dificultad consistirá en reconocer qué resultado es pertinente, qué hipótesis lo activa y qué conclusión podemos obtener legítimamente sin importar herramientas posteriores. Conviene trabajar con este protocolo:

$$\boxed{\text{dominio} \to \text{propiedad local} \to \text{hipótesis globales} \to \text{teorema autorizado} \to \text{conclusión} \to \text{auditoría}}$$

La pregunta ya no será solamente

¿es continua esta función?,

sino también

> **¿continua dónde, con respecto a qué dominio, y qué puede concluirse después de saberlo?**

Estación 1 — El dominio decide qué significa continuidad. Consideremos $A = \{-1\} \cup [0,4]$ y la función $f\colon A \to \mathbb{R}$ definida por

$$f(-1) = 7, \qquad f(x) = \sqrt{x} \quad (0 \le x \le 4).$$

Problema. Analizar la continuidad de f en los puntos

$$-1, \qquad 0, \qquad 4,$$

y explicar por qué los tres casos requieren lecturas locales distintas aunque pertenezcan a una sola función.

En -1: continuidad por aislamiento. El punto -1 es aislado en A. Por ejemplo, $V_{1/2}(-1) \cap A = \{-1\}$. Por Proposición 5.1, **toda** función definida sobre A es continua en -1, independientemente del valor que tome allí. El número $f(-1) = 7$ no necesita estar relacionado con los valores de $\sqrt{x}$ sobre $[0,4]$, porque alrededor de -1 no existen otros puntos del dominio que puedan acercarse al centro.

En 0: continuidad unilateral producida por el dominio. En §5.9 demostramos que $\sqrt{}\colon [0,\infty) \to [0,\infty)$ es continua. Su restricción a $[0,4]$ es continua en 0, pero nuestra función tiene el dominio algo mayor A, que contiene también el punto aislado -1. Por eso debemos excluir localmente esa otra componente. No debemos pedir información sobre números negativos arbitrariamente próximos a 0: no existen en el dominio. El único punto negativo de A es -1, situado a distancia 1.

La condición relevante es

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in A: \quad |x| < \delta \Longrightarrow |f(x) - f(0)| < \varepsilon.$$

Tomamos

$$\delta = \min\left\{\frac{1}{2}, \varepsilon^2\right\}.$$

Si $x \in A$ y $|x| < \delta$, entonces $x \neq -1$ y por tanto $x \in [0,4]$. Además $0 \le x < \varepsilon^2$, de donde $|f(x) - f(0)| = \sqrt{x} < \varepsilon$. Así, f es continua en 0 relativamente al dominio completo A.

En 4: el otro extremo. En 4 ocurre la situación simétrica respecto de la geometría del intervalo. Los puntos del dominio próximos a 4 solo pueden llegar desde la izquierda. Como $\sqrt{x}$ es continua en su dominio, la restricción a $[0,4]$ es continua también en 4. No necesitamos extender artificialmente la función a números mayores que 4 para hablar de continuidad de la restricción.

Lectura de la estación. Los tres puntos muestran tres geometrías locales distintas:

-1 : punto aislado,
0 : extremo izquierdo de una componente intervalar,
4 : extremo derecho de una componente intervalar.

La continuidad no pertenece a una fórmula desnuda. Pertenece a una **función con dominio** y se interpreta relativamente a ese dominio. Antes de calcular límites o manipular expresiones debemos preguntar:

¿qué puntos del dominio pueden acercarse realmente al centro?

Estación 2 — Composición: por qué el centro deja de ser peligroso. En §4.9 aprendimos que la composición de límites funcionales tenía una dificultad delicada: el límite exterior ignoraba su punto central, de modo que era necesario controlar que la función interior no golpeara precisamente ese punto. Comparemos ahora dos funciones exteriores.

Definamos

$$\phi(y) = \begin{cases} 0, & y \neq 0, \\ 1, & y = 0, \end{cases}$$

y $\psi(y) = y^2 + 1$.

Sea además

$$g(x) = 0 \qquad (x \in \mathbb{R}).$$

Problema. Explicar por qué la composición con g se comporta de manera radicalmente distinta según usemos ϕ o ψ.

El caso de ϕ. La función ϕ tiene límite perforado

$$\lim_{y \to 0} \phi(y) = 0,$$

pero $\phi(0) = 1$.

Por Proposición 5.1, ϕ no es continua en 0. Como $g(x) = 0$ para todo x, la composición es $(\phi \circ g)(x) = \phi(0) = 1$.

Por tanto

$$\lim_{x \to a} (\phi \circ g)(x) = 1$$

para cualquier a, y no 0.

Esto es exactamente el tipo de fenómeno que obligó a introducir la condición de no impacto en Teorema 4.5.

El caso de ψ. La función $\psi(y) = y^2 + 1$ es continua en 0, y g es continua en todo punto porque es constante. Por Proposición 5.2, $\psi \circ g$ es continua.

En efecto, $(\psi \circ g)(x) = 1$ para todo x.

La función interior golpea el punto central 0 **en cada entrada**, pero eso ya no produce ninguna dificultad.

¿Qué cambió? En el límite perforado exterior, el antecedente era $0 < |y - 0| < \eta$. Para continuidad de ψ en 0, en cambio, basta $|y - 0| < \eta$. El caso $y = 0$ está incluido y perfectamente controlado.

Por tanto,

continuidad exterior $\Longrightarrow$ el valor central deja de ser una zona ciega.

Esta estación muestra por qué no debemos memorizar que «en continuidad la composición siempre funciona» como una regla aislada. Funciona porque la definición exterior controla también el centro.

Estación 3 — IVT garantiza existencia sin entregar una fórmula. Consideremos $p(x) = x^5 + x - 1$.

Problema. Demostrar que la ecuación $x^5 + x - 1 = 0$ tiene una única solución en $(0, 1)$ y localizarla en un intervalo más pequeño, sin derivadas y sin intentar resolver algebraicamente la ecuación de quinto grado.

Existencia. Como p es un polinomio, es continua en $\mathbb{R}$ por Corolario 5.1.

Además, $p(0) = -1 < 0$ y $p(1) = 1 > 0$.

Por Corolario 5.2 existe al menos un punto $c \in (0, 1)$ tal que $p(c) = 0$.

La conclusión es existencial. El IVT no nos entrega una fórmula cerrada para c.

Una primera localización. Probemos valores racionales sencillos.

Para $x = 3/4$,

$$p\left(\frac{3}{4}\right) = \left(\frac{3}{4}\right)^5 + \frac{3}{4} - 1 = \frac{243}{1024} - \frac{256}{1024} = -\frac{13}{1024} < 0.$$

Para $x = 4/5$,

$$p\left(\frac{4}{5}\right) = \left(\frac{4}{5}\right)^5 + \frac{4}{5} - 1 = \frac{1024}{3125} - \frac{1}{5} = \frac{399}{3125} > 0.$$

Otra aplicación de Bolzano proporciona

$$\boxed{\frac{3}{4} < c < \frac{4}{5}.}$$

Si quisiéramos una localización cada vez más precisa, §5.5 nos autoriza a continuar mediante bisección.

Unicidad sin derivadas. Sean $0 \le x < y \le 1$.

Entonces

$$\begin{aligned} p(y) - p(x) &= y^5 - x^5 + y - x \\ &= (y - x)(y^4 + y^3x + y^2x^2 + yx^3 + x^4 + 1). \end{aligned}$$

Todos los términos del segundo factor son no negativos y aparece además el término 1. Por tanto, $p(y) - p(x) > 0$. Así, p es estrictamente creciente en $[0, 1]$. Una función estrictamente creciente no puede tomar el valor 0 en dos puntos distintos.

La raíz es única.

Lectura de la estación. Aquí intervinieron tres tareas distintas:

IVT $\Rightarrow$ existencia,
evaluaciones + IVT $\Rightarrow$ localización,
monotonía $\Rightarrow$ unicidad.

No debemos confundirlas. Saber que existe una solución no equivale a calcularla; encerrarla en un intervalo tampoco demuestra por sí solo que sea única.

Estación 4 — Compactitud: cuándo los controles locales se vuelven globales. Sea $f(x) = \frac{1}{3-x}$.

Comparemos dos dominios: $A = [-2, 2]$ y $B = (-2, 3)$.

Problema. Determinar qué conclusiones globales permite la continuidad en cada caso y explicar exactamente qué papel desempeña la compactitud.

Sobre $A = [-2, 2]$. El denominador satisface

$$3 - x \geq 1 > 0 \qquad (x \in [-2, 2]),$$

por lo que f es continua en todo A. Como A es un intervalo cerrado y acotado, Teorema 5.4 nos dice que es compacto. Entonces Teorema 5.5 garantiza que f es acotada y alcanza un máximo y un mínimo. En este ejemplo podemos identificarlos además por orden. Si $x < y < 3$, entonces $3 - x > 3 - y > 0$, y al tomar recíprocos positivos, $\frac{1}{3-x} < \frac{1}{3-y}$. Por tanto f es estrictamente creciente sobre A y $\min_A f = f(-2) = \frac{1}{5}$,

$$\max_A f = f(2) = 1.$$

Sobre $B = (-2, 3)$. La misma fórmula define una función continua en B. Sin embargo, B no es compacto. De hecho, f ni siquiera es acotada superiormente allí.

Sea $M > 0$. Elegimos $x = 3 - \frac{1}{M+1}$.

Entonces $x \in (-2, 3)$ y $f(x) = M + 1 > M$.

Por tanto no existe una cota superior global. La continuidad puntual sigue siendo cierta; lo que falla es el mecanismo que permite extraer de infinitos controles locales una cantidad finita de controles que cubran todo el dominio.

La arquitectura local–global. Para una función continua en $[a, b]$, cada punto proporciona un entorno en el cual la función está localmente controlada. La compactitud permite pasar de

un entorno para cada punto

a

finitos entornos que cubren todo $[a, b]$.

Una vez el problema se vuelve finito, podemos tomar máximos de finitísimas cotas o mínimos de finitísimos radios. Ese mecanismo es el corazón común de varios resultados del capítulo:

$$\boxed{\text{continuidad local} + \text{compactitud} \Longrightarrow \text{control global finito.}}$$

EVT y Heine–Cantor utilizan ese motor de maneras distintas.

Estación 5 — El mismo cuadrado: continuidad puntual frente a uniformidad. Consideremos nuevamente $f(x) = x^2$.

Sabemos que es continua en $\mathbb{R}$. Pero §5.8 mostró que no es uniformemente continua en todo $\mathbb{R}$.

Ahora restrinjamos el dominio a $[0, 1]$.

Problema. Demostrar directamente que la restricción es uniformemente continua y explicar por qué el argumento ya no depende del centro.

Sean $x, y \in [0, 1]$. Entonces $|x^2 - y^2| = |x - y|\,|x + y|$.

Como $0 \leq x, y \leq 1$, tenemos $|x + y| \leq 2$.

Por tanto, $|x^2 - y^2| \leq 2|x - y|$.

Sea $\varepsilon > 0$. Tomemos $\delta = \frac{\varepsilon}{2}$.

Si $x, y \in [0, 1]$ y $|x - y| < \delta$, entonces $|x^2 - y^2| \leq 2|x - y| < 2\delta = \varepsilon$.

Así, x^2 es uniformemente continua en $[0, 1]$.

Compare ahora los dos controles. En §5.1, para demostrar continuidad puntual de x^2 en un centro arbitrario $a \in \mathbb{R}$, usamos un radio del tipo

$$\delta(a, \varepsilon) = \min\left\{1, \frac{\varepsilon}{2|a|+1}\right\}.$$

El radio se deteriora cuando $|a|$ crece. En $[0,1]$, en cambio, tenemos una cota global $|x+y| \leq 2$, válida simultáneamente para **todos** los pares del dominio. Por eso podemos elegir $\delta(\varepsilon) = \frac{\varepsilon}{2}$ sin mencionar un centro. La diferencia puede verse como

continuidad puntual: una familia de radios δ_a,
continuidad uniforme: un solo radio sirve en todo el dominio.

Heine–Cantor afirma que, aunque no encontremos una estimación algebraica tan cómoda como la anterior, la compactitud de $[a,b]$ garantiza que una función continua siempre admite ese control uniforme.

Estación 6 — Una inversa continua que no necesitamos escribir. Consideremos

$$F\colon \mathbb{R} \to \mathbb{R}, \qquad F(x) = x^3 + x.$$

Problema. Demostrar que F posee una inversa definida en todo $\mathbb{R}$ y que esa inversa es continua, sin obtener una fórmula explícita para F^{-1}.

Paso 1: continuidad. F es un polinomio. Por Corolario 5.1 es continua en $\mathbb{R}$.

Paso 2: monotonía estricta. Sean $x < y$. Entonces

$$\begin{aligned} F(y) - F(x) &= y^3 - x^3 + y - x \\ &= (y-x)(y^2 + xy + x^2 + 1). \end{aligned}$$

Además,

$$y^2 + xy + x^2 = \left(x + \frac{y}{2}\right)^2 + \frac{3y^2}{4} \geq 0.$$

Por tanto, $y^2 + xy + x^2 + 1 > 0$.

Como $y - x > 0$, concluimos $F(y) > F(x)$.

Así, F es estrictamente creciente y, en particular, inyectiva.

Paso 3: sobreyectividad mediante IVT. Sea $t \in \mathbb{R}$ arbitrario. Queremos demostrar que existe x tal que $F(x) = t$.

Tomemos $M = |t| + 1 > 0$.

Entonces

$$F(M) = M^3 + M > M > |t| \geq t,$$

mientras que

$$F(-M) = -M^3 - M < -M < -|t| \leq t.$$

La función $x \longmapsto F(x) - t$ es continua y toma signos opuestos en $-M$ y M. Por Bolzano existe $c \in (-M, M)$ tal que $F(c) - t = 0$.

Es decir, $F(c) = t$.

Como t era arbitrario, F es sobreyectiva sobre $\mathbb{R}$. Ya sabemos que es inyectiva. Por tanto, $F\colon \mathbb{R} \to \mathbb{R}$ es biyectiva.

Paso 4: continuidad de la inversa. Ahora se activa Teorema 5.7. F es continua e inyectiva sobre el intervalo $\mathbb{R}$. En consecuencia, su inversa $F^{-1}\colon \mathbb{R} \to \mathbb{R}$ es continua. No hemos resuelto explícitamente la ecuación cúbica $x^3 + x = t$.

No lo necesitábamos.

La teoría nos permite concluir

$$\boxed{\text{para cada } t \in \mathbb{R} \text{ existe un único } x, \text{ y ese } x \text{ depende continuamente de } t.}$$

Lectura de la estación. La cadena completa es

$$\boxed{\text{polinomio} \to \text{continuidad} \to \text{monotonía estricta} \to \text{inyectividad}}$$

junto con

$$\boxed{\text{IVT} + \text{valores a ambos lados de cualquier nivel} \to \text{sobreyectividad.}}$$

Entonces

$$\boxed{\text{biyectividad} + \text{continuidad sobre un intervalo} \to \text{inversa continua.}}$$

Este es un ejemplo importante de cómo un teorema estructural puede decirnos mucho más que una fórmula explícita.

Auditoría transversal — ¿qué herramienta responde a qué pregunta? Las seis estaciones pueden condensarse en una tabla de decisión.

Pregunta	Herramienta principal	Hipótesis que deben auditarse	Conclusión típica
¿Es continua en un punto concreto?	definición / Proposición 5.1 / Teorema 5.1	dominio, punto aislado o de acumulación	control local
¿Puedo componer funciones continuas?	Proposición 5.2	continuidad interior y exterior en el valor correcto	continuidad de la composición
¿Debe existir una solución entre dos puntos?	IVT / Bolzano	intervalo + continuidad + valor intermedio o cambio de signo	existencia
¿Puedo pasar de controles locales a una conclusión en todo $[a, b]$?	compactitud / EVT	dominio compacto + continuidad	acotación y extremos alcanzados
¿Puede elegirse una sola δ para todo el dominio?	definición uniforme / Heine–Cantor	control global directo o continuidad sobre $[a, b]$	continuidad uniforme

Pregunta	Herramienta principal	Hipótesis que deben auditarse	Conclusión típica
¿La inversa existe y es continua?	inyectividad + IVT + Teorema 5.7	dominio intervalar, continuidad, inyectividad; sobreyectividad sobre el codominio pretendido	inversa continua

La tabla no sustituye las demostraciones. Su función es impedir dos errores frecuentes:

1. aplicar un teorema antes de verificar sus hipótesis;
2. pedirle a un teorema una conclusión que no promete.

IVT, por ejemplo, prueba existencia pero no da automáticamente unicidad. EVT garantiza extremos alcanzados sobre un intervalo cerrado, pero no calcula dónde están. Heine–Cantor garantiza uniformidad, pero no necesariamente ofrece el mejor radio posible. El teorema de inversa continua garantiza continuidad de la inversa, pero primero debemos saber sobre qué imagen o codominio esa inversa está realmente definida.

El mapa local–global del capítulo. Podemos ahora leer todo `T1-C06` como una sola cadena conceptual. Comenzamos con un punto a y una condición local: $|x-a|<\delta \Longrightarrow |f(x)-f(a)|<\varepsilon$. Después aprendimos que la geometría del dominio decide qué aproximaciones son posibles. Al movernos sobre un intervalo, la continuidad local impide saltos de valores. Al trabajar sobre un intervalo cerrado, la compactitud permite reemplazar una familia potencialmente infinita de controles locales por una familia finita. Esa finitud produce cotas globales, extremos alcanzados y radios uniformes.

Finalmente, combinada con inyectividad, la continuidad sobre intervalos impide cambios arbitrarios de orientación y hace continua la operación de deshacer la función.

La trayectoria completa es

$$\boxed{\text{punto} \to \text{entorno} \to \text{intervalo} \to \text{recubrimiento finito} \to \text{control global.}}$$

Pero hay una segunda trayectoria, igualmente importante:

$$\boxed{\begin{array}{c}\text{continuidad local}\\ \to \text{IVT}\\ \to \text{existencia}\\ \to \text{bisección / monotonía}\\ \to \text{localización, unicidad e inversión.}\end{array}}$$

Y una tercera:

$$\boxed{\text{continuidad punto a punto} + \text{compactitud} \to \text{continuidad uniforme.}}$$

En estas consecuencias intervienen dos datos diferentes: la continuidad controla las imágenes cerca de cada punto y la estructura del dominio permite extender o coordinar esos controles. Conviene conservar ambas condiciones al aplicar los resultados.

Antes de pasar a los ejercicios. El lector debería poder ahora distinguir con precisión las siguientes afirmaciones:

1. una función puede ser continua en un punto aislado sin que formulemos allí un límite perforado;
2. continuidad de una composición no necesita la condición de no impacto de §4.9 porque el centro exterior está controlado;
3. IVT garantiza valores intermedios, pero no por sí solo unicidad;
4. la continuidad sobre un intervalo abierto o no compacto no garantiza acotación ni alcanzamiento de extremos;
5. continuidad puntual no significa continuidad uniforme;
6. compactitud de $[a, b]$ convierte controles locales en controles finitos globales;
7. continuidad + inyectividad sobre un intervalo fuerza monotonía estricta e inversa continua;
8. una inversa puede conocerse cualitativamente —existencia, unicidad y continuidad— sin disponer de una fórmula cerrada.

Con esto termina la parte expositiva del capítulo. En los ejercicios habrá que reconocer qué estructura del dominio utiliza cada teorema, combinar los controles disponibles y construir contraejemplos cuando falte una hipótesis. Las diferencias entre existencia, unicidad y regularidad seguirán determinando qué conclusión está justificada.

5.11. Ejercicios y soluciones

Llegamos al banco final de `T1-C06`. Las secciones anteriores construyeron continuidad primero como condición local y después como fuente de resultados globales sobre intervalos: criterio secuencial, álgebra y composición, discontinuidades, valor intermedio, bisección, compactitud, valores extremos, continuidad uniforme e inversas continuas. Ahora las herramientas dejan de aparecer separadas por subtítulos. Antes de comenzar cada problema conviene auditar cuatro preguntas:

1. ¿la afirmación es local o global?;
2. ¿qué geometría tiene el dominio: punto aislado, intervalo, intervalo cerrado, dominio no compacto?;
3. ¿qué hipótesis exige exactamente el teorema que queremos usar?;
4. ¿se pide existencia, unicidad, localización, una cota o una propiedad uniforme?

El banco contiene exactamente cuarenta ejercicios:

$$7A + 7B + 7C + 6D + 5E + 5F + 3G = 40.$$

Todos son originales y pueden resolverse con los capítulos publicados desde los números reales hasta el presente. No se necesitan derivadas, teorema de Rolle, teorema del valor medio diferencial, integrales, series, Bolzano–Weierstrass, teoría sistemática de subsucesiones, criterio de Cauchy, espacios métricos ni topología general.

Nivel A — Reconocimiento y comprensión. Ejercicio A1. Continuidad relativa al dominio. Sea $A = \{-1\} \cup [0, 2]$ y define $f : A \to \mathbb{R}$ por

$$f(-1) = 50, \qquad f(x) = x^2 \quad (0 \le x \le 2).$$

Decide si f es continua en -1, en 0 y en 2. En cada caso explica qué puntos del dominio pueden acercarse al centro.

Ejercicio A2. Aislado, acumulación o exterior. Considera

$$A = \{0\} \cup \left\{ \frac{1}{n} : n \geq 1 \right\} \cup [2, 3].$$

Clasifica cada punto como punto aislado de A, punto de acumulación de A, ambas cosas si fuera posible, o punto que no pertenece a A:

$$0, \qquad 1, \qquad \frac{1}{2}, \qquad 2, \qquad 3, \qquad 4.$$

Indica en cuáles de los puntos del dominio toda función $A \to \mathbb{R}$ es automáticamente continua por aislamiento.

Ejercicio A3. Límite existente no significa continuidad. Sea

$$f(x) = \begin{cases} x^2, & x \neq 1, \\ 5, & x = 1. \end{cases}$$

Calcula

$$\lim_{x \to 1} f(x)$$

y decide si f es continua en 1. ¿Qué único cambio en el valor puntual repararía la continuidad sin modificar el límite perforado?

Ejercicio A4. Tres maneras de fallar. Clasifica el defecto de continuidad indicado en cada caso como **removible**, **salto** u **oscilación persistente**.

1. $$r(x) = \begin{cases} \dfrac{x^2 - 1}{x - 1}, & x \neq 1, \\ 100, & x = 1; \end{cases}$$
   estudiar $x = 1$.
2. $$s(x) = \begin{cases} -1, & x < 0, \\ 1, & x \geq 0; \end{cases}$$
   estudiar $x = 0$.
3. $$q(x) = \begin{cases} 1, & x \in \mathbb{Q}, \\ -1, & x \notin \mathbb{Q}; \end{cases}$$
   estudiar $x = 0$.

Justifica cada clasificación mediante límites o sucesiones apropiadas.

Ejercicio A5. Qué promete cada teorema. Para cada afirmación indica qué teorema responde directamente a la conclusión pedida: **IVT**, **EVT**, ambos o ninguno. Si otro teorema también tiene sus hipótesis satisfechas pero no produce esa conclusión, indícalo.

1. $f : [0, 1] \to \mathbb{R}$ es continua y $f(0) < 0 < f(1)$; se quiere demostrar que existe una raíz.
2. $f : (0, 1) \to \mathbb{R}$ es continua; se quiere afirmar que alcanza máximo y mínimo.
3. $f : [-2, 2] \to \mathbb{R}$ es continua; se quiere afirmar que es acotada y alcanza extremos.

4. $f : [a, b] \to \mathbb{R}$ es continua y y está entre $f(a)$ y $f(b)$; se quiere demostrar que y pertenece a la imagen.
5. $f : [0, 1] \to \mathbb{R}$ satisface $f(0) < 0 < f(1)$, pero no se conoce continuidad; se quiere demostrar que tiene una raíz.

En los casos donde un teorema no sea aplicable, identifica la hipótesis ausente.

Ejercicio A6. Dos órdenes de cuantificadores. Compara $\forall a \in A\ \forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in A$ con $\forall \varepsilon > 0\ \exists \delta > 0\ \forall x, y \in A$.

1. ¿Cuál corresponde a continuidad punto a punto y cuál a continuidad uniforme?
2. ¿En cuál de las dos puede depender δ del centro a?
3. Explica por qué la segunda afirmación es, en ese sentido, globalmente más exigente.

Ejercicio A7. Recubrimientos y compactitud.

1. Verifica que
$$U_1 = (-1, 3/5), \qquad U_2 = (2/5, 2)$$
forman un recubrimiento abierto de $[0, 1]$. ¿Hay un subrecubrimiento finito?
2. Para $n \geq 2$, define
$$V_n = \left(\frac{1}{n}, 1\right).$$
Demuestra que $\{V_n : n \geq 2\}$ recubre $(0, 1)$, pero ninguna subfamilia finita lo recubre.
3. Explica por qué el segundo apartado no contradice la compactitud de los intervalos cerrados demostrada en Teorema 5.4.

Nivel B — Aplicación directa. Ejercicio B1. Polinomios y racionales. Demuestra que $f(x) = \frac{x^3-2x+1}{x^2+1}$ es continua en todo $\mathbb{R}$. Audita explícitamente el denominador y calcula $f(1)$.

Ejercicio B2. Valor absoluto y composición. Demuestra que $g(x) = |x^2 - 4x + 3|$ es continua en $\mathbb{R}$. Calcula además $g(1)$ y $g(3)$.

Ejercicio B3. Ajustar una unión por tramos. Sea
$$f_a(x) = \begin{cases} 2x + a, & x \leq 1, \\ x^2 + 1, & x > 1. \end{cases}$$
Determina el único valor de a para el cual f_a es continua en 1. Justifica por qué con ese valor la función es continua en todo $\mathbb{R}$.

Ejercicio B4. Criterio secuencial en acción. Usa Teorema 5.1 para demostrar que $f(x) = \frac{1}{1+x^2}$ es continua en 0. La prueba debe comenzar con una sucesión arbitraria $x_n \to 0$ y concluir que $f(x_n) \to f(0)$.

Ejercicio B5. Una raíz sin fórmula. Demuestra que la ecuación $x^3 + x - 1 = 0$ tiene al menos una solución en $(0, 1)$. Indica con precisión qué teorema utilizas y qué hipótesis verificas.

Ejercicio B6. Extremos sin derivadas. Considera $f(x) = x^4 - 2x^2$ en $[-2, 2]$.

1. Explica por qué f alcanza máximo y mínimo.
2. Sin derivadas, determina esos valores y los puntos donde se alcanzan usando la sustitución $t = x^2$.

Ejercicio B7. Uniformidad directa. Demuestra directamente desde Definición 5.4 que $f(x) = 3x - 2$ es uniformemente continua en $\mathbb{R}$. Da una elección explícita de δ en función de ε.

Nivel C — Combinación estructural. Ejercicio C1. Álgebra más composición. Demuestra que

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

es continua en todo $\mathbb{R}$. La solución debe identificar las funciones básicas y las operaciones que preservan continuidad.

Ejercicio C2. Tres parámetros y dos uniones. Define

$$g(x) = \begin{cases} ax + b, & x < 1, \\ x^2, & 1 \le x \le 2, \\ 3x + c, & x > 2. \end{cases}$$

Determina $a, b, c \in \mathbb{R}$ de modo que g sea continua en $\mathbb{R}$ y además $g(0) = 0$.

Ejercicio C3. Existencia y unicidad con IVT y orden. Sea $p(x) = x^3 + x - 1$.

1. Demuestra que existe una raíz c en
$$\left(\frac{1}{2}, \frac{3}{4}\right).$$
2. Demuestra algebraicamente que p es estrictamente creciente en $\mathbb{R}$.
3. Concluye que esa raíz es la única raíz real.

Ejercicio C4. Bisección y error. Aplica bisección a $p(x) = x^3 + x - 1$ partiendo de $I_0 = [0, 1]$.

1. Calcula los primeros cuatro intervalos conservados I_1, I_2, I_3, I_4.
2. Si I_n es el intervalo tras n bisecciones y m_n su punto medio, demuestra que la raíz c satisface $|m_n - c| \le \frac{1}{2^{n+1}}$.
3. Determina cuántas bisecciones bastan para garantizar un error menor que 10^{-3}.

Ejercicio C5. De acotación local a acotación global. Sea $f : [0, 1] \to \mathbb{R}$ continua.

1. Para cada $a \in [0, 1]$, usa continuidad con $\varepsilon = 1$ para obtener un entorno relativo en el cual $|f(x)| < |f(a)| + 1$.
2. Explica por qué esos entornos forman un recubrimiento abierto de $[0, 1]$.
3. Usa compactitud para extraer una subcubierta finita y deducir una constante M tal que
$$|f(x)| \le M \qquad (0 \le x \le 1).$$

Ejercicio C6. Heine–Cantor como herramienta de transporte. Sea $p(x) = x^4 - 3x$ en $[0, 1]$.

1. Demuestra que p es continua en $[0, 1]$ y concluye por Teorema 5.6 que es uniformemente continua allí.
2. Sean (x_n) y (y_n) sucesiones en $[0, 1]$ tales que $|x_n - y_n| \to 0$. Demuestra, usando la continuidad uniforme obtenida en el primer apartado, que $|p(x_n) - p(y_n)| \to 0$.

Ejercicio C7. Una inversa continua sin fórmula cerrada. Sea $F(x) = x^3 + x$.

1. Demuestra que F es continua y estrictamente creciente en $\mathbb{R}$.
2. Demuestra que $F(\mathbb{R}) = \mathbb{R}$ usando IVT.
3. Concluye que existe una inversa $F^{-1} : \mathbb{R} \to \mathbb{R}$ y que es continua, sin intentar obtener una fórmula explícita para ella.

Nivel D — Inversión y diagnóstico. Ejercicio D1. ¿Límite implica continuidad? Diagnostica la afirmación:

> Si $\lim_{x \to a} f(x)$ existe como número real, entonces f es continua en a.

Si es falsa, da un contraejemplo y formula una versión correcta.

Ejercicio D2. ¿Intervalo acotado basta? Diagnostica:

> Toda función continua en un intervalo acotado es acotada.

Si es falsa, identifica qué condición geométrica falta y da un contraejemplo.

Ejercicio D3. ¿Continua implica uniformemente continua? Diagnostica:

> Toda función continua es uniformemente continua en su dominio.

Si es falsa, refútala directamente con una función continua sobre $\mathbb{R}$.

Ejercicio D4. ¿Cambio de signo basta? Diagnostica:

> Si $f(a) < 0 < f(b)$, entonces existe $c \in (a, b)$ tal que $f(c) = 0$.

Determina la hipótesis ausente y construye un ejemplo que muestre que no puede omitirse.

Ejercicio D5. ¿Acotación más continuidad asegura máximo? Diagnostica:

> Toda función continua y acotada alcanza su supremo.

Si es falsa, da un contraejemplo e identifica la hipótesis que EVT utiliza para asegurar alcanzamiento.

Ejercicio D6. ¿Monotonía estricta implica continuidad? Diagnostica:

> Toda función estrictamente creciente en un intervalo es continua.

Si es falsa, construye una función estrictamente creciente con un salto y verifica ambas propiedades.

Nivel E — Contraejemplos. Ejercicio E1. Continua en un intervalo abierto y no acotada. Construye una función continua en $(0,1)$ que no esté acotada. Verifica continuidad y no acotación.

Ejercicio E2. Acotada, continua y sin extremos alcanzados. Construye una función continua y acotada sobre un intervalo no compacto que no alcance ni su supremo ni su ínfimo. Determina ambos valores extremos no alcanzados.

Ejercicio E3. Continua en $\mathbb{R}$ pero no uniforme. Construye una función polinómica distinta de x^2 que sea continua en $\mathbb{R}$ pero no uniformemente continua. La refutación de uniformidad debe ser cuantitativa.

Ejercicio E4. Bolzano falla sin continuidad. Construye una función $f : [-1,1] \to \mathbb{R}$ tal que $f(-1) < 0 < f(1)$, pero $f(x) \neq 0$ para todo $x \in [-1,1]$. Explica exactamente qué hipótesis de Corolario 5.2 falla.

Ejercicio E5. Monotonía con salto. Construye una función estrictamente creciente $f : \mathbb{R} \to \mathbb{R}$ que sea discontinua en 0. Verifica rigurosamente la monotonía en los tres casos $x < y < 0$, $0 \leq x < y$ y $x < 0 \leq y$.

Nivel F — Descubrimiento guiado. Ejercicio F1. Reconstruir el criterio secuencial de continuidad. Sea $f : A \to \mathbb{R}$ y $a \in A$.

1. Supón que f es continua en a y que $x_n \in A$, $x_n \to a$. Partiendo de un $\varepsilon > 0$, encadena la elección de δ con la elección de N y demuestra $f(x_n) \to f(a)$.
2. Supón ahora que f no es continua en a. Escribe la negación cuantificada de continuidad.
3. Usando $\delta_n = 1/(n+1)$, construye una sucesión $x_n \to a$ para la cual $f(x_n) \not\to f(a)$.
4. Concluye la equivalencia de Teorema 5.1 y explica por qué aquí no es necesario imponer $x_n \neq a$.

Ejercicio F2. Reconstruir IVT desde un supremo. Sea $f : [a,b] \to \mathbb{R}$ continua y supón $f(a) < y < f(b)$.

Define $S = \{x \in [a,b] : f(x) \leq y\}$.

1. Demuestra que S es no vacío y está acotado superiormente.
2. Sea $c = \sup S$. Explica por qué $a < c < b$.
3. Supón $f(c) < y$ y usa continuidad para obtener puntos de S mayores que c.
4. Supón $f(c) > y$ y combina continuidad con la caracterización aproximativa del supremo para obtener una contradicción.
5. Concluye que $f(c) = y$.

Ejercicio F3. Reconstruir la compactitud de $[a,b]$. Sea $\mathcal{U}$ un recubrimiento abierto de $[a,b]$ y define

$$S = \{x \in [a,b] : [a,x] \text{ admite un subrecubrimiento finito de } \mathcal{U}\}.$$

1. Demuestra que S es no vacío y está acotado superiormente.
2. Sea $c = \sup S$. Toma $U \in \mathcal{U}$ con $c \in U$ y usa apertura para obtener un radio $r > 0$ alrededor de c contenido en U.

3. Demuestra que si $c < b$, entonces puede cubrirse finitamente un intervalo $[a, c+\eta]$ con $\eta > 0$, contradiciendo que c sea cota superior de S.
4. Concluye que $c = b$ y termina cuidadosamente la prueba de que $[a, b]$ admite un subrecubrimiento finito.

Ejercicio F4. Reconstruir Heine–Cantor. Sea $f : [a, b] \to \mathbb{R}$ continua y fija $\varepsilon > 0$.

1. Para cada $x \in [a, b]$, usa continuidad con tolerancia $\varepsilon/2$ para obtener un radio $r_x > 0$.
2. Comprueba que los intervalos $V_{r_x/2}(x)$ forman un recubrimiento abierto de $[a, b]$.
3. Extrae una subcubierta finita con centros $x_1, \ldots, x_m$ y define
$$\delta = \frac{1}{2} \min_{1 \le i \le m} r_{x_i}.$$
4. Si $u, v \in [a, b]$ y $|u - v| < \delta$, demuestra que ambos quedan dentro del entorno completo de algún mismo centro x_i.
5. Concluye $|f(u) - f(v)| < \varepsilon$.

Ejercicio F5. Bisección hasta una precisión prescrita. Aplica bisección a $f(x) = x^2 - 2$ sobre $[1, 2]$.

1. Construye los intervalos $I_1, \ldots, I_6$ que conservan el cambio de signo.
2. Demuestra que cada I_n contiene una raíz y tiene longitud 2^{-n}.
3. Si m_6 es el punto medio de I_6, demuestra que $|m_6 - \sqrt{2}| < 10^{-2}$.
4. Explica qué parte del argumento demuestra existencia de la raíz y qué parte proporciona localización numérica.

Nivel G — Desafíos. Ejercicio G1. Separación uniforme de cero. Sea $f : [a, b] \to \mathbb{R}$ continua y supón que

$$f(x) \neq 0 \qquad \text{para todo } x \in [a, b].$$

Demuestra que existe $m > 0$ tal que

$$|f(x)| \geq m \qquad \text{para todo } x \in [a, b].$$

La prueba debe usar EVT y no derivadas.

Ejercicio G2. La imagen es un intervalo cerrado concreto. Sea $f : [a, b] \to \mathbb{R}$ continua. Demuestra que existen $m, M \in \mathbb{R}$ tales que $f([a, b]) = [m, M]$.

La prueba debe combinar EVT e IVT y justificar las dos inclusiones.

Ejercicio G3. Punto fijo en un intervalo. Sea $f : [a, b] \to [a, b]$ continua.

1. Demuestra que existe $c \in [a, b]$ tal que $f(c) = c$.
2. Si además f es estrictamente decreciente, demuestra que ese punto fijo es único.

Usa únicamente continuidad, orden e IVT.

Soluciones. Las soluciones siguen exactamente el orden de los ejercicios. En A–C se hace explícita la herramienta elegida; en D–E se comprueba que cada contraejemplo conserva las hipótesis que pretende conservar; en F se reconstruyen las pruebas centrales sin importar resultados posteriores; y en G se combinan los teoremas globales del capítulo.

Soluciones del nivel A. **Solución A1.**

En -1, el punto es aislado. Por ejemplo, $V_{1/2}(-1) \cap A = \{-1\}$. Por Proposición 5.1, toda función definida sobre A es continua en ese punto, independientemente del valor 50. En 0, los puntos del dominio que pueden aproximarse suficientemente son los de $[0, 2]$ por la derecha. El punto -1 queda excluido tomando, por ejemplo, un radio menor que $1/2$. Como $f(x) = x^2$ sobre esa componente y x^2 es continua, f es continua en 0 relativamente a A. En 2, solo pueden aproximarse puntos de $[0, 2]$ desde la izquierda. De nuevo, la restricción de x^2 es continua. Por tanto, f es continua en los tres puntos.

La diferencia está en la geometría local: en -1 no hay vecinos del dominio arbitrariamente próximos; en 0 y 2 sí los hay, pero solo desde el lado permitido por A.

Solución A2.

- 0 pertenece a A y es punto de acumulación porque $1/n \to 0$.
- $1 = 1/1$ pertenece a A y es aislado: existe un radio pequeño que no contiene otro punto de la sucesión ni puntos de $[2, 3]$.
- $1/2 = 1/2$ también es aislado por la misma razón.
- 2 pertenece a $[2, 3]$ y es punto de acumulación, con aproximaciones desde la derecha.
- 3 pertenece a $[2, 3]$ y es punto de acumulación, con aproximaciones desde la izquierda.
- $4 \notin A$, así que no es punto aislado de A; tampoco es punto de acumulación porque, por ejemplo, un entorno de radio $1/2$ alrededor de 4 no corta a A.

Entre los puntos listados que pertenecen a A, la continuidad automática para **toda** función se obtiene en 1 y $1/2$, porque son aislados. En $0, 2, 3$ una función arbitraria puede ser discontinua.

Solución A3.

En un entorno perforado de 1 la función coincide con x^2. Por tanto,

$$\lim_{x \to 1} f(x) = 1.$$

Pero $f(1) = 5$.

Como 1 es punto de acumulación de $\mathbb{R}$, Proposición 5.1 da

$$f \text{ continua en } 1 \iff \lim_{x \to 1} f(x) = f(1).$$

La igualdad falla, así que f no es continua en 1.

Para repararla basta redefinir $f(1) = 1$.

El límite perforado no cambia porque no depende del valor puntual.

Solución A4.

1. Para $x \neq 1$, $r(x) = \frac{(x-1)(x+1)}{x-1} = x + 1$, de modo que
$$\lim_{x \to 1} r(x) = 2 \neq 100 = r(1).$$
El defecto es **removible**: redefinir $r(1) = 2$ produce continuidad.

2. Para s,
$$\lim_{x\to 0^-} s(x) = -1, \qquad \lim_{x\to 0^+} s(x) = 1.$$
Los laterales finitos existen pero son distintos. Hay un **salto**.

3. Para q, por densidad de $\mathbb{Q}$ y de $\mathbb{R}\setminus\mathbb{Q}$ podemos escoger una sucesión racional $r_n \to 0$ y una irracional $u_n \to 0$. Entonces
$$q(r_n) = 1, \qquad q(u_n) = -1.$$
Las imágenes tienen comportamientos incompatibles, así que no existe un único límite en 0. El fallo es una **oscilación persistente** entre dos valores en toda escala.

Solución A5.

1. Se aplica **IVT**, y más precisamente Bolzano Corolario 5.2: dominio intervalar cerrado, continuidad y cambio de signo. Se concluye existencia de una raíz.
2. **Ninguno** de los dos teoremas globales en la forma requerida. El dominio $(0, 1)$ no es compacto; EVT no puede invocarse para extremos alcanzados. IVT no responde a esa pregunta.
3. Se aplica **EVT**: continuidad sobre el intervalo cerrado $[-2, 2]$ garantiza acotación y alcanzamiento de máximo y mínimo. IVT también está disponible para preguntas de valores intermedios, pero no es el teorema que produce los extremos.
4. Se aplica **IVT** directamente: y está entre los valores extremos de los puntos a, b y la función es continua en todo el intervalo.
5. **Ninguno** permite concluir la raíz: falta continuidad en $[0, 1]$. El cambio de signo aislado no basta.

Solución A6.

La primera estructura cuantificada corresponde a continuidad punto a punto. Más precisamente, para cada centro a y tolerancia ε puede elegirse un radio $\delta = \delta(a, \varepsilon)$. La segunda corresponde a continuidad uniforme. Primero se fija ε y después se exige un único $\delta(\varepsilon)$ que funcione simultáneamente para **todos** los pares $x, y \in A$ suficientemente próximos. Por tanto, solo en la continuidad ordinaria el radio puede depender del centro. La continuidad uniforme es más exigente porque prohíbe que la escala necesaria se deteriore al movernos por el dominio.

Solución A7.

1. Ambos conjuntos son abiertos. Si $x \in [0, 1]$ y $x < 3/5$, entonces $x \in U_1$; si $x \geq 3/5$, entonces $x > 2/5$ y $x \in U_2$. Así, $[0, 1] \subseteq U_1 \cup U_2$. El propio recubrimiento ya es finito, de modo que constituye un subrecubrimiento finito.

2. Sea $x \in (0, 1)$. Por arquimedianidad existe n suficientemente grande para que $1/n < x$, y entonces $x \in V_n$. Por tanto, los V_n recubren $(0, 1)$.

   Consideremos ahora una subfamilia finita $V_{n_1}, \ldots, V_{n_k}$ y sea $N = \max\{n_1, \ldots, n_k\}$. Como los intervalos crecen con n, su unión es
$$\left(\frac{1}{N}, 1\right).$$
El punto $1/(N + 1)$ pertenece a $(0, 1)$ pero no a esa unión. No hay subrecubrimiento finito.

3. No hay contradicción: Teorema 5.4 afirma la compactitud de **intervalos cerrados** $[a, b]$. El dominio $(0, 1)$ no es cerrado y este recubrimiento exhibe precisamente su falta de compactitud.

Soluciones del nivel B. **Solución B1.**

El numerador $x^3 - 2x + 1$ y el denominador $x^2 + 1$ son polinomios, luego son continuos en $\mathbb{R}$ por Corolario 5.1. Además, $x^2 + 1 \geq 1 > 0$ para todo $x \in \mathbb{R}$. El denominador nunca se anula, así que la ley del cociente de Proposición 5.2 es aplicable en todo punto. Por tanto, f es continua en $\mathbb{R}$.

Finalmente, $f(1) = \frac{1-2+1}{2} = 0$.

Solución B2.

El polinomio $h(x) = x^2 - 4x + 3$ es continuo en $\mathbb{R}$. La función valor absoluto es continua, y Proposición 5.2 garantiza que la composición $g = |\cdot| \circ h$ es continua en $\mathbb{R}$.

Además,

$$g(1) = |1 - 4 + 3| = 0, \qquad g(3) = |9 - 12 + 3| = 0.$$

Solución B3.

Cada tramo es polinómico y, por tanto, continuo en el interior de su región. Solo debemos auditar la unión $x = 1$. Como el primer tramo incluye 1, $f_a(1) = 2 + a$. El límite por la izquierda también vale $2 + a$. Por la derecha,

$$\lim_{x \to 1^+} (x^2 + 1) = 2.$$

La continuidad exige $2 + a = 2$, de donde

$$\boxed{a = 0}.$$

Con ese valor, ambos laterales y el valor puntual coinciden. No quedan otros puntos de unión, así que f_0 es continua en todo $\mathbb{R}$.

Solución B4.

Sea (x_n) una sucesión arbitraria de números reales tal que $x_n \to 0$. Por el álgebra de límites secuenciales,

$$x_n^2 \to 0, \qquad 1 + x_n^2 \to 1.$$

Como el límite del denominador es $1 \neq 0$, la ley del cociente da $\frac{1}{1+x_n^2} \to 1$.

Pero $f(0) = 1$.

Por tanto, toda sucesión $x_n \to 0$ satisface $f(x_n) \to f(0)$. El criterio secuencial Teorema 5.1 implica que f es continua en 0.

Solución B5.

Sea $p(x) = x^3 + x - 1$.

Es un polinomio, luego es continuo en $[0, 1]$. Además,

$$p(0) = -1 < 0, \qquad p(1) = 1 > 0.$$

Por Bolzano Corolario 5.2 existe $c \in (0, 1)$ tal que $p(c) = 0$.

Es decir, $c^3 + c - 1 = 0$.

El teorema garantiza **existencia**; este argumento todavía no afirma unicidad ni proporciona una fórmula para c.

Solución B6.

La función es polinómica, por tanto continua en el intervalo cerrado $[-2, 2]$. Por EVT Teorema 5.5 alcanza máximo y mínimo. Para localizarlos sin derivadas, ponemos $t = x^2$.

Como $x \in [-2, 2]$, tenemos $t \in [0, 4]$, y $f(x) = t^2 - 2t = (t - 1)^2 - 1$.

El mínimo es -1, alcanzado cuando $t = 1$, es decir, en $x = \pm 1$. Para el máximo, en $t \in [0, 4]$ la cantidad $(t - 1)^2$ no supera 9: la mayor distancia a 1 se obtiene en $t = 4$. Así, máx $f = 9 - 1 = 8$, alcanzado cuando $t = 4$, es decir, en $x = \pm 2$.

Solución B7.

Sea $\varepsilon > 0$. Elegimos $\delta = \frac{\varepsilon}{3}$.

Si $x, y \in \mathbb{R}$ y $|x - y| < \delta$, entonces

$$|f(x) - f(y)| = |3x - 2 - (3y - 2)| = 3|x - y| < 3\delta = \varepsilon.$$

El radio depende solo de ε y no de un centro particular. Por Definición 5.4, f es uniformemente continua en $\mathbb{R}$.

Soluciones del nivel C. **Solución C1.**

Partimos de la identidad $x \mapsto x$, que es continua. Por el álgebra de continuidad, x^2 y luego

$$x^2 - 1, \qquad x^2 + 1$$

son continuas en $\mathbb{R}$. La composición con valor absoluto produce $|x^2 - 1|$, también continua. Además, $x^2 + 1 \geq 1 > 0$ para todo x. Por tanto el cociente

$$\frac{|x^2 - 1|}{x^2 + 1}$$

está definido y es continuo en todo $\mathbb{R}$ por Proposición 5.2.

Solución C2.

La condición $g(0) = 0$ actúa sobre el primer tramo: $g(0) = b = 0$. En $x = 1$, el valor del tramo central es $g(1) = 1$. El límite izquierdo del primer tramo debe coincidir con él: $a + b = 1$.

Como $b = 0$, obtenemos $a = 1$.

En $x = 2$, el valor central es $g(2) = 4$. El límite por la derecha vale $6 + c$, así que continuidad exige $6 + c = 4$, de donde $c = -2$.

Por tanto,

$$\boxed{a = 1, \qquad b = 0, \qquad c = -2}.$$

Cada tramo es polinómico y las dos uniones han sido auditadas, de modo que la función resultante es continua en $\mathbb{R}$.

Solución C3.

Primero,

$$p\left(\frac{1}{2}\right) = \frac{1}{8} + \frac{1}{2} - 1 = -\frac{3}{8} < 0.$$

Y

$$p\left(\frac{3}{4}\right) = \frac{27}{64} + \frac{3}{4} - 1 = \frac{27 + 48 - 64}{64} = \frac{11}{64} > 0.$$

Como p es polinómica y, por tanto, continua, Bolzano produce

$$c \in \left(\frac{1}{2}, \frac{3}{4}\right)$$

tal que $p(c) = 0$.

Para la monotonía, sean $x < y$. Entonces

$$\begin{aligned} p(y) - p(x) &= (y^3 - x^3) + (y - x) \\ &= (y - x)(y^2 + xy + x^2 + 1). \end{aligned}$$

El primer factor es positivo. Además,

$$y^2 + xy + x^2 = \left(y + \frac{x}{2}\right)^2 + \frac{3x^2}{4} \geq 0,$$

de modo que $y^2 + xy + x^2 + 1 > 0$.

Por tanto $p(y) - p(x) > 0$: p es estrictamente creciente. Una función estrictamente creciente no puede tomar el mismo valor 0 en dos puntos distintos. La raíz encontrada es, por consiguiente, la única raíz real.

Solución C4.

Partimos de

$$p(0) = -1 < 0, \qquad p(1) = 1 > 0.$$

- Primer punto medio: $1/2$. Como $p(1/2) = -3/8 < 0$, conservamos $I_1 = [1/2, 1]$.
- Segundo punto medio: $3/4$. Como $p(3/4) = 11/64 > 0$, conservamos $I_2 = [1/2, 3/4]$.
- Tercer punto medio: $5/8$. Tenemos

  $$p(5/8) = \frac{125}{512} + \frac{5}{8} - 1 = -\frac{67}{512} < 0,$$

  así que $I_3 = [5/8, 3/4]$.

- Cuarto punto medio: 11/16. Como

$$p(11/16) = \frac{1331}{4096} + \frac{11}{16} - 1 = \frac{51}{4096} > 0,$$

obtenemos $I_4 = [5/8, 11/16]$.

Después de n bisecciones, la longitud es $|I_n| = 2^{-n}$. La raíz c permanece dentro de I_n, y el punto medio m_n está a distancia como máximo la mitad de esa longitud: $|m_n - c| \leq \frac{|I_n|}{2} = \frac{1}{2^{n+1}}$.

Queremos $\frac{1}{2^{n+1}} < 10^{-3}$.

Como $2^{10} = 1024 > 1000$, basta tomar

$$\boxed{n = 9}.$$

En efecto, $2^{-10} < 10^{-3}$.

Solución C5.

Fijemos $a \in [0, 1]$. Por continuidad en a, tomando $\varepsilon = 1$ existe $r_a > 0$ tal que

$$x \in [0, 1], \qquad |x - a| < r_a \Longrightarrow |f(x) - f(a)| < 1.$$

Entonces

$$|f(x)| \leq |f(x) - f(a)| + |f(a)| < |f(a)| + 1.$$

Los intervalos abiertos $V_{r_a/2}(a)$ con $a \in [0, 1]$ cubren $[0, 1]$. Por compactitud existe una subcubierta finita correspondiente a puntos $a_1, \ldots, a_m$.

Definimos

$$M = 1 + \max_{1 \leq i \leq m} |f(a_i)|.$$

Sea $x \in [0, 1]$. Como la familia finita cubre el intervalo, existe i tal que $|x - a_i| < \frac{r_{a_i}}{2} < r_{a_i}$. Por el control local asociado a a_i, $|f(x)| < |f(a_i)| + 1 \leq M$. Así, un conjunto infinito de cotas locales se reduce, mediante compactitud, a un máximo sobre una familia finita:

$$|f(x)| \leq M \qquad (0 \leq x \leq 1).$$

Solución C6.

El polinomio $p(x) = x^4 - 3x$ es continuo en $\mathbb{R}$, luego también en $[0, 1]$. Como $[0, 1]$ es compacto, Heine–Cantor Teorema 5.6 implica que p es uniformemente continua en ese intervalo. Ahora sea $\varepsilon > 0$. Por uniformidad existe $\delta > 0$ tal que para cualesquiera $u, v \in [0, 1]$, $|u - v| < \delta \Longrightarrow |p(u) - p(v)| < \varepsilon$. Como $|x_n - y_n| \to 0$, existe N tal que para $n \geq N$, $|x_n - y_n| < \delta$.

Por tanto, para todo $n \geq N$, $|p(x_n) - p(y_n)| < \varepsilon$.

Esto demuestra $|p(x_n) - p(y_n)| \to 0$.

El punto clave es que el mismo δ funciona sin saber alrededor de qué centro se encuentran x_n e y_n.

Solución C7.

F es un polinomio, por tanto continua en $\mathbb{R}$.

Si $x < y$, entonces

$$F(y) - F(x) = (y - x)(y^2 + xy + x^2 + 1) > 0$$

por la misma estimación del ejercicio C3. Así, F es estrictamente creciente y, en particular, inyectiva. Para demostrar sobreyectividad, fijemos $y \in \mathbb{R}$ y tomemos $R = |y| + 1 > 0$.

Entonces

$$F(R) = R^3 + R > R > |y| \geq y$$

y

$$F(-R) = -R^3 - R < -R < -|y| \leq y.$$

Como F es continua en $[-R, R]$, IVT garantiza un $c \in [-R, R]$ tal que $F(c) = y$.

Por tanto, $F(\mathbb{R}) = \mathbb{R}$.

La función es biyectiva $\mathbb{R} \to \mathbb{R}$. Como además es continua e inyectiva sobre el intervalo $\mathbb{R}$, Teorema 5.7 asegura que $F^{-1} : \mathbb{R} \to \mathbb{R}$ es continua. No necesitamos una fórmula cerrada para conocer esta propiedad.

Soluciones del nivel D. **Solución D1.**

La afirmación es falsa. Consideremos

$$f(x) = \begin{cases} x, & x \neq 0, \\ 1, & x = 0. \end{cases}$$

Entonces

$$\lim_{x \to 0} f(x) = 0,$$

pero $f(0) = 1$.

Por Proposición 5.1, f no es continua en 0. Una versión correcta es: si a es punto de acumulación del dominio y a pertenece al dominio, entonces

$$f \text{ continua en } a \iff \lim_{x \to a} f(x) = f(a).$$

La mera existencia del límite no fuerza su coincidencia con el valor puntual.

Solución D2.

La afirmación es falsa. El intervalo $(0, 1)$ es acotado, pero no compacto. La función $f(x) = \frac{1}{x}$ es continua en $(0, 1)$ y no está acotada superiormente: dado $M > 0$, escogemos

$$0 < x < \text{mín}\left\{1, \frac{1}{M}\right\},$$

y entonces $f(x) > M$.

La condición geométrica que permite pasar de continuidad a acotación global en este capítulo es la compactitud; en particular, continuidad sobre un intervalo cerrado y acotado $[a, b]$.

Solución D3.

Es falsa. Tomemos $f(x) = x^2$ en $\mathbb{R}$. Es continua, pero no uniformemente continua. Fijemos $\varepsilon = 1$. Sea $\delta > 0$ arbitrario. Elegimos n natural con $\frac{1}{n} < \delta$ y tomamos

$$x = n, \qquad y = n + \frac{1}{n}.$$

Entonces $|x - y| = \frac{1}{n} < \delta$, pero

$$|f(y) - f(x)| = \left(n + \frac{1}{n}\right)^2 - n^2 = 2 + \frac{1}{n^2} > 1.$$

Por tanto ningún δ funciona globalmente para $\varepsilon = 1$.

Solución D4.

Falta continuidad en todo el intervalo entre a y b.

Por ejemplo, en $[-1, 1]$ definimos

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

Entonces

$$f(-1) = -1 < 0 < 1 = f(1),$$

pero f nunca toma el valor 0. La función tiene un salto en 0. Precisamente esa discontinuidad impide aplicar Bolzano.

Solución D5.

La afirmación es falsa. Sea

$$f : (0, 1) \to \mathbb{R}, \qquad f(x) = x.$$

Es continua y acotada: $0 < f(x) < 1$.

Además, $\sup f((0, 1)) = 1$, pero no existe $x \in (0, 1)$ con $f(x) = 1$. EVT no usa solo continuidad y acotación de los valores; usa continuidad sobre un **dominio compacto**, aquí un intervalo cerrado $[a, b]$. La compactitud es la estructura que fuerza el alcanzamiento.

Solución D6.

Es falsa. Definimos

$$f(x) = \begin{cases} x, & x < 0, \\ x + 1, & x \geq 0. \end{cases}$$

Verifiquemos crecimiento estricto. Si $x < y < 0$, entonces $f(x) = x < y = f(y)$. Si $0 \leq x < y$, entonces

$$f(x) = x + 1 < y + 1 = f(y).$$

Si $x < 0 \leq y$, entonces

$$f(x) = x < 0 < 1 \leq y + 1 = f(y).$$

Por tanto f es estrictamente creciente en $\mathbb{R}$.

Sin embargo,

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

por lo que es discontinua en 0. La monotonía estricta garantiza inyectividad, no continuidad.

Soluciones del nivel E. **Solución E1.**

Un ejemplo es

$$f(x) = \frac{1}{x(1-x)}, \qquad 0 < x < 1.$$

El denominador es positivo y no nulo en $(0,1)$, de modo que f es racional y continua en todo su dominio. No está acotada. Por ejemplo, tomando $x_n = 1/(n+2)$ obtenemos $x_n \in (0,1)$ y

$$f(x_n) = \frac{1}{x_n(1-x_n)} > \frac{1}{x_n} = n + 2 \to +\infty.$$

La continuidad sobre un intervalo abierto no obliga a acotación global.

Solución E2.

Tomemos

$$f : (-2, 3) \to \mathbb{R}, \qquad f(x) = x.$$

Es continua y acotada: $-2 < f(x) < 3$.

Por densidad de los reales dentro del intervalo, los valores pueden acercarse arbitrariamente a ambos extremos, así que

$$\text{ínf}\, f((-2,3)) = -2, \qquad \sup f((-2,3)) = 3.$$

Ninguno se alcanza porque -2 y 3 no pertenecen al dominio. Por tanto, la función no tiene mínimo ni máximo.

Solución E3.

Podemos tomar $f(x) = x^3$.

Es polinómica y, por tanto, continua en $\mathbb{R}$. Fijemos $\varepsilon = 1$. Dado $\delta > 0$, escogemos n con $\frac{1}{n^2} < \delta$ y ponemos

$$x = n, \qquad y = n + \frac{1}{n^2}.$$

Entonces $|x - y| = \frac{1}{n^2} < \delta$, pero

$$\begin{aligned} y^3 - x^3 &= 3n^2\frac{1}{n^2} + 3n\frac{1}{n^4} + \frac{1}{n^6} \\ &= 3 + \frac{3}{n^3} + \frac{1}{n^6} > 1. \end{aligned}$$

Así, para la tolerancia fija $\varepsilon = 1$ ningún radio global funciona. x^3 no es uniformemente continua en $\mathbb{R}$.

Solución E4.
Definimos

$$f(x) = \begin{cases} -2, & x < 0, \\ 3, & x \geq 0. \end{cases}$$

Entonces

$$f(-1) = -2 < 0 < 3 = f(1),$$

pero la imagen de f es solo $\{-2, 3\}$, de modo que 0 nunca se alcanza. La hipótesis que falla en Corolario 5.2 es la continuidad sobre $[-1, 1]$: la función tiene un salto en 0.

Solución E5.
Un ejemplo es

$$f(x) = \begin{cases} x, & x < 0, \\ x + 2, & x \geq 0. \end{cases}$$

Si $x < y < 0$, entonces $f(x) = x < y = f(y)$. Si $0 \leq x < y$, entonces

$$f(x) = x + 2 < y + 2 = f(y).$$

Si $x < 0 \leq y$, entonces

$$f(x) = x < 0 < 2 \leq y + 2 = f(y).$$

Así, f es estrictamente creciente.
Pero

$$\lim_{x \to 0^-} f(x) = 0$$

mientras $f(0) = 2$.
Por tanto, es discontinua en 0.

Soluciones del nivel F. **Solución F1.**
Supongamos primero que f es continua en a y sea (x_n) una sucesión de puntos de A con $x_n \to a$. Fijado $\varepsilon > 0$, continuidad proporciona $\delta > 0$ tal que

$$x \in A, \qquad |x - a| < \delta \Longrightarrow |f(x) - f(a)| < \varepsilon.$$

Como $x_n \to a$, existe N tal que $n \geq N \Longrightarrow |x_n - a| < \delta$.
Por tanto, $n \geq N \Longrightarrow |f(x_n) - f(a)| < \varepsilon$, y así $f(x_n) \to f(a)$.
Para la recíproca, neguemos continuidad. Existe $\varepsilon_0 > 0$ tal que

$$\forall \delta > 0 \; \exists x \in A : \qquad |x - a| < \delta \quad \text{y} \quad |f(x) - f(a)| \geq \varepsilon_0.$$

Aplicando esta afirmación con $\delta_n = \frac{1}{n+1}$, escogemos $x_n \in A$ con $|x_n - a| < \frac{1}{n+1}$ y $|f(x_n) - f(a)| \geq \varepsilon_0$. Por sándwich, $x_n \to a$, pero las imágenes no pueden converger a $f(a)$ porque permanecen a distancia al menos ε_0. Esto demuestra la contraposición de la recíproca y, por tanto, la equivalencia. No necesitamos imponer $x_n \neq a$: en la dirección directa el centro está

permitido por la definición de continuidad; en la construcción de la discontinuidad, la desigualdad $|f(x_n) - f(a)| \geq \varepsilon_0 > 0$ impide automáticamente que $x_n = a$.

Solución F2.

Como $f(a) < y$, tenemos $a \in S$, de modo que S es no vacío. Además $S \subseteq [a, b]$, así que está acotado superiormente por b. Por completitud existe $c = \sup S$. Como $f(a) < y$ y f es continua en a, existe un pequeño tramo a la derecha de a donde todavía $f(x) < y$; por tanto S contiene puntos mayores que a, así que $c > a$. Análogamente, como $f(b) > y$ y f es continua en b, existe un tramo a la izquierda de b donde $f(x) > y$, lo que obliga a $c < b$.

Supongamos primero $f(c) < y$.

Tomamos $\varepsilon = \frac{y - f(c)}{2} > 0$.

Por continuidad existe $\delta > 0$ tal que si $|x - c| < \delta$, entonces $|f(x) - f(c)| < \varepsilon$. En particular, para algún $x > c$ suficientemente próximo y todavía dentro de $[a, b]$,

$$f(x) < f(c) + \varepsilon = \frac{f(c) + y}{2} < y.$$

Entonces $x \in S$ y $x > c$, contradicción con que c sea cota superior.

Supongamos ahora $f(c) > y$.

Tomamos $\varepsilon = \frac{f(c) - y}{2} > 0$.

Por continuidad existe $\delta > 0$ tal que $|x - c| < \delta \implies f(x) > y$. Como $c = \sup S$, la caracterización aproximativa del supremo garantiza un $s \in S$ con $c - \delta < s \leq c$. No puede ser $s = c$, pues $f(c) > y$ y entonces $c \notin S$. Por tanto $|s - c| < \delta$, y la continuidad implica $f(s) > y$, contradiciendo $s \in S$. Ambas desigualdades estrictas son imposibles. Luego

$$\boxed{f(c) = y}.$$

Esta es la arquitectura del IVT: continuidad local + completitud mediante supremo $\Rightarrow$ valor intermedio global.

Solución F3.

Sea $\mathcal{U}$ un recubrimiento abierto de $[a, b]$. El punto a pertenece a algún $U_a \in \mathcal{U}$. El intervalo degenerado $[a, a] = \{a\}$ queda cubierto por ese único miembro de la cobertura. Por tanto, $a \in S$, y S es no vacío. Además $S \subseteq [a, b]$, luego está acotado superiormente por b.

Sea $c = \sup S$.

Escogemos $U \in \mathcal{U}$ con $c \in U$. Por apertura existe $r > 0$ tal que $V_r(c) \subseteq U$. Supongamos $c < b$. Elegimos $\eta > 0$ tan pequeña que $0 < \eta < \min\left\{\frac{r}{2}, b - c\right\}$. Por la propiedad aproximativa del supremo existe $s \in S$ con $c - \frac{r}{2} < s \leq c$. Como $s \in S$, existe una subfamilia finita de $\mathcal{U}$ que cubre $[a, s]$. Además, el conjunto U cubre todo el tramo desde s hasta $c + \eta$, porque ambos extremos quedan dentro de $V_r(c)$. Añadiendo U a la familia finita obtenemos un subrecubrimiento finito de $[a, c + \eta]$. Por tanto $c + \eta \in S$, contradicción con que c sea cota superior. Luego $c = b$. Falta todavía demostrar que $b \in S$, pues $c = \sup S$ por sí solo no lo garantiza. Escogemos $U_b \in \mathcal{U}$ con $b \in U_b$. Por apertura existe $r > 0$ tal que $V_r(b) \subseteq U_b$.

Como $b = \sup S$, existe $s \in S$ con $b - \frac{r}{2} < s \leq b$.

Una familia finita cubre $[a, s]$ y U_b cubre $[s, b]$. Su unión finita cubre todo $[a, b]$. Así, todo recubrimiento abierto de $[a, b]$ posee un subrecubrimiento finito: el intervalo es compacto.

Solución F4.

Fijemos $\varepsilon > 0$. Para cada $x \in [a,b]$, por continuidad existe $r_x > 0$ tal que

$$z \in [a,b], \qquad |z-x| < r_x \Longrightarrow |f(z)-f(x)| < \frac{\varepsilon}{2}.$$

Los abiertos $V_{r_x/2}(x)$ cubren $[a,b]$. Por compactitud extraemos una subcubierta finita con centros $x_1, \ldots, x_m$.

Definimos

$$\delta = \frac{1}{2} \min_{1 \le i \le m} r_{x_i} > 0.$$

Sean $u, v \in [a,b]$ con $|u-v| < \delta$.

Como los entornos reducidos cubren el intervalo, existe i tal que $|u-x_i| < \frac{r_{x_i}}{2}$.

Además,

$$|v-x_i| \le |v-u| + |u-x_i| < \delta + \frac{r_{x_i}}{2} \le r_{x_i}.$$

Así, tanto u como v quedan dentro del entorno completo donde la oscilación respecto de $f(x_i)$ es menor que $\varepsilon/2$. Por la desigualdad triangular,

$$\begin{aligned} |f(u)-f(v)| &\le |f(u)-f(x_i)| + |f(x_i)-f(v)| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

El δ obtenido depende solo de ε. Por tanto, f es uniformemente continua.

Solución F5.

Tenemos

$$f(1) = -1 < 0, \qquad f(2) = 2 > 0.$$

Por Bolzano existe una raíz en $[1,2]$, y la bisección mantiene un cambio de signo en cada intervalo.

Los pasos son:

1. $m_0 = 3/2$ y $f(3/2) = \frac{1}{4} > 0$, luego $I_1 = [1, 3/2]$.
2. $m_1 = 5/4$ y
$$f(5/4) = \frac{25}{16} - 2 = -\frac{7}{16} < 0,$$
luego $I_2 = [5/4, 3/2]$.
3. $m_2 = 11/8$ y
$$f(11/8) = \frac{121}{64} - 2 = -\frac{7}{64} < 0,$$
luego $I_3 = [11/8, 3/2]$.
4. $m_3 = 23/16$ y
$$f(23/16) = \frac{529}{256} - 2 = \frac{17}{256} > 0,$$
luego $I_4 = [11/8, 23/16]$.

5. $m_4 = 45/32$ y
$$f(45/32) = \frac{2025}{1024} - 2 = -\frac{23}{1024} < 0,$$
luego $I_5 = [45/32, 23/16]$.
6. $m_5 = 91/64$ y
$$f(91/64) = \frac{8281}{4096} - 2 = \frac{89}{4096} > 0,$$
luego $I_6 = [45/32, 91/64]$.

Cada paso conserva un cambio de signo, por lo que Bolzano asegura que I_n contiene una raíz. Además, cada bisección divide la longitud por 2: $|I_n| = 2^{-n}$. El punto medio de I_6 es

$$m_6 = \frac{1}{2}\left(\frac{45}{32} + \frac{91}{64}\right) = \frac{181}{128}.$$

Como $\sqrt{2} \in I_6$,

$$|m_6 - \sqrt{2}| \leq \frac{|I_6|}{2} = \frac{1}{128} < \frac{1}{100}.$$

La **existencia** proviene de continuidad + cambio de signo + Bolzano. La **localización numérica** proviene del procedimiento de bisección y de la cota geométrica para la longitud de los intervalos.

Soluciones del nivel G. **Solución G1.**

La función $g(x) = |f(x)|$ es continua en $[a, b]$. Por EVT existe $x_0 \in [a, b]$ tal que g alcanza allí su mínimo: $m = g(x_0) = \min_{x\in[a,b]} |f(x)|$. Como, por hipótesis, $f(x) \neq 0$ para todo x, en particular $f(x_0) \neq 0$.

Por tanto, $m = |f(x_0)| > 0$.

Por la definición de mínimo,

$$|f(x)| \geq m \qquad \text{para todo } x \in [a, b].$$

La conclusión es más fuerte que decir que cada valor individual es no nulo: la compactitud permite obtener una **separación positiva común** respecto de cero.

Solución G2.

Por EVT existen puntos $x_m, x_M \in [a, b]$ tales que

$$m = f(x_m) = \min f([a, b]), \qquad M = f(x_M) = \max f([a, b]).$$

De la definición de mínimo y máximo se sigue inmediatamente que $f([a, b]) \subseteq [m, M]$. Para la inclusión inversa, sea $y \in [m, M]$. Si $y = m$ o $y = M$, ya sabemos que se alcanza en x_m o x_M.

Supongamos $m < y < M$.

Los puntos x_m y x_M pueden aparecer en cualquier orden. Consideremos el intervalo cerrado cuyos extremos son esos dos puntos. La función es continua en él y toma en sus extremos los valores m y M. Por IVT existe un punto c entre x_m y x_M tal que $f(c) = y$.

Así, $[m, M] \subseteq f([a, b])$.

Concluimos

$$\boxed{f([a, b]) = [m, M]}.$$

EVT aporta los extremos **alcanzados**; IVT rellena todos los valores entre ellos.

Solución G3.

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

Como f y la identidad son continuas, g es continua en $[a, b]$. La condición $f([a, b]) \subseteq [a, b]$ implica, en particular, $f(a) \geq a$, y por tanto $g(a) = f(a) - a \geq 0$.

También $f(b) \leq b$, de modo que $g(b) = f(b) - b \leq 0$.

Si alguno de estos valores es 0, ya tenemos un punto fijo en un extremo. En caso contrario, $g(a) > 0 > g(b)$, y Bolzano produce $c \in (a, b)$ con $g(c) = 0$. En todos los casos existe $c \in [a, b]$ tal que $f(c) = c$. Supongamos ahora que f es estrictamente decreciente y que existen dos puntos fijos distintos $c < d$. Por decrecimiento estricto, $f(c) > f(d)$. Pero, por ser puntos fijos,

$$f(c) = c, \qquad f(d) = d.$$

Entonces $c > d$, contradicción. Por tanto, el punto fijo es único.

Capítulo 6

La derivada y la aproximación lineal local

En los ejemplos periódicos usaremos las funciones triangulares σ y κ definidas en C03, con $\omega = 2$. Sus cotas $|\sigma|, |\kappa| \leq 1$ y los valores $\sigma(\omega/2 + n\omega) = (-1)^n$ proceden directamente de aquella definición por tramos; no se presupone trigonometría.

Los límites describen cómo se aproximan los valores de una función a un número; la continuidad exige que esa aproximación concuerde con el valor tomado en el centro. Queda por comparar la escala de la variación de la función con la del incremento de su argumento. Si a es un punto de acumulación del dominio y f es continua en a, sabemos que

$$f(a+h) \longrightarrow f(a) \qquad \text{cuando} \qquad h \longrightarrow 0,$$

con h restringido a incrementos no nulos para los que $a + h$ pertenece al dominio. Pero esta información deja abierta una pregunta más fina. Dos funciones pueden ser continuas en el mismo punto y, sin embargo, cambiar allí de maneras muy diferentes. Cerca de 0, por ejemplo,

$$x, \qquad x^2, \qquad |x|$$

se aproximan todas a 0 cuando $x \to 0$, pero no lo hacen con la misma geometría ni con la misma escala de variación. La continuidad responde a una pregunta del tipo

¿se hace pequeña la variación de la función cuando se hace pequeño el incremento?

El cálculo diferencial comienza cuando refinamos la pregunta:

¿cuál es la parte principal de esa variación cuando observamos la función a escala muy pequeña?

Para hacer esa comparación necesitamos una cantidad que relacione el cambio de salida con el de entrada. El cociente incremental permitirá estudiar esa relación en puntos próximos, y su límite, cuando exista como número real, definirá la derivada. El recorrido de este capítulo será

secante

→ cociente incremental

→ límite

→ derivada

→ tangente

→ aproximación afín local

→ error pequeño frente al incremento.

Ese límite medirá una **tasa instantánea de cambio** y proporcionará el coeficiente de la **aproximación lineal de primer orden**. Las pendientes de las secantes y las velocidades medias permitirán construir el cociente; el control del residuo precisará después en qué sentido la aproximación es local. Comenzaremos por la pregunta más elemental: ¿cómo puede una razón de cambio calculada entre **dos** puntos contener información acerca del comportamiento de una función en **un solo** punto?

6.1. De la secante a la tasa instantánea

Medir un cambio exige comparar dos estados. Supongamos que una cantidad y depende de una variable x mediante $y = f(x)$. Si queremos describir cuánto ha cambiado y, necesitamos comparar dos valores de la variable. Tomemos un punto de referencia a y un incremento h tal que $a + h$ también pertenezca al dominio. Así obtenemos otro punto cercano $a + h$. El número $h = (a + h) - a$ mide el **incremento de la variable de entrada**. El correspondiente cambio en la salida es $f(a + h) - f(a)$. Conviene leer estas dos cantidades en paralelo:

$$\begin{aligned} \text{incremento de entrada :} &\qquad h, \\ \text{variación de salida :} &\qquad f(a+h) - f(a). \end{aligned}$$

Si $h > 0$, el segundo punto está a la derecha de a; si $h < 0$, está a la izquierda. En ambos casos, mientras $h \neq 0$, tenemos dos entradas distintas y podemos comparar los cambios. Una diferencia aislada como $f(a + h) - f(a)$ no dice todavía cuán rápido está cambiando la función. Una variación de 10 unidades puede ser enorme si ocurrió durante un incremento de entrada de 0.001, pero muy pequeña si ocurrió durante un incremento de 10^6. Para medir una **razón de cambio** debemos comparar la variación de salida con el tamaño orientado del incremento de entrada. Esto conduce al cociente

$$\frac{f(a+h) - f(a)}{h}, \qquad h \neq 0.$$

Este será el objeto central de la sección.

Lectura de la fórmula
En $\frac{f(a+h)-f(a)}{h}$, el numerador mide cuánto cambió la salida y el denominador mide cuánto cambió la entrada. El cociente pregunta:

> ¿cuántas unidades de salida cambian, en promedio, por cada unidad de entrada durante este desplazamiento?

La palabra **promedio** es importante. Todavía estamos comparando dos puntos distintos.

La misma razón de cambio tiene una lectura geométrica. Consideremos la gráfica de f y los puntos $P = (a, f(a))$ y

$$Q_h = (a + h, f(a + h)), \qquad h \neq 0.$$

Como P y Q_h son distintos, determinan una única recta. Esa recta es una **secante** a la gráfica en los dos puntos considerados.

Su pendiente es

$$\frac{f(a+h) - f(a)}{(a+h) - a} = \frac{f(a+h) - f(a)}{h}.$$

Así, el cociente anterior no es una construcción arbitraria. Es exactamente la pendiente de la recta que une los dos puntos de la gráfica.

Si escribimos

$$Q_a(h) := \frac{f(a+h) - f(a)}{h}, \qquad h \neq 0,$$

solo estamos dando un nombre temporal a esa pendiente secante. No estamos definiendo todavía la derivada. La secante que pasa por P y Q_h puede escribirse como $S_h(x) = f(a) + Q_a(h)(x-a)$. Todas estas rectas pasan por el mismo punto P. Lo único que cambia con h es su pendiente. Esto sugiere una estrategia geométrica. Si hacemos que Q_h se acerque a P, es decir, si hacemos $h \to 0$, podemos observar qué ocurre con las pendientes de las secantes. No necesitamos hablar de un misterioso «límite de rectas». Ya sabemos estudiar límites de números. Basta preguntar si los números $Q_a(h)$ se aproximan a algún valor cuando h se aproxima a 0.

Antes de seguir
Imagina que para incrementos cada vez menores las pendientes de las secantes son

$$1.9, \quad 1.99, \quad 1.999, \quad 2.001, \quad 2.01.$$

¿Qué número parece estar emergiendo? ¿Qué parte de nuestra teoría previa permite sustituir esta impresión visual por una afirmación matemática rigurosa?
La respuesta a la segunda pregunta es: **la teoría de límites**.

Un primer laboratorio: la parábola. Veamos el mecanismo en una función suficientemente sencilla para que toda la estructura quede visible.

Ejemplo 6.1. **Secantes de $f(x) = x^2$ y estabilización de sus pendientes.** Fijemos un punto $a \in \mathbb{R}$ y consideremos $f(x) = x^2$.

Para $h \neq 0$, la pendiente de la secante que une (a, a^2) con $(a+h, (a+h)^2)$ es $\frac{(a+h)^2 - a^2}{h}$.

Desarrollando el numerador,

$$(a+h)^2 - a^2 = a^2 + 2ah + h^2 - a^2 = 2ah + h^2.$$

Como estamos trabajando con $h \neq 0$, podemos factorizar y cancelar h: $\frac{2ah+h^2}{h} = \frac{h(2a+h)}{h} = 2a + h$. Por tanto, cada incremento no nulo produce una pendiente secante $Q_a(h) = 2a + h$. Ahora el comportamiento cuando h se hace pequeño es transparente:

$$2a + h \longrightarrow 2a \qquad (h \to 0).$$

Tomemos, por ejemplo, $a = 1$. Entonces $Q_1(h) = 2 + h$.

Algunos valores son:

h	$Q_1(h)$
1	3
0.1	2.1
0.01	2.01
0.001	2.001
−0.001	1.999
−0.01	1.99
−0.1	1.9

h	$Q_1(h)$

Las pendientes obtenidas desde la derecha y desde la izquierda se concentran alrededor de 2. La observación importante no es la tabla finita. La igualdad exacta $Q_1(h) = 2 + h$ permite usar la teoría de límites y concluir rigurosamente que

$$\lim_{h\to 0} Q_1(h) = 2.$$

Más generalmente,

$$\lim_{h\to 0} Q_a(h) = 2a.$$

Todavía no hemos dado un nombre formal a este límite. Por ahora solo hemos descubierto que, en la parábola, las pendientes de las secantes poseen una pendiente límite bien determinada alrededor de cada punto.

El ejemplo contiene varias ideas que conviene separar. Primero, el cociente $\frac{f(a+h)-f(a)}{h}$ solo está definido para $h \neq 0$. La cancelación anterior fue legítima precisamente porque trabajábamos en un entorno **perforado** de 0. Segundo, después de cancelar obtenemos una expresión que sí tiene un comportamiento sencillo cuando $h \to 0$. El límite no exige sustituir $h = 0$ en la expresión original. Tercero, no bastó observar que las secantes «parecen estabilizarse». La teoría de límites transformó esa intuición en una afirmación exacta.

Error frecuente: poner $h = 0$ demasiado pronto
El cociente incremental de la parábola en a es inicialmente $\frac{(a+h)^2-a^2}{h}$. Si sustituimos $h = 0$, obtenemos $\frac{0}{0}$, que no es un número real. Esto **no** significa que el proceso haya fracasado. Significa que estamos ante un problema de límite. Para $h \neq 0$ simplificamos el cociente a $2a + h$ y después estudiamos qué ocurre cuando h **se aproxima** a 0. El límite analiza valores arbitrariamente próximos al centro; no necesita evaluar la expresión en el centro.

Cambiar de escala sin perder la información principal. La variable h cumple aquí un papel semejante al radio δ de nuestros capítulos de límites y continuidad: controla la escala con la que observamos la función alrededor de a. Si $|h|$ es grande, el segundo punto puede estar lejos y la secante describe un comportamiento bastante global. Si reducimos $|h|$, los dos puntos están cada vez más próximos y la secante examina la gráfica a una escala cada vez más local. Podemos visualizar el proceso como

$$|h| \text{ grande} \quad \longrightarrow \quad |h| \text{ pequeño} \quad \longrightarrow \quad Q_h \text{ próximo a } P.$$

Pero hay una sutileza fundamental. Al acercar Q_h a P, la secante pierde en el límite sus **dos puntos distintos**: cuando $h = 0$, ambos puntos coincidirían. Por eso no podemos definir la pendiente buscada usando directamente dos puntos coincidentes. La información debe sobrevivir de otra manera. Lo que podemos conservar es el comportamiento límite de las pendientes $\frac{f(a+h)-f(a)}{h}$. Así aparece una característica esencial del cálculo diferencial:

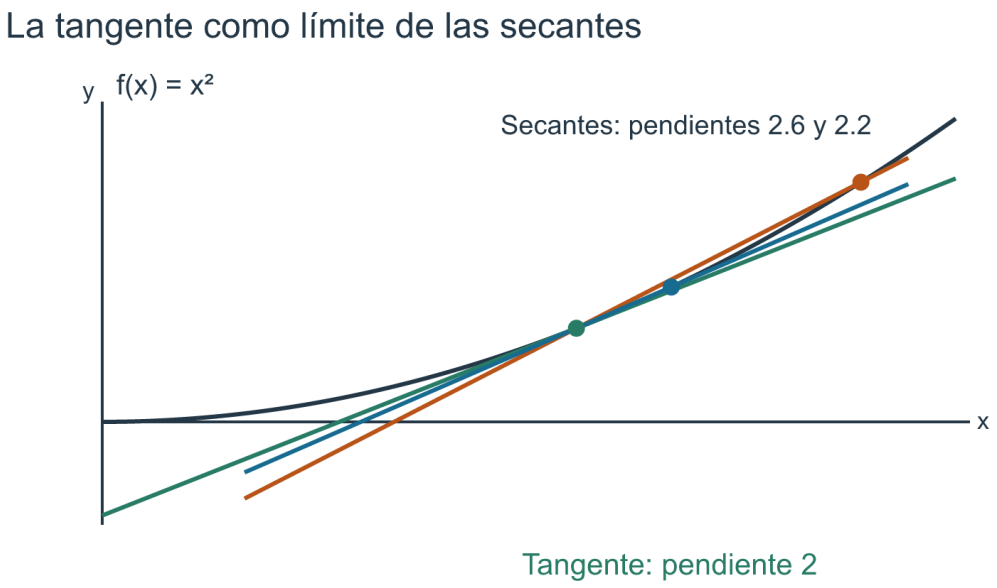


FIGURA 6.1. Las secantes de $f(x) = x^2$ desde $x = 1$ tienen pendiente $2 + h$. Al disminuir $h > 0$, sus pendientes se acercan a 2, pendiente de la tangente $y = 2x - 1$.

para obtener información en un punto, estudiamos primero información entre puntos cercanos y después tomamos un límite.

La lectura cinemática: de velocidad media a velocidad instantánea. La misma estructura aparece si f describe una posición que cambia con el tiempo. Sea $s(t)$ la posición de una partícula en el instante t. Durante el intervalo que va de t a $t + h$, el desplazamiento es $s(t + h) - s(t)$, y la duración orientada del intervalo es h. La **velocidad media** durante ese intervalo es entonces

$$\frac{s(t+h) - s(t)}{h}, \qquad h \neq 0.$$

La fórmula es exactamente el mismo cociente que apareció en la pendiente de la secante.

Geométricamente:

$$\frac{f(a+h) - f(a)}{h} = \text{pendiente secante.}$$

Cinemáticamente: $\frac{s(t+h)-s(t)}{h} = \text{velocidad media.}$

En ambos casos estamos midiendo una razón de cambio sobre un intervalo no degenerado. Pero un velocímetro pretende responder otra pregunta: ¿cuál es la velocidad **en el instante** t?

No podemos resolverla tomando literalmente un intervalo de duración cero, porque $\frac{s(t)-s(t)}{0}$ no está definido. La estrategia correcta vuelve a ser la misma: calcular velocidades medias en

intervalos cada vez más pequeños y preguntar si esas velocidades se aproximan a un único número. Por ejemplo, si $s(t) = t^2$, entonces, para $h \neq 0$, $\frac{s(t+h)-s(t)}{h} = \frac{(t+h)^2-t^2}{h} = 2t + h$. Al reducir $|h|$, la velocidad media se aproxima a $2t$. De nuevo, el valor asociado al instante no se obtiene evaluando el cociente en $h = 0$, sino estudiando su límite cuando $h \to 0$. Esta coincidencia entre geometría y movimiento no es accidental. Ambas situaciones están expresando una misma estructura matemática: **la tasa de cambio local de una función**.

Tasa media primero, tasa instantánea después. Para dos puntos distintos x_1 y x_2 del dominio de una función, la razón $\frac{f(x_2)-f(x_1)}{x_2-x_1}$ mide la tasa media de cambio de f entre ambos puntos. Si fijamos $x_1 = a$ y escribimos $x_2 = a + h$, obtenemos exactamente $\frac{f(a+h)-f(a)}{h}$. La palabra «media» no implica aquí que estemos calculando una media aritmética de muchos valores. Significa que estamos condensando todo el cambio ocurrido entre dos entradas en una única razón global para ese intervalo. Para asociar una tasa a un solo instante, consideramos tasas medias en intervalos cada vez más pequeños. El límite, cuando existe, permite obtener un número independiente del incremento concreto con el que se aproxima el segundo instante:

$$\text{tasa media en un intervalo} \quad \xrightarrow{h\to 0} \quad \text{candidato a tasa instantánea en el punto.}$$

La palabra **candidato** importa. No toda función producirá necesariamente un límite finito y único. Puede ocurrir que las pendientes provenientes de ambos lados se aproximen a números diferentes, que crezcan sin cota, que oscilen o que el punto ni siquiera admita aproximaciones no triviales dentro del dominio. Todas esas posibilidades serán estudiadas antes de que podamos hablar con seguridad de derivabilidad.

No estamos usando infinitesimales como números misteriosos. Es frecuente describir informalmente el cálculo diciendo que h se vuelve «infinitamente pequeño». Esa frase puede servir como intuición inicial, pero no será nuestro fundamento lógico. En la teoría que ya hemos construido, escribir $h \to 0$ no significa introducir un número real no nulo cuyo tamaño sea «menor que cualquier real positivo». Significa estudiar el comportamiento de una expresión para valores reales $h \neq 0$ que pueden elegirse tan próximos a 0 como sea necesario. La diferencia es importante. Durante todo el proceso $h \neq 0$, de modo que el cociente incremental está definido. Lo que hacemos es preguntar si existe un número L tal que los valores $\frac{f(a+h)-f(a)}{h}$ puedan forzarse a quedar arbitrariamente próximos a L haciendo $|h|$ suficientemente pequeño.

Ésta es exactamente la clase de pregunta que aprendimos a responder con la definición rigurosa de límite.

La transición conceptual
La derivada no aparecerá porque «dividimos por un infinitesimal». Aparecerá porque construimos una función del incremento,

$$h \longmapsto \frac{f(a+h)-f(a)}{h}, \qquad h \neq 0,$$

y estudiamos si esa función posee un límite finito cuando $h \to 0$.

Una misma estructura, tres lenguajes. Hemos llegado al mismo objeto desde tres perspectivas. En lenguaje algebraico, tenemos el cociente incremental $\frac{f(a+h)-f(a)}{h}$. En lenguaje geométrico, tenemos la pendiente de la secante que une

$$(a, f(a)) \quad \text{y} \quad (a+h, f(a+h)).$$

En lenguaje de tasas de cambio, tenemos la razón media entre variación de salida e incremento de entrada.

Podemos resumirlo así:

$$\boxed{\begin{array}{c}\text{cociente incremental}\\ \Updownarrow \\ \text{pendiente de una secante}\\ \Updownarrow \\ \text{tasa media de cambio}\end{array}}$$

La siguiente operación también es común a las tres lecturas: $h \to 0$.

- geométricamente, el segundo punto de la secante se aproxima al primero;
- cinemáticamente, el intervalo de tiempo se hace cada vez más corto;
- analíticamente, estudiamos el límite del cociente incremental.

Si ese proceso produce un único número real, habremos encontrado una cantidad local asociada al punto a.

Qué falta todavía. El ejemplo de x^2 mostró un caso favorable, pero una definición matemática no puede construirse a partir de un dibujo convincente ni de una función bien comportada. Antes de dar el siguiente paso debemos responder con precisión varias preguntas.

1. ¿Qué hipótesis sobre el dominio permiten que h se aproxime realmente a 0?
2. ¿Qué significa exactamente que el cociente incremental tenga un límite?
3. ¿Por qué, si ese límite existe, hay un único valor posible?
4. ¿Es equivalente acercarse mediante $h \to 0$ a escribir directamente $x \to a$?
5. ¿Qué ocurre en el extremo de un intervalo, donde quizá solo podemos acercarnos desde un lado?
6. ¿Debemos aceptar $+\infty$ o $-\infty$ como valores de la nueva cantidad?
7. ¿Cómo podemos demostrar rigurosamente que el límite no existe cuando las pendientes se comportan de manera incompatible?

Estas preguntas muestran que ya estamos preparados para la definición, pero también que la definición debe formularse con cuidado. La idea obtenida en esta sección puede condensarse en una sola cadena:

$$\boxed{\begin{array}{c}\text{comparar dos puntos}\\ \to \text{formar una tasa media}\\ \to \text{reducir la escala}\\ \to \text{buscar un límite en un punto.}\end{array}}$$

En la sección siguiente convertiremos esa cadena en una noción matemática exacta: la **derivada de una función en un punto**.

6.2. Derivada en un punto: definición, unicidad y criterio secuencial

La sección anterior terminó con una pregunta precisa. Fijados un punto a y una función f, podemos formar para puntos cercanos el cociente incremental

$$\frac{f(a+h)-f(a)}{h}, \qquad h \neq 0,$$

y preguntar si esos valores se aproximan a un único número cuando $h \to 0$. La definición debe conservar las condiciones que hacen posible formar el cociente y tomar su límite. Tendremos que atender a tres aspectos:

1. **el dominio desde el que nos acercamos a** a;
2. **la existencia de un límite real finito**;
3. **la unicidad del número obtenido**.

La derivada no será, por tanto, un símbolo nuevo añadido a una fórmula. Será una propiedad local de una función en un punto, construida enteramente a partir de la teoría de límites del capítulo 4.

Antes de definir: ¿desde dónde puede acercarse el segundo punto? Sea $f\colon A \to \mathbb{R}$ y sea $a \in A$.

El cociente compara $f(a)$ con valores tomados en argumentos distintos de a. Necesitamos, por tanto, que a pertenezca a A y sea un **punto de acumulación de** A, para disponer de incrementos admisibles arbitrariamente pequeños. Esta condición garantiza que existen puntos de $A \setminus \{a\}$ arbitrariamente próximos a a y, por tanto, que el cociente $\frac{f(x)-f(a)}{x-a}$ puede estudiarse mediante un límite relativo al dominio. Si a fuera un punto aislado de A, no habría información incremental arbitrariamente próxima al centro. En ese caso no asignaremos una derivada a f en a.

Por qué no definimos derivadas en puntos aislados
En un punto aislado no existen puntos del dominio, distintos de a, suficientemente próximos a a. La teoría de límites del capítulo 4 fue construida precisamente para evitar conclusiones vacías en esta situación. La derivada pretende medir **cómo cambia** la función al desplazarnos desde a. Si el dominio no permite desplazamientos arbitrariamente pequeños, esa información no existe.

Cuando a es un punto interior de A, podemos acercarnos desde ambos lados. Si el dominio permite acercarse solo desde uno, el límite relativo conserva esa geometría del dominio. En §6.4 separaremos explícitamente las nociones de derivada derecha e izquierda; por ahora mantendremos la formulación general relativa al dominio.

La definición.

Definición 6.1. **Diferenciabilidad y derivada en un punto.** Sea $f\colon A \to \mathbb{R}$ y sea $a \in A$ un punto de acumulación de A.

Decimos que f es **diferenciable en** a **respecto del dominio** A si existe un número real L tal que

$$\lim_{\substack{x \to a \\ x \in A}} \frac{f(x)-f(a)}{x-a} = L.$$

Ese número, necesariamente único, se llama **derivada de** f **en** a **respecto de** A y se escribe

$$f'_A(a) = L.$$

Cuando el dominio está claro por el contexto escribiremos simplemente

$$f'(a).$$

El conjunto de puntos del dominio en los que f es diferenciable será

$$\mathrm{Dom}(f') = \{a \in A : a \text{ es punto de acumulación de } A \text{ y } f'_A(a) \text{ existe como número real}\}.$$

Así, cuando este conjunto no es vacío, la **función derivada** es la aplicación

$$f' \colon \mathrm{Dom}(f') \to \mathbb{R}, \qquad a \longmapsto f'(a).$$

La definición contiene varias decisiones que conviene hacer visibles. Primero, el cociente se evalúa para $x \neq a$. El punto central no pertenece al dominio del cociente incremental porque allí aparecería una división por cero. Segundo, el límite es **relativo a** A. No pedimos información sobre puntos que no pertenecen al dominio de f. Tercero, exigimos que el límite sea un número de $\mathbb{R}$. Si el cociente crece sin cota hacia $+\infty$ o $-\infty$, podemos obtener una geometría interesante, pero bajo esta definición f no es diferenciable en a. Cuarto, la diferenciabilidad es una propiedad **puntual**. Afirmar que f es diferenciable en a no afirma todavía nada sobre la existencia de derivadas en puntos vecinos.

Leer la definición como una cadena de condiciones
Para preguntar si $f'(a)$ existe, audita en este orden:

$$\boxed{\begin{array}{c} a \in A \\ \longrightarrow a \text{ punto de acumulación de } A \\ \longrightarrow \text{cociente incremental definido para } x \neq a \\ \longrightarrow \text{límite real finito.} \end{array}}$$

Si falla una de estas etapas, la derivada no queda definida por `#def-t1-0042`.

La forma con incrementos h. La definición anterior utiliza un punto variable x. En §6.1 trabajamos, en cambio, con un incremento h y escribimos $x = a + h$. Ambas formas describen exactamente el mismo proceso, pero conviene demostrarlo en lugar de tratarlo como una sustitución simbólica automática. Para un punto fijo $a \in A$, definamos el conjunto de incrementos admisibles $H_a := \{h \in \mathbb{R} : a + h \in A\}$. Entonces $0 \in H_a$ porque $a \in A$, y si a es punto de acumulación de A, 0 es punto de acumulación de H_a. Además, la correspondencia $x \longleftrightarrow h = x - a$ transforma los puntos $x \in A \setminus \{a\}$ próximos a a en incrementos $h \in H_a \setminus \{0\}$ próximos a 0. Por eso el cociente puede escribirse como $\frac{f(x)-f(a)}{x-a} = \frac{f(a+h)-f(a)}{h}$. Esta identidad algebraica es sencilla. Lo importante es que también se conserva la forma de aproximación al punto.

Proposición 6.1. **Equivalencia de las dos formas del cociente incremental y unicidad de la derivada.** Sea $f \colon A \to \mathbb{R}$ y sea $a \in A$ un punto de acumulación de A. Para $L \in \mathbb{R}$, son equivalentes:

$$\lim_{\substack{x \to a \\ x \in A}} \frac{f(x) - f(a)}{x - a} = L,$$

y

$$\lim_{\substack{h\to 0\\ h\in H_a}} \frac{f(a+h)-f(a)}{h} = L,$$

donde $H_a = \{h \in \mathbb{R} : a + h \in A\}$ y en ambos cocientes se entiende que $x \neq a$ o, equivalentemente, $h \neq 0$. En particular, si la derivada existe, su valor es único.

Demostración. Consideremos primero la transformación $h = x - a$.

Cuando $x \in A$ y $x \neq a$, tenemos $h \in H_a$ y $h \neq 0$. Además,

$$x \to a \quad \Longleftrightarrow \quad x - a \to 0 \quad \Longleftrightarrow \quad h \to 0.$$

Y para esos puntos $\frac{f(x)-f(a)}{x-a} = \frac{f(a+h)-f(a)}{h}$.

Por tanto, la primera función de aproximación no es más que la segunda escrita en la coordenada desplazada $h = x - a$. El límite relativo a A en a existe y vale L exactamente cuando el límite relativo a H_a en 0 existe y vale L. Para la unicidad, si dos números reales L y M fueran valores de la derivada en a, el mismo cociente incremental tendría límite L y límite M en el mismo punto. Por la unicidad del límite funcional demostrada en `T1-C05`, necesariamente $L = M$. Así, cuando la derivada existe, no hay elección posible para su valor. □

La demostración separa dos cuestiones diferentes.

- La equivalencia de las formas $x \to a$ y $h \to 0$ proviene de una **traslación de coordenadas**.
- La unicidad de $f'(a)$ no es una nueva propiedad misteriosa: hereda directamente la **unicidad del límite**.

La definición en lenguaje ε–δ. Como la derivada es un límite funcional, podemos abrir inmediatamente su contenido cuantificado. Decir $f'_A(a) = L$ significa que para todo $\varepsilon > 0$ existe $\delta > 0$ tal que, para todo $x \in A$, $0 < |x - a| < \delta$ implica

$$\left|\frac{f(x)-f(a)}{x-a} - L\right| < \varepsilon.$$

En lenguaje de incrementos, equivale a afirmar que para todo $\varepsilon > 0$ existe $\delta > 0$ tal que, para todo incremento admisible h, $0 < |h| < \delta$ implica

$$\left|\frac{f(a+h)-f(a)}{h} - L\right| < \varepsilon.$$

La estructura lógica es la ya conocida:

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall h \in H_a : \quad 0 < |h| < \delta \Longrightarrow \left|\frac{f(a+h)-f(a)}{h} - L\right| < \varepsilon.$$

La novedad no está en los cuantificadores. Está en **qué función** sometemos al control ε–δ: el cociente incremental.

Qué controla realmente la derivada
La condición anterior no exige que $f(a + h) - f(a)$ sea simplemente pequeño. La continuidad ya controlaba eso. La derivada exige algo más preciso: que la **razón** $\frac{f(a+h)-f(a)}{h}$ sea arbitrariamente próxima a un número fijo L cuando la escala $|h|$ se hace pequeña.

Primeros cálculos desde la definición. El objetivo de este capítulo no es construir todavía un catálogo de reglas. Antes de aprender a derivar mecánicamente sumas, productos o composiciones, necesitamos ver qué dice la definición en los casos más simples.

Ejemplo 6.2. **Funciones constante y afín desde la definición.** Consideremos primero una función constante $f(x) = c$ sobre un dominio $A \subseteq \mathbb{R}$, y sea $a \in A$ un punto de acumulación de A. Para todo incremento admisible $h \neq 0$, $\frac{f(a+h)-f(a)}{h} = \frac{c-c}{h} = 0$.

Por tanto, $f'_A(a) = 0$.

Ahora sea $g(x) = mx + b$.

Entonces, para $h \neq 0$,

$$\begin{aligned}\frac{g(a+h)-g(a)}{h} &= \frac{m(a+h)+b-(ma+b)}{h}\\ &= \frac{mh}{h}\\ &= m.\end{aligned}$$

El cociente incremental es ya constante en todo entorno perforado admisible de 0. En consecuencia, $g'_A(a) = m$. Así, la derivada recupera exactamente la pendiente de una función afín y da pendiente nula a una función constante.

Este ejemplo merece una lectura conceptual. Para una función afín, todas las secantes tienen la misma pendiente. No hace falta que las pendientes **se estabilicen** al reducir la escala: ya están estabilizadas desde el comienzo. En una función no afín, como x^2, la pendiente de la secante depende de h y solo al tomar un límite aparece una pendiente local. Ésta es una primera diferencia entre una función exactamente afín y una función que, más adelante, será aproximadamente afín a pequeña escala.

El criterio secuencial para diferenciabilidad. La definición ε–δ es la formulación básica, pero el capítulo 4 nos dio otro lenguaje completamente equivalente para los límites funcionales: las sucesiones. Apliquémoslo ahora no a f directamente, sino a la función cociente incremental

$$q_a(x) := \frac{f(x)-f(a)}{x-a}, \qquad x \in A \setminus \{a\}.$$

Preguntar si f es diferenciable en a es exactamente preguntar si $q_a(x)$ tiene un límite real cuando $x \to a$ dentro de A.

Proposición 6.2. **Criterio secuencial para diferenciabilidad.** Sea $f\colon A \to \mathbb{R}$ y sea $a \in A$ un punto de acumulación de A. Sea $L \in \mathbb{R}$. Entonces son equivalentes:

1. f es diferenciable en a y $f'_A(a) = L$;
2. para toda sucesión (x_n) de puntos de $A \setminus \{a\}$ tal que
$$x_n \to a,$$
se cumple
$$\frac{f(x_n)-f(a)}{x_n-a} \longrightarrow L.$$

Equivalentemente, en lenguaje de incrementos: para toda sucesión de incrementos admisibles $h_n \in H_a \setminus \{0\}$ con $h_n \to 0$,

$$\frac{f(a+h_n)-f(a)}{h_n} \longrightarrow L.$$

Demostración. Consideremos de nuevo $q_a(x) = \frac{f(x)-f(a)}{x-a}$ sobre $A \setminus \{a\}$.
Por definición, $f'_A(a) = L$ si y solo si

$$\lim_{\substack{x\to a \\ x\in A}} q_a(x) = L.$$

El criterio secuencial para límites funcionales de `T1-C05` afirma que este límite existe y vale L si y solo si, para toda sucesión (x_n) del dominio perforado que converja a a, $q_a(x_n) \to L$. Sustituyendo la definición de q_a obtenemos exactamente $\frac{f(x_n)-f(a)}{x_n-a} \to L$. La versión con h_n se obtiene mediante $h_n = x_n - a$ y la equivalencia de `#prp-t1-0035`. □

Este criterio tiene dos usos distintos. Para **demostrar** que una derivada existe, hay que controlar el cociente para toda sucesión admisible que se acerque al punto. En muchos ejemplos sencillos será más directo calcular el límite funcional. Para **demostrar que una derivada no existe**, en cambio, basta encontrar dos sucesiones admisibles que se acerquen al mismo punto y produzcan comportamientos incompatibles del cociente incremental. Ese método será central en §6.3.

Cómo refutar una derivada con sucesiones
Si encontramos $x_n \to a$ y $y_n \to a$, con $x_n, y_n \in A \setminus \{a\}$, tales que $\frac{f(x_n)-f(a)}{x_n-a} \to L_1$ y $\frac{f(y_n)-f(a)}{y_n-a} \to L_2$, con $L_1 \neq L_2$, entonces $f'(a)$ no existe. No estamos usando una regla especial de derivación: estamos negando la existencia de un límite mediante el criterio secuencial.

Tres pruebas de estrés para la definición. Una definición matemática se comprende mejor cuando intentamos hacerla fallar.

1. El centro pertenece al dominio, pero se excluye del cociente. Para hablar de $f'(a)$ necesitamos conocer $f(a)$, pues aparece en el numerador. Sin embargo, el cociente $\frac{f(x)-f(a)}{x-a}$ solo se evalúa cuando $x \neq a$. No hay contradicción. La función original debe estar definida en el centro, mientras que la función auxiliar q_a vive en un dominio perforado alrededor de ese centro.

2. Un punto aislado no tiene derivada bajo nuestra definición. Supongamos, por ejemplo, que $A = \{0\} \cup [1, \infty)$ y consideramos cualquier función $f\colon A \to \mathbb{R}$. El punto 0 pertenece al dominio, pero es aislado. No existen puntos de $A \setminus \{0\}$ arbitrariamente próximos a 0. Por tanto, `#def-t1-0042` no asigna $f'_A(0)$. La ausencia de puntos cercanos no constituye una prueba de regularidad; constituye ausencia de información diferencial.

3. Una pendiente que escapa a infinito no es una derivada real. Nuestra definición exige $f'_A(a) \in \mathbb{R}$.

Un cociente incremental que tiende a $+\infty$ o $-\infty$ no tiene límite **real finito**. Ese comportamiento puede recibir una descripción geométrica propia, pero no proporciona una derivada de acuerdo con la definición adoptada.

Derivada en un punto y función derivada no son lo mismo. La notación $f'(a)$ designa un número asociado a un punto concreto, siempre que el límite correspondiente exista. La notación f' designa, en cambio, una nueva función cuyo dominio es

$$\mathrm{Dom}(f') = \{a \in A : f \text{ es diferenciable en } a\}.$$

Estas dos afirmaciones tienen niveles lógicos diferentes:

$$\boxed{f \text{ es diferenciable en } a}$$

es una afirmación puntual, mientras que

$$\boxed{f \text{ es diferenciable en todos los puntos de un conjunto } E}$$

es una familia de afirmaciones puntuales, una por cada $a \in E$. Por el momento estudiamos la existencia de la derivada en un punto. La continuidad de la función derivada constituye otra propiedad y requerirá un análisis adicional; la existencia del número $f'(a)$ no permite darla por supuesta.

Lo que la definición todavía no nos regala. Hemos formalizado la derivada, pero no hemos obtenido ninguna regla de cálculo. De `#def-t1-0042` no se sigue automáticamente una tabla como

$$(f+g)' = f' + g', \qquad (fg)' = f'g + fg', \qquad (g \circ f)' = (g' \circ f)f'.$$

Esas afirmaciones requieren teoremas y pertenecen a `T1-C08`. En este capítulo seguiremos trabajando primero **desde la definición**. Esta disciplina tiene una finalidad: aprender qué mecanismo debe verificarse antes de convertir la derivación en una técnica algebraica eficiente. Por el momento, el núcleo conceptual es exactamente éste:

$$\boxed{f'_A(a) = L \quad \Longleftrightarrow \quad \frac{f(a+h)-f(a)}{h} \longrightarrow L \text{ para incrementos admisibles } h \to 0.}$$

Y, por el criterio secuencial,

$$\boxed{f'_A(a) = L \quad \Longleftrightarrow \quad \frac{f(a+h_n)-f(a)}{h_n} \longrightarrow L \text{ para toda sucesión admisible } h_n \to 0,}$$

con $h_n \neq 0$.

La sección siguiente utilizará estas dos formulaciones como herramientas de trabajo. Calcularemos derivadas sencillas directamente desde la definición y, con la misma definición, aprenderemos a demostrar que una derivada **no** existe.

6.3. Calcular desde la definición y aprender a fallar

Ya tenemos una definición precisa de derivada y dos lenguajes equivalentes para usarla. Si f está definida alrededor de un punto a, podemos estudiar directamente

$$\frac{f(a+h)-f(a)}{h}, \qquad h \neq 0,$$

y preguntar si ese cociente tiene un límite real cuando $h \to 0$. Éste será todavía nuestro único mecanismo de cálculo. No usaremos reglas de suma, producto, cociente, potencias o composición. Esas reglas llegarán en `T1-C08` y deberán ser demostradas. Por ahora queremos adquirir una habilidad más básica: **mirar una función, construir su cociente incremental y decidir qué le ocurre cuando la escala se reduce**. La misma definición servirá para dos tareas aparentemente opuestas:

hacer aparecer un límite y demostrar que ningún límite puede existir.

Esa simetría es importante. Saber derivar no consiste solo en obtener números; también consiste en reconocer cuándo la información local de la función es incompatible con una única pendiente.

Un protocolo de cálculo desde la definición. Antes de trabajar ejemplos conviene separar las operaciones que realmente intervienen. Fijado un punto a, seguiremos este esquema:

1. calcular exactamente $f(a+h)-f(a)$;
2. formar el cociente incremental para $h \neq 0$;
3. simplificar **solo bajo la condición** $h \neq 0$;
4. estudiar el límite de la expresión simplificada cuando $h \to 0$;
5. si el límite existe como número real, identificarlo con $f'(a)$.

La parte algebraica ocurre en los pasos 1–3. La parte analítica ocurre en el paso 4. Confundirlas suele producir dos errores opuestos: intentar sustituir $h=0$ antes de simplificar o, al revés, creer que una simplificación algebraica por sí sola ya demuestra la existencia del límite.

La pregunta que debe guiar el cálculo
No preguntes primero «¿qué fórmula de derivación conozco?». Pregunta:

¿en qué se convierte exactamente el cociente incremental para $h \neq 0$?

Una vez respondida esa pregunta, la teoría de límites hace el resto.

La parábola, ahora sí como cálculo de derivada. En §6.1 ya descubrimos que las pendientes secantes de x^2 se estabilizan. Ahora podemos reinterpretar ese mismo cálculo mediante la definición formal introducida en §6.2.

Ejemplo 6.3. **Derivada de $f(x)=x^2$ en un punto arbitrario desde la definición.** Sea

$$f(x)=x^2, \qquad x \in \mathbb{R},$$

y fijemos $a \in \mathbb{R}$.
Para $h \neq 0$,

$$\begin{aligned} \frac{f(a+h)-f(a)}{h} &= \frac{(a+h)^2-a^2}{h} \\ &= \frac{2ah+h^2}{h} \\ &= 2a+h. \end{aligned}$$

La cancelación de h es legítima porque estamos trabajando en el dominio perforado $h \neq 0$. Ahora sí tomamos el límite:

$$\lim_{h\to 0}(2a+h)=2a.$$

Por `#def-t1-0042`,

$$\boxed{f'(a) = 2a.}$$

Como a era arbitrario, la función es diferenciable en todo punto real y su función derivada queda dada, en este caso concreto, por $f'(x) = 2x$. Este resultado ha sido obtenido **sin** utilizar una regla para derivar potencias.

El cálculo anterior contiene un patrón que volverá a aparecer muchas veces: el numerador parece producir una indeterminación si uno intenta poner $h = 0$, pero para $h \neq 0$ contiene un factor h que puede separarse. Después de cancelar ese factor queda una expresión cuyo límite es ordinario. Podemos describir la arquitectura del cálculo así:

$$\boxed{\frac{(a+h)^2 - a^2}{h} \xrightarrow{h\neq 0} 2a + h \xrightarrow{h\to 0} 2a.}$$

La primera flecha es álgebra. La segunda es análisis.

Un grado más: x^3 sin regla de potencias. Repitamos deliberadamente el procedimiento con $g(x) = x^3$.

Fijado $a \in \mathbb{R}$ y para $h \neq 0$,

$$\begin{aligned}\frac{g(a+h) - g(a)}{h} &= \frac{(a+h)^3 - a^3}{h} \\ &= \frac{3a^2h + 3ah^2 + h^3}{h} \\ &= 3a^2 + 3ah + h^2.\end{aligned}$$

Por tanto,

$$\lim_{h\to 0} \left(3a^2 + 3ah + h^2\right) = 3a^2,$$

y así

$$\boxed{g'(a) = 3a^2.}$$

De nuevo, no hemos usado una regla general. Hemos demostrado el resultado para esta función mediante una identidad algebraica exacta y las leyes de límites ya disponibles. Conviene observar qué fue esencial y qué no. No era importante «recordar» la expansión del cubo como un truco de derivación. Lo importante era transformar $(a+h)^3 - a^3$ en una expresión donde el factor h que necesariamente aparece al comparar dos valores pueda hacerse visible y cancelarse para $h \neq 0$.

Factorizar no es tomar el límite
En los ejemplos de x^2 y x^3, la factorización elimina una obstrucción algebraica. Pero la derivada aparece **después**, al demostrar un límite.
No debemos confundir

simplificar el cociente

con

demostrar que el cociente tiene límite.

La definición también debe saber decir «no». Los ejemplos anteriores admitieron simplificaciones que produjeron un límite real. Para decidir si ese procedimiento funciona en otros casos debemos examinar también cocientes cuyos límites laterales discrepan o cuyos valores no se estabilizan.

Eso es falso.

La diferenciabilidad exige que **todos** los modos admisibles de acercarse al punto produzcan el mismo comportamiento límite del cociente. Por eso el criterio secuencial de `#prp-t1-0036` es especialmente eficaz para refutar una derivada: basta encontrar dos sucesiones que se aproximen al mismo punto y produzcan límites incompatibles. El valor absoluto proporciona el ejemplo elemental decisivo.

EJEMPLO 6.4. $f(x) = |x|$ **en** 0**: continuidad sin derivabilidad.** Consideremos $f(x) = |x|$.

Sabemos que f es continua en 0 y que $f(0) = 0$. Para estudiar la derivabilidad debemos examinar otra cantidad:

$$\frac{f(h) - f(0)}{h} = \frac{|h|}{h}, \qquad h \neq 0.$$

Si $h > 0$, entonces $|h| = h$ y por tanto $\frac{|h|}{h} = 1$.

Si $h < 0$, entonces $|h| = -h$ y $\frac{|h|}{h} = -1$.

Ahora usemos dos sucesiones de incrementos admisibles:

$$h_n = \frac{1}{n}, \qquad k_n = -\frac{1}{n}.$$

Ambas satisfacen

$$h_n \to 0, \qquad k_n \to 0,$$

pero los correspondientes cocientes incrementales cumplen $\frac{|h_n|}{h_n} = 1 \longrightarrow 1$, mientras que $\frac{|k_n|}{k_n} = -1 \longrightarrow -1$. Si $f'(0)$ existiera, `#prp-t1-0036` obligaría a que **toda** sucesión admisible produjera el mismo límite. Como hemos obtenido dos valores distintos, $1 \neq -1$, concluimos que

$$\boxed{f'(0) \text{ no existe.}}$$

Así, $|x|$ es continua en 0 pero no diferenciable allí.

Este argumento merece ser leído lógicamente. No hemos demostrado la inexistencia de la derivada examinando «muchos» valores de h. Hemos usado la negación exacta del criterio universal.

La diferenciabilidad exigiría:

$$\text{para toda sucesión } h_n \to 0, \quad \frac{f(h_n) - f(0)}{h_n} \to L$$

para un mismo L. Para negar esa afirmación basta exhibir dos sucesiones que obliguen a valores incompatibles.

En otras palabras,

$$\boxed{\text{dos caminos incompatibles bastan para destruir un límite único.}}$$

No necesitamos todavía introducir el vocabulario formal de derivada derecha e izquierda. En §6.4 daremos nombre a esos comportamientos. Aquí basta observar que los incrementos positivos y negativos producen información incompatible.

Continuidad y diferenciabilidad están mirando escalas distintas. El ejemplo de $|x|$ permite ver una diferencia conceptual importante. Para continuidad en 0 estudiamos $|h|-|0| = |h|$, y esta cantidad tiende a 0. Para diferenciabilidad estudiamos la variación **dividida por la escala del incremento**: $\frac{|h|}{h}$. La variación absoluta se hace pequeña, pero su razón con el incremento conserva una diferencia de signo que no desaparece al aumentar el zoom. Podemos resumir el contraste así:

$$\boxed{\begin{array}{lcl} |h| & \longrightarrow & 0, \\ \dfrac{|h|}{h} & \not\longrightarrow & \text{un único número real.} \end{array}}$$

Esto muestra concretamente por qué conocer la continuidad no basta para conocer la derivabilidad. La primera controla que el cambio sea pequeño; la segunda exige además que ese cambio tenga una **parte proporcional estable de primer orden**.

Cambiar de fórmula no implica perder la derivabilidad. El fracaso de $|x|$ podría sugerir otra regla falsa: «si una función cambia de fórmula en un punto, entonces allí no puede ser diferenciable».

También es falso.

Lo decisivo no es cuántas fórmulas aparecen en la definición, sino qué ocurre con el cociente incremental.

Ejemplo 6.5. $f(x) = x|x|$ **en 0: cambio de fórmula sin pérdida de derivabilidad.** Consideremos $f(x) = x|x|$.

Esta función puede escribirse por tramos como

$$f(x) = \begin{cases} x^2, & x \geq 0, \\ -x^2, & x < 0. \end{cases}$$

En particular, $f(0) = 0$.

Para $h \neq 0$, $\frac{f(h)-f(0)}{h} = \frac{h|h|}{h} = |h|$.

Ahora el comportamiento es completamente distinto del ejemplo anterior:

$$|h| \longrightarrow 0 \qquad (h \to 0).$$

Por tanto,

$$\boxed{f'(0) = 0.}$$

También podemos verificarlo con el criterio secuencial. Si (h_n) es **cualquier** sucesión no nula con $h_n \to 0$, entonces $\frac{f(h_n)-f(0)}{h_n} = |h_n| \to 0$. La forma de la función cambia al atravesar 0, pero toda sucesión admisible produce el mismo límite del cociente incremental.

Comparemos los dos ejemplos centrales:

función	cociente incremental en 0	comportamiento
$\lvert x\rvert$	$\dfrac{\lvert h\rvert}{h}$	produce 1 para $h>0$ y -1 para $h<0$
$x\lvert x\rvert$	$\lvert h\rvert$	tiende a 0 desde cualquier sucesión admisible

La diferencia entre ambas funciones es pequeña en apariencia, pero decisiva a escala diferencial. Multiplicar $|x|$ por x añade precisamente un factor que hace desaparecer la incompatibilidad de primer orden. La comparación puede hacerse directamente en los cocientes: el denominador fija la escala respecto de la cual se mide la variación de la función.

Dos funciones por tramos pueden comportarse de maneras opuestas. Los ejemplos anteriores permiten desmontar una intuición gráfica demasiado rápida. Una función por tramos puede presentar un cambio visible de expresión algebraica y, sin embargo, ser diferenciable en la interfaz. Otra puede ser perfectamente continua y aun así no poseer una derivada allí. Por tanto, ninguna de estas frases constituye un criterio válido:

- «la fórmula cambia, luego no hay derivada»;
- «la gráfica no tiene salto, luego hay derivada»;
- «la función se ve suave, luego el límite existe».

El criterio válido sigue siendo el mismo:

construir el cociente incremental y estudiar su límite.

Cómo elegir una estrategia: cálculo o refutación. En los ejemplos de esta sección han aparecido dos arquitecturas distintas. Cuando queremos **demostrar existencia**, buscamos una expresión exacta que permita controlar todos los incrementos admisibles: $\frac{f(a+h)-f(a)}{h} \longrightarrow L$.

Así procedimos para x^2, x^3 y $x|x|$.

Cuando queremos **demostrar inexistencia**, puede ser mucho más eficiente construir testigos incompatibles:

$$h_n \to 0, \qquad k_n \to 0,$$

pero

$$\frac{f(a+h_n)-f(a)}{h_n} \to L_1, \qquad \frac{f(a+k_n)-f(a)}{k_n} \to L_2, \qquad L_1 \neq L_2.$$

Así procedimos para $|x|$.

Podemos convertirlo en un pequeño mapa de decisión:

formar el cociente incremental

↓

¿puedo simplificarlo y controlar todos los h?

↙ ↘

sí — no o sospecho incompatibilidad

↓ ↓

calcular el límite — buscar sucesiones testigo.

Errores que estos ejemplos permiten diagnosticar.

Cuatro errores frecuentes
1. Sustituir $h = 0$ en el cociente original.
La derivada es un límite en un entorno perforado. El cociente no necesita estar definido en $h = 0$.
2. Comprobar solo incrementos positivos.
En un punto interior de un dominio como $\mathbb{R}$, la definición exige controlar todos los incrementos suficientemente pequeños, no solo los de un signo.
3. Usar una regla de derivación que todavía no ha sido demostrada.
Saber por experiencia que $(x^3)' = 3x^2$ no reemplaza la prueba desde la definición que estamos practicando aquí.
4. Decidir por la apariencia de la gráfica.
Un dibujo puede sugerir qué investigar, pero la existencia o inexistencia de la derivada es una afirmación sobre un límite.

Hay además una disciplina algebraica que conviene conservar. Factorizar, expandir o racionalizar no son rituales del cálculo. Son herramientas que se usan solo cuando ayudan a revelar el comportamiento del cociente para $h \neq 0$. En x^2 y x^3 la factorización aparece naturalmente; en otros ejemplos puede ser necesaria otra transformación. La estrategia debe responder a la estructura del numerador, no a una receta memorizada.

Tres niveles de lectura de un cálculo de derivada. Un cálculo completo puede leerse en tres niveles.

Nivel algebraico. ¿Qué identidad exacta convierte el cociente en una forma manejable?

Nivel analítico. ¿Qué teorema de límites justifica el paso $h \to 0$?

Nivel lógico. ¿Estoy probando existencia para todos los acercamientos admisibles o refutándola mediante testigos incompatibles?

Por ejemplo, para x^2:

$$\underbrace{\frac{(a+h)^2 - a^2}{h} = 2a + h}_{\text{álgebra para } h\neq 0} \quad \Longrightarrow \quad \underbrace{2a + h \to 2a}_{\text{límite}} \quad \Longrightarrow \quad \underbrace{f'(a) = 2a}_{\text{definición}}.$$

Para $|x|$:

$$\underbrace{\frac{|h|}{h}}_{\text{cociente}} \xrightarrow{h_n = 1/n} 1, \qquad \underbrace{\frac{|h|}{h}}_{\text{mismo cociente}} \xrightarrow{k_n = -1/n} -1,$$

y la incompatibilidad de esos límites refuta la existencia de una derivada única.

Qué hemos aprendido antes de tener reglas. Sin disponer todavía de un álgebra de derivadas, ya podemos distinguir cuatro situaciones conceptuales:

1. una función afín, cuyo cociente incremental es constante;
2. una función como x^2 o x^3, cuyo cociente depende de h pero converge tras simplificar;
3. una función como $|x|$, cuyo cociente conserva comportamientos incompatibles al acercarse al punto;
4. una función por tramos como $x|x|$, cuya expresión cambia pero cuyo cociente sí converge.

El punto común a las cuatro situaciones es que la derivabilidad no está codificada en la apariencia superficial de la fórmula. Está codificada en el **límite del cambio relativo a la escala**. La sección puede resumirse mediante una sola oposición:

existencia de derivada	$\Longleftrightarrow$	un único comportamiento límite del cociente,
no derivabilidad	$\Longleftarrow$	dos comportamientos límite incompatibles.

En §6.4 daremos una organización más fina a las maneras en que este límite puede fallar. Introduciremos derivadas derecha e izquierda, distinguiremos una esquina de una pendiente que crece sin cota y separaremos ambos fenómenos de la oscilación persistente del cociente incremental.

6.4. Derivadas laterales, extremos, esquinas y pendientes verticales

En §6.3 vimos que la derivabilidad puede fallar porque distintos modos de acercarse al punto producen comportamientos incompatibles del cociente incremental. El ejemplo de $|x|$ ya contenía la idea esencial: los incrementos positivos y negativos aportaban información diferente. Ahora daremos un nombre preciso a esas dos direcciones y separaremos tres fenómenos que un dibujo puede confundir fácilmente:

esquina	vs.	pendiente vertical	vs.	oscilación persistente.

Los tres impiden una derivada bilateral real finita, pero lo hacen por razones matemáticamente distintas.

Acercarse al punto desde un solo lado. Sea $f\colon A \to \mathbb{R}$, y sea $a \in A$.

Para hablar de comportamiento diferencial desde la derecha necesitamos que existan puntos de A mayores que a arbitrariamente próximos a a. Es decir, a debe ser punto de acumulación de $A \cap (a, \infty)$. Análogamente, para hablar de comportamiento desde la izquierda necesitamos que a sea punto de acumulación de $A \cap (-\infty, a)$. La idea es exactamente la misma que para los límites laterales del capítulo 4, pero aplicada al cociente incremental.

Definición 6.2. **Derivadas laterales.** Sea $f\colon A \to \mathbb{R}$ y sea $a \in A$.

Si a es punto de acumulación de $A \cap (a, \infty)$, decimos que f tiene **derivada derecha** en a respecto de A si existe como número real finito el límite

$$f'_{A,+}(a) := \lim_{\substack{x \to a \\ x \in A,\ x > a}} \frac{f(x) - f(a)}{x - a}.$$

Equivalentemente, en lenguaje de incrementos,

$$f'_{A,+}(a) = \lim_{\substack{h \to 0^+ \\ a+h \in A}} \frac{f(a+h) - f(a)}{h}.$$

Si a es punto de acumulación de $A \cap (-\infty, a)$, decimos que f tiene **derivada izquierda** en a respecto de A si existe como número real finito el límite

$$f'_{A,-}(a) := \lim_{\substack{x \to a \\ x \in A,\ x < a}} \frac{f(x) - f(a)}{x - a},$$

o, equivalentemente,

$$f'_{A,-}(a) = \lim_{\substack{h\to 0^- \\ a+h\in A}} \frac{f(a+h)-f(a)}{h}.$$

Cuando el dominio sea claro escribiremos simplemente $f'_+(a)$ y $f'_-(a)$.

Hay dos detalles de la definición que conviene destacar. Primero, una derivada lateral sigue siendo una **derivada definida mediante un límite**. No consiste en sustituir un signo en una fórmula ni en mirar solo una parte del dibujo. Segundo, bajo nuestra convención, una derivada lateral también debe ser un número real finito. Si el cociente incremental tiende a $+\infty$ o $-\infty$ por uno de los lados, describiremos ese comportamiento explícitamente, pero no escribiremos que la derivada lateral vale infinito.

El criterio bilateral por derivadas laterales. En un punto interior de un intervalo —o, más generalmente, en un punto donde el dominio acumula por ambos lados— la derivada ordinaria debe reconciliar la información derecha e izquierda. La razón es estructural. El cociente incremental $q_a(x) = \frac{f(x)-f(a)}{x-a}$ posee un límite bilateral exactamente cuando sus dos límites laterales existen y coinciden.

Proposición 6.3. **Criterio lateral para la derivada bilateral.** Sea $f\colon A \to \mathbb{R}$ y sea $a \in A$. Supongamos que a es punto de acumulación de A tanto por la izquierda como por la derecha.

Entonces f es diferenciable en a respecto de A si y solo si existen las dos derivadas laterales como números reales y

$$f'_{A,-}(a) = f'_{A,+}(a).$$

En ese caso,

$$f'_A(a) = f'_{A,-}(a) = f'_{A,+}(a).$$

Demostración. Definamos, para $x \in A \setminus \{a\}$, $q_a(x) := \frac{f(x)-f(a)}{x-a}$.

Por `#def-t1-0042`, f es diferenciable en a exactamente cuando existe el límite bilateral

$$\lim_{\substack{x\to a \\ x\in A}} q_a(x)$$

como número real finito.

Como A acumula en a por ambos lados, el criterio de límites laterales demostrado en `T1-C05` afirma que este límite bilateral existe y vale L si y solo si existen los dos límites laterales de q_a y ambos valen L. Pero esos dos límites son precisamente

$$f'_{A,-}(a) \qquad \text{y} \qquad f'_{A,+}(a).$$

Por tanto, la derivada bilateral existe si y solo si ambas derivadas laterales existen como números reales y coinciden. En ese caso los tres valores son iguales. □

La prueba es deliberadamente corta porque no necesita una teoría nueva: el resultado se hereda directamente del teorema bilateral $\leftrightarrow$ laterales para límites funcionales.

La hipótesis de dos lados no es ornamental
El criterio anterior supone que el dominio acumula en a **por ambos lados**. Si $A = [0, \infty)$ y $a = 0$, no existe aproximación dentro de A desde la izquierda. En ese caso la derivada relativa al dominio de `#def-t1-0042` es naturalmente unilateral y coincide con la derivada derecha cuando ésta existe. No debemos exigir un límite lateral desde puntos que no pertenecen al dominio.

Extremos de intervalos: la geometría del dominio importa. Consideremos la restricción

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

En 0 solo existen incrementos admisibles $h \geq 0$. Para $h > 0$, $\frac{f(h)-f(0)}{h} = \frac{h^2}{h} = h$, y por tanto

$$\lim_{h\to 0^+} h = 0.$$

Así, $f'_{[0,\infty)}(0) = f'_+(0) = 0$.

No hay ninguna derivada izquierda que verificar respecto de este dominio, porque 0 no es punto de acumulación de $[0, \infty)$ por la izquierda. Esto muestra por qué la frase «para que exista una derivada deben coincidir las derivadas izquierda y derecha» necesita una condición previa. Es correcta en un punto donde ambas aproximaciones existen dentro del dominio; no es un criterio universal para extremos.

Una esquina: dos pendientes finitas, pero incompatibles. Regresemos a $f(x) = |x|$. En §6.3 calculamos $\frac{f(h)-f(0)}{h} = \frac{|h|}{h}$.

Ahora podemos expresar el diagnóstico en el nuevo lenguaje:

$$f'_-(0) = -1, \qquad f'_+(0) = 1.$$

Las dos derivadas laterales existen y son finitas, pero no coinciden. Por `#prp-t1-0037`, $f'(0)$ no existe. A este patrón local lo llamaremos una **esquina**: las pendientes laterales se aproximan a dos números reales distintos. El punto importante no es la palabra geométrica, sino la firma analítica:

$$\boxed{\text{esquina en } a \quad \Longrightarrow \quad f'_-(a), f'_+(a) \in \mathbb{R} \text{ existen, pero } f'_-(a) \neq f'_+(a).}$$

La implicación se usa aquí como descripción del patrón que estamos llamando esquina; no pretende clasificar toda posible geometría no suave a partir de una figura.

Pendiente vertical: el cociente no se estabiliza en un número real. Una situación distinta ocurre cuando las pendientes secantes crecen sin cota al acercarnos al punto. Consideremos la función de raíz cuadrada con signo

$$g(x) = \begin{cases} \sqrt{x}, & x \geq 0, \\ -\sqrt{-x}, & x < 0. \end{cases}$$

La continuidad de $\sqrt{\ }$ en $[0, \infty)$ ya está disponible desde `T1-C06`, de modo que $g(x) \to 0 = g(0)$ cuando $x \to 0$. En 0, para $h \neq 0$,

$$\frac{g(h)-g(0)}{h}=\frac{1}{\sqrt{|h|}}.$$

Como $\sqrt{|h|}\to 0^+$ cuando $h\to 0$,

$$\frac{1}{\sqrt{|h|}}\longrightarrow +\infty.$$

Geométricamente, las pendientes de las secantes se hacen arbitrariamente grandes y la gráfica presenta una **tangente vertical** en el sentido geométrico usual, la recta $x=0$. Pero bajo `#def-t1-0042` esto no significa $g'(0)=+\infty$. Tal igualdad no pertenece a nuestra definición: $+\infty$ no es un número real. La conclusión correcta es

$$\boxed{g'(0)\text{ no existe como derivada real finita.}}$$

El fenómeno es diferente de una esquina. En una esquina existen dos pendientes laterales reales incompatibles; aquí no aparece ninguna pendiente real finita porque el cociente escapa sin cota.

Un mismo «no hay derivada» puede esconder mecanismos distintos.

Ejemplo 6.6. **Esquina, pendiente vertical y oscilación: tres fallos diferentes.** Compararemos tres funciones en 0.

1. Esquina:

$$f(x)=|x|.$$

Tenemos

$$f'_-(0)=-1,\qquad f'_+(0)=1.$$

Los dos límites laterales del cociente existen como números reales, pero son distintos.

2. Pendiente vertical:

$$g(x)=\begin{cases}\sqrt{x}, & x\geq 0,\\ -\sqrt{-x}, & x<0.\end{cases}$$

Su cociente incremental es

$$\frac{g(h)-g(0)}{h}=\frac{1}{\sqrt{|h|}}\longrightarrow +\infty.$$

No hay una derivada real finita, aunque geométricamente las secantes sugieren la recta vertical $x=0$.

3. Oscilación persistente: definamos

$$p(x)=\begin{cases}x\sigma(1/x), & x\neq 0,\\ 0, & x=0.\end{cases}$$

La función es continua en 0 porque $|x\sigma(1/x)|\leq |x|\longrightarrow 0$.

Sin embargo, para $h\neq 0$, $\frac{p(h)-p(0)}{h}=\sigma(1/h)$.

Incluso restringiéndonos a $h>0$, tomemos

$$h_n = \frac{1}{\frac{\omega}{2} + 2\omega n}, \qquad k_n = \frac{1}{\frac{3\omega}{2} + 2\omega n}.$$

Entonces

$$h_n \to 0^+, \qquad k_n \to 0^+,$$

pero

$$\sigma(1/h_n) = 1, \qquad \sigma(1/k_n) = -1.$$

Por el criterio secuencial, ni siquiera existe la derivada derecha de p en 0. Los tres casos concluyen «no existe una derivada bilateral real finita», pero por mecanismos distintos:

fenómeno	comportamiento del cociente incremental
esquina	dos límites laterales finitos distintos
pendiente vertical	crecimiento sin cota
oscilación	no hay estabilización ni siquiera lateral

Este cuadro es más informativo que una única etiqueta «no diferenciable». El cálculo diferencial no solo pregunta si un límite existe: cuando falla, conviene identificar **cómo** falla.

Cuidado con la palabra «cúspide». En dibujos elementales se usa a veces la palabra **cúspide** para diversos perfiles puntiagudos con pendientes no acotadas. Esa terminología geométrica no será nuestra herramienta de clasificación principal. Analíticamente distinguiremos siempre el comportamiento concreto del cociente incremental. Por ejemplo, puede ocurrir que $\frac{f(a+h)-f(a)}{h} \to +\infty$ desde ambos lados, o que tienda a $-\infty$ desde un lado y a $+\infty$ desde el otro. Ninguno de estos casos produce una derivada real finita, aunque sus dibujos puedan recibir nombres geométricos diferentes. La disciplina del capítulo será describir primero los límites que realmente ocurren y solo después añadir una interpretación gráfica.

Las derivadas laterales no son dos números que siempre existan. Es tentador pensar que toda función posee automáticamente una «pendiente por la izquierda» y otra «por la derecha», y que el único problema consiste en compararlas.

Eso también es falso.

Para que $f'_+(a)$ exista necesitamos un límite real finito del cociente cuando $h \to 0^+$. El ejemplo oscilatorio anterior muestra que este límite puede no existir. Lo mismo puede ocurrir por la izquierda. Así, en un punto donde el dominio acumula por ambos lados, la auditoría correcta es:

¿existe $f'_-(a)$ como real finito?
¿existe $f'_+(a)$ como real finito?
si ambas existen, ¿son iguales?

Solo una respuesta afirmativa a las tres preguntas produce una derivada bilateral.

Un mapa de diagnóstico local. Podemos organizar los comportamientos vistos hasta ahora como un árbol de decisión.

Partimos del cociente incremental $Q_a(h) = \frac{f(a+h)-f(a)}{h}$.

Si el dominio admite incrementos de ambos signos, preguntamos primero por cada lado:

$Q_a(h)$ cuando $h \to 0^-$	$Q_a(h)$ cuando $h \to 0^+$
$\downarrow$	$\downarrow$
¿límite real finito?	¿límite real finito?

A partir de allí:

- si ambos límites finitos existen y son iguales, hay derivada bilateral;
- si ambos existen pero son distintos, aparece el patrón de esquina;
- si el cociente escapa a $\pm\infty$, hay pendiente no acotada, no una derivada real;
- si el cociente oscila sin límite, falla incluso la estabilización lateral correspondiente.

En un extremo del dominio, en cambio, solo se audita el lado que realmente pertenece al dominio.

Qué no hemos usado. Toda la sección descansa únicamente en:

1. la definición de derivada de §6.2;
2. los límites laterales de `T1-C05`;
3. el criterio bilateral $\leftrightarrow$ laterales para límites funcionales;
4. el criterio secuencial;
5. estimaciones y álgebra elementales.

No hemos usado reglas de suma, producto, cociente o cadena para derivadas. Tampoco hemos usado Rolle, el teorema del valor medio ni ninguna caracterización global de monotonía. Eso preserva la dirección deductiva del capítulo.

Hacia una relación estructural con la continuidad. Las secciones anteriores han mostrado que una función puede ser continua y, sin embargo, fallar en ser diferenciable: $|x|$ y la función oscilatoria $x\sigma(1/x)$ en 0 son ejemplos distintos de ese fenómeno. La pregunta inversa es más profunda:

si una función es diferenciable en a, ¿debe ser continua en a?

En §6.5 demostraremos que la respuesta es afirmativa. El resultado no necesitará ninguna regla de derivación: surgirá directamente de la definición y del álgebra de límites.

6.5. Diferenciabilidad implica continuidad

Las secciones anteriores han separado cuidadosamente dos preguntas locales. La **continuidad** en a pregunta si

$$f(x) - f(a) \longrightarrow 0 \qquad (x \to a),$$

mientras que la **diferenciabilidad** pregunta si la variación, después de dividirla por la escala $x-a$, posee un límite real finito: $\frac{f(x)-f(a)}{x-a} \longrightarrow f'(a)$. A primera vista son condiciones distintas. Sin embargo, la segunda contiene más información que la primera. Si conocemos cómo se comporta

la variación **por unidad de incremento**, entonces podemos reconstruir la variación misma multiplicando de nuevo por el incremento. La pregunta guía de esta sección es, por tanto,

¿por qué controlar la variación hasta primer orden obliga ya a controlar el valor de la función?

Antes de demostrar: qué relación debemos esperar. Ya sabemos que la continuidad no garantiza diferenciabilidad. La función $|x|$ es continua en 0, pero su cociente incremental tiene comportamientos laterales incompatibles.

Así que una equivalencia

$$\text{continuidad} \iff \text{diferenciabilidad}$$

no puede ser verdadera.

Queda abierta la dirección contraria:

$$\text{diferenciabilidad} \implies \text{continuidad}.$$

¿Por qué debería funcionar?

Si f es diferenciable en a, entonces cerca de a el cociente $\frac{f(x)-f(a)}{x-a}$ se aproxima a un número real. En particular, no puede crecer sin control ni oscilar con amplitud ilimitada. Pero $f(x) - f(a) = (x-a)\frac{f(x)-f(a)}{x-a}$ para $x \neq a$. El primer factor tiende a 0; el segundo tiende a un número real finito. Esa identidad contiene ya toda la prueba.

Arquitectura antes de la demostración
La estrategia es

$$\begin{array}{c} \text{diferenciabilidad} \\ \Downarrow \\ \dfrac{f(x)-f(a)}{x-a} \to f'(a) \in \mathbb{R} \\ \Downarrow \\ (x-a)\dfrac{f(x)-f(a)}{x-a} \to 0 \cdot f'(a) = 0 \\ \Downarrow \\ f(x) - f(a) \to 0 \\ \Downarrow \\ \text{continuidad}. \end{array}$$

La prueba no necesita ninguna regla de derivación. Solo necesita la definición de derivada y el álgebra de límites ya construida.

El teorema fundamental de esta sección.

TEOREMA 6.1. **Diferenciabilidad implica continuidad.** Sea $f\colon A \to \mathbb{R}$ y sea $a \in A$ un punto de acumulación de A. Si f es diferenciable en a respecto del dominio A, entonces f es continua en a respecto de A.

Demostración. Supongamos que f es diferenciable en a. Entonces existe el número real $L = f'_A(a)$ y, por definición,

$$\lim_{\substack{x\to a\\ x\in A}} \frac{f(x)-f(a)}{x-a} = L.$$

Para todo $x \in A$ con $x \neq a$ tenemos la identidad exacta $f(x) - f(a) = (x-a)\frac{f(x)-f(a)}{x-a}$.

Cuando $x \to a$ dentro de A, $x - a \longrightarrow 0$ y $\frac{f(x)-f(a)}{x-a} \longrightarrow L$.

Por la regla del producto para límites funcionales,

$$\lim_{\substack{x\to a\\ x\in A}} (f(x)-f(a)) = 0 \cdot L = 0.$$

Por tanto,

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

que es precisamente la continuidad de f en a relativa a su dominio. □

La demostración es corta, pero conviene no confundir brevedad con trivialidad. Su mecanismo depende de una decisión anterior muy importante: la derivada fue definida como un **límite real finito**. Si el cociente incremental escapara a infinito, no podríamos aplicar simplemente el producto de límites reales como acabamos de hacerlo.

Leer la prueba sin memorizarla. La identidad central $f(x) - f(a) = (x-a)\frac{f(x)-f(a)}{x-a}$ puede reconstruirse preguntando qué queremos demostrar.

Para continuidad necesitamos controlar $f(x) - f(a)$.

Para diferenciabilidad ya controlamos $\frac{f(x)-f(a)}{x-a}$.

La forma natural de conectar ambas expresiones es multiplicar el cociente por aquello que habíamos dividido: $(x-a)\frac{f(x)-f(a)}{x-a}$. La hipótesis controla el cociente incremental y el objetivo concierne al incremento de la función. Escribir este último como producto permite aplicar exactamente la información disponible.

El factor que hace el trabajo

La diferenciabilidad controla el cociente incremental: $\frac{f(x)-f(a)}{x-a} \to L$. La continuidad aparece porque el incremento de entrada satisface $x - a \to 0$.

Por eso

$$\underbrace{f(x)-f(a)}_{\text{variación de salida}} = \underbrace{x-a}_{\to 0}\ \underbrace{\frac{f(x)-f(a)}{x-a}}_{\to L}.$$

La variación de salida queda forzada a tender a 0.

La misma prueba descomprimida en lenguaje ε–δ. El argumento con álgebra de límites es completo. Pero descomprimirlo permite ver cuantitativamente dónde aparece el control.

Sea $L = f'_A(a)$.

Como $\frac{f(x)-f(a)}{x-a} \to L$, podemos exigir, por ejemplo, que para x suficientemente próximo a a y $x \neq a$,

$$\left|\frac{f(x)-f(a)}{x-a}-L\right|<1.$$

Entonces, por desigualdad triangular,

$$\left|\frac{f(x)-f(a)}{x-a}\right|\leq |L|+1.$$

Es decir: la existencia de una derivada finita fuerza al cociente incremental a quedar **localmente acotado**. Ahora, para esos mismos puntos,

$$|f(x)-f(a)|=|x-a|\left|\frac{f(x)-f(a)}{x-a}\right|\leq (|L|+1)|x-a|.$$

Dado $\varepsilon>0$, basta exigir además $|x-a|<\frac{\varepsilon}{|L|+1}$.

Combinando ambas restricciones sobre $|x-a|$, obtenemos $|f(x)-f(a)|<\varepsilon$. Y si $x=a$, la desigualdad es automática porque $|f(a)-f(a)|=0$. Así se recupera directamente la definición de continuidad. Esta versión hace visible una idea útil:

derivabilidad $\Longrightarrow$ cociente incremental localmente acotado $\Longrightarrow$ $|f(x)-f(a)|\leq C|x-a|$ cerca de a.

No estamos introduciendo aquí una teoría nueva de regularidad. Solo estamos leyendo con mayor detalle la información que ya contiene la existencia de la derivada.

Una consecuencia inmediata: la discontinuidad descarta la derivabilidad. El teorema anterior puede leerse por contraposición.

Corolario 6.1. **Una discontinuidad excluye diferenciabilidad.** Sea $f\colon A\to\mathbb{R}$ y sea $a\in A$ un punto de acumulación de A. Si f no es continua en a respecto de A, entonces f no es diferenciable en a respecto de A.

Demostración. Si f fuera diferenciable en a, el teorema anterior obligaría a que fuera continua en a. Por contraposición, la discontinuidad impide la diferenciabilidad. □

Este corolario es una herramienta de **descarte**, no un procedimiento para calcular derivadas. Si una función tiene un salto, una oscilación que destruye continuidad o un valor central incompatible con su límite, ya sabemos que no puede ser diferenciable allí. No hace falta estudiar después su cociente incremental. Pero debe leerse con cuidado. El corolario afirma

$$\text{discontinua en } a \Longrightarrow \text{no diferenciable en } a.$$

No afirma la recíproca.

La conversa es falsa.

Ejemplo 6.7. **Continuidad necesaria, pero no suficiente: $f(x)=|x|$ en 0.** Consideremos $f(x)=|x|$.

La función es continua en 0 porque

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

Sin embargo, en §6.3 calculamos el cociente incremental

$$\frac{f(h)-f(0)}{h}=\frac{|h|}{h}, \qquad h\neq 0,$$

y en §6.4 identificamos sus derivadas laterales:

$$f'_-(0)=-1, \qquad f'_+(0)=1.$$

Como los dos límites laterales finitos son distintos, la derivada bilateral no existe.
Por tanto,

$$\boxed{|x| \text{ es continua en 0, pero no diferenciable en 0.}}$$

Este ejemplo refuta la conversa del teorema.

La relación lógica correcta es entonces

$$\boxed{\text{diferenciabilidad en } a \Longrightarrow \text{continuidad en } a,}$$

pero

$$\boxed{\text{continuidad en } a \not\Longrightarrow \text{diferenciabilidad en } a.}$$

En otras palabras, la continuidad es una **condición necesaria** para la diferenciabilidad, pero no es suficiente.

Qué ocurre en un extremo del dominio. Como todas nuestras definiciones son relativas al dominio, el teorema no exige que a sea un punto interior de A. Por ejemplo, si $f\colon [0,\infty)\to\mathbb{R}$ es diferenciable en 0 según la aproximación permitida por su dominio, el cociente incremental solo usa $h>0$ suficientemente pequeño. La misma identidad $f(h)-f(0)=h\frac{f(h)-f(0)}{h}$ muestra que

$$f(h)\to f(0) \qquad (h\to 0^+).$$

Eso es exactamente continuidad relativa a $[0,\infty)$ en 0. No necesitamos inventar valores de la función a la izquierda del dominio para obtener la conclusión.

Dos errores lógicos que conviene bloquear.

No invertir ni fortalecer el teorema
Error 1. «Si f es continua en a, entonces f es diferenciable en a.»
Falso: $|x|$ en 0 lo refuta.
Error 2. «Si f es diferenciable en a, entonces f es diferenciable cerca de a.»
Tampoco se sigue. El teorema es puntual: de la derivabilidad en **ese punto** solo deducimos continuidad en **ese punto**.

La segunda advertencia será especialmente importante más adelante. Una derivada en un punto no convierte automáticamente a la función en una función diferenciable sobre un intervalo, ni autoriza todavía conclusiones sobre monotonía, extremos o comportamiento global.

Una lectura secuencial del mismo resultado. También podemos reconstruir el teorema usando sucesiones, sin introducir una prueba nueva esencial. Supongamos que f es diferenciable en a y tomemos cualquier sucesión (x_n) de puntos de A tal que $x_n \to a$. Si algunos términos satisfacen $x_n = a$, entonces para esos índices $f(x_n) - f(a) = 0$.

Para los índices restantes,

$$f(x_n) - f(a) = (x_n - a)\frac{f(x_n) - f(a)}{x_n - a}.$$

Sobre esa parte de la sucesión, el primer factor tiende a 0 y, por el criterio secuencial de diferenciabilidad, el segundo tiende a $f'_A(a)$. Así, los términos no centrales de $f(x_n) - f(a)$ tienden a 0, mientras que los términos centrales son exactamente 0. En conjunto, $f(x_n) - f(a) \to 0$, y así $f(x_n) \to f(a)$. El criterio secuencial de continuidad recupera la misma conclusión. Esta lectura confirma que la prueba no depende de una forma particular de aproximación: funciona para **todo** acercamiento permitido por el dominio.

Qué hemos ganado. Antes de esta sección teníamos dos nociones locales estudiadas por separado. Ahora sabemos que están jerarquizadas:

diferenciabilidad es una condición local más fuerte que continuidad.

La razón estructural puede condensarse en una sola identidad:

$$\boxed{f(x) - f(a) = (x - a)\frac{f(x) - f(a)}{x - a}.}$$

La derivada controla el segundo factor; la proximidad $x \to a$ hace desaparecer el primero. Todavía, sin embargo, seguimos leyendo $f'(a)$ principalmente como un número obtenido de pendientes secantes. En §6.6 convertiremos ese número en un objeto geométrico preciso: la **recta tangente**. La definición será local y no dependerá de cuántas veces una recta intersecte globalmente la gráfica.

6.6. La recta tangente: una definición local, no una figura

Hasta ahora la derivada ha aparecido como un número: el límite de las pendientes de secantes cuando el segundo punto de la gráfica se aproxima al primero. Esa construcción sugiere inmediatamente una recta, pero conviene separar con cuidado la intuición geométrica de la definición matemática.

La pregunta es ahora:

si $f'(a)$ existe, ¿qué recta representa exactamente esa pendiente local?

La respuesta será sencilla en su fórmula, pero importante en su interpretación. La recta tangente no se definirá por «tocar» la gráfica, por quedar de un solo lado de ella ni por tener una única intersección. Se definirá a partir de la información local contenida en la derivada.

De una pendiente local a una recta. Supongamos que $f\colon A \to \mathbb{R}$ es diferenciable en un punto $a \in A$ respecto del dominio A. Entonces ya conocemos dos datos: $P = (a, f(a))$ y $f'_A(a)$. El primero fija un punto del plano. El segundo fija una pendiente real finita. Por la geometría elemental de la recta, existe una única recta que pasa por P y tiene esa pendiente. Su ecuación, en forma punto-pendiente, es $y - f(a) = f'_A(a)(x - a)$.

Equivalentemente, $y = f(a) + f'_A(a)(x-a)$.

Esta observación conduce a la definición.

DEFINICIÓN 6.3. **Recta tangente asociada a una derivada finita.** Sea $f\colon A \to \mathbb{R}$ y sea $a \in A$ un punto de acumulación de A. Supongamos que f es diferenciable en a respecto de A.

La **recta tangente a la gráfica de f en** $(a, f(a))$, respecto de ese dominio, es la recta de ecuación

$$\boxed{T_a(x) = f(a) + f'_A(a)(x-a).}$$

Cuando el dominio está claro por el contexto escribiremos simplemente

$$T_a(x) = f(a) + f'(a)(x-a).$$

La definición conserva exactamente la información que produjo la derivada:

$$\begin{aligned} \text{punto de paso}: &\quad (a, f(a)), \\ \text{pendiente}: &\quad f'_A(a). \end{aligned}$$

No añade ninguna condición global sobre la gráfica.

Por qué esta recta está vinculada a las secantes. En §6.1 escribimos la secante determinada por los puntos

$$(a, f(a)) \quad \text{y} \quad (a+h, f(a+h))$$

como $S_h(x) = f(a) + Q_a(h)(x-a)$, con $Q_a(h) = \frac{f(a+h)-f(a)}{h}$.

Si f es diferenciable en a, entonces $Q_a(h) \longrightarrow f'_A(a)$ para incrementos admisibles $h \to 0$. Por tanto, el coeficiente que cambia en la familia de secantes —su pendiente— converge precisamente a la pendiente de T_a.

Para cualquier x fijo,

$$S_h(x) - T_a(x) = \big(Q_a(h) - f'_A(a)\big)(x-a),$$

y de aquí $S_h(x) \longrightarrow T_a(x)$ cuando $h \to 0$ dentro del dominio admisible. En este sentido preciso, las secantes aproximan a la tangente: sus pendientes convergen y, en cada x fijo, también lo hacen sus valores. La afirmación se refiere a esos controles concretos y no requiere una teoría abstracta de convergencia de rectas.

Qué significa aquí «límite de secantes»
Todas las secantes S_h pasan por el mismo punto $(a, f(a))$. Lo que varía es su pendiente. La derivabilidad afirma exactamente que pendiente de $S_h \longrightarrow f'_A(a)$. La tangente es la recta que pasa por el mismo punto y posee esa pendiente límite.

Una tangente puede cruzar la gráfica. Una imagen escolar frecuente sugiere que una recta tangente «toca» la curva pero no la atraviesa. Esa descripción puede funcionar en algunos dibujos, pero no puede servir como definición. El ejemplo más simple ya está disponible con una función que calculamos desde la definición.

Ejemplo 6.8. **La tangente a $f(x) = x^3$ en 0 cruza la gráfica.** Consideremos $f(x) = x^3$.

En §6.3 obtuvimos desde el cociente incremental que $f'(a) = 3a^2$.

En particular,

$$f(0) = 0, \qquad f'(0) = 0.$$

Por `#def-t1-0044`, la recta tangente en el origen es $T_0(x) = f(0) + f'(0)(x - 0) = 0$. Es decir, la tangente es el eje horizontal $y = 0$.

Pero la función satisface

$$x^3 < 0 \quad \text{si } x < 0,$$

y

$$x^3 > 0 \quad \text{si } x > 0.$$

Por tanto, al atravesar $x = 0$, la gráfica pasa de un lado de la recta $y = 0$ al otro. La recta tangente **cruza** la gráfica en el punto de tangencia. Así queda refutada la regla informal

«una tangente no puede atravesar la curva».

La derivada no contiene ninguna condición que obligue a la gráfica a permanecer de un solo lado de la tangente. Solo determina el comportamiento de primer orden de las pendientes secantes alrededor del punto.

Una tangente puede volver a intersectar la gráfica lejos del punto. También es falsa otra descripción frecuente: «la tangente es una recta que tiene exactamente un punto en común con la gráfica». Consideremos, sin introducir un nuevo objeto numerado, $p(x) = x^2(x - 10)$.

En $a = 0$ tenemos $p(0) = 0$ y, para $h \neq 0$, $\frac{p(h)-p(0)}{h} = \frac{h^2(h-10)}{h} = h(h - 10)$.

Como $h(h - 10) \longrightarrow 0$, se sigue directamente de la definición que $p'(0) = 0$. La tangente en el origen es entonces $y = 0$.

Sin embargo, $p(10) = 10^2(10 - 10) = 0$.

La misma recta tangente vuelve a intersectar la gráfica en $(10, 0)$. Nada de esto modifica la tangencia en el origen. El punto $x = 10$ está fuera de la información que interviene en el límite $h \to 0$. Este ejemplo deja una separación importante:

tangencia en a es una noción local, no un conteo global de intersecciones.

Lo local y lo global no deben confundirse. Para determinar T_a necesitamos solamente $f(a)$ y $f'_A(a)$.

El segundo dato se obtiene observando el cociente incremental para puntos arbitrariamente próximos a a. Por eso cambiar la función lejos de a puede alterar drásticamente la gráfica completa sin cambiar la pendiente tangente en a. Podemos organizar la distinción así:

información relevante para la tangente en a	información que no forma parte de la definición
el valor $f(a)$	cuántas intersecciones hay lejos de a
el límite del cociente incremental en a	si la gráfica queda globalmente arriba o abajo de la recta
la pendiente finita $f'_A(a)$	la forma completa de la gráfica

información relevante para la tangente en a	información que no forma parte de la definición
el comportamiento arbitrariamente próximo a a	lo que ocurre en regiones alejadas del punto

Más adelante formalizaremos con precisión esta idea de **localidad** de la derivada. Por ahora basta observar que la definición de tangente hereda la naturaleza local del límite que define $f'(a)$.

El caso de un extremo del dominio. La definición tampoco exige que a sea interior del dominio.

Supongamos, por ejemplo, que

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

En $a = 0$, los incrementos admisibles son $h > 0$. Desde el dominio,

$$\frac{f(h) - f(0)}{h} = h \longrightarrow 0 \qquad (h \to 0^+).$$

Así, $f'_{[0,\infty)}(0) = 0$, y la recta tangente determinada por nuestra definición es $T_0(x) = 0$. La evidencia diferencial proviene solo del lado disponible del dominio, pero la recta obtenida es una recta ordinaria del plano. No necesitamos extender artificialmente la función a $x < 0$ para definirla.

Pendiente vertical y nuestra definición de tangente. En §6.4 distinguimos un cociente incremental que tiende a infinito de una derivada real finita. Esa decisión tiene ahora una consecuencia geométrica explícita.

`#def-t1-0044` define una recta tangente mediante el número real $f'_A(a)$.

Si el cociente incremental crece sin cota, entonces $f'_A(a)$ no existe bajo nuestra definición y, por tanto, tampoco existe una recta tangente **en el sentido de `#def-t1-0044`**. Geométricamente puede ser útil decir que la gráfica presenta una **tangente vertical** y señalar la recta $x = a$, pero esa es una extensión terminológica distinta. No debemos escribir $f'(a) = \infty$ ni confundir una pendiente vertical con una derivada real.

Tres definiciones falsas de tangente
No adoptaremos ninguna de estas reglas:

1. **«La tangente toca la gráfica pero no la cruza.»** Falso: x^3 en 0.
2. **«La tangente tiene exactamente una intersección con la gráfica.»** Falso: $x^2(x-10)$ en 0 tiene la tangente $y = 0$, que vuelve a cortar la gráfica en $x = 10$.
3. **«Cualquier recta que visualmente parezca ajustarse a la curva es tangente.»** Insuficiente: la pendiente debe ser exactamente la derivada definida mediante el límite del cociente incremental.

La ecuación de la tangente todavía no es toda la historia. Hemos convertido la derivada en una recta precisa:

$$\boxed{T_a(x) = f(a) + f'_A(a)(x - a).}$$

Sabemos además de dónde proviene su pendiente: es el límite de las pendientes secantes. Pero todavía queda una pregunta más profunda.

¿Por qué esta recta no es solo una traducción geométrica del número $f'(a)$, sino un **modelo local de la función**?

Para responder necesitamos comparar directamente $f(a+h)$ con $T_a(a+h) = f(a) + f'_A(a)h$ y medir el error entre ambos a la escala del incremento h. Ese será el objetivo de §6.7. Allí pasaremos de

pendiente tangente

a

aproximación lineal local de primer orden,

sin usar Taylor ni ninguna regla de derivación posterior.

6.7. Diferenciabilidad como aproximación lineal local

La sección anterior definió la recta tangente a partir de la derivada y dejó abierta una pregunta más profunda. Si $T_a(x) = f(a) + f'_A(a)(x-a)$,

¿en qué sentido matemático esta recta **aproxima** realmente a la función cerca de a?

Decir solamente que ambas pasan por $(a, f(a))$ es demasiado débil: infinitas rectas pasan por ese punto. Decir que tienen la misma pendiente local ya es más informativo, pero todavía queremos convertir esa idea en una afirmación directa sobre los valores de f.

La pregunta guía será

¿qué información más profunda contiene el número $f'(a)$ además de una pendiente?

La respuesta será que la derivada determina la **parte lineal principal** del cambio de la función. Después de retirar esa parte lineal, queda un error que se vuelve despreciable frente a la propia escala del incremento.

De la continuidad al primer orden. Supongamos que f es continua en a. Entonces

$$f(a+h) - f(a) \longrightarrow 0 \qquad (h \to 0),$$

con h restringido, como siempre, a incrementos admisibles por el dominio. Esta afirmación dice que la variación total se hace pequeña. Pero no dice **a qué velocidad** se hace pequeña respecto de h.

Por ejemplo, las cantidades

$$h, \qquad h^2, \qquad |h|^{1/2}$$

tienden todas a 0 cuando $h \to 0$. Sin embargo, su tamaño relativo frente a $|h|$ es muy distinto:

$$\frac{|h|}{|h|} = 1, \qquad \frac{|h|^2}{|h|} = |h| \to 0, \qquad \frac{|h|^{1/2}}{|h|} = \frac{1}{|h|^{1/2}} \to +\infty.$$

Así, «el error tiende a 0» no basta para identificar un comportamiento de primer orden. La diferenciabilidad exige algo más fino. Si $L = f'_A(a)$, queremos separar $f(a+h) - f(a)$ en una parte principal Lh y un residuo mucho menor que h.

El residuo después de retirar la parte lineal. Fijemos $L \in \mathbb{R}$ y definamos, para todo incremento admisible h, $r(h) := f(a+h) - f(a) - Lh$. Entonces tenemos la identidad exacta $f(a+h) = f(a) + Lh + r(h)$. No hemos aproximado nada todavía: simplemente hemos nombrado la diferencia entre el valor real de la función y el modelo afín $f(a) + Lh$. La cuestión decisiva es el tamaño de $r(h)$ cuando $h \to 0$. Si solo sabemos que $r(h) \to 0$, el residuo puede seguir siendo del mismo orden que h. Por ejemplo, $r(h) = h$ tiende a 0, pero $\frac{r(h)}{h} = 1$. Para que Lh sea realmente la parte principal de primer orden necesitamos una condición más fuerte: $\frac{r(h)}{h} \longrightarrow 0$. Esto significa que, comparado con el incremento h, el residuo se vuelve proporcionalmente insignificante.

La notación $o(h)$.

Definición 6.4. **Pequeño-o de primer orden.** Sea $H \subseteq \mathbb{R}$ un conjunto para el cual 0 es punto de acumulación, y sea $r\colon H \to \mathbb{R}$ una función definida al menos en un entorno relativo de 0.

Escribimos

$$r(h) = o(h) \qquad (h \to 0,\ h \in H)$$

si

$$\lim_{\substack{h \to 0 \\ h \in H}} \frac{r(h)}{h} = 0.$$

Equivalentemente, para todo $\varepsilon > 0$ existe $\delta > 0$ tal que, para todo $h \in H$,

$$0 < |h| < \delta \quad \Longrightarrow \quad |r(h)| < \varepsilon |h|.$$

Cuando el conjunto de incrementos admisibles está claro por el contexto, abreviaremos simplemente

$$r(h) = o(h).$$

La definición debe leerse como una comparación de escalas:

$$\boxed{r(h) = o(h) \quad \Longleftrightarrow \quad \text{el residuo es despreciable frente a } |h| \text{ cuando } h \to 0.}$$

Qué significa —y qué no significa— $o(h)$
La expresión $o(h)$ **no designa un número especial** ni una función universal llamada «o». En una fórmula como $f(a+h) = f(a) + Lh + o(h)$, la notación afirma que existe un término residual $r(h)$ que satisface $r(h) = o(h)$. Tampoco basta con que $r(h) \to 0$. La condición característica es la comparación $\frac{r(h)}{h} \to 0$.

En este capítulo no desarrollaremos un cálculo asintótico general. Usaremos $o(h)$ únicamente para expresar con precisión la aproximación de primer orden asociada a la derivada.

El teorema de linealización local. Ahora podemos reformular la diferenciabilidad sin hablar primero de cocientes incrementales.

TEOREMA 6.2. **Caracterización de la diferenciabilidad por linealización local.** Sea $f\colon A \to \mathbb{R}$, sea $a \in A$ un punto de acumulación de A y sea $L \in \mathbb{R}$. Denotemos por

$$H_a = \{h \in \mathbb{R} : a + h \in A\}$$

el conjunto de incrementos admisibles.

Son equivalentes:

1. f es diferenciable en a respecto de A y
$$f'_A(a) = L;$$
2. existe un residuo r tal que, para $h \in H_a$ suficientemente próximo a 0,
$$f(a+h) = f(a) + Lh + r(h)$$
y
$$r(h) = o(h) \qquad (h \to 0,\ h \in H_a).$$

Por tanto podemos escribir de forma abreviada

$$\boxed{f'_A(a) = L \quad \Longleftrightarrow \quad f(a+h) = f(a) + Lh + o(h)}$$

cuando $h \to 0$ a través de los incrementos admisibles del dominio.

Demostración. Supongamos primero que $f'_A(a) = L$.

Por la forma incremental de la derivada,

$$\frac{f(a+h) - f(a)}{h} \longrightarrow L \qquad (h \to 0,\ h \in H_a).$$

Definamos $r(h) = f(a+h) - f(a) - Lh$.

Para $h \neq 0$, $\frac{r(h)}{h} = \frac{f(a+h)-f(a)}{h} - L$.

Al tomar el límite, $\frac{r(h)}{h} \longrightarrow L - L = 0$.

Luego $r(h) = o(h)$, y la identidad $f(a+h) = f(a) + Lh + r(h)$ da la linealización buscada.

Recíprocamente, supongamos que $f(a+h) = f(a) + Lh + r(h)$ con $r(h) = o(h)$.

Para $h \neq 0$ podemos dividir por h: $\frac{f(a+h)-f(a)}{h} = L + \frac{r(h)}{h}$.

Como $\frac{r(h)}{h} \longrightarrow 0$, obtenemos $\frac{f(a+h)-f(a)}{h} \longrightarrow L$.

Por la definición de derivada, $f'_A(a) = L$.

Esto prueba ambas implicaciones. □

La misma afirmación escrita con cuantificadores. La notación $o(h)$ es compacta, pero no debe ocultar su contenido lógico. Aplicando directamente `#def-t1-0045`, el teorema anterior dice que $f'_A(a) = L$ si y solo si $\forall \varepsilon > 0\ \exists \delta > 0$ tal que, para todo $h \in H_a$,

$$0 < |h| < \delta \quad \Longrightarrow \quad |f(a+h) - f(a) - Lh| < \varepsilon |h|.$$

Esta desigualdad merece una lectura pausada. El término $f(a+h) - f(a) - Lh$ es el error que queda después de retirar la variación lineal prevista por Lh. La cota $|f(a+h) - f(a) - Lh| < \varepsilon|h|$

dice que podemos hacer ese error menor que **cualquier fracción prescrita de la escala** $|h|$, siempre que observemos suficientemente cerca de a.

No estamos afirmando solamente $|f(a+h) - f(a) - Lh| < \varepsilon$.

Eso sería un control absoluto. La diferenciabilidad proporciona un control más fuerte y escalado por el propio incremento.

Continuidad frente a diferenciabilidad
La continuidad proporciona $f(a+h) - f(a) \longrightarrow 0$, mientras que la diferenciabilidad proporciona la afirmación más fuerte $f(a+h) - f(a) - f'_A(a)h = o(h)$. La segunda afirmación contiene información de escala que la primera no posee.

La tangente aparece ahora como modelo, no solo como pendiente. Si $L = f'_A(a)$, la recta tangente satisface $T_a(a+h) = f(a) + Lh$. Por tanto el residuo es exactamente $r(h) = f(a+h) - T_a(a+h)$. El teorema de linealización se convierte en

$$\boxed{f(a+h) - T_a(a+h) = o(h).}$$

Esta fórmula da un significado preciso a la frase «la tangente aproxima localmente a la gráfica». No estamos midiendo aquí la distancia euclídea mínima desde un punto de la gráfica hasta la recta. Comparamos los valores de ambas expresiones en la **misma coordenada horizontal** $a+h$. La diferencia vertical, dividida por el tamaño de la escala horizontal, tiende a cero:

$$\frac{|f(a+h) - T_a(a+h)|}{|h|} \longrightarrow 0.$$

Así, al observar la función a escalas cada vez menores, la discrepancia entre la gráfica y su modelo tangente se vuelve insignificante frente al propio desplazamiento horizontal. La misma identidad puede leerse desde las pendientes. Como $f(a+h) - f(a) = Lh + r(h)$, para $h \neq 0$ tenemos $\frac{f(a+h)-f(a)}{h} = L + \frac{r(h)}{h}$.

Por tanto $\frac{f(a+h)-f(a)}{h} - L = \frac{r(h)}{h} \longrightarrow 0$.

La interpretación mediante secantes y la interpretación mediante linealización son, pues, dos lecturas de la misma estructura.

Un laboratorio exacto con x^2.

Ejemplo 6.9. **La parábola: el error exacto es h^2.** Consideremos $f(x) = x^2$ y fijemos $a \in \mathbb{R}$.

Ya demostramos directamente desde la definición que $f'(a) = 2a$.

Evaluemos la función en $a+h$:

$$f(a+h) = (a+h)^2 = a^2 + 2ah + h^2.$$

Como $f(a) = a^2$, obtenemos la descomposición exacta

$$\boxed{f(a+h) = f(a) + 2ah + h^2.}$$

La parte lineal es $2ah = f'(a)h$, y el residuo es $r(h) = h^2$.

Ahora $\frac{r(h)}{h} = h \longrightarrow 0$, así que $h^2 = o(h)$.

Por tanto $f(a+h) = f(a) + f'(a)h + o(h)$.

La recta tangente en a es $T_a(x) = a^2 + 2a(x-a)$. Al evaluar en $x = a+h$, $T_a(a+h) = a^2 + 2ah$, de modo que el error vertical es exactamente $f(a+h) - T_a(a+h) = h^2$.

Además,

$$\frac{|f(a+h) - T_a(a+h)|}{|h|} = |h| \longrightarrow 0.$$

La versión cuantificada también es transparente. Dado $\varepsilon > 0$, basta elegir, por ejemplo, $\delta = \varepsilon$. Si $0 < |h| < \delta$, entonces $|h^2| = |h|^2 < \varepsilon|h|$.

Así vemos simultáneamente la derivada, la tangente y el control de primer orden del error.

El ejemplo contiene una advertencia importante. Para x^2 el residuo h^2 es no negativo, de modo que la gráfica queda por encima de su tangente. Esa propiedad particular **no forma parte de la definición de diferenciabilidad** y no debe generalizarse todavía. El estudio sistemático de la forma de las gráficas y de la convexidad pertenece a capítulos posteriores.

Un test conceptual: ¿es pequeño o es de orden menor? Consideremos tres posibles residuos:

$$r_1(h) = h, \qquad r_2(h) = h^2, \qquad r_3(h) = h|h|.$$

Los tres tienden a 0. Pero $\frac{r_1(h)}{h} = 1$, mientras que $\frac{r_2(h)}{h} = h \to 0$ y $\frac{r_3(h)}{h} = |h| \to 0$.

Por tanto

$$r_1(h) \neq o(h), \qquad r_2(h) = o(h), \qquad r_3(h) = o(h).$$

La pregunta correcta no es solamente

> ¿el residuo se hace pequeño?

sino

> ¿se hace pequeño **en comparación con el incremento**?

Esa segunda pregunta es la que detecta el primer orden.

La formulación respeta el dominio. Todo el argumento anterior conserva la convención relativa al dominio introducida en §6.2. Si $f\colon A \to \mathbb{R}$, los incrementos se toman en $H_a = \{h : a + h \in A\}$. Si a es un extremo de un intervalo, quizá solo existan incrementos positivos o solo negativos suficientemente pequeños. La condición $r(h) = o(h)$ se interpreta entonces mediante ese acercamiento unilateral permitido por H_a. No necesitamos extender artificialmente la función fuera de su dominio para hablar de linealización relativa.

Qué hemos ganado al reformular la derivada. La definición original decía $\frac{f(a+h)-f(a)}{h} \longrightarrow f'_A(a)$.

Ahora podemos decir exactamente lo mismo de otra manera:

$$\boxed{f(a+h) = f(a) + f'_A(a)h + o(h).}$$

Esta forma separa tres niveles de información:

$$\begin{array}{rcl} f(a) & : & \text{valor de referencia,} \\ f'_A(a)h & : & \text{variación lineal de primer orden,} \\ o(h) & : & \text{error despreciable frente a la escala } h. \end{array}$$

La derivada deja así de ser solamente la pendiente de una recta y aparece como el coeficiente que organiza el comportamiento local de la función hasta primer orden. Todavía queda una pregunta lógica importante. Si una pendiente L produce un residuo $o(h)$, ¿podría otra pendiente distinta M producir también un residuo $o(h)$?

Responderemos esa pregunta en §6.8. Allí daremos un significado exacto a la afirmación de que la aproximación afín de primer orden es **única**, sin introducir Taylor, mínimos cuadrados ni ninguna teoría posterior.

6.8. Por qué la aproximación de primer orden es única

En §6.7 vimos que la diferenciabilidad en a puede expresarse mediante una descomposición

$$f(a+h) = f(a) + Lh + r(h), \qquad r(h) = o(h).$$

La parte Lh recoge la variación de primer orden y el residuo es despreciable frente a h. Pero esta formulación deja una pregunta decisiva:

> si una pendiente L produce un error $o(h)$, ¿podría otra pendiente distinta M producir también un error $o(h)$?

Si la respuesta fuese afirmativa, la idea de «aproximación lineal de primer orden» sería ambigua. Una misma función admitiría varias pendientes igualmente válidas a esa escala. La respuesta es negativa. La pendiente de primer orden está forzada por la función.

Dos modelos candidatos no pueden sobrevivir a la misma escala. Supongamos que, para los incrementos admisibles h alrededor de a, tenemos simultáneamente $f(a+h) = f(a) + Lh + r_L(h)$ y $f(a+h) = f(a) + Mh + r_M(h)$, con

$$r_L(h) = o(h) \qquad \text{y} \qquad r_M(h) = o(h).$$

Las dos expresiones describen el mismo número $f(a+h)$. Por tanto podemos restarlas: $Lh + r_L(h) = Mh + r_M(h)$, de donde $(L-M)h = r_M(h) - r_L(h)$.

Mientras $h \neq 0$, dividimos por h: $L - M = \frac{r_M(h)}{h} - \frac{r_L(h)}{h}$.

Ahora aparece el punto clave. Los dos cocientes de la derecha tienden a 0. Por álgebra de límites, $L - M = 0$.

Así, $L = M$.

No queda ninguna libertad en el coeficiente lineal.

Teorema 6.3. **Unicidad del modelo afín de primer orden.** Sea $f\colon A \to \mathbb{R}$, sea $a \in A$ un punto de acumulación de A y consideremos los incrementos admisibles

$$H_a = \{h \in \mathbb{R} : a + h \in A\}.$$

Supongamos que existen $L, M \in \mathbb{R}$ y residuos r_L, r_M tales que, para $h \in H_a$ suficientemente próximo a 0,

$$f(a+h) = f(a) + Lh + r_L(h), \qquad r_L(h) = o(h),$$

y

$$f(a+h) = f(a) + Mh + r_M(h), \qquad r_M(h) = o(h).$$

Entonces

$$L = M.$$

En particular, si f es diferenciable en a, el único modelo afín de la forma

$$x \longmapsto f(a) + m(x-a)$$

cuyo error respecto de f es $o(x-a)$ cuando $x \to a$ dentro de A es

$$T_a(x) = f(a) + f'_A(a)(x-a).$$

Demostración. Restando las dos representaciones obtenemos

$$(L-M)h = r_M(h) - r_L(h).$$

Para todo incremento admisible no nulo suficientemente pequeño,

$$L - M = \frac{r_M(h)}{h} - \frac{r_L(h)}{h}.$$

Como

$$\frac{r_L(h)}{h} \longrightarrow 0 \qquad \text{y} \qquad \frac{r_M(h)}{h} \longrightarrow 0,$$

el miembro derecho tiende a 0. El miembro izquierdo es constante, así que necesariamente

$$L - M = 0.$$

Por tanto $L = M$. □

La prueba contiene una idea que conviene retener: **si elegimos una pendiente incorrecta, el error conserva inevitablemente una parte de tamaño proporcional a** h. Esa parte no puede esconderse dentro de un término $o(h)$.

Qué significa aquí «mejor aproximación». Ya podemos precisar una expresión frecuente del cálculo: «la tangente es la mejor aproximación lineal de la función cerca del punto». En este capítulo, la palabra **mejor** no significa que hayamos minimizado una distancia global ni una suma de cuadrados. Tampoco significa que la tangente sea la recta que visualmente parezca ajustarse mejor a una gráfica.

Significa algo exacto:

la tangente es la única recta por $(a, f(a))$ cuyo error es $o(\lvert x-a\rvert)$.

Equivalentemente, entre todas las rectas $P_m(x) = f(a) + m(x-a)$, solo una puede satisfacer

$$\frac{f(x) - P_m(x)}{x-a} \longrightarrow 0 \qquad (x \to a,\ x \in A).$$

Por `#thm-t1-0026`, su pendiente es necesariamente $m = f'_A(a)$.

El sentido preciso de mejor
No comparamos rectas en todo el dominio. No medimos un error total sobre un intervalo. No resolvemos un problema de optimización. Comparamos el **orden local del error** cuando $x \to a$. La tangente es privilegiada porque elimina toda componente lineal restante del error.

Qué ocurre si forzamos una pendiente equivocada. Volvamos a la parábola $f(x) = x^2$ alrededor de a. Sabemos exactamente que $f(a+h) = a^2 + 2ah + h^2$. La pendiente correcta es $2a$. Supongamos, en cambio, que intentamos usar $m = 2a + c$, con $c \neq 0$.

El modelo propuesto sería $a^2 + (2a + c)h$.

Su residuo exacto es

$$\begin{aligned} r_c(h) &= f(a+h) - a^2 - (2a+c)h \\ &= h^2 - ch. \end{aligned}$$

Este error sí tiende a 0 cuando $h \to 0$. Pero eso no basta. Dividiendo por h, $\frac{r_c(h)}{h} = h - c \longrightarrow -c \neq 0$.

Por tanto $r_c(h) \neq o(h)$.

La pendiente equivocada deja una huella lineal $-ch$ que sobrevive exactamente a la escala de primer orden. Este cálculo muestra por qué la condición $o(h)$ es mucho más discriminante que pedir simplemente que el error tienda a cero.

Mismo primer orden, errores diferentes. La unicidad del coeficiente lineal no significa que dos funciones con el mismo primer orden deban ser iguales. Pueden diferir en términos que sean pequeños frente a h.

Ejemplo 6.10. **Dos funciones indistinguibles a primer orden, con residuos distintos.** Consideremos $f(x) = x + x^2$ y $g(x) = x + x|x|$.

En $a = 0$ ambas toman el mismo valor: $f(0) = g(0) = 0$.

Además, para un incremento h, $f(h) = h + h^2$ y $g(h) = h + h|h|$.

Ahora $\frac{h^2}{h} = h \longrightarrow 0$ y $\frac{h|h|}{h} = |h| \longrightarrow 0$.

Por tanto

$$h^2 = o(h) \qquad \text{y} \qquad h|h| = o(h).$$

Las dos funciones admiten así la misma linealización de primer orden: $f(h) = h + o(h)$,

$$g(h) = h + o(h).$$

Por `#thm-t1-0025`, ambas son diferenciables en 0 y $f'(0) = g'(0) = 1$. Sin embargo, sus residuos no son iguales: $h^2 \neq h|h|$ para $h < 0$. La derivada detecta el comportamiento de primer orden, pero no registra toda la información de escalas menores.

Este ejemplo permite formular una consecuencia general sin introducir nueva notación formal. Si $f, g : A \to \mathbb{R}$ son diferenciables en a respecto de un mismo dominio A, con a punto de acumulación de A, y

$$f(a) = g(a), \qquad f'(a) = g'(a) = L,$$

entonces podemos escribir $f(a+h) = f(a) + Lh + r_f(h)$ y $g(a+h) = g(a) + Lh + r_g(h)$, con

$$\frac{r_f(h)}{h} \to 0, \qquad \frac{r_g(h)}{h} \to 0.$$

Restando,

$$f(a+h) - g(a+h) = r_f(h) - r_g(h).$$

Y al dividir por h,

$$\frac{f(a+h) - g(a+h)}{h} = \frac{r_f(h)}{h} - \frac{r_g(h)}{h} \longrightarrow 0.$$

En este sentido preciso, dos funciones con el mismo valor y la misma derivada en un punto son **indistinguibles a primer orden** alrededor de ese punto. Eso no significa que sean iguales cerca de a. Significa que su diferencia es demasiado pequeña, comparada con $|h|$, para ser detectada en la escala lineal.

El residuo puede contener información que todavía no estamos clasificando. En `#exm-t1-0071` aparecieron dos residuos distintos:

$$h^2 \qquad \text{y} \qquad h|h|.$$

Ambos son $o(h)$. Para la teoría de este capítulo eso es suficiente: los dos pertenecen a una escala menor que la lineal. Podríamos intentar distinguir entre residuos cuadráticos, cúbicos o de otros órdenes. Esa pregunta es legítima, pero todavía no corresponde desarrollarla sistemáticamente. Las derivadas superiores y la fórmula de Taylor aparecerán más adelante. Aquí la frontera conceptual es deliberada:

$$\boxed{\text{primer orden} = \text{parte lineal única} + \text{resto } o(h).}$$

No necesitamos clasificar más finamente el resto para comprender qué significa ser diferenciable.

Una aproximación numérica que no usa Taylor. La linealización puede producir aproximaciones numéricas aun antes de disponer de una teoría general de Taylor. Pero debemos distinguir con cuidado entre una aproximación respaldada por una identidad exacta y una regla general de estimación del error. Para $f(x) = x^2$ alrededor de $a = 3$, la identidad exacta es $(3+h)^2 = 9 + 6h + h^2$. La parte de primer orden es $9 + 6h$. Si queremos aproximar $(3.01)^2$, tomamos $h = 0.01$ y obtenemos $9 + 6(0.01) = 9.06$. La identidad exacta muestra que el error es $h^2 = 0.0001$, de modo que $(3.01)^2 = 9.0601$. Aquí no hemos usado Taylor ni el teorema del valor medio. Conocíamos de antemano el residuo exacto h^2.

En una función diferenciable arbitraria, la información $f(a+h) = f(a) + f'(a)h + o(h)$ garantiza que el error es pequeño **relativamente a** $|h|$, pero por sí sola no proporciona una fórmula numérica universal para ese error. Obtener cotas más explícitas requerirá herramientas adicionales.

Qué hemos demostrado realmente. La cadena lógica de §§6.7–6.8 puede resumirse así:

$$\boxed{\begin{array}{c} f \text{ diferenciable en } a \\ \Downarrow \\ f(a+h) = f(a) + f'_A(a)h + o(h) \\ \Downarrow \\ \text{la parte lineal de primer orden es única.} \end{array}}$$

Por eso la derivada no es solo **una** pendiente compatible con el comportamiento local. Es **la única pendiente** que deja un error despreciable frente al incremento. La noción de tangente, la derivada y la linealización quedan ahora unificadas:

$$\boxed{\text{derivada} \iff \text{pendiente tangente} \iff \text{coeficiente lineal único de primer orden.}}$$

Podemos reunir los valores de las derivadas en una función definida sobre los puntos donde existen. La dependencia del comportamiento próximo a cada centro permite además precisar en qué sentido la derivada es local.

6.9. La derivada como función y la localidad del concepto

Hasta ahora hemos usado la notación $f'(a)$ para describir un número asociado a un punto concreto. Pero, si repetimos la misma pregunta en muchos puntos del dominio, obtenemos algo nuevo: una función cuyos valores son precisamente esas derivadas. Esta transición parece pequeña, $a \longmapsto f'(a)$, pero cambia la escala de nuestras preguntas. Ya no preguntamos solamente

> ¿cuál es la derivada de f en este punto?

sino también

> ¿en qué puntos existe la derivada y cómo se comportan entre sí los valores obtenidos?

La primera pregunta sigue siendo puntual. La segunda empieza a mirar la derivada como un objeto funcional.

Del valor $f'(a)$ a la función f'. En §6.2 definimos el conjunto

$$\mathrm{Dom}(f') = \{a \in A : a \text{ es punto de acumulación de } A \text{ y } f'_A(a) \text{ existe como número real}\}.$$

Cuando este conjunto no es vacío, podemos reunir todos los valores de la derivada en la aplicación

$$f' \colon \mathrm{Dom}(f') \to \mathbb{R}, \qquad a \longmapsto f'_A(a).$$

Conviene detenerse en el dominio. En general, $\mathrm{Dom}(f') \subseteq A$, y la inclusión puede ser estricta. Que f esté definida en un punto no significa que sea diferenciable allí. El ejemplo de $|x|$ en 0 ya mostró exactamente esa diferencia. Por eso hay tres objetos que no debemos confundir:

$$\boxed{A, \qquad \mathrm{Dom}(f'), \qquad f'.}$$

El primero es el dominio de la función original. El segundo selecciona los puntos donde existe una derivada real finita según nuestra definición. El tercero asigna a cada uno de esos puntos el número correspondiente. Cuando digamos que f es **diferenciable en un conjunto** $E \subseteq A$,

entenderemos simplemente que $E \subseteq \mathrm{Dom}(f')$. Esta convención no añade una nueva teoría: abrevia una afirmación punto por punto.

La derivada solo mira lo que ocurre cerca del punto. La definición de derivada está escrita mediante un límite. Y los límites poseen una propiedad fundamental que ya conocemos: cambiar una función lejos del punto de estudio no altera su límite en ese punto. La misma idea debe heredarse por la derivada. Supongamos que queremos decidir qué ocurre en a. El cociente incremental solo utiliza $f(a)$ y valores $f(x)$ para x arbitrariamente próximo a a. Por tanto, lo que haga la función lejos de a no debería importar. Esta intuición se formaliza en la siguiente proposición.

Proposición 6.4. **Localidad de la derivada.** Sean

$$f, g\colon A \to \mathbb{R}$$

y sea $a \in A$ un punto de acumulación de A. Supongamos que existe $\delta_0 > 0$ tal que

$$x \in A, \qquad |x-a| < \delta_0 \Longrightarrow f(x) = g(x).$$

Entonces

$$f \text{ es diferenciable en } a \iff g \text{ es diferenciable en } a.$$

Cuando estas condiciones se cumplen,

$$f'_A(a) = g'_A(a).$$

Demostración. Como $a \in A$ y $|a-a| = 0 < \delta_0$, la hipótesis implica primero que

$$f(a) = g(a).$$

Además, si

$$x \in A, \qquad 0 < |x-a| < \delta_0,$$

entonces

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

Por tanto, en ese entorno perforado relativo al dominio,

$$\frac{f(x)-f(a)}{x-a} = \frac{g(x)-g(a)}{x-a}.$$

Los dos cocientes incrementales coinciden para todos los puntos de A suficientemente próximos a a, salvo el propio centro. Por la localidad del límite, uno de estos cocientes posee un límite real cuando $x \to a$ dentro de A si y solo si el otro lo posee, y en tal caso ambos límites son iguales.

Así,

$$f'_A(a) = \lim_{\substack{x\to a\\ x\in A}} \frac{f(x)-f(a)}{x-a} = \lim_{\substack{x\to a\\ x\in A}} \frac{g(x)-g(a)}{x-a} = g'_A(a).$$

Esto prueba simultáneamente la equivalencia de la diferenciabilidad y la igualdad de las derivadas. □

La arquitectura de la prueba es corta porque toda la carga conceptual ya estaba contenida en la teoría de límites:

$$\boxed{\text{coincidencia local de } f \text{ y } g \Longrightarrow \text{coincidencia local de cocientes} \Longrightarrow \text{mismo límite.}}$$

La derivada es, por tanto, una propiedad genuinamente **local**.

Local no significa «solo importa el entorno perforado». Hay una sutileza importante en la proposición anterior. Pedimos que f y g coincidan en un entorno de a **incluido el valor en el centro**. No basta con que coincidan para $x \neq a$.

Consideremos $f(x) = x$ para todo $x \in \mathbb{R}$ y definamos

$$g(x) = \begin{cases} x, & x \neq 0, \\ 1, & x = 0. \end{cases}$$

Las funciones coinciden en todo punto distinto de 0. Sin embargo,

$$f(0) = 0 \qquad \text{y} \qquad g(0) = 1.$$

Para f, el cociente incremental en 0 es

$$\frac{f(h) - f(0)}{h} = \frac{h}{h} = 1, \qquad h \neq 0,$$

de modo que $f'(0) = 1$.

Para g,

$$\frac{g(h) - g(0)}{h} = \frac{h - 1}{h} = 1 - \frac{1}{h}, \qquad h \neq 0,$$

y esta expresión no posee límite real cuando $h \to 0$. Así, modificar solamente el valor central puede destruir la diferenciabilidad. La palabra **local** significa que solo importa lo que ocurre en algún entorno suficientemente pequeño del punto; no significa que podamos olvidar el propio valor $f(a)$.

Un uso inmediato: reconstruir la derivada de $|x|$ fuera de la esquina. La localidad permite obtener información sin desarrollar todavía reglas generales de derivación.

Consideremos $f(x) = |x|$.

Ya sabemos que f no es diferenciable en 0.

Tomemos ahora un punto $a > 0$.

Podemos elegir, por ejemplo, $\delta = \frac{a}{2}$.

Si $|x - a| < \delta$, entonces

$$x > a - \frac{a}{2} = \frac{a}{2} > 0.$$

En ese entorno, $|x| = x$.

Por localidad, $|x|$ tiene en a la misma derivada que la función afín $x \mapsto x$. Como esta última tiene derivada 1 desde la definición,

$$f'(a) = 1 \qquad (a > 0).$$

Análogamente, si $a < 0$, podemos elegir un entorno suficientemente pequeño contenido en $(-\infty, 0)$. Allí $|x| = -x$, y por localidad

$$f'(a) = -1 \qquad (a < 0).$$

Junto con la no diferenciabilidad en 0, obtenemos $\mathrm{Dom}(f') = \mathbb{R} \setminus \{0\}$ y

$$f'(x) = \begin{cases} -1, & x < 0, \\ 1, & x > 0. \end{cases}$$

No hemos usado una regla de derivación para el valor absoluto. Hemos usado únicamente:

1. derivadas de funciones afines ya calculadas desde la definición;
2. la igualdad local de las funciones;
3. `#prp-t1-0038`.

Este ejemplo muestra para qué sirve pensar en la derivada como función: podemos preguntar simultáneamente por su dominio y por sus valores.

«f es diferenciable» y «f' es continua» son afirmaciones distintas. Una vez que aparece la función derivada, surge una tentación natural: mezclar la existencia de f' con la continuidad de f'.

Son preguntas diferentes.

Afirmar que f es diferenciable en un punto a significa estudiar el límite

$$\lim_{\substack{x \to a \\ x \in A}} \frac{f(x) - f(a)}{x - a}.$$

En cambio, preguntar si f' **es continua en** a significa aplicar la definición de continuidad relativa al dominio $\mathrm{Dom}(f')$. Cuando $a \in \mathrm{Dom}(f')$ es además punto de acumulación de ese dominio, esto equivale a estudiar otro límite:

$$\lim_{\substack{x \to a \\ x \in \mathrm{Dom}(f')}} f'(x)$$

y compararlo con $f'(a)$.

Las dos preguntas operan en niveles distintos:

f diferenciable en a	:	comportamiento de f a primer orden en a,
f' continua en a	:	comportamiento de las derivadas en puntos próximos a a.

Que $f'(a)$ exista nos entrega **un valor** de la función derivada. No establece por sí solo cómo se relaciona ese valor con $f'(x)$ cuando x varía. En este capítulo no supondremos ni utilizaremos que la existencia de la derivada implique continuidad de la función derivada. Tampoco introduciremos todavía una clasificación sistemática de funciones según la regularidad de f'. Eso pertenece a una capa posterior de la teoría.

Dos errores de nivel
De $f'(a) > 0$ no podemos concluir que $f'(x) > 0$ para todos los x cercanos a a. La primera afirmación habla de **un solo valor** de la función derivada. Para transferir información de $f'(a)$ a valores vecinos de f' necesitaríamos hipótesis adicionales. Y aun una afirmación sobre el signo de f' en muchos puntos no debe convertirse todavía, sin prueba, en un teorema global de monotonía. Esa conexión será establecida más adelante mediante los teoremas del valor medio.

Esta distinción es importante porque evita convertir la notación f' en una caja negra. Primero existe una derivada punto por punto; solo después podemos estudiar propiedades de la nueva función formada por esos valores.

La localidad también fija una estrategia de cálculo. La proposición de localidad no es solamente un dato teórico. Sugiere una estrategia reutilizable. Para calcular una derivada en a, podemos preguntar:

> ¿coincide la función, en algún entorno de a, con otra función cuyo comportamiento diferencial ya conocemos?

Si la respuesta es afirmativa, no necesitamos analizar la fórmula global. Por ejemplo, una función por tramos puede tener una expresión complicada lejos de a y, sin embargo, coincidir cerca de a con una función afín. En tal caso, la derivada en a queda determinada por ese tramo local. Esto explica por qué una modificación remota de la gráfica no cambia la tangente ni la linealización en el punto. Podemos resumir la idea como

la derivada en a depende del germen local de la función en a, no de su comportamiento global.

No necesitamos formalizar aquí la palabra *germen*. La frase expresa solamente la propiedad probada en `#prp-t1-0038`: dos funciones que coinciden en un entorno del punto son indistinguibles para la derivada en ese punto.

Qué queda fuera de esta sección. La aparición de la función f' abre varias preguntas que deliberadamente no resolveremos todavía. En `T1-C08` desarrollaremos las reglas que permiten construir nuevas derivadas a partir de derivadas conocidas:

- suma y escalares;
- producto;
- cociente;
- regla de la cadena.

En `T1-C09` estudiaremos sistemáticamente derivadas de funciones elementales, inversas e implícitas. En `T1-C10` aparecerán Rolle y los teoremas del valor medio, que conectarán información sobre la derivada con el comportamiento de la función a lo largo de un intervalo. Y las derivadas superiores pertenecen a capítulos posteriores. La frontera actual puede escribirse así:

C07: qué es $f'(a)$, qué información local contiene y cómo se organiza en f'
$\Downarrow$
C08–C10: cómo calcular f' sistemáticamente y qué consecuencias globales produce.

Cierre conceptual. Durante las primeras secciones del capítulo, la derivada apareció como un límite de pendientes secantes. Después se convirtió en pendiente tangente y, finalmente, en el coeficiente único de la aproximación de primer orden. Ahora añadimos una última perspectiva:

f' es una nueva función, pero cada uno de sus valores se determina localmente.

La cadena conceptual queda ampliada:

cociente incremental
$\to$ derivada en un punto
$\to$ tangente
$\to$ linealización
$\to$ unicidad de primer orden
$\to$ función derivada y localidad.

Con esto la teoría básica de C07 está conceptualmente construida. Antes de pasar al banco de ejercicios, §6.10 funcionará como laboratorio de reconstrucción: no añadirá teoría nueva, sino que obligará a reconocer qué razonamientos producen realmente cada una de las ideas desarrolladas en el capítulo.

6.10. Laboratorio: reconstruir el pensamiento diferencial

La teoría básica de este capítulo ya está construida. Sabemos pasar de secantes a cocientes incrementales, definir la derivada como límite, diagnosticar fallos laterales, obtener la tangente, expresar la diferenciabilidad como aproximación de primer orden y reconocer que la derivada es local. Pero conocer una lista de resultados no garantiza todavía que sepamos **reconstruir el pensamiento que los produce**. Ése es el objetivo de este laboratorio. No introduciremos teoría nueva. Tampoco añadiremos fórmulas para memorizar. Trabajaremos, en cambio, sobre ocho movimientos cognitivos que deberían poder recuperarse incluso si olvidáramos temporalmente los enunciados del capítulo. La pregunta que gobernará toda la sección es:

si desaparecieran las fórmulas de nuestras notas, ¿podríamos volver a construir la idea de derivada?

Cada estación tendrá tres momentos:

1. **reconstrucción**: partir de una pregunta y producir el objeto matemático necesario;
2. **auditoría**: localizar el paso lógico delicado;
3. **autoexplicación**: formular con palabras propias qué se ha aprendido y qué no se ha demostrado.

La meta no es rapidez. La meta es que el mecanismo conceptual quede disponible para problemas nuevos.

Estación 1 — Reconstruir la derivada desde una familia de secantes. Supongamos que olvidamos por un momento la definición formal de derivada. Conservemos únicamente una pregunta geométrica:

¿cómo describir la inclinación local de la gráfica de f en el punto $(a, f(a))$?

Con un solo punto no podemos calcular una pendiente. Necesitamos un segundo punto de la gráfica. Tomemos

$$(a+h, f(a+h)), \qquad h \neq 0.$$

La pendiente de la secante que une ambos puntos es $\frac{f(a+h)-f(a)}{h}$. Hasta aquí no hay derivada. Hay una **familia de pendientes**, una para cada incremento admisible $h \neq 0$. El siguiente movimiento intelectual consiste en reducir la escala. Queremos que el segundo punto se acerque al primero, pero sin hacerlo coincidir con él mientras calculamos la pendiente. Por eso preguntamos por el comportamiento de $\frac{f(a+h)-f(a)}{h}$ cuando $h \to 0$. Si las pendientes secantes se estabilizan alrededor de un único número real L, ese número contiene precisamente la información local que buscábamos. La reconstrucción completa es entonces

dos puntos de la gráfica
⇓
pendiente secante
⇓
segundo punto se aproxima al primero
⇓
límite de pendientes
⇓
derivada, si el límite real existe.

Obsérvese que la definición no aparece como una convención arbitraria. Es la respuesta rigurosa a un problema previo.

Autoexplicación
Sin mirar la definición de §6.2, intenta responder:

- ¿por qué aparecen **dos** valores de la función en el numerador?;
- ¿por qué dividimos por el incremento de entrada?;
- ¿por qué h debe ser distinto de 0 durante el cálculo?;
- ¿por qué finalmente hacemos $h \to 0$?;
- ¿por qué exigimos un límite **real finito**?

Si puedes reconstruir esas cinco respuestas, la fórmula deja de ser una pieza aislada de memoria.

La misma arquitectura puede leerse sin geometría. Si f representa una magnitud que cambia con el tiempo, el mismo cociente mide una tasa media y el límite busca una tasa instantánea. El objeto matemático es el mismo; cambia únicamente la interpretación.

Estación 2 — Separar continuidad, derivabilidad y derivadas laterales. Uno de los errores más frecuentes consiste en tratar continuidad, derivabilidad y existencia de derivadas laterales como si fueran tres nombres para una misma propiedad.

No lo son.

La continuidad en a pregunta si

$$f(x) \to f(a) \qquad (x \to a).$$

La derivabilidad pregunta si el cambio de la función, después de dividirse por la escala $x-a$, posee un límite real finito: $\frac{f(x)-f(a)}{x-a} \longrightarrow L$. Las derivadas laterales, cuando el dominio acumula por ambos lados, separan ese segundo problema en dos procesos:

$$x \to a^- \qquad \text{y} \qquad x \to a^+.$$

La relación lógica correcta puede organizarse así:

$$\boxed{\text{derivabilidad} \Longrightarrow \text{continuidad},}$$

pero la conversa es falsa. Además, en un punto interior con aproximación desde ambos lados,

$$\boxed{\text{derivabilidad bilateral} \Longleftrightarrow \text{derivadas laterales finitas existentes e iguales.}}$$

El valor absoluto en 0 sigue siendo el laboratorio más económico. La función $f(x)=|x|$ es continua en 0, porque $|x| \to 0$. Sin embargo, los cocientes laterales son

$$\frac{|h|}{h} = \begin{cases} -1, & h<0, \\ 1, & h>0. \end{cases}$$

Cada lado tiene un límite real, pero los dos límites son distintos. Por tanto la continuidad no basta para fabricar una derivada. Podemos comparar varios mecanismos ya vistos:

Comportamiento en el punto	Continuidad	Laterales del cociente	Derivada real bilateral
x^2 en 0	sí	existen y coinciden	sí
$\|x\|$ en 0	sí	existen, pero difieren	no
$x\sigma(1/x)$ en 0, con valor 0	sí	el cociente oscila	no
pendiente que crece sin cota	puede haber continuidad	no hay límite lateral real finito	no

La tabla no sustituye las definiciones. Su función es mostrar que distintos fallos deben diagnosticarse con distintas preguntas.

Orden de diagnóstico
Ante un punto problemático, no preguntes primero «¿es derivable?». Descompón el problema:

1. ¿el punto pertenece al dominio y es punto de acumulación?;
2. ¿la función es continua allí?;
3. ¿qué ocurre con el cociente por cada lado disponible?;
4. ¿los límites laterales son reales finitos?;
5. si existen ambos, ¿coinciden?

Este orden evita confundir discontinuidad, esquina, pendiente no acotada y oscilación.

En un extremo del dominio, la auditoría cambia: solo exigimos la dirección desde la cual el dominio realmente se aproxima al punto. La geometría del dominio forma parte del problema.

Estación 3 — Por qué no se evalúa el cociente incremental en $h = 0$. Consideremos otra vez el cociente $Q(h) = \frac{f(a+h)-f(a)}{h}$. El impulso de «sustituir $h = 0$» produce $\frac{f(a)-f(a)}{0}$, que no está definido. El error no consiste solamente en haber obtenido una división por cero. Es más profundo: se está confundiendo **evaluación** con **límite**. La derivada no pregunta por el valor de Q en 0. Pregunta por el comportamiento de $Q(h)$ para incrementos no nulos arbitrariamente pequeños. La distinción es exactamente la misma que ya utilizamos en teoría de límites:

valor en el centro $\neq$ comportamiento cuando nos aproximamos al centro.

Por eso una expresión puede no estar definida en $h = 0$ y, sin embargo, poseer un límite perfectamente determinado cuando $h \to 0$.

Para la parábola,

$$Q(h) = \frac{(a+h)^2 - a^2}{h}, \qquad h \neq 0.$$

No necesitamos asignar ningún valor al cociente original en $h = 0$. Toda la información relevante vive en el entorno perforado.

Diagnóstico del error
La frase

> «para hallar la derivada pongo $h = 0$»

es incorrecta.
La frase correcta es

> «estudio valores con $h \neq 0$ y determino si el cociente posee un límite cuando h se aproxima a 0».

La diferencia entre ambas frases contiene buena parte de la lógica del cálculo.

Estación 4 — Cancelar h: qué estamos haciendo realmente. En el ejemplo de x^2 aparece $\frac{(a+h)^2-a^2}{h} = \frac{2ah+h^2}{h}$.

Factorizamos: $\frac{h(2a+h)}{h}$.

Y cancelamos h para obtener $2a + h$.

¿Por qué es legítima la cancelación si después queremos hacer $h \to 0$?

Porque durante la cancelación estamos trabajando bajo la hipótesis $h \neq 0$.

La igualdad exacta es

$$\frac{(a+h)^2 - a^2}{h} = 2a + h \qquad \text{para todo } h \neq 0.$$

No hemos demostrado que las dos expresiones sean la misma función en $h = 0$. De hecho, la primera ni siquiera está definida allí. Lo que hemos demostrado es algo suficiente para los límites: **coinciden en un entorno perforado de** 0. Como el límite solo inspecciona valores arbitrariamente próximos al centro y no exige usar el valor central, podemos reemplazar una expresión por la otra al calcular el límite.

La cadena correcta es

$h \neq 0$ porque estamos en el dominio perforado
$\Downarrow$
cancelación algebraica legítima
$\Downarrow$
$Q(h) = 2a + h$ para $h \neq 0$ cercano a 0
$\Downarrow$
$\lim_{h\to 0} Q(h) = 2a.$

El límite no vuelve retroactivamente ilegal la cancelación. Al contrario: la teoría de límites está diseñada precisamente para permitir que trabajemos en dominios perforados.

Pregunta de control
¿Qué afirmación es más precisa?

1. «Cancelamos h y luego ponemos $h = 0$.»
2. «Para $h \neq 0$ simplificamos el cociente; después calculamos el límite de la expresión equivalente en el entorno perforado.»

La segunda. En ella no aparece ninguna división por cero ni ninguna sustitución ilegítima.

Estación 5 — La simetría no puede sustituir la derivada ordinaria. Alguien podría proponer un atajo. En vez de estudiar $\frac{f(a+h)-f(a)}{h}$, podría comparar puntos simétricos alrededor de a y formar $\frac{f(a+h)-f(a-h)}{2h}$. Esta expresión puede ser útil en otros contextos, pero **no puede reemplazar nuestra definición de derivada**. El valor absoluto muestra por qué.

Tomemos $f(x) = |x|$ en $a = 0$. Para $h \neq 0$, $\frac{f(h)-f(-h)}{2h} = \frac{|h|-|-h|}{2h} = 0$.

Así, el cociente simétrico tiene el límite 0. Sin embargo, ya sabemos que la derivada ordinaria de $|x|$ en 0 no existe, porque

$$\lim_{h\to 0^-} \frac{|h|}{h} = -1 \qquad \text{y} \qquad \lim_{h\to 0^+} \frac{|h|}{h} = 1.$$

¿Qué ocurrió?

El cociente simétrico combinó información de ambos lados de una manera que hizo desaparecer la incompatibilidad. Las dos pendientes laterales -1 y 1 se compensaron en la comparación simétrica. Pero la derivada ordinaria exige algo más fuerte: que **cada** cociente incremental, para todos los incrementos admisibles suficientemente pequeños, se aproxime al mismo número.

Por eso

estabilización de un promedio simétrico $\not\Longrightarrow$ existencia de la derivada ordinaria.

Prueba de estrés de una definición
Cuando aparezca una propuesta alternativa para definir un concepto, no preguntes solo si funciona en funciones suaves. Pruébala contra un caso límite conocido. Aquí $|x|$ cumple esa función: conserva suficiente simetría para engañar al cociente propuesto, pero mantiene una esquina que la derivada ordinaria debe detectar.

Este ejercicio de diagnóstico es más general que el caso concreto: una definición adecuada debe distinguir los fenómenos que la teoría pretende distinguir.

Estación 6 — ¿Qué dice exactamente $r(h) = o(h)$? La notación

$$r(h) = o(h) \qquad (h \to 0)$$

puede parecer una etiqueta sofisticada si se memoriza sin abrirla. Su contenido es únicamente $\frac{r(h)}{h} \to 0$. La idea no es simplemente que $r(h) \to 0$. Eso sería demasiado débil. También $r(h) = 3h$ tiende a 0, pero $\frac{r(h)}{h} = 3$, por lo que $3h$ **no** es $o(h)$. La condición $o(h)$ compara dos escalas. Dice que el error $r(h)$ se vuelve despreciable **respecto de la escala lineal** h.

Podemos auditar algunos casos:

$$r(h) = h^2 \quad \Longrightarrow \quad \frac{r(h)}{h} = h \to 0,$$

así que $h^2 = o(h)$.

También

$$r(h) = h|h| \quad \Longrightarrow \quad \frac{r(h)}{h} = |h| \to 0,$$

por lo que $h|h| = o(h)$.

En cambio, $r(h) = |h|$ no satisface la condición bilateral, porque

$$\frac{|h|}{h} = \begin{cases} -1, & h < 0, \\ 1, & h > 0. \end{cases}$$

La pregunta correcta ante cualquier candidato es, por tanto:

$$\boxed{\text{¿qué ocurre con } \frac{r(h)}{h} \text{ cuando } h \to 0?}$$

Esta lectura permite recuperar la linealización sin memorizarla. Si f es diferenciable en a con derivada L, la afirmación $f(a+h) = f(a)+Lh+o(h)$ dice que, después de restar el valor central y el término lineal principal, lo que queda es pequeño **en comparación con** h. Equivalentemente, para tolerancias arbitrarias, $|f(a+h) - f(a) - Lh| < \varepsilon|h|$ cuando h es suficientemente pequeño y admisible.

Autoexplicación

Completa mentalmente la frase:

«El error tiende a cero» no basta para ser de orden menor que h, porque...

La respuesta esencial es: **debemos comparar el error con h dividiendo por h.** Solo si esa razón tiende a 0 el error es despreciable a escala lineal.

Estación 7 — Reconstruir diferenciabilidad ⇒ continuidad. Supongamos que hemos olvidado la demostración del teorema, pero recordamos las dos definiciones. Queremos probar continuidad en a. Eso significa que debemos lograr $f(x) - f(a) \to 0$.

¿Qué información entrega la diferenciabilidad?

Nos dice que $\frac{f(x)-f(a)}{x-a} \to f'(a)$, con un límite real finito. Ahora hay que encontrar un puente entre ambas expresiones. Para $x \neq a$, $f(x) - f(a) = (x-a)\frac{f(x)-f(a)}{x-a}$. La estructura de la prueba aparece por sí sola:

- el primer factor tiende a 0;
- el segundo factor tiende al número real $f'(a)$;
- por el álgebra de límites, el producto tiende a 0.

Por tanto, $f(x) - f(a) \to 0$, y de ahí $f(x) \to f(a)$.

La continuidad ha sido reconstruida. El punto delicado vuelve a ser el dominio perforado. La identidad de factorización se escribe para $x \neq a$, pero eso es exactamente lo que necesita un límite. No existe ninguna laguna al no usar el centro durante la factorización. Podemos reconstruir el mismo razonamiento con la linealización. Si

$$f(a+h) - f(a) = f'(a)h + r(h), \qquad r(h) = o(h),$$

entonces $f'(a)h \to 0$ y, como $r(h)/h \to 0$, también $r(h) = h\,(r(h)/h) \to 0$. Así la variación total tiende a 0. Las dos pruebas son lecturas diferentes de una misma idea:

diferenciabilidad controla la variación a una escala más fina que la continuidad.

Reconstrucción, no recitación
Si olvidas la prueba, no intentes recordar una secuencia de líneas. Pregunta:

1. ¿cuál es el objetivo de continuidad?;
2. ¿qué expresión controla la diferenciabilidad?;
3. ¿cómo puedo transformar una en la otra?

La factorización surge entonces como una necesidad lógica, no como un truco que debamos memorizar.

Y la conversa sigue siendo falsa: $|x|$ es continua en 0 pero no diferenciable allí. La diferencia entre necesidad y suficiencia debe permanecer visible.

Estación 8 — De $f'(a) > 0$ a información puntual, no a monotonía. Supongamos ahora que $f'(a) > 0$.

¿Qué podemos concluir **solo** desde la definición?

Como $\frac{f(a+h)-f(a)}{h} \longrightarrow f'(a)$, podemos elegir una tolerancia menor que el propio valor límite. Por ejemplo, $\varepsilon = \frac{f'(a)}{2}$. Para h suficientemente pequeño, no nulo y admisible,

$$\left|\frac{f(a+h)-f(a)}{h} - f'(a)\right| < \frac{f'(a)}{2}.$$

De aquí se obtiene $\frac{f(a+h)-f(a)}{h} > \frac{f'(a)}{2} > 0$.

Ésta es una conclusión rigurosa.

Si $h > 0$, entonces $f(a+h) > f(a)$.

Si $h < 0$, al multiplicar por un número negativo cambia el sentido de la desigualdad y obtenemos $f(a+h) < f(a)$. Así, respecto del **centro fijo** a, la función queda localmente por

debajo a la izquierda y por encima a la derecha. Pero esto no demuestra todavía que f sea creciente en un entorno de a.

¿Por qué no?

Porque crecer en un intervalo exige comparar **dos puntos arbitrarios** $x < y$ del intervalo y demostrar $f(x) \leq f(y)$. La información obtenida desde $f'(a) > 0$ compara, en cambio, cada punto cercano únicamente con el centro a:

$$f(a+h) \quad \text{frente a} \quad f(a).$$

No compara entre sí dos puntos situados ambos a la derecha de a, ni dos puntos situados ambos a la izquierda. La diferencia lógica puede verse así:

$$\boxed{\begin{array}{c} f'(a) > 0 \\ \Downarrow \\ \text{signo de cocientes con centro fijo } a \end{array}}$$

no es todavía

$$\boxed{\forall x < y \text{ cercanos}, \qquad f(x) \leq f(y).}$$

Convertir información sobre derivadas en afirmaciones de monotonía sobre intervalos requerirá un mecanismo nuevo. Ese mecanismo aparecerá más adelante mediante los teoremas del valor medio.

No adelantar un teorema futuro
Del signo de **una derivada puntual** no debemos fabricar una conclusión global o local entre pares arbitrarios. La pregunta de auditoría es siempre:

> ¿qué puntos está comparando realmente la desigualdad que acabo de demostrar?

Si todos los cocientes tienen el mismo centro a, la conclusión sigue siendo puntual.

Síntesis — Ocho preguntas que deben sobrevivir al capítulo. Después de recorrer el laboratorio, la teoría diferencial básica puede reconstruirse mediante ocho preguntas.

1. **¿De dónde sale el cociente incremental?** De comparar la variación de salida con el incremento de entrada entre dos puntos.
2. **¿Qué distingue continuidad de derivabilidad?** La derivabilidad controla la variación después de dividir por la escala del incremento.
3. **¿Por qué no evaluamos en $h = 0$?** Porque la derivada es un límite sobre un entorno perforado, no el valor del cociente en el centro.
4. **¿Por qué podemos cancelar h?** Porque la cancelación se realiza donde $h \neq 0$, y la coincidencia en un entorno perforado basta para conservar el límite.
5. **¿Por qué un cociente simétrico no sustituye la derivada?** Porque puede ocultar discrepancias entre los dos lados.
6. **¿Qué significa $o(h)$?** Que el error dividido por h tiende a 0; el error es despreciable frente a la escala lineal.
7. **¿Por qué diferenciabilidad implica continuidad?** Porque la variación es el producto de un incremento que tiende a 0 por un cociente que posee límite finito, o equivalentemente porque es término lineal más error de orden menor.

8. **¿Por qué $f'(a) > 0$ no demuestra monotonía?** Porque la definición compara puntos cercanos con un centro fijo, no todos los pares de puntos de un intervalo.

Estas preguntas forman un mapa cognitivo del capítulo:

construir
→ distinguir
→ diagnosticar
→ justificar
→ poner a prueba
→ interpretar
→ reconstruir
→ delimitar.

Reconstruir estas respuestas exige volver a las definiciones y a las hipótesis de cada resultado. Así podremos distinguir el cálculo de una derivada de las afirmaciones adicionales sobre continuidad, aproximación o monotonía que se quieren deducir de ella. El banco de ejercicios reúne esas dificultades en problemas de reconocimiento, demostración y construcción de contraejemplos. Cada solución deberá justificar la conclusión con la información que realmente proporciona el enunciado.

6.11. Ejercicios y soluciones

Llegamos al banco final de `T1-C07`. El capítulo ha construido la derivada desde el cociente incremental y ha mostrado que una misma idea puede leerse como tasa instantánea, pendiente tangente y coeficiente de la aproximación afín local de primer orden. En estos problemas no se utilizarán todavía reglas generales de derivación. Antes de comenzar cada ejercicio conviene auditar cuatro preguntas:

1. ¿el problema pide calcular un cociente incremental, estudiar su límite o reconstruir una linealización?;
2. ¿el punto es interior del dominio, extremo o punto donde el dominio cambia de forma?;
3. ¿la conclusión es puntual —centrada en un solo punto— o pretende comparar pares arbitrarios en un entorno?;
4. ¿la herramienta necesaria pertenece realmente a este capítulo, o estaríamos anticipando una regla de `T1-C08` o un teorema del valor medio de `T1-C10`?

El banco contiene exactamente cuarenta ejercicios:

$$7A + 7B + 7C + 6D + 5E + 5F + 3G = 40.$$

Todos son originales y pueden resolverse usando los capítulos publicados desde los números reales hasta el presente. Las reglas sistemáticas de derivación, la regla de la cadena, la derivación de funciones elementales como repertorio, Rolle, los teoremas del valor medio, Taylor y Newton quedan fuera.

Nivel A — Reconocimiento y comprensión. Ejercicio A1. Incrementos y pendiente secante. Sea $f(x) = x^2$. Fija $a \in \mathbb{R}$ y un incremento $h \neq 0$.

1. Identifica el incremento de entrada Δx entre a y $a + h$.
2. Identifica la variación de salida Δy.
3. Escribe la pendiente de la secante que une $(a, f(a))$ con $(a + h, f(a + h))$.
4. Simplifica esa pendiente sin tomar todavía ningún límite.

Ejercicio A2. La derivada de una constante desde la definición. Sea $A \subseteq \mathbb{R}$, sea $a \in A$ un punto de acumulación de A y sea

$$f(x) = c \qquad (x \in A),$$

donde $c \in \mathbb{R}$ es fijo. Demuestra directamente desde la definición que f es diferenciable en a y que $f'_A(a) = 0$.

Ejercicio A3. La identidad desde la definición. Para $f(x) = x$, demuestra directamente, sin usar ninguna regla de derivación, que $f'(a) = 1$ para todo $a \in \mathbb{R}$.

Ejercicio A4. El cuadrado desde la definición. Sea $f(x) = x^2$. Calcula $f'(a)$ para un punto arbitrario $a \in \mathbb{R}$ partiendo únicamente de $\frac{f(a+h)-f(a)}{h}$.

No cites una regla para derivar potencias.

Ejercicio A5. Tres objetos que no deben confundirse. Sea f diferenciable en a. Clasifica cada expresión como **pendiente de una secante**, **derivada en el punto**, **pendiente de la tangente**, **recta tangente** o **ninguna de las anteriores**:

1. $$\frac{f(a + h) - f(a)}{h}, \qquad h \neq 0;$$
2. $$\lim_{h \to 0} \frac{f(a + h) - f(a)}{h};$$
3. $f'(a)$;
4. $T_a(x) = f(a) + f'(a)(x - a)$;
5. $f(a)$.

Explica por qué dos de las respuestas anteriores representan el mismo número, aunque procedan de descripciones conceptuales diferentes.

Ejercicio A6. Tangente a la parábola. Usando únicamente el cálculo directo de la derivada de x^2, escribe la recta tangente a $f(x) = x^2$ en el punto de abscisa a. Simplifica su ecuación.

Ejercicio A7. Reconocer un error de orden menor. Decide cuáles de las siguientes funciones de h satisfacen

$$r(h) = o(h) \qquad (h \to 0) :$$

$$h^2, \qquad h|h|, \qquad |h|, \qquad h + h^2.$$

Justifica cada respuesta dividiendo por h para $h \neq 0$.

Nivel B — Aplicación directa. Ejercicio B1. El cubo desde la definición. Para $f(x) = x^3$, calcula $f'(a)$ en un punto arbitrario a usando solamente el cociente incremental y álgebra elemental.

Ejercicio B2. El recíproco desde la definición. Sea $f(x) = \frac{1}{x}$ y fija $a \neq 0$. Calcula $f'(a)$ directamente desde la definición. Explica por qué, para h suficientemente pequeño, el denominador $a + h$ permanece distinto de 0.

Ejercicio B3. Cambiar de fórmula sin perder diferenciabilidad. Define $f(x) = x|x|$. Demuestra desde la definición que f es diferenciable en 0 y calcula $f'(0)$.

Ejercicio B4. Las dos pendientes de una esquina. Para $f(x) = |x|$, calcula la derivada derecha y la derivada izquierda en 0. Decide si existe la derivada bilateral.

Ejercicio B5. Un extremo del dominio. Considera

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

1. Calcula la derivada derecha de f en 0.
2. Considera la extensión $F : \mathbb{R} \to \mathbb{R}$, $F(x) = x^2$. Calcula su derivada bilateral en 0 desde la definición.
3. Explica por qué los dos cálculos coinciden en este caso, aunque conceptualmente no sean la misma pregunta.

Ejercicio B6. Refutar un límite con dos sucesiones. Usa las sucesiones

$$h_n = \frac{1}{n}, \qquad k_n = -\frac{1}{n}$$

para demostrar, mediante el criterio secuencial, que $|x|$ no es diferenciable en 0.

Ejercicio B7. Un salto excluye diferenciabilidad. Sea $f : A \to \mathbb{R}$ y sea $a \in A$ un punto de acumulación. Supón que f es discontinua en a; por ejemplo, porque presenta un salto allí. Justifica que f no puede ser diferenciable en a.

Indica con precisión qué resultado del capítulo estás usando y qué forma lógica tiene el argumento.

Nivel C — Combinación estructural. Ejercicio C1. Pegar dos fórmulas con diferenciabilidad. Para parámetros $\alpha, \beta, \gamma, \delta \in \mathbb{R}$, define

$$f(x) = \begin{cases} \alpha x + \beta, & x \leq 0, \\ x^2 + \gamma x + \delta, & x > 0. \end{cases}$$

Determina todas las relaciones entre los parámetros que hacen a f diferenciable en 0. Organiza la solución en dos controles separados: continuidad y coincidencia de los cocientes laterales.

Ejercicio C2. Localidad de la derivada. Sean $f, g : A \to \mathbb{R}$ y sea $a \in A$ un punto de acumulación. Supón que existe $\delta_0 > 0$ tal que

$$x \in A, \quad |x - a| < \delta_0 \Longrightarrow f(x) = g(x).$$

Demuestra directamente que f es diferenciable en a si y solo si g lo es y que, cuando existen, $f'_A(a) = g'_A(a)$.

Ejercicio C3. Linealización cuando el valor central es cero. Supón que f es diferenciable en a y que $f(a) = 0$.

Demuestra que

$$f(a + h) = f'(a)h + o(h) \qquad (h \to 0).$$

Ejercicio C4. Derivada nula y error sublineal. Sea f diferenciable en a. Demuestra la equivalencia

$$f'(a) = 0 \iff f(a + h) - f(a) = o(h) \qquad (h \to 0).$$

Ejercicio C5. Mismo primer orden, distinto residuo. Considera

$$f(x) = x^2, \qquad g(x) = x^2 + x^3.$$

En $a = 0$:

1. calcula directamente $f(0)$, $g(0)$, $f'(0)$ y $g'(0)$;
2. escribe la linealización de primer orden de ambas funciones;
3. calcula los residuos respecto de esa linealización;
4. explica en qué sentido las funciones tienen el mismo comportamiento de primer orden aunque no sean iguales cerca de 0.

Ejercicio C6. Reconstruir diferenciabilidad implica continuidad. Supón que $f(a+h) = f(a) + Lh + o(h)$.

Demuestra directamente, sin volver al cociente incremental, que f es continua en a.

Ejercicio C7. La derivada depende del dominio. Sea $f(x) = |x|$.

1. Considera f como función $\mathbb{R} \to \mathbb{R}$ y estudia su diferenciabilidad en 0.
2. Restringe ahora la misma regla a

   $$g : [0, \infty) \to \mathbb{R}, \qquad g(x) = |x|.$$

   Calcula la derivada relativa de g en 0.
3. Explica por qué no hay contradicción entre ambas respuestas.

Nivel D — Inversión y diagnóstico. Ejercicio D1. Reparar una implicación falsa. Un estudiante escribe:

> «Si f es continua en a, entonces $f(a+h)-f(a)\to 0$. Como además $h\to 0$, el cociente $\frac{f(a+h)-f(a)}{h}$ también debe tener límite; por tanto, toda función continua es diferenciable».

Identifica exactamente el paso inválido y refuta la conclusión mediante una función concreta.

Ejercicio D2. El cociente simétrico no es la derivada. Para una función f definida cerca de a, considera el cociente simétrico $\frac{f(a+h)-f(a-h)}{2h}$.

Demuestra que, para $f(x)=|x|$ y $a=0$, este cociente tiene límite cuando $h\to 0$, pero la derivada ordinaria de f en 0 no existe. Explica qué información pierde el promedio simétrico.

Ejercicio D3. Una tangente no se reconoce por el dibujo global. Considera primero $f(x)=x^3$.

1. Calcula desde la definición $f'(0)$ y escribe la tangente en 0.
2. Demuestra que la gráfica cruza esa tangente en el punto de tangencia.

Ahora considera $g(x)=x^2(x-1)$.

3. Calcula $g'(0)$ directamente desde el cociente incremental.
4. Comprueba que la tangente en 0 vuelve a intersectar la gráfica en $x=1$.

Explica por qué estos dos ejemplos invalidan las caracterizaciones «una tangente no cruza la curva» y «una tangente solo puede encontrarse con la gráfica en el punto de tangencia».

Ejercicio D4. Pendiente vertical no significa derivada finita. Sea $f(x)=\sqrt[3]{x}$.

1. Justifica que f es continua en 0.
2. Estudia el cociente incremental en 0.
3. Decide si $f'(0)$ existe según la definición de este capítulo.
4. Interpreta geométricamente el comportamiento del cociente.

Ejercicio D5. Una derivada derecha no determina una extensión. Sea

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

1. Calcula la derivada derecha de f en 0.
2. Construye dos extensiones $F_1,F_2:\mathbb{R}\to\mathbb{R}$ que coincidan con f en $[0,\infty)$, pero tales que F_1 sea diferenciable en 0 y F_2 no lo sea.
3. Explica por qué una derivada lateral en un extremo no autoriza a atribuir una derivada bilateral a una extensión arbitraria.

Ejercicio D6. No sustituir $h=0$ en el cociente. Un estudiante intenta hallar $f'(a)$ escribiendo $f'(a)=\frac{f(a+0)-f(a)}{0}=\frac{0}{0}$ y concluye que «la derivada nunca existe».

Repara el razonamiento. Explica:

1. dónde está definido el cociente incremental;
2. qué estudia realmente el límite $h\to 0$;
3. por qué una simplificación algebraica válida para $h\neq 0$ puede determinar el límite.

Nivel E — Contraejemplos. Ejercicio E1. Continuidad sin diferenciabilidad. Da una función continua en 0 que no sea diferenciable en 0 y verifica ambas afirmaciones desde las definiciones o criterios ya demostrados.

Ejercicio E2. Dos pendientes laterales finitas y distintas. Construye una función continua por tramos en 0 cuyas derivadas laterales en 0 existan como números reales pero sean diferentes. Verifica continuidad y calcula ambos límites laterales.

Ejercicio E3. Cociente incremental no acotado. Construye una función continua en 0 cuyo cociente incremental en 0 sea no acotado en todo entorno perforado de 0. Verifica ambas propiedades.

Ejercicio E4. Derivada simétrica existente, derivada ordinaria inexistente. Construye una función para la cual

$$\lim_{h\to 0} \frac{f(h) - f(-h)}{2h}$$

exista, pero

$$\lim_{h\to 0} \frac{f(h) - f(0)}{h}$$

no exista. Justifica los dos hechos.

Ejercicio E5. Racionales e irracionales producen pendientes distintas. Define $f(0) = 0$, y, para $x \neq 0$,

$$f(x) = \begin{cases} x, & x \in \mathbb{Q}, \\ 2x, & x \notin \mathbb{Q}. \end{cases}$$

1. Demuestra que f es continua en 0.
2. Demuestra que f no es diferenciable en 0 usando sucesiones racionales e irracionales que converjan a 0.

Nivel F — Descubrimiento guiado. Ejercicio F1. Descubrir el residuo. Supón que

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

Sin citar directamente el teorema de linealización, realiza los siguientes pasos:

1. define $r(h) = f(a+h) - f(a) - Lh$;
2. calcula $r(h)/h$ para $h \neq 0$;
3. demuestra que $r(h) = o(h)$;
4. reescribe $f(a+h)$ usando Lh y $r(h)$.

¿Qué información estructural sobre la función has recuperado?

Ejercicio F2. Del control del error a la continuidad. Supón que para cada $\varepsilon > 0$ existe $\delta > 0$ tal que

$$0 < |h| < \delta \Longrightarrow |f(a+h) - f(a) - Lh| < \varepsilon|h|.$$

Demuestra directamente que f es continua en a. No cites como caja negra «diferenciabilidad implica continuidad»: reconstruye la estimación que lo hace verdadero.

Ejercicio F3. Reconstruir el criterio bilateral. Sea a un punto interior del dominio de f. Supón que existen como números reales las derivadas laterales

$$f'_-(a), \qquad f'_+(a).$$

Demuestra que $f'(a)$ existe si y solo si $f'_-(a) = f'_+(a)$, y que, en ese caso, los tres valores coinciden.

Ejercicio F4. La única recta que puede pegarse suavemente a la parábola. Fija $a \in \mathbb{R}$ y una recta $\ell(x) = mx + b$.

Define

$$F(x) = \begin{cases} \ell(x), & x < a, \\ x^2, & x \geq a. \end{cases}$$

Determina los únicos valores de m y b para los cuales F es diferenciable en a. Interpreta la recta obtenida.

Ejercicio F5. Diferenciable con error mayor que cuadrático. Sea $f(x) = |x|^{3/2}$.

1. Demuestra que f es diferenciable en 0 y que $f'(0) = 0$.
2. Su linealización en 0 es la función nula. Demuestra que el residuo $r(h) = |h|^{3/2}$ no satisface una cota cuadrática local: no existen $C > 0$ y $\delta > 0$ tales que
$$|r(h)| \leq C|h|^2 \qquad (0 < |h| < \delta).$$
3. Explica por qué la diferenciabilidad solo exige $r(h) = o(h)$ y no un error de orden cuadrático.

Nivel G — Desafío. Ejercicio G1. Una derivada positiva da información con centro fijo. Supón que f es diferenciable en a y $f'(a) > 0$.

1. Demuestra que existe $\delta > 0$ tal que para todo h admisible con $0 < |h| < \delta$ se cumple $\frac{f(a+h)-f(a)}{h} > 0$.
2. Deduce que, para esos incrementos, $h > 0 \Longrightarrow f(a+h) > f(a)$, y $h < 0 \Longrightarrow f(a+h) < f(a)$.
3. Explica cuidadosamente por qué esto **no** demuestra que f sea creciente en un entorno de a.

Ejercicio G2. Mismo valor y misma derivada como equivalencia de primer orden. Sean $f, g : A \to \mathbb{R}$ diferenciables en un punto de acumulación $a \in A$, respecto del mismo dominio A. Todos los límites siguientes se entienden a través de incrementos admisibles de A. Demuestra que son equivalentes las siguientes afirmaciones:

1. $$f(a) = g(a) \qquad \text{y} \qquad f'(a) = g'(a);$$
2. $$f(a+h) - g(a+h) = o(h) \qquad (h \to 0).$$

La dirección inversa requiere justificar primero la igualdad de los valores centrales y después la igualdad de las derivadas.

Ejercicio G3. Diferenciable en un punto, irregular en todos los demás. Define $f : \mathbb{R} \to \mathbb{R}$ por $f(0) = 0$, y, para $x \neq 0$,

$$f(x) = \begin{cases} x^2, & x \in \mathbb{Q}, \\ 0, & x \notin \mathbb{Q}. \end{cases}$$

1. Demuestra que f es continua en 0.
2. Demuestra que f es diferenciable en 0 y calcula $f'(0)$.
3. Sea $c \neq 0$. Usa la densidad de $\mathbb{Q}$ y de $\mathbb{R} \setminus \mathbb{Q}$ para construir dos sucesiones que converjan a c y produzcan comportamientos incompatibles de f.
4. Concluye que f es discontinua en todo $c \neq 0$ y, por tanto, no es diferenciable en ninguno de esos puntos.

Soluciones.

Soluciones del nivel A. **Solución A1.**

Entre las entradas a y $a+h$, $\Delta x = (a+h) - a = h$.

Como $f(x) = x^2$,

$$\Delta y = f(a+h) - f(a) = (a+h)^2 - a^2.$$

La pendiente de la secante es

$$\frac{\Delta y}{\Delta x} = \frac{(a+h)^2 - a^2}{h}, \qquad h \neq 0.$$

Desarrollando y factorizando,

$$(a+h)^2 - a^2 = 2ah + h^2 = h(2a+h).$$

Por tanto,

$$\boxed{\frac{f(a+h) - f(a)}{h} = 2a + h} \qquad (h \neq 0).$$

Todavía no hemos tomado un límite: $2a+h$ es la pendiente de una secante concreta determinada por el incremento h.

Solución A2.

Para todo $x \in A$ con $x \neq a$, $\frac{f(x)-f(a)}{x-a} = \frac{c-c}{x-a} = 0$.

El cociente incremental es, por tanto, idénticamente nulo en el dominio perforado relativo a A. Luego

$$\lim_{\substack{x \to a \\ x \in A}} \frac{f(x) - f(a)}{x - a} = 0.$$

Por definición,

$$\boxed{f'_A(a) = 0}.$$

La hipótesis de que a sea punto de acumulación garantiza que el límite describe información incremental no vacía.

Solución A3.

Si $f(x) = x$, entonces para $h \neq 0$,

$$\frac{f(a+h) - f(a)}{h} = \frac{(a+h) - a}{h} = \frac{h}{h} = 1.$$

Así,

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h} = 1,$$

y por definición

$$\boxed{f'(a) = 1}.$$

El cálculo no depende del valor de a.

Solución A4.

Para $h \neq 0$, $\frac{f(a+h)-f(a)}{h} = \frac{(a+h)^2-a^2}{h}$.

Factorizando, $(a+h)^2 - a^2 = h(2a+h)$, de modo que $\frac{(a+h)^2-a^2}{h} = 2a + h$.

Ahora sí tomamos el límite:

$$\lim_{h \to 0} (2a + h) = 2a.$$

Por tanto,

$$\boxed{f'(a) = 2a}.$$

La fórmula se ha obtenido desde la definición, no desde una regla general para potencias.

Solución A5.

1. $\frac{f(a+h)-f(a)}{h}$ con $h \neq 0$ es la **pendiente de una secante**.
2. $$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h}$$
   es la **derivada en el punto**, siempre que el límite exista como número real.

3. $f'(a)$ es también la **derivada en el punto** y, por definición de tangente, es la **pendiente de la tangente**.
4. $T_a(x) = f(a) + f'(a)(x-a)$ es la **recta tangente**.
5. $f(a)$ es el valor de la función en el centro; no es, por sí mismo, una pendiente ni una recta.

Los apartados 2 y 3 representan el mismo número porque $f'(a)$ es precisamente el nombre asignado al límite del cociente incremental. Ese mismo número adquiere una lectura geométrica adicional como pendiente de la tangente.

Solución A6.
Para $f(x) = x^2$ ya obtuvimos directamente $f'(a) = 2a$.
La recta tangente es

$$T_a(x) = f(a) + f'(a)(x-a) = a^2 + 2a(x-a).$$

Simplificando,

$$T_a(x) = a^2 + 2ax - 2a^2 = 2ax - a^2.$$

Por tanto,

$$\boxed{T_a(x) = 2ax - a^2}.$$

Solución A7.
Debemos estudiar $r(h)/h$ para $h \neq 0$.
Para $r(h) = h^2$, $\frac{h^2}{h} = h \to 0$, así que $h^2 = o(h)$.
Para $r(h) = h|h|$, $\frac{h|h|}{h} = |h| \to 0$, por lo que $h|h| = o(h)$.
Para $r(h) = |h|$,

$$\frac{|h|}{h} = \begin{cases} 1, & h > 0, \\ -1, & h < 0, \end{cases}$$

y el límite bilateral no existe. Luego $|h|$ no es $o(h)$.
Finalmente, $\frac{h+h^2}{h} = 1 + h \to 1$, no a 0. Por tanto, $h + h^2$ tampoco es $o(h)$.
En resumen,

$$\boxed{h^2,\ h|h| \text{ son } o(h);}$$

$$\boxed{|h|,\ h + h^2 \text{ no lo son.}}$$

Soluciones del nivel B. **Solución B1.**
Para $h \neq 0$, $\frac{(a+h)^3-a^3}{h} = \frac{a^3+3a^2h+3ah^2+h^3-a^3}{h}$.
Factorizando h, $\frac{(a+h)^3-a^3}{h} = 3a^2 + 3ah + h^2$.
Al hacer $h \to 0$, $3a^2 + 3ah + h^2 \longrightarrow 3a^2$.
Por tanto,

$$\boxed{f'(a) = 3a^2}.$$

Solución B2.
Como $a \neq 0$, podemos elegir por ejemplo $|h| < \frac{|a|}{2}$.
Entonces

$$|a + h| \geq |a| - |h| > \frac{|a|}{2} > 0,$$

de modo que $a + h \neq 0$.
Para esos $h \neq 0$,

$$\frac{\frac{1}{a+h} - \frac{1}{a}}{h} = \frac{a - (a + h)}{a(a + h)h} = \frac{-h}{a(a + h)h} = -\frac{1}{a(a + h)}.$$

Como $a + h \to a$, $-\frac{1}{a(a+h)} \longrightarrow -\frac{1}{a^2}$.
Así,

$$\boxed{f'(a) = -\frac{1}{a^2}}.$$

Solución B3.
Tenemos $f(0) = 0$. Para $h \neq 0$, $\frac{f(h)-f(0)}{h} = \frac{h|h|}{h} = |h|$.
Como $|h| \to 0$, se sigue que el cociente incremental posee límite real y

$$\boxed{f'(0) = 0}.$$

El cambio de fórmula implícito en el valor absoluto no impide la diferenciabilidad porque, después de dividir por h, queda un término que tiende a 0 desde ambos lados.

Solución B4.
Como $f(0) = 0$, para $h \neq 0$, $\frac{f(h)-f(0)}{h} = \frac{|h|}{h}$.
Si $h > 0$, $\frac{|h|}{h} = 1$, por lo que $f'_+(0) = 1$.
Si $h < 0$, $\frac{|h|}{h} = -1$, y entonces $f'_-(0) = -1$.
Los dos límites laterales existen como números reales pero son distintos. Por el criterio lateral,

$$\boxed{f'(0) \text{ no existe.}}$$

Solución B5.

En el dominio $[0,\infty)$ solo son admisibles incrementos $h>0$ cerca de 0. Por tanto,

$$f'_+(0)=\lim_{h\to 0^+}\frac{h^2-0}{h}=\lim_{h\to 0^+}h=0.$$

Para la extensión $F(x)=x^2$ sobre todo $\mathbb{R}$,

$$F'(0)=\lim_{h\to 0}\frac{h^2}{h}=\lim_{h\to 0}h=0.$$

Los valores coinciden:

$$\boxed{f'_+(0)=F'(0)=0}.$$

Pero las preguntas son distintas. La primera utiliza únicamente puntos del dominio a la derecha de 0; la segunda exige compatibilidad de los cocientes desde ambos lados. La coincidencia ocurre porque la extensión concreta x^2 posee el mismo comportamiento desde la izquierda.

Solución B6.

El cociente incremental de $|x|$ en 0 es $Q(h)=\frac{|h|}{h}$.

Para $h_n=\frac{1}{n}$, tenemos $h_n\to 0$ y $Q(h_n)=1$ para todo n.

Para $k_n=-\frac{1}{n}$, tenemos $k_n\to 0$ y $Q(k_n)=-1$.

Si $Q(h)$ tuviera un límite real cuando $h\to 0$, el criterio secuencial obligaría a que ambos valores sucesionales convergieran al mismo número. Pero $1\neq -1$.

Por tanto,

$$\boxed{|x| \text{ no es diferenciable en } 0}.$$

Solución B7.

El capítulo demostró

$$f \text{ diferenciable en } a \Longrightarrow f \text{ continua en } a.$$

Su contraposición es

$$f \text{ discontinua en } a \Longrightarrow f \text{ no diferenciable en } a.$$

Como la hipótesis afirma precisamente que f es discontinua en a, podemos aplicar esa contraposición y concluir

$$\boxed{f'(a) \text{ no existe como derivada real finita.}}$$

El argumento es lógico, no un cálculo del cociente: usamos una condición necesaria para diferenciabilidad.

Soluciones del nivel C. **Solución C1.**

Para que f sea diferenciable en 0, primero debe ser continua allí.

Como $f(0) = \beta$, mientras que

$$\lim_{x\to 0^+} f(x) = \delta,$$

la continuidad exige

$$\boxed{\beta = \delta}.$$

Bajo esta condición, calculamos los cocientes laterales.

Por la izquierda, para $h < 0$, $\frac{f(h)-f(0)}{h} = \frac{\alpha h+\beta-\beta}{h} = \alpha$.

Así, $f'_-(0) = \alpha$.

Por la derecha, para $h > 0$ y usando $\delta = \beta$, $\frac{f(h)-f(0)}{h} = \frac{h^2+\gamma h+\delta-\beta}{h} = h + \gamma$, de modo que $f'_+(0) = \gamma$. La derivada bilateral existe exactamente cuando ambos valores coinciden:

$$\boxed{\alpha = \gamma}.$$

Por tanto,

$$\boxed{f \text{ es diferenciable en } 0 \iff \beta = \delta \text{ y } \alpha = \gamma.}$$

Solución C2.

La hipótesis incluye el centro, así que $f(a) = g(a)$.

Además, cuando

$$x \in A, \qquad 0 < |x - a| < \delta_0,$$

tenemos $f(x) = g(x)$. Por tanto, $\frac{f(x)-f(a)}{x-a} = \frac{g(x)-g(a)}{x-a}$ en un entorno perforado relativo al dominio. Dos funciones que coinciden en algún entorno perforado de un punto tienen el mismo comportamiento límite allí. En consecuencia, uno de esos cocientes posee límite real si y solo si el otro lo posee y, en tal caso, los límites son iguales.

Así,

$$\boxed{f \text{ diferenciable en } a \iff g \text{ diferenciable en } a,}$$

y, cuando ocurre,

$$\boxed{f'_A(a) = g'_A(a)}.$$

Solución C3.

Por la caracterización de la diferenciabilidad mediante linealización, $f(a+h) = f(a) + f'(a)h + o(h)$.

Como $f(a) = 0$, $f(a + h) = f'(a)h + o(h)$.

Por tanto,

$$\boxed{f(a + h) = f'(a)h + o(h).}$$

La hipótesis sobre el valor central elimina únicamente el término constante de la aproximación afín.

Solución C4.

Si $f'(a) = 0$, la linealización da $f(a+h) = f(a)+0\cdot h+o(h)$, por lo que $f(a+h)-f(a) = o(h)$.

Recíprocamente, supongamos $f(a+h) - f(a) = o(h)$.

Para $h \neq 0$, $\frac{f(a+h)-f(a)}{h} = \frac{o(h)}{h}$.

Por definición de $o(h)$, el lado derecho tiende a 0. Entonces $f'(a) = 0$.

Así,

$$\boxed{f'(a) = 0 \iff f(a+h) - f(a) = o(h).}$$

Solución C5.

En 0, $f(0) = g(0) = 0$.

Para f, $\frac{f(h)-f(0)}{h} = h \to 0$.

Para g, $\frac{g(h)-g(0)}{h} = \frac{h^2+h^3}{h} = h + h^2 \to 0$.

Por tanto, $f'(0) = g'(0) = 0$.

La linealización de primer orden de ambas funciones es $T(h) = 0$.

Los residuos son $r_f(h) = h^2$, y $r_g(h) = h^2 + h^3$.

Ambos satisfacen

$$\frac{r_f(h)}{h} \to 0, \qquad \frac{r_g(h)}{h} \to 0,$$

de modo que ambos son $o(h)$. Las funciones no son iguales cerca de 0 porque $g(h) - f(h) = h^3$, pero esa diferencia también es $o(h)$. Así, a escala de primer orden comparten el mismo valor central y el mismo coeficiente lineal, aunque sus términos de orden menor sean distintos.

Solución C6.

Escribamos $f(a + h) - f(a) = Lh + r(h)$, donde $r(h) = o(h)$.

Como $r(h)/h \to 0$, existe una función $\varepsilon(h) \to 0$ tal que $r(h) = h\varepsilon(h)$ para $h \neq 0$ suficientemente pequeño. Entonces $f(a+h)-f(a) = h\big(L+\varepsilon(h)\big)$. El factor h tiende a 0, mientras que $L+\varepsilon(h) \to L$ es acotado cerca de 0. Por álgebra de límites, $f(a + h) - f(a) \to 0$. Por tanto, $f(a + h) \to f(a)$, y f es continua en a.

Solución C7.

Como función sobre $\mathbb{R}$, $\frac{|h|-0}{h} = \frac{|h|}{h}$.

El límite por la derecha es 1 y por la izquierda es -1. Por tanto, $|x|$ no es diferenciable en 0 sobre $\mathbb{R}$.

En cambio, para

$$g : [0, \infty) \to \mathbb{R}, \qquad g(x) = |x| = x,$$

los incrementos admisibles cerca de 0 satisfacen $h > 0$. Así,

$$g'_{[0,\infty)}(0) = \lim_{h\to 0^+} \frac{h}{h} = 1.$$

No hay contradicción. La derivada relativa es un límite relativo al dominio. Sobre $\mathbb{R}$ debemos compatibilizar ambos lados; sobre $[0,\infty)$ solo existe aproximación desde la derecha.

Soluciones del nivel D. **Solución D1.**

La continuidad proporciona $f(a+h)-f(a)\to 0$.

Pero de

$$u(h)\to 0, \qquad v(h)\to 0$$

no se sigue que $u(h)/v(h)$ tenga límite. La regla del cociente para límites requiere que el límite del denominador sea distinto de 0. Aquí precisamente $h\to 0$, de modo que la división no está protegida por esa regla. Un contraejemplo es $f(x)=|x|$ en $a=0$. La función es continua porque $|x|\to 0$, pero

$$\frac{|h|}{h}=\begin{cases}1, & h>0,\\ -1, & h<0,\end{cases}$$

no posee límite bilateral.

Por tanto, continuidad no implica diferenciabilidad.

Solución D2.

Para $f(x)=|x|$ y $a=0$, $\frac{f(h)-f(-h)}{2h}=\frac{|h|-|-h|}{2h}$.

Como $|-h|=|h|$, el numerador es 0 y, para $h\neq 0$, $\frac{|h|-|-h|}{2h}=0$. Por tanto, el cociente simétrico tiene límite 0. Sin embargo, el cociente incremental ordinario es $\frac{|h|-0}{h}=\frac{|h|}{h}$, cuyos límites laterales son 1 y -1. La derivada ordinaria no existe. El cociente simétrico compara simultáneamente los valores a ambos lados y puede hacer que comportamientos opuestos se cancelen. Por eso no sustituye el control unilateral necesario para el límite ordinario.

Solución D3.

Para $f(x)=x^3$, $\frac{f(h)-f(0)}{h}=h^2\to 0$.

Así, $f'(0)=0$, y la tangente es $y=0$.

Pero

$$x^3<0 \quad (x<0), \qquad x^3>0 \quad (x>0).$$

La gráfica pasa de un lado de la recta $y=0$ al otro: la tangente **cruza** la gráfica en el punto de tangencia.

Ahora, $g(x)=x^2(x-1)=x^3-x^2$, y $g(0)=0$. Para $h\neq 0$,

$$\frac{g(h)-g(0)}{h}=\frac{h^2(h-1)}{h}=h(h-1)\to 0.$$

Por tanto, la tangente en 0 vuelve a ser $y=0$. Las intersecciones con la gráfica satisfacen $x^2(x-1)=0$, así que aparecen en

$$x=0 \qquad \text{y} \qquad x=1.$$

Los dos ejemplos muestran que la tangencia es una noción local de primer orden. Ni «no cruzar» ni «intersectar una sola vez» pueden servir como definición.

Solución D4.
Primero comprobamos la continuidad en 0. Dado $\varepsilon > 0$, si $|x| < \varepsilon^3$, entonces

$$|f(x) - f(0)| = |\sqrt[3]{x}| = |x|^{1/3} < \varepsilon.$$

Por tanto,

$$f(x) \to 0 = f(0) \qquad (x \to 0).$$

Para $h \neq 0$, $\frac{f(h)-f(0)}{h} = \frac{\sqrt[3]{h}}{h}$.
Como $h = (\sqrt[3]{h})^3$, obtenemos

$$\frac{\sqrt[3]{h}}{h} = \frac{1}{(\sqrt[3]{h})^2} = \frac{1}{|h|^{2/3}}.$$

Cuando $h \to 0$, $\frac{1}{|h|^{2/3}} \to +\infty$.
El cociente incremental no converge a un número real finito. Por tanto, según la definición del capítulo,

$$\boxed{f'(0) \text{ no existe.}}$$

Geométricamente, el crecimiento sin cota de las pendientes secantes sugiere una tangente vertical. Pero una tangente vertical en ese sentido geométrico no convierte $+\infty$ en una derivada real.

Solución D5.
Para $h > 0$, $\frac{f(h)-f(0)}{h} = \frac{h}{h} = 1$.
Así, $f'_+(0) = 1$.
Una primera extensión es

$$F_1(x) = x \qquad (x \in \mathbb{R}).$$

Entonces $F_1'(0) = 1$.
Una segunda extensión es

$$F_2(x) = |x| = \begin{cases} -x, & x < 0, \\ x, & x \geq 0. \end{cases}$$

Esta coincide con f para $x \geq 0$, pero en 0 sus derivadas laterales son

$$F'_{2,-}(0) = -1, \qquad F'_{2,+}(0) = 1.$$

Por tanto, $F_2'(0)$ no existe.
La información del dominio $[0, \infty)$ no contiene ninguna condición sobre cómo debe comportarse una extensión a la izquierda de 0. Una derivada derecha no determina ese dato ausente.

Solución D6.

El cociente incremental $\frac{f(a+h)-f(a)}{h}$ solo se considera para $h \neq 0$. En $h = 0$ aparece una división por cero, así que no hay valor del cociente que debamos evaluar allí. La derivada pregunta por el límite

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

es decir, por el comportamiento de la expresión en un entorno perforado de 0.

Por ejemplo, para $f(x) = x^2$,

$$\frac{(a+h)^2-a^2}{h} = 2a+h \qquad (h \neq 0).$$

La igualdad es válida exactamente donde se necesita para estudiar el límite. Aunque el cociente original no esté definido en $h = 0$, $2a+h \to 2a$. Por tanto, la derivada puede existir perfectamente. El límite no exige evaluar la expresión original en el punto excluido.

Soluciones del nivel E. **Solución E1.**

Tomemos $f(x) = |x|$.

Como $|f(x) - f(0)| = |x|$, se tiene

$$f(x) \to 0 = f(0) \qquad (x \to 0),$$

así que f es continua en 0. Pero para $h \neq 0$, $\frac{f(h)-f(0)}{h} = \frac{|h|}{h}$, que vale 1 si $h > 0$ y -1 si $h < 0$. Los límites laterales no coinciden.

Por tanto,

$$\boxed{|x| \text{ es continua pero no diferenciable en } 0.}$$

Solución E2.

Un ejemplo es

$$f(x) = \begin{cases} 2x, & x < 0, \\ x, & x \geq 0. \end{cases}$$

En ambos lados,

$$f(x) \to 0 = f(0) \qquad (x \to 0),$$

por lo que f es continua en 0.

Para $h < 0$, $\frac{f(h)-f(0)}{h} = \frac{2h}{h} = 2$, así que $f'_-(0) = 2$.

Para $h > 0$, $\frac{f(h)-f(0)}{h} = \frac{h}{h} = 1$, por lo que $f'_+(0) = 1$.

Ambas derivadas laterales existen y son finitas, pero $2 \neq 1$. Así, la derivada bilateral no existe.

Solución E3.

Tomemos $f(x) = \sqrt{|x|}$.

La función es continua en 0 porque $\sqrt{|x|} \to 0 = f(0)$.

Para $h \neq 0$,

$$\frac{f(h) - f(0)}{h} = \frac{\sqrt{|h|}}{h}.$$

Su valor absoluto es

$$\left| \frac{\sqrt{|h|}}{h} \right| = \frac{1}{\sqrt{|h|}}.$$

Cuando $h \to 0$,

$$\frac{1}{\sqrt{|h|}} \to +\infty.$$

Por tanto, el cociente incremental es no acotado en todo entorno perforado de 0 y no puede converger a una derivada real finita.

Solución E4.

De nuevo podemos tomar $f(x) = |x|$.

Para $h \neq 0$, $\frac{f(h)-f(-h)}{2h} = \frac{|h|-|-h|}{2h} = 0$.

Así, la derivada simétrica existe y vale 0.

En cambio, $\frac{f(h)-f(0)}{h} = \frac{|h|}{h}$ vale 1 para $h > 0$ y -1 para $h < 0$. Por tanto, la derivada ordinaria no existe. El ejemplo muestra que la existencia del promedio simétrico es estrictamente más débil que la existencia del límite ordinario del cociente incremental.

Solución E5.

Para todo x, $|f(x)| \leq 2|x|$.

Como $2|x| \to 0$ cuando $x \to 0$, el teorema del sándwich da $f(x) \to 0 = f(0)$.

Así, f es continua en 0.

Para $h \neq 0$,

$$\frac{f(h) - f(0)}{h} = \begin{cases} 1, & h \in \mathbb{Q}, \\ 2, & h \notin \mathbb{Q}. \end{cases}$$

Por densidad existen sucesiones de racionales no nulos $q_n \to 0$ y de irracionales $r_n \to 0$. A lo largo de ellas, $\frac{f(q_n)}{q_n} = 1$, mientras que $\frac{f(r_n)}{r_n} = 2$. El criterio secuencial impide que el cociente tenga un único límite. Por tanto,

$$\boxed{f'(0) \text{ no existe.}}$$

Soluciones del nivel F. **Solución F1.**

Definimos $r(h) = f(a+h) - f(a) - Lh$.

Para $h \neq 0$, $\frac{r(h)}{h} = \frac{f(a+h)-f(a)}{h} - L$.

Por hipótesis, $\frac{f(a+h)-f(a)}{h} \to L$, de modo que $\frac{r(h)}{h} \to 0$.

Esto significa exactamente $r(h) = o(h)$.

Reordenando la definición de r,

$$\boxed{f(a+h) = f(a) + Lh + o(h).}$$

Hemos recuperado la estructura de primer orden: la variación de f se descompone en una parte lineal principal Lh y un error despreciable frente a h.

Solución F2.

Queremos demostrar que $f(a+h) - f(a) \to 0$.

Tomemos primero $\varepsilon = 1$ en la hipótesis. Existe $\delta_1 > 0$ tal que, si $0 < |h| < \delta_1$, entonces $|f(a+h) - f(a) - Lh| < |h|$.

Por desigualdad triangular,

$$|f(a+h) - f(a)| \leq |L||h| + |f(a+h) - f(a) - Lh| < (|L| + 1)|h|.$$

Dada una tolerancia $\eta > 0$, elijamos

$$\delta = \min\left\{\delta_1, \frac{\eta}{|L|+1}\right\}.$$

Entonces

$$0 < |h| < \delta \Longrightarrow |f(a+h) - f(a)| < \eta.$$

Para $h = 0$ la diferencia es exactamente 0. Por tanto, $f(a+h) \to f(a)$, y f es continua en a.

Solución F3.

Consideremos el cociente incremental $Q(x) = \frac{f(x)-f(a)}{x-a}$ sobre el dominio perforado. La derivada izquierda es el límite de $Q(x)$ cuando $x \to a^-$ y la derecha es el límite cuando $x \to a^+$. La teoría previa de límites laterales establece que un límite bilateral existe como número real si y solo si existen ambos límites laterales y son iguales.

Por tanto,

$$f'(a) \text{ existe } \iff f'_-(a) \text{ y } f'_+(a) \text{ existen y son iguales.}$$

Si $f'_-(a) = f'_+(a) = L$, entonces el límite bilateral del mismo cociente es L, y

$$\boxed{f'(a) = f'_-(a) = f'_+(a) = L.}$$

Solución F4.

Para que F sea diferenciable en a, debe ser continua allí. Como el valor del tramo derecho es $F(a) = a^2$, la continuidad desde la izquierda exige $ma + b = a^2$. La derivada izquierda de la recta es, desde el cociente incremental, m. La derivada derecha del tramo x^2 en a es $2a$. La diferenciabilidad exige que ambas coincidan: $m = 2a$. Sustituyendo en la condición de continuidad, $2a^2 + b = a^2$, de donde $b = -a^2$.

La única recta es

$$\boxed{\ell(x) = 2ax - a^2.}$$

Pero ésta es exactamente $a^2 + 2a(x - a)$, la recta tangente a x^2 en $x = a$.

Solución F5.

Como $f(0) = 0$, para $h \neq 0$, $\frac{f(h)-f(0)}{h} = \frac{|h|^{3/2}}{h}$.

Escribiendo $h = \operatorname{sgn}(h)|h|$, $\frac{|h|^{3/2}}{h} = \operatorname{sgn}(h)|h|^{1/2}$.

Su valor absoluto es $|h|^{1/2} \to 0$. Por tanto,

$$\boxed{f'(0) = 0}.$$

La linealización es la función nula y el residuo es $r(h) = |h|^{3/2}$. Supongamos que existieran $C > 0$ y $\delta > 0$ tales que $|h|^{3/2} \leq C|h|^2$ para todo $0 < |h| < \delta$. Dividiendo por $|h|^{3/2} > 0$, $1 \leq C|h|^{1/2}$. Pero el lado derecho puede hacerse menor que 1 tomando $|h|$ suficientemente pequeño, contradicción.

Equivalentemente,

$$\frac{|h|^{3/2}}{|h|^2} = \frac{1}{|h|^{1/2}} \to +\infty.$$

La diferenciabilidad solo exige $\frac{r(h)}{h} \to 0$, es decir, $r(h) = o(h)$. No exige una cota cuadrática.

Soluciones del nivel G. **Solución G1.**

Sea $L = f'(a) > 0$.

Por definición, $\frac{f(a+h)-f(a)}{h} \to L$.

Tomemos $\varepsilon = \frac{L}{2} > 0$.

Existe $\delta > 0$ tal que, para todo incremento admisible con $0 < |h| < \delta$, se cumple

$$\left|\frac{f(a+h) - f(a)}{h} - L\right| < \frac{L}{2}.$$

Entonces

$$\frac{f(a+h) - f(a)}{h} > L - \frac{L}{2} = \frac{L}{2} > 0.$$

Si $h > 0$, multiplicar por h conserva el signo: $f(a+h) - f(a) > 0$, así que $f(a+h) > f(a)$. Si $h < 0$, multiplicar por h invierte el signo: $f(a+h) - f(a) < 0$, y por tanto $f(a+h) < f(a)$. Sin embargo, todas estas comparaciones tienen el mismo centro a. No hemos demostrado que, para puntos arbitrarios $x < y$ cercanos a a, $f(x) \leq f(y)$. Esa afirmación exige comparar pares

variables, no únicamente cada punto con a. Por eso el resultado es puntual y no constituye todavía un teorema de monotonía en un entorno.

Solución G2.

Supongamos primero $f(a) = g(a)$ y $f'(a) = g'(a) = L$.

Las linealizaciones son $f(a+h) = f(a) + Lh + r_f(h)$,

$$g(a+h) = g(a) + Lh + r_g(h),$$

con

$$r_f(h) = o(h), \qquad r_g(h) = o(h).$$

Restando,

$$f(a+h) - g(a+h) = r_f(h) - r_g(h).$$

Como $\frac{r_f(h)-r_g(h)}{h} = \frac{r_f(h)}{h} - \frac{r_g(h)}{h} \to 0$, obtenemos $f(a+h) - g(a+h) = o(h)$.

Recíprocamente, supongamos

$$d(h) := f(a+h) - g(a+h) = o(h).$$

Entonces $\frac{d(h)}{h} \to 0$.

En particular, $d(h) \to 0$. Como f y g son diferenciables en a, ambas son continuas allí; por tanto,

$$d(h) = f(a+h) - g(a+h) \to f(a) - g(a).$$

La unicidad del límite da $f(a) - g(a) = 0$, es decir, $f(a) = g(a)$. Ahora, para $h \neq 0$, $\frac{d(h)}{h} = \frac{f(a+h)-f(a)}{h} - \frac{g(a+h)-g(a)}{h}$, porque los valores centrales son iguales. Al tomar límites, $0 = f'(a) - g'(a)$.

Así, $f'(a) = g'(a)$.

Concluimos la equivalencia completa.

Solución G3.

Para todo x, $|f(x)| \leq x^2$.

Como $x^2 \to 0$ cuando $x \to 0$, el sándwich da $f(x) \to 0 = f(0)$. Por tanto, f es continua en 0.

Para $h \neq 0$,

$$\frac{f(h) - f(0)}{h} = \begin{cases} h, & h \in \mathbb{Q}, \\ 0, & h \notin \mathbb{Q}. \end{cases}$$

En ambos casos,

$$\left| \frac{f(h) - f(0)}{h} \right| \leq |h|.$$

Como $|h| \to 0$,

$$\boxed{f'(0) = 0}.$$

Ahora sea $c \neq 0$. Por densidad existen sucesiones

$$q_n \in \mathbb{Q}, \qquad q_n \to c,$$

y

$$r_n \notin \mathbb{Q}, \qquad r_n \to c.$$

A lo largo de los racionales, $f(q_n) = q_n^2 \to c^2$. A lo largo de los irracionales, $f(r_n) = 0 \to 0$.

Como $c \neq 0$, $c^2 \neq 0$.

Por el criterio secuencial, f no puede ser continua en c. Así, f es discontinua en todo punto no nulo.

Finalmente, diferenciabilidad implica continuidad. Por contraposición,

$$\boxed{f \text{ no es diferenciable en ningún } c \neq 0.}$$

El ejemplo muestra que la diferenciabilidad en un solo punto no impone regularidad diferencial en un entorno completo.

Capítulo 7

Álgebra de derivadas y regla de la cadena

En el capítulo anterior aprendimos a determinar la derivada de una función *en un punto*. Si a pertenece al dominio A de f y es punto de acumulación de A, la existencia de $f'_A(a)$ significa que el cociente de incrementos tiene un límite real finito. Aprendimos también a interpretar ese número como el coeficiente de una aproximación afín cuyo error es despreciable frente al incremento.

Hasta aquí, las derivadas se han obtenido principalmente desde su definición. Pero las funciones que queremos estudiar rara vez llegan aisladas: se suman, se multiplican, se dividen y se componen. Por ejemplo, incluso cuando conocemos las derivadas de f y de g, todavía debemos justificar qué podemos afirmar sobre $f+g$, fg, f/g $\circ$ $g \circ f$.

Para derivar una combinación de funciones necesitaremos relacionar su incremento con los incrementos de las componentes. En la composición habrá que seguir dos escalas: el cambio del argumento y el cambio que este produce en la función interior. La prueba de la regla de la cadena deberá admitir que ese segundo cambio pueda anularse. La ruta de construcción será

cociente incremental
$\to$ combinaciones lineales
$\to$ producto
$\to$ recíproco y cociente
$\to$ potencias y racionales
$\to$ composición y cadena.

El producto permite ver la dificultad inicial: conocer dos pendientes no determina la pendiente del producto mediante una multiplicación de esos números.

7.1. Por qué las reglas de derivación necesitan demostración

Una conjetura natural que no resiste la definición. Consideremos dos funciones $f, g : \mathbb{R} \to \mathbb{R}$ diferenciables y formemos su producto puntual: $(fg)(x) = f(x)g(x)$. Si conocemos $f'(a)$ y $g'(a)$, ¿bastará multiplicar ambas pendientes para obtener la pendiente de fg? Es tentador escribir

$$\underbrace{(fg)'(a)}_{\text{pendiente del producto}} \stackrel{?}{=} \underbrace{f'(a)g'(a)}_{\text{producto de las pendientes}} .$$

La igualdad tiene una apariencia razonable: como el producto se forma multiplicando valores, podría parecer que su derivada se forma multiplicando derivadas. Pero la apariencia de una fórmula no demuestra que describa el límite que define su miembro izquierdo.

Ejemplo 7.1. **Dos pendientes iguales a** 1 **cuyo producto tiene pendiente** 2**.** Sean

$$f(x) = x, \qquad g(x) = x, \qquad x \in \mathbb{R},$$

y fijemos $a = 1$. Desde la definición de derivada, para todo $h \neq 0$, $\frac{f(1+h)-f(1)}{h} = \frac{(1+h)-1}{h} = 1$. Por tanto, $f'(1) = 1$; exactamente el mismo cálculo da $g'(1) = 1$. Si nuestra conjetura fuera correcta, tendríamos $(fg)'(1) = f'(1)g'(1) = 1$. Sin embargo, el producto es la función $p(x) = (fg)(x) = x^2$. Su cociente incremental en 1 es

$$\begin{aligned} \frac{p(1+h) - p(1)}{h} &= \frac{(1+h)^2 - 1}{h} \\ &= \frac{2h + h^2}{h} \\ &= 2 + h, \qquad h \neq 0. \end{aligned}$$

La cancelación se efectuó **antes** de tomar el límite y solo para $h \neq 0$; no se ha sustituido $h = 0$ en una fracción indefinida. Al hacer $h \to 0$ obtenemos

$$\boxed{(fg)'(1) = 2 \neq 1 = f'(1)g'(1).}$$

Este único caso refuta la presunta identidad universal $(fg)' = f'g'$. No hemos utilizado ninguna regla de derivación para el producto: hemos calculado las tres derivadas directamente mediante cocientes incrementales.

Detengámonos en el paso que determina el resultado. Cuando ambos factores pasan de 1 a $1+h$, el producto pasa de 1 a $(1+h)(1+h) = 1+2h+h^2$. La variación total es $2h+h^2$, no h^2. A escala de primer orden, los dos términos lineales aportan $2h$, mientras h^2 se vuelve despreciable frente a h. Este mecanismo elemental sugiere qué debe estudiar la verdadera regla del producto; **no constituye todavía su demostración para funciones arbitrarias**, que realizaremos en §7.3.

Diagnóstico: una fórmula falsa, no una excepción de cálculo
El problema no consiste en que $f(x) = g(x) = x$ sea un caso excepcional o en que hayamos elegido mal el punto. Una afirmación formulada para *todas* las funciones diferenciables queda refutada por un solo par admisible que incumpla su conclusión. Lo que falla es identificar la variación de un producto con el producto de las variaciones.

Recuperar el objeto que realmente estamos derivando. El capítulo anterior definió la derivada a partir de un **límite de cocientes de diferencias**. Para recordarlo con precisión, sea $u : A \to \mathbb{R}$, con $a \in A$ punto de acumulación, y consideremos el conjunto de incrementos admisibles $H_a = \{h \in \mathbb{R} : a + h \in A\}$. Si existe un límite real finito, entonces

$$u'_A(a) = \lim_{\substack{h \to 0 \\ h \in H_a \setminus \{0\}}} \frac{u(a+h) - u(a)}{h}.$$

En particular, para $u = fg$ debemos comenzar por

$$\frac{f(a+h)g(a+h) - f(a)g(a)}{h}, \qquad h \in H_a \setminus \{0\}.$$

No hemos demostrado que esta expresión pueda sustituirse por el producto de los cocientes incrementales de f y de g. Sería necesario justificar primero una identidad algebraica que permitiera hacerlo, y el ejemplo anterior demuestra que la pretendida identidad no puede ser válida en general.

Tampoco debemos confundir este problema con el álgebra de límites estudiada en el capítulo 4. Si $F(h) \to L$ y $G(h) \to M$, sabemos que $F(h)G(h) \to LM$. Ese teorema permite pasar al límite **una vez que hemos expresado correctamente** el cociente de diferencias; no nos dice que el cociente del incremento de fg sea el producto de los cocientes de sus factores.

La estrategia matemática será, por tanto, siempre la misma: escribir la diferencia que corresponde a la función nueva, encontrar una transformación algebraica válida para incrementos admisibles no nulos y solo entonces utilizar los teoremas de límites. En §7.2 veremos que las combinaciones lineales admiten una descomposición inmediata; el producto requerirá una idea adicional.

La segunda lectura: qué ocurre con las aproximaciones locales. Hay otra manera de apreciar la dificultad, sin inventar una regla. El capítulo 6 demostró que, cuando f es diferenciable en a,

$$f(a+h) = f(a) + f'_A(a)h + r_f(h), \qquad r_f(h) = o(h),$$

para incrementos admisibles. Existe una expresión análoga para g. Esta fórmula enseña a separar **valor central, variación lineal y error menor que el incremento**. En nuestro ejemplo, al tomar $a = 1$, ambas aproximaciones son exactas:

$$f(1+h) = 1 + h, \qquad g(1+h) = 1 + h.$$

El incremento del producto depende también de los valores centrales de los factores, además de sus pendientes. Las aproximaciones con residuo permitirán identificar esos términos y separar los que contribuyen al primer orden de los que desaparecen al dividir por el incremento.

Antes de aplicar una regla: cuatro preguntas. De aquí en adelante, una operación con funciones deberá pasar por una auditoría breve:

1. **¿Qué función nueva hemos formado y en qué dominio está definida?** Una suma o un producto de funciones con dominios distintos solo tiene sentido donde ambas estén definidas; el cociente exige, además, un denominador no nulo.
2. **¿En qué punto calculamos y qué incrementos son admisibles?** El punto debe pertenecer al dominio pertinente y ser punto de acumulación para que exista una derivada en nuestro sentido. En extremos de dominio trabajaremos con límites relativos, no con extensiones supuestas.
3. **¿Qué hipótesis conocemos realmente?** Que dos funciones sean diferenciables en un punto es un dato que podemos usar; que su combinación sea diferenciable y satisfaga cierta fórmula es precisamente lo que habrá que demostrar.
4. **¿Qué identidad y qué límite justifican la conclusión?** La transformación debe ser válida antes del paso al límite y no puede introducir divisiones por cantidades posiblemente nulas.

Estas preguntas distinguen dos actividades que suelen mezclarse en un cálculo rutinario: **probar una regla** y **aplicarla después de haber comprobado sus hipótesis**. Una prueba de suficiencia tampoco demostraría, por sí sola, que las mismas hipótesis fueran necesarias.

Antes de seguir
Reconstruye sin consultar el ejemplo: si $p(x) = x^2$, ¿por qué el cociente incremental en 1 vale $2 + h$ y no 1? Señala exactamente dónde es lícito cancelar h y qué límite se usa después. A continuación, explica en una frase por qué el teorema del límite de un producto no permite afirmar $(fg)' = f'g'$.

El programa demostrativo del capítulo. La primera regla que investigaremos será la de las **combinaciones lineales**, porque sus incrementos se separan mediante la propia distributividad de los números reales. Después estudiaremos el producto y descubriremos qué término intermedio permite controlar sus dos factores. El recíproco y el cociente exigirán, además, demostrar que un denominador no nulo en el centro permanece no nulo cerca de él. Con esas herramientas construiremos las derivadas de potencias enteras, polinomios y funciones racionales. Por último abordaremos las composiciones. En ellas, el incremento de la función interior puede anularse incluso cuando el incremento de entrada es distinto de cero. Una demostración correcta de la regla de la cadena deberá funcionar también en ese caso, sin introducir una división ilegítima.

Así, el compromiso de este capítulo puede formularse en una sola frase:

Primero justificar la transformación; después utilizar la fórmula.

La siguiente sección comienza por la situación algebraicamente más sencilla: demostrar, desde la definición, cómo se comporta la derivada frente a la suma y la multiplicación por escalares.

7.2. La derivación respeta combinaciones lineales

¿Qué hace distinta a una suma? En §7.1 vimos que no hay razón para identificar el incremento de un producto con el producto de los incrementos. La suma plantea una situación diferente. Si una función cambia en cierta cantidad y otra cambia en otra cantidad, la variación de su suma es *exactamente* la suma de esas dos variaciones. No hay términos adicionales que controlar. La misma observación vale cuando multiplicamos previamente cada función por un número fijo. Sean $f, g : A \to \mathbb{R}$, con $a \in A$ punto de acumulación de A, y sean $\alpha, \beta \in \mathbb{R}$. Definimos sobre el **mismo dominio** A la función

$$u(x) = \alpha f(x) + \beta g(x), \qquad x \in A.$$

¿Qué información necesitamos para afirmar que u es diferenciable en a? No basta advertir que la fórmula tiene la forma de una combinación lineal: debemos mostrar que su cociente incremental posee un límite real finito. El trabajo algebraico, esta vez, es directo.

Teorema 7.1. **Teorema — Linealidad de la derivada.** Sean $A \subseteq \mathbb{R}$, $a \in A$ un punto de acumulación de A y $f, g : A \to \mathbb{R}$ diferenciables en a relativamente a A. Para cualesquiera escalares $\alpha, \beta \in \mathbb{R}$, la función $u = \alpha f + \beta g : A \to \mathbb{R}$ es diferenciable en a y satisface

$$\boxed{u'_A(a) = \alpha f'_A(a) + \beta g'_A(a).}$$

Demostración. Consideremos el conjunto de incrementos admisibles $H_a = \{h \in \mathbb{R} : a+h \in A\}$. Para $h \in H_a \setminus \{0\}$, la definición puntual de u y las propiedades distributivas de los números reales dan la identidad exacta

$$\begin{aligned}\frac{u(a+h)-u(a)}{h} &= \frac{\alpha f(a+h)+\beta g(a+h)-\alpha f(a)-\beta g(a)}{h}\\ &= \alpha\frac{f(a+h)-f(a)}{h}+\beta\frac{g(a+h)-g(a)}{h}.\end{aligned}$$

Por hipótesis, los dos cocientes del último miembro convergen a $f'_A(a)$ y $g'_A(a)$, respectivamente, cuando $h \to 0$ a través de $H_a \setminus \{0\}$. El álgebra de límites funcionales, demostrada en el capítulo 4, permite multiplicarlos por los escalares fijos y sumarlos. En consecuencia,

$$\lim_{\substack{h\to 0\\ h\in H_a\setminus\{0\}}} \frac{u(a+h)-u(a)}{h} = \alpha f'_A(a)+\beta g'_A(a) \in \mathbb{R}.$$

El límite es finito, por lo que u es diferenciable en a según la definición del capítulo 6, y el número obtenido es precisamente su derivada. □

Lectura de la prueba. El primer paso es un hecho de álgebra de *valores*: la diferencia de una combinación lineal se separa en las diferencias de sus componentes. El segundo paso es un hecho de *análisis*: cada cociente tiene un límite por una hipótesis explícita. Solo después se usa el álgebra de límites. La existencia de la derivada de u es una **conclusión**; no se presupone para manipular el cociente. Tampoco hemos usado la regla del producto, que aún no está demostrada.

Suma, resta y multiplicación por constantes: una sola regla. El teorema comprende varias fórmulas familiares, sin necesidad de repetir la demostración. Si $\alpha = \beta = 1$, $(f+g)'_A(a) = f'_A(a) + g'_A(a)$.

Si $\alpha = 1$ y $\beta = -1$, obtenemos $(f-g)'_A(a) = f'_A(a) - g'_A(a)$.

Si $\beta = 0$, la combinación es αf, y el resultado se reduce a $(\alpha f)'_A(a) = \alpha f'_A(a)$. El caso $\alpha = 0$ no requiere dividir por α ni tratarlo como una excepción: el teorema lo incluye. Una función constante c tiene derivada 0 por cálculo directo desde la definición (capítulo 6), pues su cociente incremental es siempre 0. Así, añadir una constante modifica el valor central y la altura de la tangente, pero no su pendiente: $(f+c)'_A(a) = f'_A(a)$. Esta última afirmación es *puntual*: solo requiere que f sea diferenciable en a. No afirma por sí misma que f sea diferenciable en los demás puntos de A.

Verificar una fórmula por dos caminos.

Ejemplo 7.2. **Una combinación de dos funciones conocidas.** Sean $f(x) = x^2$ y $g(x) = x$ sobre $\mathbb{R}$, y formemos $u(x) = 3f(x) - 2g(x) = 3x^2 - 2x$.

Fijemos un $a \in \mathbb{R}$ cualquiera. **Primera ruta: desde la definición.** Para $h \neq 0$,

$$\begin{aligned}\frac{u(a+h)-u(a)}{h} &= \frac{3(a+h)^2-2(a+h)-(3a^2-2a)}{h}\\ &= \frac{6ah+3h^2-2h}{h}\\ &= 6a+3h-2.\end{aligned}$$

El límite cuando $h \to 0$ existe y vale $6a - 2$. Por tanto, $u'(a) = 6a - 2$.

Segunda ruta: aplicar el teorema recién demostrado. Los cocientes incrementales de las funciones de partida, calculados en C07, dan $f'(a) = 2a$ y $g'(a) = 1$. Como ambas son diferenciables en a y comparten dominio,

$$u'(a) = 3f'(a) - 2g'(a) = 3(2a) - 2(1) = 6a - 2.$$

Las dos rutas coinciden. La primera verifica el cálculo directamente; la segunda ilustra cómo un teorema demostrado ahorra rehacer expansiones y límites cada vez. La segunda ruta **no** sustituye la prueba general: solo es legítima porque Teorema 7.1 ya ha sido establecido.

¿Por qué también funcionan las combinaciones finitas? El teorema trata dos funciones, pero una suma finita no debe quedar como una extrapolación informal. Sean $f_1, \dots, f_n : A \to \mathbb{R}$ diferenciables en un mismo punto de acumulación $a \in A$, y $c_1, \dots, c_n \in \mathbb{R}$, con $n \geq 1$. Se afirma que

$$\left(\sum_{k=1}^{n} c_k f_k\right)'_A(a) = \sum_{k=1}^{n} c_k f'_{k,A}(a).$$

Demostración por inducción sobre n. Para $n = 1$, es el caso de multiplicación escalar del teorema. Supongamos demostrada la afirmación para n funciones. La suma parcial $s_n = \sum_{k=1}^{n} c_k f_k$ es entonces diferenciable en a y tiene derivada $\sum_{k=1}^{n} c_k f'_{k,A}(a)$. Aplicando Teorema 7.1 a s_n y f_{n+1} con coeficientes 1 y c_{n+1}, concluimos que $s_n + c_{n+1} f_{n+1}$ es diferenciable y que

$$\begin{aligned}(s_n + c_{n+1} f_{n+1})'_A(a) &= s'_{n,A}(a) + c_{n+1} f'_{n+1,A}(a) \\ &= \sum_{k=1}^{n+1} c_k f'_{k,A}(a).\end{aligned}$$

Esto cierra el paso inductivo. Se trata de una **suma finita**; no hemos afirmado que se pueda intercambiar derivación y una serie infinita, problema que requerirá hipótesis adicionales en el Tomo II. Si todas las funciones son diferenciables en cada punto de un conjunto $E \subseteq A$ cuyos puntos sean de acumulación de A, podemos reunir las igualdades puntuales en una igualdad de funciones sobre E:

$$D(\alpha f + \beta g) = \alpha Df + \beta Dg \qquad \text{en } E.$$

Aquí Df es la función derivada donde está definida; D no introduce una segunda definición de derivada. La fórmula funcional **no** extiende automáticamente las derivadas a puntos fuera de E.

Las hipótesis no son decoración. El dominio común del teorema garantiza que los tres cocientes —el de f, el de g y el de $\alpha f + \beta g$— se calculan mediante los mismos incrementos admisibles. Si dos funciones vienen dadas sobre dominios A y B distintos, primero se define su suma sobre $C = A \cap B$. Solo podemos aplicar la regla cuando $a \in C$ es punto de acumulación de C y ambas restricciones a C son diferenciables en a.

Por ejemplo, tomemos

$$A = \{0\} \cup \{1/n : n \in \mathbb{N},\ n \geq 1\}, \qquad B = \{0\} \cup \{-1/n : n \in \mathbb{N},\ n \geq 1\}.$$

El punto 0 es de acumulación de ambos conjuntos. Las funciones constantes nulas sobre A y sobre B son diferenciables en 0 con derivada relativa 0. Sin embargo, $A \cap B = \{0\}$, cuyo único punto es aislado; de acuerdo con nuestra definición, **la suma definida exclusivamente en la intersección no tiene derivada relativa en** 0. No es una falla de la fórmula algebraica, sino la desaparición de los incrementos con los que se define el límite.

En cambio, en el extremo $a = 0$ de $A = [0, \infty)$, dos funciones diferenciables relativamente a A sí satisfacen el teorema con incrementos $h > 0$; no hace falta inventar extensiones ni suponer derivadas bilaterales.

Qué se ha probado y qué no
La diferenciabilidad de **ambas** funciones sobre el mismo dominio en a es una hipótesis suficiente para que su combinación lineal sea diferenciable allí. No es una condición necesaria: por ejemplo, $f(x) = |x|$ y $g(x) = -|x|$ no son diferenciables en 0, mientras $f + g = 0$ sí lo es. Por tanto, de la diferenciabilidad de una suma no se deduce que cada sumando sea diferenciable. Tampoco se deduce ninguna regla para fg o para $g \circ f$; esas operaciones tienen incrementos de otra forma.

Reconstruir la estrategia. La identidad de incrementos es algebraica; el paso a la derivada utiliza los límites de los cocientes de cada función. Un coeficiente cero elimina el término correspondiente, y la inducción extiende la regla de dos sumandos a cualquier número finito. Son estas justificaciones las que permiten reutilizar la fórmula. En un producto, el incremento no se separa inmediatamente en dos diferencias independientes. Introduciremos un término intermedio para obtener una identidad útil y utilizaremos continuidad para controlar el factor que conserva un argumento variable.

7.3. La regla del producto: separar una variación en dos

Una diferencia que no se deja repartir directamente. En §7.1 calculamos, desde la definición, que el producto de dos funciones con pendiente 1 puede tener pendiente 2. Ya sabemos, por tanto, que la fórmula $(fg)' = f'g'$ es falsa en general. El problema no se resuelve memorizando una expresión diferente: necesitamos descubrir una **identidad algebraica verdadera** para el incremento del producto. Sean $f, g : A \to \mathbb{R}$ y $a \in A$. Para un incremento h tal que $a + h \in A$, escribamos abreviadamente

$$\Delta f = f(a + h) - f(a), \qquad \Delta g = g(a + h) - g(a).$$

El producto cambia desde $f(a)g(a)$ hasta $f(a+h)g(a+h)$. Si intentamos describir esa variación solo mediante $\Delta f\, \Delta g$, perdemos el efecto de los valores iniciales. Una manera de recuperar toda la información consiste en efectuar el cambio **en dos etapas**: primero modificamos un factor y después el otro. Introducimos el valor intermedio $f(a + h)g(a)$, que pertenece a la misma aritmética real y no requiere ninguna hipótesis analítica:

$$\begin{aligned} f(a+h)g(a+h) - f(a)g(a) &= f(a+h)g(a+h) - f(a+h)g(a) \\ &\quad + f(a+h)g(a) - f(a)g(a) \\ &= f(a+h)\big(g(a+h) - g(a)\big) \\ &\quad + g(a)\big(f(a+h) - f(a)\big). \end{aligned}$$

Esta identidad vale para **todos** los incrementos admisibles, incluso $h = 0$. Solo al dividirla por h exigiremos $h \neq 0$. La prueba general podrá construirse ahora con resultados de límites que ya conocemos.

Teorema 7.2. **Teorema — Regla del producto.** Sean $A \subseteq \mathbb{R}$, $a \in A$ un punto de acumulación de A y $f, g : A \to \mathbb{R}$ diferenciables en a relativamente a A. Entonces el producto puntual $p = fg : A \to \mathbb{R}$ es diferenciable en a y

$$\boxed{(fg)'_A(a) = f'_A(a)g(a) + f(a)g'_A(a).}$$

Demostración. Sea

$$H_a = \{h \in \mathbb{R} : a + h \in A\}.$$

Para cada $h \in H_a \setminus \{0\}$, la identidad anterior y la división legítima por h proporcionan

$$\begin{aligned}\frac{p(a+h) - p(a)}{h} &= f(a+h)\frac{g(a+h) - g(a)}{h} \\ &\quad + g(a)\frac{f(a+h) - f(a)}{h}.\end{aligned}$$

Identifiquemos **cada límite antes de combinarlo**. Por hipótesis de diferenciabilidad,

$$\frac{f(a+h) - f(a)}{h} \longrightarrow f'_A(a), \qquad \frac{g(a+h) - g(a)}{h} \longrightarrow g'_A(a)$$

cuando $h \to 0$ a través de $H_a \setminus \{0\}$. Además, la diferenciabilidad de f implica su continuidad relativa en a, por Teorema 6.1; por tanto,

$$f(a+h) \longrightarrow f(a).$$

El número $g(a)$ permanece fijo. Aplicamos ahora, y solo ahora, el álgebra de límites a los dos sumandos del cociente incremental:

$$\begin{aligned}\lim_{\substack{h\to 0 \\ h\in H_a\setminus\{0\}}} \frac{p(a+h) - p(a)}{h} &= f(a)g'_A(a) + g(a)f'_A(a) \\ &= f'_A(a)g(a) + f(a)g'_A(a).\end{aligned}$$

El resultado es un número real finito. En consecuencia, el límite que **define** $p'_A(a)$ existe y coincide con el miembro derecho de la fórmula. □

Auditoría de la prueba. La identidad algebraica produce dos términos, no el producto de dos cocientes incrementales. Las hipótesis aportan los límites de esos cocientes. El factor adicional $f(a+h)$ tiene límite porque diferenciabilidad implica continuidad, un resultado demostrado antes de este capítulo. La existencia de $(fg)'_A(a)$ es la conclusión de la prueba, no una premisa encubierta.

Leer la fórmula sin perder los valores centrales. La igualdad del teorema tiene dos contribuciones: el cambio de f ponderado por el valor de g en el centro, y el cambio de g ponderado por el valor de f en el centro. Conviene leerla como

$$\boxed{\text{variación de primer orden del producto} = \text{primer factor variable} + \text{segundo factor variable}.}$$

El orden de los sumandos es indiferente, pero sus **factores no son intercambiables arbitrariamente**: en cada término aparece una derivada y el *valor* del otro factor. La fórmula no contiene $f'_A(a)g'_A(a)$ como único término.

Ejemplo 7.3. **Verificación directa en dos funciones polinómicas.** Tomemos $f(x) = x^2$, $g(x) = x + 1$ sobre $\mathbb{R}$ y fijemos un $a \in \mathbb{R}$. Por el capítulo 6, $f'(a) = 2a$ y $g'(a) = 1$. El teorema, cuyas hipótesis se cumplen, da

$$\begin{aligned}(fg)'(a) &= 2a(a+1) + a^2 \\ &= 3a^2 + 2a.\end{aligned}$$

Comprobemos el mismo resultado **sin aplicar la regla del producto**. Como $p(x) = x^3 + x^2$, expandimos directamente para $h \neq 0$:

$$\begin{aligned}\frac{p(a+h) - p(a)}{h} &= \frac{(a+h)^3 + (a+h)^2 - a^3 - a^2}{h} \\ &= 3a^2 + 2a + (3a+1)h + h^2.\end{aligned}$$

Al tomar el límite obtenemos $p'(a) = 3a^2 + 2a$. Los dos caminos coinciden. Esta comprobación particular no reemplaza la demostración universal: muestra cómo utilizarla y cómo contrastar un cálculo cuando existe una ruta independiente.

¿Qué ocurre si un factor vale cero? El teorema **no exige** que $f(a)$ ni $g(a)$ sean distintos de cero. Si $f(a) = 0$, la fórmula se simplifica a $(fg)'_A(a) = f'_A(a)g(a)$. Si además $g(a) = 0$, entonces $(fg)'_A(a) = 0$, aunque las derivadas de los factores puedan ser no nulas. Por ejemplo, $f(x) = g(x) = x$ en $a = 0$ tiene $f'(0) = g'(0) = 1$, pero $(x^2)'(0) = 0$. Esto no contradice el contraejemplo de §7.1, que utilizó el **mismo par de funciones en otro punto**, $a = 1$. Los valores centrales son parte indispensable de la regla.

De dos factores a un producto finito. La regla también permite derivar un producto de un número finito de factores sin inventar una fórmula nueva para cada cantidad. Si $f_1, \dots, f_n : A \to \mathbb{R}$ son diferenciables en a, con $n \geq 2$, entonces el producto $P = \prod_{j=1}^{n} f_j$ es diferenciable en a y

$$\boxed{P'_A(a) = \sum_{k=1}^{n} f'_{k,A}(a) \prod_{\substack{1 \leq j \leq n \\ j \neq k}} f_j(a).}$$

Demostración por inducción. Para $n = 2$ es el teorema recién probado. Supongamos conocida la afirmación para n factores y pongamos $P_n = \prod_{j=1}^{n} f_j$. Por la hipótesis inductiva, P_n es diferenciable. La regla del producto aplicada a $P_{n+1} = P_n f_{n+1}$ produce

$$\begin{aligned}P'_{n+1,A}(a) &= P'_{n,A}(a) f_{n+1}(a) + P_n(a) f'_{n+1,A}(a) \\ &= \sum_{k=1}^{n+1} f'_{k,A}(a) \prod_{\substack{1 \leq j \leq n+1 \\ j \neq k}} f_j(a).\end{aligned}$$

La última igualdad reúne los n términos inductivos y el término correspondiente a $k = n + 1$. Así queda justificada la fórmula para todo $n \geq 2$. **No hemos probado** una regla para productos infinitos; la inducción solo permite un número finito de factores.

Una segunda lectura mediante la aproximación afín. Podemos comprobar conceptualmente por qué aparecen precisamente los dos términos del teorema. La linealización demostrada en el capítulo 6 permite escribir, para incrementos admisibles,

$$f(a+h) = f(a) + f'_A(a)h + r_f(h),$$
$$g(a+h) = g(a) + g'_A(a)h + r_g(h),$$

con $r_f(h) = o(h)$ y $r_g(h) = o(h)$. Multiplicamos las expresiones y sustraemos $f(a)g(a)$. Los términos lineales en h son exactamente $\big(f'_A(a)g(a) + f(a)g'_A(a)\big)h$.

¿Por qué lo demás es $o(h)$? Los términos $f(a)r_g(h)$ y $g(a)r_f(h)$ son $o(h)$; los productos de h por un residuo y el producto de residuos también son $o(h)$ porque $r_f(h)/h \to 0$, $r_g(h)/h \to 0$ y, por consiguiente, ambos residuos tienden a cero. El término $f'_A(a)g'_A(a)h^2$ es asimismo $o(h)$. La interpretación afín confirma así la misma fórmula sin asumir ninguna regla posterior. No la necesitamos para establecer el teorema —la prueba por cocientes ya lo hizo—, pero revela por qué el producto de dos aproximaciones de primer orden conserva dos contribuciones lineales.

Dominios distintos, extremos y límites de la inferencia. Para aplicar la regla sobre un **dominio común** debe haber incrementos admisibles arbitrariamente pequeños en ese dominio. Si $f : A \to \mathbb{R}$ y $g : B \to \mathbb{R}$, el producto tiene dominio $C = A \cap B$; en $a \in C$ verificaremos que a es punto de acumulación de C y que las restricciones $f|_C$ y $g|_C$ son diferenciables relativamente a C. La diferenciabilidad en los dominios originales no garantiza por sí sola la acumulación en su intersección. En un extremo, por ejemplo $a = 0$ para $A = [0, \infty)$, la prueba permanece válida con incrementos admisibles $h > 0$. No hemos supuesto derivadas bilaterales ni prolongaciones fuera de A.

La conclusión no se puede invertir
Que fg sea diferenciable en a no implica que **ambos** factores lo sean. En $\mathbb{R}$, la función $f(x) = |x|$ no es diferenciable en 0, mientras que $g(x) = 0$ y $fg = 0$ sí lo son. La regla demostrada es una **condición suficiente** para la diferenciabilidad del producto, no una caracterización necesaria de los factores.

Antes de continuar. Reconstruye la prueba sin mirar la fórmula: ¿qué valor agregamos y sustraemos para separar el incremento del producto? ¿Qué límite exige invocar derivabilidad ⇒ continuidad? ¿Por qué la expresión tiene sentido incluso si $f(a) = 0$ o $g(a) = 0$? ¿Qué cambia si las funciones se han definido sobre dominios distintos? Identificar estas cuatro decisiones prepara el siguiente problema: **el recíproco y el cociente exigirán además controlar que un denominador no se anule cerca del punto**. Esa será la tarea de §7.4.

7.4. Recíprocos y cocientes: la hipótesis que impide dividir por cero

Antes de derivar, determinar dónde existe la función. Para formar el recíproco tenemos que excluir los ceros de g: la expresión $1/g(x)$ solo define un número real cuando $g(x) \neq 0$. La primera comprobación consiste en identificar el dominio de la función que queremos derivar. Sean $g : A \to \mathbb{R}$ y $a \in A$, donde $A \subseteq \mathbb{R}$ y a es punto de acumulación de A. Definamos el dominio natural del recíproco: $A_g = \{x \in A : g(x) \neq 0\}$. Aunque $g(a) \neq 0$ garantiza que $a \in A_g$, todavía debemos justificar que hay puntos de A_g arbitrariamente próximos a a y que allí los denominadores no se anulan. La diferenciabilidad de g proporciona precisamente el control que necesitamos.

La no anulación local no es una suposición adicional. Supongamos que g es diferenciable relativamente a A en a y que $g(a) \neq 0$. Por el teorema «diferenciabilidad implica continuidad» (Teorema 6.1), g es continua relativamente a A en a. Apliquemos la definición de continuidad

con la tolerancia positiva $\varepsilon = |g(a)|/2$. Existe $\delta > 0$ tal que, para todo $x \in A$ con $|x-a| < \delta$, $|g(x) - g(a)| < \frac{|g(a)|}{2}$. Por la desigualdad triangular inversa,

$$|g(x)| \geq |g(a)| - |g(x) - g(a)| > \frac{|g(a)|}{2} > 0.$$

En particular, $g(x) \neq 0$ para todos esos puntos. Hemos obtenido una **cota inferior cuantitativa** para el valor absoluto del denominador; no hemos supuesto que el cociente exista fuera del dominio A_g. Además, A y A_g coinciden dentro del entorno relativo $A \cap (a-\delta, a+\delta)$. Como a es punto de acumulación de A, también lo es de A_g. Esta coincidencia local permite usar la derivada de g relativa a A al calcular límites a través de A_g. No hemos extendido la función a puntos nuevos: solo hemos reconocido que la restricción $g|_{A_g}$ conserva, cerca de a, los mismos incrementos admisibles.

Proposición 7.1. **Proposición — Derivada del recíproco.** Sean $A \subseteq \mathbb{R}$, $a \in A$ un punto de acumulación y $g : A \to \mathbb{R}$ diferenciable en a relativamente a A. Si $g(a) \neq 0$, entonces a es punto de acumulación de $A_g = \{x \in A : g(x) \neq 0\}$, la función

$$r : A_g \to \mathbb{R}, \qquad r(x) = \frac{1}{g(x)},$$

es diferenciable en a relativamente a A_g y satisface

$$\boxed{r'_{A_g}(a) = -\frac{g'_A(a)}{g(a)^2}.}$$

Demostración. La continuidad de g y la condición $g(a) \neq 0$ garantizan, por el argumento precedente, que existe $\delta > 0$ tal que $g(a+h) \neq 0$ cuando $a+h \in A$ y $|h| < \delta$. También garantizan que a es punto de acumulación de A_g.

Para cada incremento $h \neq 0$ admisible suficientemente pequeño, la identidad algebraica

$$\frac{1}{g(a+h)} - \frac{1}{g(a)} = \frac{g(a) - g(a+h)}{g(a+h)g(a)}$$

es válida porque **ambos denominadores son no nulos**. Dividiendo ahora por h obtenemos

$$\frac{r(a+h) - r(a)}{h} = -\frac{g(a+h) - g(a)}{h} \, \frac{1}{g(a+h)g(a)}.$$

El primer factor converge a $-g'_A(a)$ por definición de derivada; la coincidencia local de A y A_g justifica el dominio del límite. Por continuidad, $g(a+h) \to g(a)$, y el límite del producto de denominadores es $g(a)^2 \neq 0$. El teorema del límite de un recíproco, que es un resultado previo sobre límites y **no** la regla de derivación que estamos probando, permite concluir

$$\lim_{\substack{h \to 0 \\ a+h \in A_g,\ h \neq 0}} \frac{1}{g(a+h)g(a)} = \frac{1}{g(a)^2}.$$

Multiplicando los dos límites se sigue que el cociente incremental de r converge a $-g'_A(a)/g(a)^2$, un número real finito. Esto demuestra la diferenciabilidad y la fórmula. □

Lectura de la prueba. La hipótesis $g(a) \neq 0$ actúa en tres lugares: define el recíproco en el centro, proporciona no anulación local mediante continuidad y permite tomar el límite del

denominador. La derivada no se obtuvo «derivando el uno y el denominador»: se obtuvo de una identidad de diferencias, seguida de límites previamente justificados.

Un recíproco en todo su dominio.

Ejemplo 7.4. **La función $r(x) = 1/(1+x^2)$.** Definamos $g(x) = 1+x^2$ sobre $\mathbb{R}$. Como $x^2 \geq 0$, se cumple $g(x) \geq 1 > 0$ para todo $x \in \mathbb{R}$. Por tanto, el dominio natural de r es todo $\mathbb{R}$. La derivada de x^2, calculada directamente en el capítulo 6, y la linealidad de §7.2 dan $g'(a) = 2a$.

Aplicando la proposición, válida en cada $a \in \mathbb{R}$, obtenemos

$$\boxed{r'(a) = -\frac{2a}{(1+a^2)^2}.}$$

En particular, $r'(0) = 0$ y $r'(1) = -1/2$. Podemos verificar el primer valor independientemente: para $h \neq 0$,

$$\frac{r(h) - r(0)}{h} = \frac{\frac{1}{1+h^2} - 1}{h} = -\frac{h}{1+h^2} \longrightarrow 0.$$

Aquí la no anulación es global y explícita. El teorema general, en cambio, solo requiere controlar el denominador cerca del punto considerado.

El cociente se construye a partir del producto y del recíproco. Ahora sean $f, g : A \to \mathbb{R}$ diferenciables relativamente a A en a, con $g(a) \neq 0$. El cociente original tiene dominio

$$A_g = \{x \in A : g(x) \neq 0\}, \qquad q : A_g \to \mathbb{R}, \quad q(x) = \frac{f(x)}{g(x)}.$$

Por el control de no anulación, A_g coincide con A suficientemente cerca de a. En particular, las restricciones de f y g a A_g tienen en a las mismas derivadas relativas que las funciones originales. Podemos escribir entonces $q = (f|_{A_g})r$, donde $r = 1/(g|_{A_g})$, y aplicar **en ese orden** la proposición recién demostrada y la regla del producto de §7.3.

Teorema 7.3. **Teorema — Regla del cociente.** Sean $A \subseteq \mathbb{R}$, $a \in A$ un punto de acumulación y $f, g : A \to \mathbb{R}$ diferenciables en a relativamente a A. Si $g(a) \neq 0$, el cociente $q = f/g$, definido en $A_g = \{x \in A : g(x) \neq 0\}$, es diferenciable en a relativamente a A_g y

$$\boxed{q'_{A_g}(a) = \frac{f'_A(a)g(a) - f(a)g'_A(a)}{g(a)^2}.}$$

Demostración. La no anulación local prueba que a es punto de acumulación de A_g y que las restricciones de f y g conservan sus derivadas en a. Por Proposición 7.1, $r = 1/g$ es diferenciable en a sobre A_g y

$$r'_{A_g}(a) = -\frac{g'_A(a)}{g(a)^2}.$$

Como $q = (f|_{A_g})r$, el teorema del producto (Teorema 7.2) proporciona tanto la existencia de $q'_{A_g}(a)$ como la igualdad

$$q'_{A_g}(a) = f'_A(a)\frac{1}{g(a)} + f(a)\left(-\frac{g'_A(a)}{g(a)^2}\right)$$
$$= \frac{f'_A(a)g(a) - f(a)g'_A(a)}{g(a)^2}.$$

Solo se ha utilizado la regla del producto **después** de disponer de la derivada del recíproco, de manera que no hay circularidad. □

Una forma útil de reconstruir la fórmula es partir de la identidad exacta

$$\frac{f(a+h)}{g(a+h)} - \frac{f(a)}{g(a)} = \frac{g(a)\big(f(a+h)-f(a)\big) - f(a)\big(g(a+h)-g(a)\big)}{g(a+h)g(a)}.$$

La identidad, válida para incrementos admisibles suficientemente pequeños, explica el **signo menos** del numerador y muestra por qué el denominador final es $g(a)^2$. Al dividir por h y aplicar los mismos límites que antes se obtiene directamente el teorema; esta lectura sirve para controlar la fórmula sin memorizarla. No hemos supuesto que $g'_A(a)$ sea distinto de cero: esa condición no interviene en absoluto.

Una simplificación algebraica no rellena un hueco del dominio.

Ejemplo 7.5. **Un cociente con discontinuidad removible.** Consideremos

$$q(x) = \frac{x^2-1}{x-1}, \qquad A_q = \mathbb{R}\setminus\{1\}.$$

Si $x \in A_q$, la factorización $x^2-1 = (x-1)(x+1)$ permite cancelar el factor $x-1$ y concluir que $q(x) = x+1$. De ello se sigue que $q'(a) = 1$ para **todo** $a \in A_q$, ya que cerca de cada uno de esos puntos el cociente coincide con una función afín. La regla recién probada lo verifica también: para $a \neq 1$,

$$q'(a) = \frac{2a(a-1) - (a^2-1)}{(a-1)^2}$$
$$= \frac{(a-1)^2}{(a-1)^2} = 1.$$

En $a = 1$ la función **no está definida**; no tiene valor ni derivada allí. La función distinta $\widetilde{q}:\mathbb{R}\to\mathbb{R}$, $\widetilde{q}(x) = x+1$, es una extensión continua de q y sí tiene $\widetilde{q}'(1) = 1$. Debemos distinguir las propiedades de la extensión de las de la función original: simplificar una expresión no cambia retroactivamente el dominio especificado.

Condiciones suficientes, extremos y errores de interpretación. La fórmula del cociente no dice que «se deriva arriba y abajo». Por ejemplo, $f(x) = x^2$ y $g(x) = x$ determinan el cociente $q(x) = x$ en $\mathbb{R}\setminus\{0\}$. En $a = 2$, su derivada es $q'(2) = 1$, mientras que $f'(2)/g'(2) = 4$. La expresión f'/g' no es la regla del cociente, ni siquiera cuando $g'(a) \neq 0$. Si a es un extremo de A, por ejemplo $A = [0,\infty)$ y $a = 0$, todas las pruebas anteriores funcionan con los incrementos admisibles de ese dominio, siempre que $g(0) \neq 0$. No afirman una derivada bilateral para prolongaciones no especificadas.

Finalmente, nuestras hipótesis son **suficientes**, no necesarias para que un cociente particular sea diferenciable. En $\mathbb{R}$, las funciones $f(x) = g(x) = 1+|x|$ no son diferenciables en 0; sin embargo, como $g(x) > 0$ en todo punto, $f/g = 1$ sí lo es. Esto no invalida el teorema: sencillamente impide invertir su implicación.

Control antes de aplicar la regla del cociente
Comprueba por orden: (1) el dominio efectivo A_g; (2) que el punto pertenece a él y es de acumulación; (3) diferenciabilidad de ambos factores en el dominio pertinente; (4) $g(a) \neq 0$; (5) los valores y las derivadas en el **mismo punto**. No añadas puntos excluidos mediante cancelaciones ni dividas por $g'(a)$.

La regla del cociente ha quedado deducida a partir de resultados anteriores. En §7.5 utilizaremos la regla del producto, junto con la linealidad, para **demostrar por inducción** la derivada de las potencias naturales y, después, la fórmula de los polinomios. Las potencias enteras negativas y las funciones racionales se desarrollarán en §7.6, una vez construidas esas herramientas.

7.5. Potencias naturales y derivación de polinomios

Una fórmula conocida todavía necesita una prueba. Las reglas establecidas hasta aquí permiten derivar sumas y productos de funciones diferenciables. Podemos, por ejemplo, obtener la derivada de $x^2 = x \cdot x$ mediante la regla del producto, pero **no** hemos demostrado todavía una fórmula para una potencia de exponente arbitrario. Escribir de inmediato $(x^n)' = nx^{n-1}$ y utilizar esa misma fórmula para justificarla sería un argumento circular.

El paso decisivo es reconocer una construcción finita: para cada entero $n \geq 1$, la función $x \mapsto x^{n+1}$ es el producto de $x \mapsto x^n$ y la función identidad. Esto permite utilizar una inducción, con una hipótesis inductiva que incluye **la existencia de la derivada**, y no solamente una igualdad formal de expresiones.

Del producto a todas las potencias naturales positivas.

Proposición 7.2. **Proposición — Derivada de una potencia natural.** Para cada entero $n \geq 1$, la función $p_n : \mathbb{R} \to \mathbb{R}$ definida por $p_n(x) = x^n$ es diferenciable en todo $a \in \mathbb{R}$ y satisface

$$\boxed{p_n'(a) = na^{n-1}.}$$

Aquí, cuando $n = 1$, la expresión a^0 designa el valor de la **función constante** $x \mapsto 1$, también en $a = 0$.

Demostración por inducción sobre n. Sea $P(n)$ la afirmación completa «p_n es diferenciable en cada punto real y $p_n'(a) = na^{n-1}$ para todo $a \in \mathbb{R}$».

Caso inicial, $n = 1$. La función $p_1(x) = x$ es la identidad. Para cualquier $a \in \mathbb{R}$ y $h \neq 0$,

$$\frac{p_1(a+h) - p_1(a)}{h} = \frac{(a+h) - a}{h} = 1.$$

El límite existe y vale $1 = 1 \cdot a^0$, según la convención recién explicitada. Por tanto, $P(1)$ es verdadera.

Paso inductivo. Supongamos verdadera $P(n)$ para cierto $n \geq 1$. Para cada $x \in \mathbb{R}$ tenemos la identidad de funciones

$$p_{n+1}(x) = p_n(x)\, p_1(x) = x^n x.$$

Ambos factores son diferenciables en cualquier a: el primero por hipótesis inductiva y el segundo por el caso inicial. La regla del producto, demostrada en Teorema 7.2, prueba **también la existencia** de la derivada de p_{n+1} y proporciona

$$\begin{aligned} p'_{n+1}(a) &= p'_n(a)p_1(a) + p_n(a)p'_1(a) \\ &= na^{n-1}a + a^n \\ &= (n+1)a^n. \end{aligned}$$

La igualdad algebraica $a^{n-1}a = a^n$ vale igualmente para $a = 0$ con $n \geq 1$ y a^0 interpretado como la función constante 1. Queda probada $P(n+1)$. El principio de inducción concluye la afirmación para todos los enteros $n \geq 1$. □

Lectura de la prueba. El paso inductivo no deriva una potencia mediante una regla que aún desconocemos: deriva un **producto** utilizando Teorema 7.2. De ese teorema obtenemos simultáneamente la existencia de la derivada y su valor. La inducción construye una familia de resultados, uno para cada exponente natural positivo; no es una autorización para sustituir n por un exponente real cualquiera.

Podemos comprobar el primer paso sin acudir a la fórmula general: $D(x^2) = D(x \cdot x) = 1 \cdot x + x \cdot 1 = 2x$. A continuación, $D(x^3) = D(x^2 \cdot x) = 2x \cdot x + x^2 \cdot 1 = 3x^2$. Son aplicaciones particulares de una demostración que ya cubre **todos** los enteros positivos, no evidencias aisladas que debamos extrapolar.

El exponente cero: una función constante, no una potencia problemática. Para construir polinomios necesitamos también el término constante. Lo definimos sin recurrir a una expresión indeterminada: la función $p_0 : \mathbb{R} \to \mathbb{R}$ es $p_0(x) = 1$ para **todo** $x \in \mathbb{R}$. Su cociente incremental es

$$\frac{p_0(a+h) - p_0(a)}{h} = \frac{1-1}{h} = 0, \qquad h \neq 0,$$

por lo que $p'_0(a) = 0$ en todo punto. Esta definición funcional permite escribir convencionalmente $x^0 = 1$ dentro de un polinomio incluso en $x = 0$; **no estamos evaluando una expresión 0^0 cuyo significado no se haya fijado**. La fórmula na^{n-1} se ha demostrado para $n \geq 1$; no se obtiene el caso $n = 0$ sustituyendo mecánicamente $n = 0$ en ella.

Si restringimos p_n o p_0 a un conjunto $A \subseteq \mathbb{R}$, sus derivadas relativas en cualquier $a \in A$ que sea punto de acumulación de A coinciden con las derivadas recién demostradas: el límite relativo utiliza solamente algunos de los incrementos para los que ya existe el límite sobre $\mathbb{R}$. Esto incluye extremos de dominio y conjuntos no intervalares, siempre que exista acumulación.

Un polinomio es una suma *finita* de monomios.

Corolario 7.1. **Corolario — Derivada de un polinomio.** Sean $m \in \mathbb{N} \cup \{0\}$ y $c_0, \ldots, c_m \in \mathbb{R}$. La función polinómica

$$p : \mathbb{R} \to \mathbb{R}, \qquad p(x) = \sum_{k=0}^{m} c_k x^k,$$

con x^0 entendido como la función constante 1, es diferenciable en todo punto real y

$$\boxed{p'(a) = \sum_{k=1}^{m} k c_k a^{k-1}.}$$

Cuando $m = 0$, la suma del miembro derecho es vacía y su valor es 0.

Demostración. El término $c_0 p_0$ es constante y tiene derivada nula. Para cada $1 \le k \le m$, la proposición Proposición 7.2 demuestra que $p_k(x) = x^k$ es diferenciable en todo punto y que $p_k'(a) = ka^{k-1}$. La regla de las combinaciones lineales finitas, establecida en §7.2 a partir de Teorema 7.1, asegura que la suma $p = \sum_{k=0}^{m} c_k p_k$ es diferenciable y que su derivada es la suma de las derivadas:

$$\begin{aligned} p'(a) &= c_0 p_0'(a) + \sum_{k=1}^{m} c_k p_k'(a) \\ &= 0 + \sum_{k=1}^{m} k c_k a^{k-1}. \end{aligned}$$

Si $m = 0$, solo interviene la función constante y el mismo razonamiento da $p'(a) = 0$. La conclusión vale para cada $a \in \mathbb{R}$. □

El orden deductivo importa. Una función polinómica no se declara diferenciable por tener una expresión algebraica familiar: hemos probado la diferenciabilidad de cada monomio y después la de su suma finita. Tampoco se ha intercambiado derivación con una **serie infinita**. El número de sumandos $m + 1$ es finito, una condición esencial para aplicar sin hipótesis adicionales el resultado de §7.2.

Si dos fórmulas algebraicas expresan la misma función polinómica, el cálculo no puede producir derivadas diferentes: en cada punto donde existe, la derivada es el límite único del cociente incremental de esa función. Por eso expandir y reunir términos semejantes puede modificar la presentación de una expresión sin modificar su derivada. La simplificación, sin embargo, no autoriza a cambiar el dominio de una función racional; ese problema ya fue examinado en §7.4.

Una verificación por dos caminos.

Ejemplo 7.6. **Un polinomio de quinto grado.** Sea

$$p(x) = 2x^5 - 3x^3 + 4x - 7, \qquad x \in \mathbb{R}.$$

El dominio es todo $\mathbb{R}$. La proposición para potencias y la linealidad finita dan, en cualquier $a \in \mathbb{R}$,

$$\begin{aligned} p'(a) &= 2(5a^4) - 3(3a^2) + 4(1) - 0 \\ &= \boxed{10a^4 - 9a^2 + 4}. \end{aligned}$$

En particular, $p'(0) = 4$ y $p'(1) = 5$. **Control independiente en $a = 0$.** Para $h \neq 0$,

$$\frac{p(h) - p(0)}{h} = \frac{2h^5 - 3h^3 + 4h}{h} = 2h^4 - 3h^2 + 4 \longrightarrow 4.$$

Segundo control, en $a = 1$. La expansión puramente algebraica, previa al límite, proporciona

$$\frac{p(1+h) - p(1)}{h} = 5 + 11h + 17h^2 + 10h^3 + 2h^4 \longrightarrow 5.$$

Los dos cocientes confirman valores particulares de la fórmula general sin utilizarla durante sus verificaciones. La prueba de diferenciabilidad para todo a, en cambio, ya está garantizada por el corolario y no depende de revisar unos cuantos puntos.

Qué hemos demostrado y qué queda fuera. Las igualdades de esta sección pueden reunirse en una identidad de funciones sobre $\mathbb{R}$: si $p(x) = \sum_{k=0}^{m} c_k x^k$, entonces $Dp(x) = \sum_{k=1}^{m} k c_k x^{k-1}$. La igualdad representa el conjunto de afirmaciones puntuales que acabamos de probar; D no es aquí una regla adicional.

Para una función diferenciable arbitraria f, la regla de productos finitos de §7.3 permite deducir también $(f^n)' = nf^{n-1}f'$ en el dominio pertinente. No la utilizaremos como premisa de la proposición sobre x^n: allí la inducción se apoyó únicamente en resultados anteriores. Su relación con la composición se retomará después, al demostrar la regla de la cadena. Tampoco hemos tratado x^{-n} para exponentes negativos, $x^{1/2}$ o x^{α} con exponente real: sus dominios e hipótesis requieren un examen específico. Los exponentes enteros negativos y las funciones racionales serán el asunto inmediato de §7.6; las potencias con otros exponentes corresponden a desarrollos posteriores.

Reconstruir la deducción antes de continuar
Explica por qué el caso inicial de la inducción debe demostrar *existencia* y *valor* de la derivada. En el paso $x^{n+1} = x^n x$, indica qué resultado permite afirmar que el nuevo producto es diferenciable. Después, localiza el único lugar de la prueba del corolario donde se utiliza que un polinomio tiene un número **finito** de términos. Finalmente, explica por qué x^0 se trata como función constante y por qué el corolario no autoriza aún la derivación de x^{-3} en 0.

Las potencias negativas requieren dividir por x, y un cociente de polinomios exige excluir los ceros de su denominador. Esas restricciones determinarán el dominio sobre el que podrán aplicarse las reglas.

7.6. Potencias enteras negativas y funciones racionales

El exponente negativo no elimina una división. En §7.5 demostramos la derivada de x^n para cada entero $n \geq 1$, y tratamos x^0 como la función constante 1. Al pasar a exponentes negativos aparece una diferencia que no podemos ocultar bajo la semejanza de las fórmulas: para $n \geq 1$, la escritura x^{-n} significa el **recíproco** de x^n. Antes de derivar, hay que decidir dónde existe ese recíproco.

Para cualquier entero $n \geq 1$, definimos

$$u_{-n} : \mathbb{R} \setminus \{0\} \longrightarrow \mathbb{R}, \qquad u_{-n}(x) = x^{-n} = \frac{1}{x^n}.$$

En un cuerpo, $x^n \neq 0$ si $x \neq 0$: es un producto finito de factores no nulos. En cambio, para $x = 0$ el denominador vale cero. Así, **el dominio es** $\mathbb{R} \setminus \{0\}$, aunque una manipulación posterior produzca una expresión aparentemente más simple. Todo punto de ese dominio es de acumulación; además, en torno a cada $a \neq 0$ podemos escoger un intervalo que no contenga el origen, por ejemplo $|x - a| < |a|/2$. La semejanza entre exponentes positivos y negativos debe, por tanto, construirse con una prueba; no se obtiene sustituyendo un entero negativo en una proposición demostrada solo para $n \geq 1$.

Del recíproco a las potencias negativas. El resultado de §7.4 ya contiene la herramienta necesaria. Si tomamos $g(x) = x^n$, sabemos por Proposición 7.2 que g es diferenciable en todo punto real y que $g'(a) = na^{n-1}$. Para $a \neq 0$ también sabemos que $g(a) = a^n \neq 0$. En consecuencia, la proposición del recíproco (Proposición 7.1) es aplicable **en ese punto**. No necesitamos inventar

una nueva regla ni recurrir a la regla de la cadena, que todavía no hemos probado. La derivada se obtiene como

$$\begin{aligned} u'_{-n}(a) &= -\frac{g'(a)}{g(a)^2} \\ &= -\frac{na^{n-1}}{a^{2n}} \\ &= -n\,a^{-n-1}, \qquad a \neq 0. \end{aligned}$$

La última igualdad utiliza únicamente las leyes de exponentes enteros **para una base no nula**. En $a = 0$ no hay fórmula que aplicar: $u_{-n}(0)$ no existe, de modo que tampoco existe una derivada de esa función en el origen. La regla de exponentes negativos es una conclusión válida en su dominio, no una autorización para añadir el punto excluido.

De los polinomios a sus cocientes. Una función racional se construye escogiendo dos polinomios reales p y q, con q no idénticamente nulo, y definiendo la función sobre el conjunto

$$D_{p,q} = \{x \in \mathbb{R} : q(x) \neq 0\}, \qquad r : D_{p,q} \to \mathbb{R}, \qquad r(x) = \frac{p(x)}{q(x)}.$$

El conjunto $D_{p,q}$ se determina con el **denominador de la expresión que define la función**; una factorización posterior puede facilitar los cálculos, pero no modifica por sí sola ese dominio. Ambos polinomios son diferenciables en toda la recta, por Corolario 7.1, y por tanto son continuos. Si $a \in D_{p,q}$, la continuidad de q y $q(a) \neq 0$ proporcionan un entorno de a donde q no se anula. Ese entorno está contenido en $D_{p,q}$: en particular, a es punto de acumulación de su dominio natural. Estamos exactamente en las hipótesis de la regla del cociente de §7.4. Podemos reunir las dos consecuencias en un enunciado que mantiene visibles sus dominios.

Corolario 7.2. **Corolario — Potencias enteras negativas y funciones racionales.** Se cumplen las siguientes afirmaciones.

1. Para cada entero $n \geq 1$, la función $u_{-n}(x) = x^{-n}$ definida en $\mathbb{R} \setminus \{0\}$ es diferenciable en todo punto $a \neq 0$ y

$$\boxed{u'_{-n}(a) = -na^{-n-1}.}$$

2. Si $p, q : \mathbb{R} \to \mathbb{R}$ son polinomios y q no es el polinomio nulo, la función $r = p/q$, definida sobre $D_{p,q} = \{x : q(x) \neq 0\}$, es diferenciable en cada $a \in D_{p,q}$ y

$$\boxed{r'(a) = \frac{p'(a)q(a) - p(a)q'(a)}{q(a)^2}.}$$

En la segunda fórmula, p' y q' son los polinomios derivados obtenidos por Corolario 7.1.

Demostración. Para la primera afirmación, sea $n \geq 1$ y sea $a \neq 0$. La proposición Proposición 7.2 da la derivabilidad de $g(x) = x^n$ y $g'(a) = na^{n-1}$. Como $g(a) = a^n \neq 0$, Proposición 7.1 demuestra la existencia de la derivada de $1/g$ en a sobre su dominio natural y proporciona

$$u'_{-n}(a) = -\frac{na^{n-1}}{a^{2n}} = -na^{-n-1}.$$

Para la segunda, fijemos $a \in D_{p,q}$. Ambos polinomios son diferenciables en a por Corolario 7.1 y $q(a) \neq 0$ por la definición del dominio. La no anulación local establecida en §7.4 asegura

que a es de acumulación de $D_{p,q}$. Aplicar Teorema 7.3 a p y q demuestra simultáneamente la diferenciabilidad de r sobre ese dominio y la fórmula indicada. Como a era arbitrario, la conclusión vale en cada punto de $D_{p,q}$. □

Auditoría de dependencias. La primera parte se apoya en potencias naturales y recíproco; la segunda, en polinomios y cociente. La regla de la cadena no aparece en ninguna prueba. La hipótesis $q \not\equiv 0$ permite hablar de un cociente racional no vacío en el sentido habitual; la derivada solo se afirma en puntos donde $q(a) \neq 0$. No se requiere que $q'(a)$ sea no nulo.

Dos cálculos y dos controles independientes.

Ejemplo 7.7. **Una potencia negativa y un cociente con puntos excluidos.** Comencemos con

$$u(x) = x^{-3} = \frac{1}{x^3}, \qquad \mathrm{Dom}(u) = \mathbb{R} \setminus \{0\}.$$

El primer apartado del corolario, con $n = 3$, da

$$\boxed{u'(x) = -3x^{-4} = -\frac{3}{x^4}, \qquad x \neq 0.}$$

Así, $u'(-2) = -3/16$. Como control **independiente** del cálculo en $a = 1$, usemos directamente el cociente incremental. Para $h \neq 0$ suficientemente pequeño y $1 + h \neq 0$,

$$\begin{aligned} \frac{u(1+h) - u(1)}{h} &= \frac{(1+h)^{-3} - 1}{h} \\ &= \frac{1 - (1+h)^3}{h(1+h)^3} \\ &= -\frac{3 + 3h + h^2}{(1+h)^3} \longrightarrow -3. \end{aligned}$$

Hemos cancelado h solo para $h \neq 0$ y comprobado por separado que el denominador restante tiende a 1, no a cero. El límite reproduce $u'(1) = -3$ sin invocar la fórmula que pretendemos controlar.

Consideremos ahora $r(x) = \frac{x^2+1}{x^2-4}$.

La ecuación $x^2 - 4 = (x - 2)(x + 2) = 0$ excluye exactamente dos puntos. Por tanto, $\mathrm{Dom}(r) = \mathbb{R} \setminus \{-2, 2\}$. Los polinomios $p(x) = x^2 + 1$ y $q(x) = x^2 - 4$ cumplen $p'(x) = 2x$ y $q'(x) = 2x$. Para todo x del dominio, la segunda parte del corolario da

$$\begin{aligned} r'(x) &= \frac{2x(x^2 - 4) - (x^2 + 1)2x}{(x^2 - 4)^2} \\ &= \boxed{-\frac{10x}{(x^2 - 4)^2}}, \qquad x \neq -2, 2. \end{aligned}$$

En particular, $r'(1) = -10/9$. También podemos comprobarlo directamente, puesto que $r(1) = -2/3$. Para $h \neq 0$ suficientemente pequeño,

$$\frac{r(1+h)-r(1)}{h} = \frac{1}{h}\left(\frac{2+2h+h^2}{-3+2h+h^2}+\frac{2}{3}\right)$$
$$= \frac{10+5h}{3(-3+2h+h^2)} \longrightarrow -\frac{10}{9}.$$

El control directo confirma la derivada en 1. En -2 y 2 la función original no está definida; ninguna de las fórmulas permite atribuirle allí un valor ni una derivada.

Simplificar una expresión y respetar la función son tareas distintas. El cálculo racional exige distinguir dos operaciones. Una **identidad algebraica** puede permitir cancelar factores cuando son no nulos; una **extensión de dominio**, en cambio, exige definir explícitamente una función nueva. En §7.4 ya analizamos $\frac{x^2-1}{x-1}$: coincide con $x+1$ solo en su dominio original, que excluye 1. Su derivada en ese punto sigue sin estar definida aunque la expresión simplificada admita una extensión diferenciable. También es incorrecto deducir de la fórmula del cociente que el denominador derivado deba satisfacer $q'(a) \neq 0$. El requisito es $q(a) \neq 0$. Por ejemplo, en el cociente de nuestro ejemplo, $q'(0) = 0$ y $q(0) = -4 \neq 0$; la regla es plenamente aplicable en 0 y produce $r'(0) = 0$.

Si se restringen p y q a un conjunto $A \subseteq \mathbb{R}$, el cociente restringido se define en $A \cap D_{p,q}$. Para hablar de su derivada relativa en a es necesario que a pertenezca a este conjunto y sea punto de acumulación de A. La condición $q(a) \neq 0$ garantiza entonces que A y $A \cap D_{p,q}$ coinciden suficientemente cerca de a, de modo que la fórmula del corolario sigue siendo válida como derivada relativa. **En un punto aislado de A no se define derivada relativa**, aunque el polinomio original tenga derivada bilateral en toda la recta.

La frontera exacta de la regla de potencias. Con §7.5 y esta sección hemos demostrado $(x^m)' = mx^{m-1}$ para enteros positivos m sobre $\mathbb{R}$, para $m = 0$ mediante la función constante, y para enteros negativos sobre $\mathbb{R} \setminus \{0\}$. Esto no demuestra la fórmula para exponentes fraccionarios o reales. Por ejemplo, antes de escribir una derivada para $\sqrt{x}$ necesitamos fijar su dominio y justificar su diferenciabilidad en los puntos pertinentes; no basta sustituir formalmente $m = 1/2$ en un teorema de exponentes enteros.

Control de lectura: primero dominio, después fórmula
Explica qué resultado garantiza $x^n \neq 0$ cuando $x \neq 0$; identifica en qué paso se usa la proposición del recíproco y por qué no interviene todavía la cadena. En el ejemplo racional, determina dónde se comprueba $q(a) \neq 0$ y por qué $q'(0) = 0$ no causa ningún problema. Por último, distingue la afirmación «una expresión simplificada tiene derivada en a» de «la función original tiene derivada en a».

La derivación de potencias enteras y racionales queda así reducida a resultados ya probados. En §7.7 estudiaremos una operación diferente: **componer funciones**. Allí el incremento de la función interior puede valer cero aunque el incremento de la variable independiente no lo sea; una demostración de la regla de la cadena deberá funcionar también en ese caso, sin dividir por una diferencia posiblemente nula.

7.7. La regla de la cadena: cuando el incremento interior puede anularse

Dos cambios de escala, una sola derivada. La suma, el producto y el cociente combinan valores de funciones calculados en una misma entrada. La composición hace algo diferente: la salida de una función se convierte en la entrada de otra. Sean $f : A \to B \subseteq \mathbb{R}$ y $g : B \to \mathbb{R}$.

Para estudiar $g \circ f$ en un punto $a \in A$ debemos seguir el recorrido $a \longmapsto b = f(a) \longmapsto g(b)$. Si desplazamos a hasta $a + h$, el incremento de la función interior es $\Delta f = f(a + h) - f(a)$. La función exterior recibe **ese** incremento, no necesariamente h. El problema consiste en relacionar la variación de g entre b y $b + \Delta f$ con la variación de f entre a y $a + h$.

La aproximación afín del capítulo 6 sugiere el resultado: una variación pequeña h produce, a primer orden, una variación $f'_A(a)h$; la función exterior transforma esta última, a primer orden, en $g'_B(b)f'_A(a)h$. Deberíamos obtener una derivada igual al producto de esas pendientes, con la exterior evaluada en $b = f(a)$. Esta interpretación orienta la prueba, pero no la reemplaza.

La división tentadora y el punto donde falla. Partamos del cociente incremental de la composición. Una manipulación habitual intenta escribir

$$\frac{g(f(a+h)) - g(f(a))}{h} \stackrel{?}{=} \frac{g(f(a+h)) - g(f(a))}{f(a+h) - f(a)} \frac{f(a+h) - f(a)}{h}.$$

Cuando $f(a + h) \neq f(a)$ y $h \neq 0$, esta identidad es correcta. **La diferenciabilidad de f no garantiza, sin embargo, que $f(a + h) \neq f(a)$ para todo incremento pequeño no nulo.** Una función puede repetir el valor central en una sucesión de puntos próximos, o incluso en todo un lado del punto. Entonces el primer factor carece de sentido: presenta denominador cero y numerador cero. Tampoco basta afirmar que su límite existe, porque la expresión propuesta no está definida para esos incrementos. La reparación adecuada consiste en dar un valor justificado al cociente exterior *cuando su entrada es exactamente el punto central*. Esto se puede hacer antes de estudiar cualquier composición.

Un cociente exterior que también tiene valor en el centro.

LEMA 7.1. **Lema — Extensión continua del cociente incremental exterior.** Sean $B \subseteq \mathbb{R}$, $b \in B$ un punto de acumulación y $g : B \to \mathbb{R}$ diferenciable en b relativamente a B. Definamos $K : B \to \mathbb{R}$ mediante

$$K(t) = \begin{cases} \dfrac{g(t) - g(b)}{t - b}, & t \in B,\ t \neq b, \\ g'_B(b), & t = b. \end{cases}$$

Entonces K es continua en b relativamente a B, y para **todo** $t \in B$ se cumple la identidad

$$\boxed{g(t) - g(b) = K(t)(t - b).}$$

Demostración. La definición de la derivada de g afirma que

$$\lim_{\substack{t \to b \\ t \in B \setminus \{b\}}} \frac{g(t) - g(b)}{t - b} = g'_B(b).$$

Para $t \neq b$, $K(t)$ coincide con el cociente que aparece en este límite; por definición, $K(b) = g'_B(b)$. Dado $\varepsilon > 0$, la existencia del límite proporciona $\delta > 0$ tal que, si $t \in B$ y $0 < |t - b| < \delta$, entonces $|K(t) - K(b)| < \varepsilon$. Si $t = b$, la misma desigualdad vale porque $|K(b) - K(b)| = 0$. Se ha probado, también en el centro, la condición de continuidad relativa de K en b. Para $t \neq b$, multiplicar la primera rama de su definición por $t - b$ recupera exactamente $g(t) - g(b)$. Para $t = b$, ambos miembros de la identidad valen cero. La igualdad es, por tanto, válida en todo B, sin excepción. □

La extensión no consiste en asignar arbitrariamente un número a una fracción $0/0$. La derivada previamente existente determina el **único valor que hace continua a K en b**: si otro valor produjera continuidad, el límite del cociente sobre $B \setminus \{b\}$ tendría que coincidir también con él, en contradicción con la unicidad del límite. Es importante distinguir el cociente original, definido solo fuera del centro, de la nueva función K, definida allí también.

La regla general y su demostración.

Teorema 7.4. **Teorema — Regla de la cadena.** Sean $A, B \subseteq \mathbb{R}$, $f : A \to B$ y $g : B \to \mathbb{R}$. Supongamos que $a \in A$ es punto de acumulación de A y que $b = f(a) \in B$ es punto de acumulación de B. Si f es diferenciable en a relativamente a A y g es diferenciable en b relativamente a B, entonces $g \circ f : A \to \mathbb{R}$ es diferenciable en a y satisface

$$\boxed{(g \circ f)'_A(a) = g'_B(f(a))\, f'_A(a).}$$

Demostración. Sea $K : B \to \mathbb{R}$ la función construida en Lema 7.1 para g y $b = f(a)$. Su identidad fundamental, aplicada al valor $t = f(a+h) \in B$, da

$$g(f(a+h)) - g(f(a)) = K(f(a+h))\big(f(a+h) - f(a)\big)$$

para **todo** incremento h admisible, incluso cuando $f(a+h) = f(a)$. Definamos $H_a = \{h \in \mathbb{R} : a + h \in A\}$. Solo dividimos por h cuando $h \in H_a \setminus \{0\}$; así obtenemos la identidad exacta

$$\frac{(g \circ f)(a+h) - (g \circ f)(a)}{h} = K(f(a+h))\, \frac{f(a+h) - f(a)}{h}.$$

Verifiquemos separadamente los dos límites del segundo miembro. Por la diferenciabilidad de f,

$$\frac{f(a+h) - f(a)}{h} \longrightarrow f'_A(a).$$

Además, diferenciabilidad implica continuidad (Teorema 6.1), de modo que $f(a+h) \to f(a) = b$ a través de A. El lema establece que K es continua en b **incluido su valor central**. La composición de funciones continuas, ya demostrada en el capítulo 5, permite concluir que

$$K(f(a+h)) \longrightarrow K(b) = g'_B(b).$$

Aquí no se exige que $f(a+h) \neq b$: cuando se produce la igualdad, $K(f(a+h)) = K(b)$ está perfectamente definido. Aplicando el teorema del límite de un producto a los dos factores obtenemos

$$\lim_{\substack{h \to 0 \\ h \in H_a \setminus \{0\}}} \frac{(g \circ f)(a+h) - (g \circ f)(a)}{h} = g'_B(b) f'_A(a) \in \mathbb{R}.$$

El cociente incremental de la composición posee, por tanto, un límite real finito. Esta es justamente la existencia de su derivada relativa en a, y su valor es el anunciado. □

Lectura de la demostración en sentido inverso. La conclusión exige demostrar que converge el cociente incremental de $g \circ f$. Para aplicar el álgebra de límites necesitábamos factorizarlo. La identidad del lema suministra esa factorización sin dividir por Δf. Para controlar el primer factor utilizamos continuidad de K y de f; para el segundo, la definición de $f'_A(a)$. La

derivada exterior aparece evaluada en $f(a)$ porque ese es el punto alrededor del cual cambia su argumento; evaluarla en a sería, en general, incorrecto o incluso carecería de sentido si $a \notin B$.

La prueba es puntual y respeta dominios arbitrarios, incluidos extremos. No presupone que f sea inyectiva, estrictamente monótona ni localmente distinta de su valor central. Las hipótesis son suficientes; no se ha demostrado su necesidad para toda composición particular.

Un polinomio compuesto: verificar por dos rutas.

Ejemplo 7.8. **La potencia cuarta de un polinomio cuadrático.** Definamos $f : \mathbb{R} \to \mathbb{R}$ y $g : \mathbb{R} \to \mathbb{R}$ por

$$f(x) = 3x^2 - 1, \qquad g(t) = t^4.$$

Ambas funciones son diferenciables en toda la recta por §7.5. Su composición es $u(x) = (3x^2-1)^4$. En un punto arbitrario a, tenemos $f'(a) = 6a$, $f(a) = 3a^2 - 1$ y $g'(f(a)) = 4(3a^2-1)^3$. La regla de la cadena demuestra la existencia de $u'(a)$ y da

$$\boxed{u'(a) = 24a(3a^2 - 1)^3.}$$

En particular, $u'(1) = 24 \cdot 2^3 = 192$. Para controlar el resultado sin usar la cadena, expandamos algebraicamente **antes** de derivar:

$$u(x) = 81x^8 - 108x^6 + 54x^4 - 12x^2 + 1.$$

El corolario polinómico Corolario 7.1, demostrado antes de la regla de la cadena, produce

$$u'(x) = 648x^7 - 648x^5 + 216x^3 - 24x.$$

Esta expresión coincide con $24x(3x^2-1)^3$ por expansión ordinaria; en $x = 1$ también da $648 - 648 + 216 - 24 = 192$. Son dos rutas lícitas que conducen a la misma derivada. La primera muestra la estructura de composición; la segunda sirve como verificación independiente y utiliza únicamente una regla anterior.

Un caso en el que dividir por el incremento interior destruye la prueba.

Ejemplo 7.9. **Una función interior que repite su valor central en todo un semieje.** Consideremos

$$f : \mathbb{R} \to [0, \infty), \qquad f(x) = \begin{cases} 0, & x \le 0, \\ x^2, & x > 0, \end{cases} \qquad g : [0, \infty) \to \mathbb{R}, \quad g(t) = t^2 + 1.$$

El punto de estudio es $a = 0$ y su imagen es $b = f(0) = 0$. Ambos son puntos de acumulación de los respectivos dominios. Para $h < 0$ tenemos $f(h) - f(0) = 0$, aunque $h \neq 0$; por ello la división tentadora por el incremento interior no está definida en **ninguno** de esos incrementos. Calculemos las derivadas necesarias desde la definición. Para $h \neq 0$,

$$\frac{f(h) - f(0)}{h} = \begin{cases} 0, & h < 0, \\ h, & h > 0, \end{cases} \longrightarrow 0.$$

Por tanto, $f'_{\mathbb{R}}(0) = 0$. Para $t > 0$, los incrementos admisibles de g en el extremo $b = 0$ son positivos y $\frac{g(t)-g(0)}{t} = \frac{t^2}{t} = t \longrightarrow 0$. Así, $g'_{[0,\infty)}(0) = 0$. El teorema recién demostrado proporciona $(g \circ f)'_{\mathbb{R}}(0) = 0 \cdot 0 = 0$ sin división ilegítima. Comprobémoslo directamente. La composición es

$$(g\circ f)(x)=\begin{cases}1, & x\le 0,\\ 1+x^4, & x>0.\end{cases}$$

Su cociente incremental en 0 vale 0 cuando $h<0$ y h^3 cuando $h>0$; ambos tienden a cero. En este ejemplo la función K del lema es $K(t)=t$ para $t>0$ y $K(0)=0$. La identidad $g(t)-g(0)=K(t)t$ es válida aun en $t=0$ y permite factorizar el cociente de la composición para **todos** los $h\neq 0$. Precisamente ahí reside la reparación de la demostración.

Qué hipótesis debemos conservar al reutilizarla. La notación compacta $(g\circ f)'(a)=g'(f(a))f'(a)$ omite información indispensable si se escribe fuera de un contexto fijado. Antes de utilizarla, hay que identificar un dominio A donde f esté definida, un conjunto B que contenga **toda** la imagen pertinente de f, y el dominio de g. Debemos comprobar además que a y $f(a)$ sean puntos de acumulación de los dominios respectivos y que las dos derivadas indicadas existan como límites reales finitos. Una función exterior definida solo en un punto aislado no tiene derivada allí bajo nuestra convención, aunque alguna composición específica resulte diferenciable.

En un extremo de A, la derivada de la composición es relativa a los incrementos disponibles. Si $f(a)$ es un extremo de B, también puede utilizarse una derivada exterior unilateral relativa a B. El teorema no atribuye derivadas bilaterales a extensiones que no se hayan definido. En cambio, puede ocurrir que $f(A)$ acumule o no en $f(a)$: la prueba no exige una derivada exterior relativa a la imagen efectiva si ya disponemos de $g'_B(f(a))$ en un dominio B apropiado.

La regla de la cadena no es una regla de cancelación
La fórmula final multiplica dos derivadas, pero no se obtiene cancelando formalmente símbolos df. Su demostración utiliza una identidad válida en todos los incrementos admisibles y dos límites justificados. En particular, no exige $f'_A(a)\neq 0$ ni que $f(a+h)-f(a)$ sea no nulo.

Reconstrucción de la prueba. Intenta demostrar el teorema partiendo solo de su conclusión: ¿qué cociente incremental debe converger?, ¿cuál es la división que no está permitida en general?, ¿cómo se define K en el punto omitido y qué propiedad justifica ese valor?, ¿qué teorema garantiza $f(a+h)\to f(a)$?, ¿en qué punto se evalúa finalmente g'? Repite después la prueba sobre el ejemplo por tramos, donde todos los incrementos negativos anulan la variación interior. Si la demostración sigue siendo válida, has identificado el mecanismo esencial. Al derivar composiciones sucesivas habrá que conservar cada punto intermedio: la derivada exterior se evalúa en la imagen de la etapa anterior, y cada factor corresponde a una de esas entradas.

7.8. Composiciones repetidas sin perder los puntos de evaluación

Una cadena tiene más de un punto de referencia. En una composición de varias funciones, cada etapa transporta la entrada a un punto nuevo. La fórmula debe incluir tanto el factor de cada derivada como su argumento correcto; calcular todos los factores en el punto inicial daría, en general, otro número. Supongamos, por ejemplo, que tenemos funciones correctamente encadenadas $A_0\xrightarrow{f_1}A_1\xrightarrow{f_2}A_2\xrightarrow{f_3}A_3$, y fijemos $x_0=a\in A_0$. Los puntos que intervienen son

$$x_1=f_1(x_0),\qquad x_2=f_2(x_1),\qquad x_3=f_3(x_2).$$

La derivada de f_1 se evalúa en x_0; la de f_2, en x_1; y la de f_3, en x_2. Escribir $f_3'(a)f_2'(a)f_1'(a)$ no está justificado: incluso podría ser una expresión indefinida si a no pertenece a A_1 o a A_2. Para tres eslabones, la regla de §7.7, aplicada primero a $f_2\circ f_1$ y después a $f_3\circ(f_2\circ f_1)$, sugiere

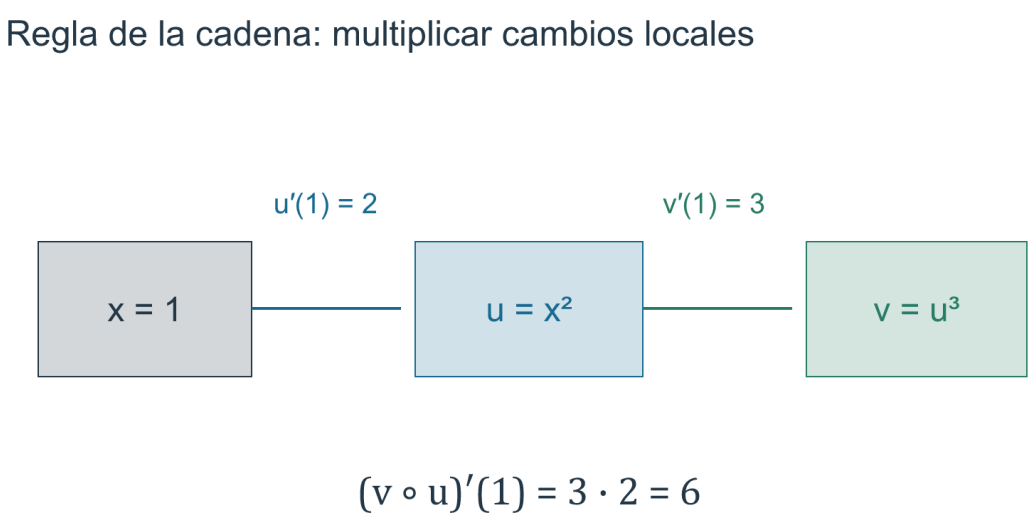


FIGURA 7.1. En $u(x) = x^2$ y $v(u) = u^3$, la entrada $x = 1$ produce $u = 1$. Los factores se evalúan en sus respectivos puntos: $(v \circ u)'(1) = v'(u(1))u'(1) = 3 \cdot 2 = 6$.

$$(f_3 \circ f_2 \circ f_1)'_{A_0}(a) = f'_{3,A_2}(x_2)\, f'_{2,A_1}(x_1)\, f'_{1,A_0}(x_0).$$

La fórmula es una consecuencia de dos aplicaciones sucesivas, no una nueva regla admitida por analogía. Para extenderla a cualquier cantidad **finita** de eslabones, debemos controlar tanto las derivadas como los dominios durante una inducción.

Teorema para una cantidad finita de composiciones.

PROPOSICIÓN 7.3. **Proposición — Regla de la cadena finita.** Sea $n \geq 1$ un entero, sean $A_0, \ldots, A_n \subseteq \mathbb{R}$ y, para $1 \leq k \leq n$, sean funciones

$$f_k : A_{k-1} \longrightarrow A_k.$$

Fijemos $x_0 = a \in A_0$ y definamos sucesivamente

$$x_k = f_k(x_{k-1}), \qquad 1 \leq k \leq n.$$

Supongamos que, para cada $1 \leq k \leq n$, el punto x_{k-1} es de acumulación de A_{k-1} y f_k es diferenciable en x_{k-1} relativamente a ese conjunto. Entonces la composición

$$F = f_n \circ f_{n-1} \circ \cdots \circ f_1 : A_0 \longrightarrow A_n$$

es diferenciable en a relativamente a A_0 y satisface

$$\boxed{F'_{A_0}(a) = \prod_{k=1}^{n} f'_{k,A_{k-1}}(x_{k-1}).}$$

Demostración por inducción. Llamemos $F_m = f_m \circ \cdots \circ f_1 : A_0 \to A_m$ para $1 \le m \le n$. Demostraremos simultáneamente que F_m es diferenciable en a y que

$$F'_{m,A_0}(a) = \prod_{k=1}^{m} f'_{k,A_{k-1}}(x_{k-1}).$$

Para $m = 1$, $F_1 = f_1$; la existencia de su derivada es una hipótesis, y el producto de un único factor es $f'_{1,A_0}(a)$. Supongamos establecida la afirmación para un $m < n$. Por la definición de los puntos intermedios,

$$F_m(a) = f_m(\cdots f_1(a)\cdots) = x_m.$$

La hipótesis inductiva garantiza que $F_m : A_0 \to A_m$ es diferenciable en a. La hipótesis correspondiente a $k = m+1$ garantiza que x_m es punto de acumulación de A_m y que $f_{m+1} : A_m \to A_{m+1}$ es diferenciable allí. Por tanto, **se cumplen las dos hipótesis de diferenciabilidad y las dos de acumulación** del teorema de la cadena Teorema 7.4. Aplicándolo a $F_{m+1} = f_{m+1} \circ F_m$, obtenemos la existencia de $F'_{m+1,A_0}(a)$ y

$$\begin{aligned} F'_{m+1,A_0}(a) &= f'_{m+1,A_m}\big(F_m(a)\big)\, F'_{m,A_0}(a) \\ &= f'_{m+1,A_m}(x_m) \prod_{k=1}^{m} f'_{k,A_{k-1}}(x_{k-1}) \\ &= \prod_{k=1}^{m+1} f'_{k,A_{k-1}}(x_{k-1}). \end{aligned}$$

Esto cierra el paso inductivo. La afirmación se sigue para $m = n$. En ningún paso hemos dividido por una derivada ni por un incremento interior: utilizamos exclusivamente la regla de §7.7, ya demostrada para incrementos interiores posiblemente nulos. □

Lectura de las hipótesis. La acumulación se exige en $x_0, \ldots, x_{n-1}$, pues en esos puntos se calculan las derivadas de $f_1, \ldots, f_n$, respectivamente. No se necesita exigir que el valor final x_n sea punto de acumulación de A_n: **no estamos derivando ninguna función adicional en** x_n. Tampoco exigimos que cada f_k sea diferenciable en todos los puntos de su dominio; bastan los puntos de evaluación indicados. La propiedad es puntual, aun cuando la función compuesta se haya definido sobre todo A_0.

La prueba tiene una frontera precisa: demuestra una regla para n **finito**. No autoriza a derivar una composición infinita ni a tomar un producto infinito de pendientes. Esos procesos requerirían nuevas nociones y controles de convergencia.

Laboratorio de cuatro eslabones: desarmar antes de derivar. Consideremos

$$H(x) = \big(2(3x-2)^2 + 1\big)^3, \qquad x \in \mathbb{R}.$$

En lugar de interpretar la expresión como un bloque indivisible, definamos cuatro funciones, cada una con dominio y codominio $\mathbb{R}$:

$$\begin{aligned} f_1(t) &= 3t - 2, \\ f_2(t) &= t^2, \\ f_3(t) &= 2t + 1, \\ f_4(t) &= t^3. \end{aligned}$$

Todas son polinómicas y, por Corolario 7.1, diferenciables en toda la recta. La composición tipada es $H = f_4 \circ f_3 \circ f_2 \circ f_1$. Fijemos un punto arbitrario a. El **árbol de dependencias** —en este caso, una cadena lineal de entradas y salidas— conserva los cuatro valores diferentes:

$$\begin{aligned} x_0 &= a, \\ x_1 &= 3a - 2, \\ x_2 &= (3a-2)^2, \\ x_3 &= 2(3a-2)^2 + 1, \\ x_4 &= \big(2(3a-2)^2+1\big)^3. \end{aligned}$$

La tabla de evaluación de las derivadas se reconstruye sin adivinar:

Eslabón	Derivada de la función	Evaluación correcta
f_1	$f_1'(t) = 3$	$f_1'(x_0) = 3$
f_2	$f_2'(t) = 2t$	$f_2'(x_1) = 2(3a-2)$
f_3	$f_3'(t) = 2$	$f_3'(x_2) = 2$
f_4	$f_4'(t) = 3t^2$	$f_4'(x_3) = 3\big(2(3a-2)^2+1\big)^2$

Como se satisfacen todas las hipótesis de Proposición 7.3, podemos multiplicar los cuatro valores. El resultado es

$$\begin{aligned} H'(a) &= f_4'(x_3) f_3'(x_2) f_2'(x_1) f_1'(x_0) \\ &= 3\big(2(3a-2)^2+1\big)^2 \cdot 2 \cdot 2(3a-2) \cdot 3 \\ &= \boxed{36(3a-2)\big(2(3a-2)^2+1\big)^2}. \end{aligned}$$

En $a = 1$, los puntos intermedios son $x_1 = 1$, $x_2 = 1$ y $x_3 = 3$, de donde $H'(1) = 3 \cdot 2 \cdot 2 \cdot 3 \cdot 3^2 = 324$. **Control independiente:** el polinomio interior completo es $u(x) = 18x^2 - 24x + 9$. Como $H = u \cdot u \cdot u$, la fórmula para productos finitos de §7.3 —demostrada *antes* de la regla de la cadena— y el corolario polinómico permiten obtener

$$H'(x) = 3u(x)^2 u'(x) = 3(18x^2 - 24x + 9)^2 (36x - 24),$$

que coincide con la expresión anterior porque $u(x) = 2(3x-2)^2 + 1$ y $36x - 24 = 12(3x-2)$. Este control usa el producto de tres funciones y la derivada de un polinomio, **no** la regla para composiciones que estamos comprobando.

Dos errores que parecen inofensivos. Evaluar todas las derivadas en a. Tomemos simplemente $f(x) = 3x - 2$, $g(t) = t^2$ y $a = 2$. Como $f(2) = 4$, la cadena da $(g \circ f)'(2) = g'(4) f'(2) = 8 \cdot 3 = 24$. El producto incorrecto $g'(2) f'(2) = 4 \cdot 3 = 12$ evalúa la derivada exterior en otro punto. Las funciones son diferenciables en toda $\mathbb{R}$: el fallo no está en las hipótesis, sino en el argumento utilizado.

Olvidar que un factor puede ser cero. Si, en alguna composición finita que satisfaga la proposición, $f_j'(x_{j-1}) = 0$, el producto de las derivadas vale cero. La regla sigue siendo válida: no hemos cancelado factores ni supuesto pendientes distintas de cero. Sin embargo, la igualdad $F'(a) = 0$ por sí sola no permite deducir qué factor se anula si desconocemos que todas las derivadas individuales existen; y aun conociéndolas, no permite afirmar que las funciones sean localmente constantes. La regla describe únicamente el comportamiento de primer orden.

Identidad funcional y dominios restringidos. Si las hipótesis de la proposición se verifican en todos los puntos de un conjunto $E \subseteq A_0$ que sean de acumulación de A_0, podemos reunir las fórmulas puntuales en una igualdad de funciones derivadas *sobre* E. Para cada $a \in E$, los puntos x_k deben recalcularse mediante las funciones de la cadena. No existe una lista fija de puntos intermedios independiente de a.

En dominios como $A_0 = [0, \infty)$, la conclusión en el extremo $a = 0$ se refiere a los incrementos admisibles de A_0. Las derivadas intermedias se toman relativamente a **sus propios dominios** A_{k-1}, que pueden ser distintos unos de otros. No se ha afirmado la existencia de derivadas bilaterales para extensiones arbitrarias ni se requiere que la imagen efectiva de cada composición parcial acumule en su valor si el codominio intermedio elegido ya satisface la hipótesis pertinente.

Rehacer la cadena sin memorizar una fórmula larga
Reconstruye la derivada de $H(x) = \big(2(3x-2)^2+1\big)^3$ partiendo del punto a, no de la expresión final: escribe $x_0, \dots, x_4$; asigna a cada función la entrada en la que se evalúa su derivada; comprueba por qué no necesitamos que x_4 sea de acumulación de un conjunto adicional. Después indica exactamente qué cambia en el paso inductivo al pasar de m a $m+1$ eslabones.

La regla de la cadena finita ya está demostrada. En §7.9 examinaremos el alcance lógico de las reglas: **que una composición sea diferenciable no implica, en general, que cada función empleada satisfaga todas las hipótesis suficientes de la regla**. Allí distinguiremos una conclusión verdadera de una aplicación injustificada del teorema.

7.9. Lo que las reglas no afirman

Un teorema condicional no es una equivalencia. Hemos demostrado reglas que permiten construir derivadas a partir de otras derivadas previamente conocidas. Su forma lógica es siempre condicional. Por ejemplo, la regla de la cadena afirma que, bajo las hipótesis de dominio, acumulación y diferenciabilidad de Teorema 7.4, **la composición es diferenciable** y su derivada tiene una expresión determinada. No afirma que esas hipótesis sean necesarias para que *alguna* composición resulte diferenciable.

Conviene separar tres preguntas: ¿la función está definida en el punto?, ¿se han verificado las hipótesis de una regla que garantiza su diferenciabilidad?, ¿podría demostrarse su diferenciabilidad por otra vía aunque esas hipótesis fallen? Una respuesta negativa a la segunda no decide la tercera. Una respuesta negativa a la primera, en cambio, impide atribuir una derivada a la función original en ese punto.

Formalmente, de una implicación $P \Rightarrow Q$ y de la verdad de Q no se deduce P. Para refutar la conversa basta encontrar un caso en el que Q sea verdadera y P falsa. El siguiente ejemplo se ocupa de la regla de la cadena, donde esta distinción resulta especialmente instructiva.

Una composición diferenciable con una función exterior no diferenciable.

Ejemplo 7.10. **El valor absoluto queda oculto al componerlo con un cuadrado.** Consideremos las funciones completas

$$f : \mathbb{R} \longrightarrow \mathbb{R}, \qquad f(x) = x^2,$$

$$g : \mathbb{R} \longrightarrow \mathbb{R}, \qquad g(t) = |t|,$$

y fijemos $a = 0$, de modo que $b = f(a) = 0$. La función interior es diferenciable en 0: su cociente incremental es $h^2/h = h$ para $h \neq 0$, luego $f'(0) = 0$. La función exterior **no** es diferenciable en $b = 0$ relativamente a su dominio declarado $\mathbb{R}$. En efecto, para $t \neq 0$,

$$\frac{g(t) - g(0)}{t} = \frac{|t|}{t} = \begin{cases} -1, & t < 0, \\ 1, & t > 0. \end{cases}$$

Los límites laterales son distintos, por lo que $g'_{\mathbb{R}}(0)$ no existe. Sin embargo, para todo $x \in \mathbb{R}$ tenemos $x^2 \geq 0$ y, por tanto, $(g \circ f)(x) = |x^2| = x^2$. Desde la definición, el cociente incremental de la composición en 0 es h y converge a 0. Así,

$$\boxed{(g \circ f)'_{\mathbb{R}}(0) = 0, \qquad g'_{\mathbb{R}}(f(0)) \text{ no existe.}}$$

La conclusión «la composición es diferenciable» es verdadera, pero **no podemos justificarla aplicando Teorema 7.4 a estas dos funciones completas**, pues falta una de sus hipótesis. Escribir $g'_{\mathbb{R}}(0) f'_{\mathbb{R}}(0) = 0$ sería ilegítimo: un factor indefinido no adquiere valor por multiplicarlo formalmente por cero.

Un matiz importante sobre los dominios. En este ejemplo, la imagen de f está contenida en $B = [0, \infty)$. Si declaramos explícitamente una función exterior distinta, a saber, la restricción $\widetilde{g} = g|_B : B \to \mathbb{R}$, entonces $\widetilde{g}(t) = t$ para $t \geq 0$. Su derivada **relativa a** B existe en 0 y vale 1, porque los incrementos admisibles allí son positivos. La regla de la cadena sí puede aplicarse a la composición tipada $\widetilde{g} \circ f$ y da

$$(\widetilde{g} \circ f)'_{\mathbb{R}}(0) = \widetilde{g}'_B(0) f'_{\mathbb{R}}(0) = 1 \cdot 0 = 0.$$

No hay contradicción: $g'_{\mathbb{R}}(0)$ y $\widetilde{g}'_B(0)$ son afirmaciones sobre **dominios diferentes**. Restringir la función exterior puede proporcionar una nueva demostración, pero no convierte retroactivamente en verdadera la hipótesis que fallaba para $g : \mathbb{R} \to \mathbb{R}$. El ejemplo refuta la conversa formulada para las funciones completas originales. La diferenciabilidad de la composición tampoco obliga a que la **función interior** sea diferenciable. Para comprobarlo sin introducir nuevas reglas, tomemos $u : \mathbb{R} \to [0, \infty)$, $u(x) = |x|$, y $v : [0, \infty) \to \mathbb{R}$, $v(t) = t^2$. La función u no es diferenciable en 0 por sus cocientes laterales -1 y 1, pero $(v \circ u)(x) = |x|^2 = x^2$ sí lo es. La diferenciabilidad de la composición no permite reconstruir por sí sola la diferenciabilidad de cada eslabón.

Las operaciones algebraicas también pueden ocultar dificultades. La regla del producto Teorema 7.2 garantiza la diferenciabilidad de fg cuando ambos factores son diferenciables en el punto pertinente. Su conversa falla incluso si **ninguno** de los dos factores es diferenciable. Sean $f(x) = g(x) = |x|$ en $\mathbb{R}$. Sus cocientes incrementales en 0 son $|h|/h$, con límites laterales distintos. Sin embargo,

$$(fg)(x) = |x|^2 = x^2, \qquad (fg)'(0) = \lim_{h \to 0} \frac{h^2}{h} = 0.$$

La regla del producto no era aplicable en 0, aunque el producto resultante tenga derivada. Esta constatación complementa el ejemplo de §7.3, donde bastaba que uno de los factores fuese la función nula. Tampoco podemos invertir la regla del cociente Teorema 7.3. Definamos $f(x) = g(x) = 1 + |x|$ para todo $x \in \mathbb{R}$. Ninguno de los dos factores es diferenciable en 0: para $h \neq 0$ su cociente incremental vale $|h|/h$. No obstante, $g(x) \geq 1$, de manera que el cociente está definido **en toda la recta**, incluida la entrada central, y satisface

$$q(x) = \frac{1+|x|}{1+|x|} = 1, \qquad q'(0) = 0.$$

La falla de la conversa no depende aquí de dividir por cero ni de rellenar un agujero: el cociente original tiene dominio $\mathbb{R}$ y es diferenciable en 0, pese a que sus dos componentes no lo sean. Lo que no está permitido es emplear la fórmula del teorema en ese punto, pues sus derivadas de partida no existen.

Estos ejemplos permiten un diagnóstico común. Que una operación produzca una función suave en determinado punto no obliga a que los objetos de partida sean individualmente suaves allí. La operación puede cancelar o transformar las dificultades; solo la definición o un teorema cuyas hipótesis se cumplan permite establecer la conclusión en cada caso.

Una hipótesis que falla y una expresión que ni siquiera está definida. La ausencia de diferenciabilidad de un factor no equivale a la ausencia de definición de la función formada. En los ejemplos anteriores, los productos, cocientes y composiciones estaban definidos en 0; por eso tenía sentido calcular sus derivadas directamente. Un denominador nulo en el centro plantea un problema diferente. Tomemos $p(x) = x$ y $q(x) = x$ sobre $\mathbb{R}$. Su cociente natural es

$$r : \mathbb{R} \setminus \{0\} \longrightarrow \mathbb{R}, \qquad r(x) = \frac{x}{x} = 1.$$

La cancelación demuestra $r(x) = 1$ **para** $x \neq 0$, no que $r(0)$ exista. La función original no tiene derivada en 0, pues 0 no pertenece a su dominio. Podemos definir otra función $\widetilde{r} : \mathbb{R} \to \mathbb{R}$ por $\widetilde{r}(x) = 1$ y obtener $\widetilde{r}'(0) = 0$, pero ese resultado pertenece a la extensión, no al cociente original. La condición $q(a) \neq 0$ de la regla no es un formalismo prescindible: en este caso delimita dónde existe la operación.

Existe otra pérdida posible de dominio. Si $f : A \to \mathbb{R}$ y $g : B \to \mathbb{R}$ son diferenciables en a relativamente a sus respectivos dominios, su producto puntual o su suma se definen sobre $C = A \cap B$. Para utilizar una regla en a es necesario que a sea de acumulación de C, no solo de A y de B por separado. Retomemos el ejemplo de §7.2:

$$A = \{0\} \cup \{1/n : n \geq 1\}, \qquad B = \{0\} \cup \{-1/n : n \geq 1\}.$$

Las funciones nulas en cada uno de esos dominios son diferenciables relativamente a ellos en 0, pero $A \cap B = \{0\}$. El producto y la suma, considerados **únicamente sobre esa intersección**, no tienen derivada relativa en 0 conforme a nuestra definición, porque el punto ha quedado aislado. No es un contraejemplo a una regla correctamente enunciada: faltan las condiciones necesarias para formular su conclusión.

Auditar una inferencia, no solamente una fórmula. Las comprobaciones anteriores se pueden ordenar según el tipo de obstáculo:

Situación	Qué es legítimo concluir
Todas las hipótesis de la regla están verificadas.	La diferenciabilidad y la fórmula se siguen del teorema.
La función resultante está definida, pero falla una hipótesis suficiente.	El teorema no decide el caso; hay que buscar una demostración independiente o verificar una restricción apropiada.
La función resultante no está definida en el punto.	No tiene derivada allí; una extensión definida expresamente es otra función.

Situación	Qué es legítimo concluir
El dominio de una operación deja aislado el punto.	No hay derivada relativa bajo la definición adoptada en el capítulo 6.

Esta distinción también impide una inferencia inversa frecuente: si una composición diferenciable tiene derivada nula, no cabe buscar automáticamente un factor nulo en la fórmula de la cadena **antes de haber verificado que dicha fórmula es aplicable**. Cuando todas sus hipótesis sí se cumplen, el producto de derivadas es legítimo y, si vale cero, al menos un factor es cero por la propiedad del producto nulo; aun así, una derivada puntual nula no demuestra que una función sea constante en ningún entorno. Las conclusiones globales sobre el comportamiento de una función requieren otros argumentos.

Tres preguntas para diagnosticar una aplicación injustificada
En el ejemplo $|x^2|$, explica por qué $g'_{\mathbb{R}}(0)$ no existe y $\widetilde{g}'_{[0,\infty)}(0)$ sí existe. Después, para $(1+|x|)/(1+|x|)$, distingue la derivada del cociente de las derivadas de sus factores. Finalmente, identifica la diferencia entre «la regla del cociente no puede aplicarse en a» y «el cociente no es una función definida en a». ¿Qué prueba o dato necesitarías en cada caso?

Con estas distinciones, el repertorio de reglas queda cerrado en su alcance exacto: sabemos construir derivadas cuando disponemos de las hipótesis suficientes, reconocer situaciones que exigen volver a la definición y detectar cuándo la expresión ni siquiera define una función en el punto. La siguiente sección reunirá estos procedimientos en un laboratorio de reconstrucción, reparación y combinación; no introducirá reglas nuevas.

7.10. Laboratorio: reconstruir, reparar y combinar

De la respuesta a la justificación. En §§7.2–7.8 demostramos las reglas de derivación y en §7.9 examinamos lo que sus conclusiones no permiten invertir. El propósito de este laboratorio no es aprender otras ocho fórmulas. Es practicar una secuencia de decisiones: **reconocer la función, localizar el punto y su dominio, elegir una identidad o un teorema, comprobar las hipótesis y justificar el límite**. Las ocho estaciones se resuelven aquí de manera guiada; los ejercicios autónomos y sus soluciones quedarán reunidos en §7.11. En cada estación conviene intentar primero la consigna, detenerse en la dificultad indicada y cotejar después la reconstrucción. Todos los cálculos se apoyan exclusivamente en los resultados ya establecidos; no introduciremos definiciones, teoremas ni identificadores de ejemplos nuevos.

Estación 1. Leer una expresión como una construcción de funciones. Consigna. Antes de derivar, identifica el dominio y las operaciones que forman

$$R(x) = \frac{(x^2+1)(2x-1)^3}{x^2-1}.$$

¿Qué regla corresponde a la operación exterior? ¿En qué orden necesitarías las otras reglas? ¿Tiene sentido solicitar $R'(1)$ después de simplificar alguna expresión?

Reconstrucción. El numerador es un *producto*: sus factores son $p(x) = x^2+1$ y $q(x) = (2x-1)^3$. El segundo factor es a su vez una *composición*: primero $u(x) = 2x-1$ y después $v(t) = t^3$. El denominador es $d(x) = x^2-1$. Así, la operación exterior es un cociente de pq entre d; no una composición de los tres términos. El dominio natural es

$$A = \{x \in \mathbb{R} : x^2 - 1 \neq 0\} = \mathbb{R} \setminus \{-1, 1\}.$$

Para calcular $R'_A(a)$ debemos escoger $a \in A$ —todos sus puntos son de acumulación—, derivar p, d mediante las fórmulas polinómicas, q mediante la regla de la cadena, pq mediante el producto y, finalmente, aplicar el cociente. No necesitamos realizar el cálculo completo para establecer esa dependencia. La solicitud $R'(1)$ carece de sentido para **esta** función: $1 \notin A$. Una eventual extensión definida en ese punto sería una función distinta y exigiría análisis propio.

Control inverso. La mera presencia de una potencia no autoriza comenzar por una supuesta «regla de potencia para todo»: debemos reconocer qué objeto está elevado a una potencia y cuáles son sus dominios. Tampoco se puede evaluar una derivada antes de decidir qué función está definida.

Estación 2. Inventar el término intermedio del producto. Consigna. Reconstruye la regla del producto sin recordarla de memoria. Para $f, g : A \to \mathbb{R}$ diferenciables en un punto de acumulación $a \in A$, empieza por la única expresión que define la derivada de fg y separa su numerador en dos variaciones aprovechables.

Reconstrucción. Para todo incremento admisible $h \neq 0$, agregamos y sustraemos el mismo término $f(a+h)g(a)$:

$$\begin{aligned} &f(a+h)g(a+h) - f(a)g(a) \\ &\quad = f(a+h)\big(g(a+h) - g(a)\big) + g(a)\big(f(a+h) - f(a)\big). \end{aligned}$$

Dividimos **solo por** $h \neq 0$:

$$\frac{(fg)(a+h) - (fg)(a)}{h} = f(a+h)\frac{g(a+h) - g(a)}{h} + g(a)\frac{f(a+h) - f(a)}{h}.$$

La diferenciabilidad de f implica $f(a+h) \to f(a)$; los dos cocientes convergen a sus derivadas. Por las leyes de los límites obtenemos

$$(fg)'_A(a) = f'_A(a)g(a) + f(a)g'_A(a).$$

Como comprobación concreta, toma $f(x) = x^2 - 1$, $g(x) = x + 2$ y $a = 1$. Aunque $f(1) = 0$, la regla da $f'(1)g(1) + f(1)g'(1) = 2 \cdot 3 + 0 = 6$. Desde la definición, con $h \neq 0$,

$$\frac{(fg)(1+h) - (fg)(1)}{h} = \frac{(2h + h^2)(3+h)}{h} = (2+h)(3+h) \longrightarrow 6.$$

Punto de reparación. El término intermedio no se «deriva»: se suma y se resta para construir una identidad exacta. En ningún momento es necesario dividir por $f(a)$, que aquí es cero.

Estación 3. Encontrar el lugar exacto de la continuidad. Consigna. En la identidad anterior, identifica qué paso sería ilegítimo si solo conociéramos el límite del cociente incremental de g. ¿De dónde proviene el límite del factor $f(a+h)$? Reconstruye además una descomposición simétrica.

Reconstrucción. Para afirmar que $f(a+h)\frac{g(a+h)-g(a)}{h} \longrightarrow f(a)g'_A(a)$ necesitamos **dos** límites. El del cociente procede de la diferenciabilidad de g; el del factor $f(a+h)$ procede de la diferenciabilidad de f mediante «derivabilidad implica continuidad» (Teorema 6.1). El otro sumando tiende a $g(a)f'_A(a)$ porque $g(a)$ es constante respecto de h. Esta identificación muestra que la continuidad no es una hipótesis suplementaria oculta: ya está garantizada por las hipótesis de la regla.

Podemos reconstruir la misma identidad usando otro término intermedio, ahora $f(a)g(a+h)$:

$$\begin{aligned} &f(a+h)g(a+h) - f(a)g(a) \\ &\quad = g(a+h)\big(f(a+h) - f(a)\big) + f(a)\big(g(a+h) - g(a)\big). \end{aligned}$$

Aquí el factor cuyo límite requiere continuidad es $g(a+h)$, y la conclusión no cambia. Las dos rutas son válidas porque **ambas funciones** son diferenciables y, por ello, continuas en a. Si solo supiéramos que g es diferenciable, no tendríamos derecho a declarar automáticamente que el producto es diferenciable: faltaría información sobre f.

Lectura inversa. Una vez escrita la fórmula de la derivada del producto, pregunta qué términos debían converger para obtener cada sumando. Esa pregunta conduce hacia atrás a la identidad algebraica y a la continuidad necesaria; evita convertir la regla en una consigna mnemotécnica.

Estación 4. Probar que un denominador permanece separado de cero. Consigna. Sean $g : A \to \mathbb{R}$, $a \in A$ punto de acumulación, $g'_A(a)$ existente y $g(a) \neq 0$. Demuestra que el cociente f/g estará definido en todos los puntos de A suficientemente cercanos a a, siempre que $f : A \to \mathbb{R}$ esté definida. No sustituyas esta prueba por «el denominador parece no anularse».

Reconstrucción. Por Teorema 6.1, g es continua en a relativamente a A. Elegimos la tolerancia positiva $\varepsilon = |g(a)|/2$. Existe $\delta > 0$ tal que, para $x \in A$ con $|x-a| < \delta$, $|g(x) - g(a)| < \frac{|g(a)|}{2}$.

La desigualdad triangular inversa proporciona

$$|g(x)| \geq |g(a)| - |g(x) - g(a)| > \frac{|g(a)|}{2} > 0.$$

En consecuencia, $g(x) \neq 0$ en ese entorno relativo y a pertenece al dominio natural $A_g = \{x \in A : g(x) \neq 0\}$; además, A_g contiene todos los puntos de A suficientemente cercanos a a. Los cocientes incrementales relativos a A_g y a A coinciden para incrementos suficientemente pequeños. Esta es la justificación local que permite usar el recíproco y el cociente.

Variación diagnóstica. Si $g(a) = 0$, la tolerancia elegida deja de ser positiva y falla el argumento. No hemos demostrado que f/g sea imposible en todo el dominio: hemos constatado que el cociente natural **no está definido en** a, de modo que carece de derivada allí. Ni una cancelación algebraica ni la existencia de $g'(a)$ restablecen por sí solas el valor ausente.

Estación 5. Reparar la falsa derivación «arriba y abajo». Consigna. Un razonamiento propone $(f/g)' = f'/g'$. Localiza el error con un caso en el que g' no sea cero, y reconstruye la fórmula correcta sin utilizar previamente la regla del cociente.

Diagnóstico. En $A = \mathbb{R} \setminus \{0\}$ tomemos $f(x) = x^2$ y $g(x) = x$. La función cociente es $q(x) = x$ en A, y por ello $q'_A(1) = 1$. En cambio, $f'(1)/g'(1) = 2/1 = 2$. La fórmula propuesta falla aun cuando ambos miembros pretendidos están definidos. No existe una identidad que identifique el incremento del cociente con el cociente de los incrementos de sus componentes.

Reconstrucción. Para funciones diferenciables en un punto de acumulación a de un dominio común, y con $g(a) \neq 0$, la estación anterior da no anulación local. Comenzamos por el recíproco, cuya diferencia exacta es

$$\frac{1}{g(a+h)} - \frac{1}{g(a)} = -\frac{g(a+h) - g(a)}{g(a+h)g(a)}.$$

Al dividir por $h \neq 0$ y pasar al límite obtenemos

$$\left(\frac{1}{g}\right)'(a) = -\frac{g'(a)}{g(a)^2}.$$

Solo ahora aplicamos la regla del producto a $f(1/g)$ y simplificamos:

$$\begin{aligned}\left(\frac{f}{g}\right)'(a) &= \frac{f'(a)}{g(a)} - \frac{f(a)g'(a)}{g(a)^2} \\ &= \boxed{\frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}}.\end{aligned}$$

Control de hipótesis. La condición decisiva es $g(a) \neq 0$, no $g'(a) \neq 0$. Por ejemplo, $g(x) = 1$ tiene derivada cero, pero $f/g = f$ está perfectamente definido y puede derivarse cuando f lo sea. La expresión f'/g' sería entonces una división por cero aunque el cociente original no presente dificultad alguna.

Estación 6. Reparar la división imposible en la regla de la cadena. Consigna. Se intenta escribir

$$\frac{g(f(a+h)) - g(f(a))}{h} = \frac{g(f(a+h)) - g(f(a))}{f(a+h) - f(a)} \frac{f(a+h) - f(a)}{h}.$$

¿Por qué la igualdad no es válida para todos los $h \neq 0$ admisibles? Construye una identidad que sí lo sea y úsala con una función interior que conserve su valor en un semieje.

Diagnóstico. Puede ocurrir $f(a+h) = f(a)$ aunque $h \neq 0$; la fracción exterior del miembro derecho queda entonces indefinida. Para evitarla, sea $b = f(a)$ y definamos sobre el dominio B de la función exterior

$$K(t) = \begin{cases} \dfrac{g(t) - g(b)}{t - b}, & t \neq b, \\ g'_B(b), & t = b. \end{cases}$$

Siempre que b sea de acumulación de B y g sea diferenciable allí, $K(t) \to g'_B(b)$ relativamente a B y K es continua en b. Además, para **todo** $t \in B$, incluido $t = b$, se cumple $g(t) - g(b) = K(t)(t - b)$. En consecuencia, para todos los incrementos $h \neq 0$ admisibles,

$$\frac{g(f(a+h)) - g(f(a))}{h} = K(f(a+h)) \frac{f(a+h) - f(a)}{h}.$$

No aparece ningún incremento interior en un denominador. La continuidad de f en a y la de K en b dan el límite del primer factor; la diferenciabilidad de f da el del segundo. Recuperamos Teorema 7.4 sin excepción para los incrementos interiores nulos.

Prueba de estrés. Definamos $f : \mathbb{R} \to [0, \infty)$ por $f(x) = 0$ si $x \leq 0$ y $f(x) = x^2$ si $x > 0$; definamos $g : [0, \infty) \to \mathbb{R}$ por $g(t) = t^2 + 1$. En $a = 0$, $f'_{\mathbb{R}}(0) = 0$ porque su cociente es 0 para $h < 0$ y h para $h > 0$; $g'_{[0,\infty)}(0) = 0$ porque su cociente exterior es t para $t > 0$. Aquí $K(t) = t$ para $t \geq 0$. Para todo $h < 0$ se cumple $f(h) - f(0) = 0$, pero la identidad corregida sigue definida. Directamente, el cociente de la composición es 0 para $h < 0$ y h^3 para $h > 0$: converge a 0, como exige la regla.

Estación 7. Auditar una composición triple y sus puntos de evaluación. Consigna. Considera las funciones tipadas

$$\begin{aligned} f &: [0,\infty) \longrightarrow [1,\infty), \quad f(x) = x^2 + 1, \\ g &: [1,\infty) \longrightarrow [0,\infty), \quad g(t) = (t-1)^2, \\ k &: [0,\infty) \longrightarrow \mathbb{R}, \qquad\quad k(s) = \frac{1}{1+s}. \end{aligned}$$

Determina la derivada relativa de $H = k \circ g \circ f$ en 0 y en 1. Indica el punto de evaluación de **cada** derivada; no evalúes las tres automáticamente en la entrada original.

Reconstrucción en el extremo. En $a = 0$ los puntos intermedios son $x_0 = 0$, $x_1 = f(0) = 1$ y $x_2 = g(1) = 0$. Son puntos de acumulación de sus dominios respectivos, y las derivadas relativas —obtenidas mediante las reglas ya demostradas, o desde sus cocientes en los extremos— son

$$f'_{[0,\infty)}(0) = 0, \qquad g'_{[1,\infty)}(1) = 0, \qquad k'_{[0,\infty)}(0) = -1.$$

La regla finita Proposición 7.3 entrega

$$H'_{[0,\infty)}(0) = k'(x_2)g'(x_1)f'(x_0) = (-1)\cdot 0 \cdot 0 = 0.$$

No sería lícito escribir $g'(0)$: 0 ni siquiera pertenece al dominio declarado de g. Tampoco hemos exigido que la salida final $H(0) = 1$ sea punto de acumulación del codominio $\mathbb{R}$ para derivar una cuarta función inexistente.

Reconstrucción en un punto interior. En $a = 1$, $x_0 = 1$, $x_1 = 2$ y $x_2 = 1$; por tanto,

$$H'_{[0,\infty)}(1) = k'(1)g'(2)f'(1) = -\frac{1}{4}\cdot 2 \cdot 2 = \boxed{-1}.$$

La expresión compuesta es $H(x) = 1/(1+x^4)$ para $x \geq 0$. Como verificación por cocientes, para $h \neq 0$ suficientemente pequeño,

$$\begin{aligned} \frac{H(1+h) - H(1)}{h} &= -\frac{(1+h)^4 - 1}{2h\big(1 + (1+h)^4\big)} \\ &\longrightarrow -\frac{4}{2\cdot 2} = -1. \end{aligned}$$

En 0, el cociente directo es $-h^3/(1+h^4) \to 0$ para $h > 0$, en concordancia con el cálculo relativo. Se han comprobado tanto el orden de evaluación como la naturaleza unilateral del extremo, sin suponer derivadas de extensiones no declaradas.

Estación 8. Reconocer dónde termina el capítulo. Consigna. Clasifica las siguientes peticiones: ¿pueden justificarse enteramente con las reglas de este capítulo o exigen un resultado que todavía no hemos demostrado?

Petición	Diagnóstico y fundamento
Derivar $1/(1+x^2)$ sobre $\mathbb{R}$.	**Sí.** El denominador es positivo; la regla del recíproco y la derivada polinómica bastan.
Derivar $(x^2+1)^3$ sobre $\mathbb{R}$.	**Sí.** La función interior es polinómica y la exterior es la potencia cúbica; aplicar cadena con evaluación en x^2+1.

Petición	Diagnóstico y fundamento
Deducir una fórmula general para la derivada de f^{-1} a partir de f'.	**No todavía.** f^{-1} denota una función inversa por composición, no el recíproco puntual $1/f$; la derivación de inversas pertenece al capítulo 8.
Derivar una relación $x^2 + y^2 = 1$ tratando y como función de x, sin fijar una rama ni probar su diferenciabilidad.	**No todavía.** Deben justificarse primero la existencia de una función pertinente y las hipótesis para derivarla; la derivación implícita pertenece al capítulo 8.
Concluir de $u'(a) > 0$ que u es creciente en todo un intervalo.	**No.** Una derivada positiva en un solo punto no autoriza tal conclusión. Los teoremas del valor medio y las aplicaciones globales de las derivadas se estudiarán en los capítulos 9 y 10.

Una frontera comprobable. La función $u(x) = x - x^2$ satisface $u'(0) = 1 > 0$ mediante las reglas polinómicas, pero no es creciente en toda $\mathbb{R}$: $u(1) = 0 > u(2) = -2$, mientras $1 < 2$. Incluso cuando una conclusión global resulta verdadera para una función particular, hace falta una prueba que corresponda a la amplitud de la conclusión. Del mismo modo, $v(x) = x^2$ tiene $v'(0) = 0$ sin ser constante en ningún entorno de 0. Las fórmulas puntuales no sustituyen a los teoremas globales.

Mapa de decisión: de la expresión a una conclusión justificada. La ruta común a las ocho estaciones puede reconstruirse en el siguiente orden, sin tomar ninguna fórmula como licencia para ignorar sus hipótesis:

1. **Definir el objeto.** Identificar función completa, dominio de cada componente y dominio efectivo de la operación. Excluir los ceros de denominadores antes de simplificar.
2. **Fijar el punto.** Comprobar pertenencia y acumulación en el dominio pertinente; distinguir derivada bilateral y derivada relativa en un extremo.
3. **Desarmar la operación exterior.** Reconocer suma, producto, recíproco, cociente o composición; para composiciones, registrar los puntos intermedios y los dominios de cada eslabón.
4. **Auditar las hipótesis.** Verificar la diferenciabilidad de las funciones que exige la regla; para un cociente, demostrar no anulación local; para la cadena, disponer de la derivada exterior en el punto correcto.
5. **Justificar la transformación.** Si se reconstruye una prueba, escribir una identidad exacta de incrementos y señalar qué continuidad y qué límites se utilizan. No dividir por cantidades que puedan anularse.
6. **Obtener y controlar el resultado.** Aplicar la fórmula únicamente en los puntos habilitados; cuando sea conveniente, contrastarla mediante un cociente incremental o una expansión algebraica independiente.
7. **Distinguir el diagnóstico final.** Hipótesis suficientes incumplidas no equivale a no diferenciabilidad; falta de valor central sí impide derivar la función original; conclusiones sobre inversas, derivación implícita o comportamiento global exigen nuevas pruebas.

Comprobación de autonomía
Ante una composición con un cociente en su interior, ¿puedes señalar el dominio antes de derivar, nombrar los puntos de evaluación, identificar la única división peligrosa de la prueba de la cadena y sustituirla por una identidad válida? ¿Puedes además explicar por qué una regla que no se aplica no demuestra que la derivada buscada no exista? Si las respuestas contienen una hipótesis y un argumento, no solo una fórmula, has recuperado el método del capítulo.

Elegir una ruta de derivación exige reconocer la operación exterior y comprobar las hipótesis en cada punto intermedio. Los cuarenta ejercicios con sus soluciones desarrolladas permiten practicar esas decisiones con un apoyo menor que el de las estaciones.

7.11. Ejercicios y soluciones

Las reglas de derivación han sido demostradas, pero emplearlas correctamente exige algo más que reconocer una fórmula: debemos definir la función, comprobar el dominio y los puntos de acumulación, identificar qué resultados están disponibles y distinguir las hipótesis suficientes de sus conversas. Este banco trabaja esas decisiones en orden creciente de autonomía. Antes de mirar cada solución, escribe la operación exterior y comprueba en qué puntos está autorizada la derivada.

Los cuarenta ejercicios están distribuidos en siete niveles, $7A{+}7B{+}7C{+}6D{+}5E{+}5F{+}3G = 40$. Cada ejercicio tiene una solución homóloga desarrollada al final de la sección. Solo se utilizan resultados de los capítulos anteriores y de las §§7.1–7.10. Ni la regla de la función inversa ni la derivación implícita, los teoremas del valor medio, Taylor o las derivadas trascendentes forman parte del repertorio autorizado.

Nivel A — Reconocimiento y comprensión. Ejercicio A1. Linealidad no significa multiplicatividad. Enuncia la regla para $D(\alpha f + \beta g)$ y explica por qué no permite concluir que $D(fg) = (Df)(Dg)$. Refuta esta última fórmula con $f(x) = g(x) = x$ en $a = 1$, calculando ambos miembros.

Ejercicio A2. Datos puntuales y combinación lineal. Sean $f, g : A \to \mathbb{R}$ diferenciables en un punto de acumulación $a \in A$, y supón que $f(a) = 2$, $g(a) = -1$, $f'_A(a) = 3$ y $g'_A(a) = 4$. Calcula $(2f - 3g)'_A(a)$ y explica qué datos se utilizan.

Ejercicio A3. Los valores centrales también intervienen. Con las funciones y los cuatro datos del ejercicio A2, calcula $(fg)'_A(a)$. Identifica los sumandos procedentes de la variación de cada factor.

Ejercicio A4. La condición del cociente. Sean $f, g : A \to \mathbb{R}$ diferenciables en a, punto de acumulación de A. ¿Qué hipótesis adicional garantiza que f/g esté definido en un entorno relativo de a y sea diferenciable **en** a? Explica por qué la condición pertinente no es $g'_A(a) \neq 0$.

Ejercicio A5. Primero el dominio. Determina el dominio natural de $r(x) = (x^2{+}1)/(x^2{-}4)$ y calcula $r'(x)$ en ese dominio. ¿Tiene sentido pedir la derivada de *esa función* en 2?

Ejercicio A6. Reconstruir la derivada cúbica. Demuestra que $D(x^3) = 3x^2$ utilizando solamente $D(x) = 1$ y la regla del producto; no cites como premisa la regla general para potencias.

Ejercicio A7. ¿Dónde se evalúa la derivada exterior? Escribe $(3x + 1)^4$ como $g \circ f$, indica los dominios y calcula la derivada en un punto arbitrario a. Explica por qué el argumento de g' es $3a + 1$ y no necesariamente a.

Nivel B — Aplicación directa. Ejercicio B1. Un polinomio. Deriva $p(x) = 3x^5 - 2x^3 + 7x - 4$ para todo $x \in \mathbb{R}$. Indica qué resultados justifican la derivación término a término.

Ejercicio B2. Un producto comprobado de dos modos. Calcula la derivada de $p(x) = (x^2+1)(x^3-2x)$ aplicando la regla del producto; expande después p y verifica independientemente la respuesta.

Ejercicio B3. Un cociente racional. Declara el dominio natural de $r(x) = (x^2+1)/(x-2)$, deriva y simplifica el numerador resultante. ¿Es legítimo evaluar $r'(2)$?

Ejercicio B4. Una potencia negativa. Deduce la derivada de $r(x) = x^{-4}$ utilizando el recíproco de x^4, con indicación del dominio. No supongas de antemano la fórmula para exponentes negativos.

Ejercicio B5. Una potencia compuesta. Calcula $D((2x^2 - 1)^5)$ identificando función interior, exterior y puntos de evaluación de sus derivadas.

Ejercicio B6. Dos capas de composición. Deriva $H(x) = ((x^2 + 1)^3 - 2)^4$. Define los valores intermedios y muestra cómo se obtiene cada factor de la derivada.

Ejercicio B7. Derivar en un extremo. Sea $A = [0, \infty)$ y sea $p : A \to \mathbb{R}$ dada por $p(x) = x^2(x + 1)$. Calcula $p'_A(0)$ mediante la regla del producto relativa a A y verifica el valor directamente con el cociente incremental derecho.

Nivel C — Combinación estructural. Ejercicio C1. Linealidad finita. Para $n \geq 1$, supón que $f_1, \ldots, f_n : A \to \mathbb{R}$ son diferenciables en el mismo punto de acumulación $a \in A$ y $c_1, \ldots, c_n \in \mathbb{R}$. Demuestra por inducción que $\sum_{k=1}^{n} c_k f_k$ es diferenciable en a y que su derivada es $\sum_{k=1}^{n} c_k f'_{k,A}(a)$.

Ejercicio C2. Producto de un número finito de factores. Para $n \geq 1$, demuestra por inducción que, si $f_1, \ldots, f_n : A \to \mathbb{R}$ son diferenciables en un punto de acumulación $a \in A$, entonces

$$\left(\prod_{k=1}^{n} f_k\right)'_A(a) = \sum_{j=1}^{n} f'_{j,A}(a) \prod_{\substack{1 \leq k \leq n \\ k \neq j}} f_k(a).$$

Convén explícitamente que el producto vacío vale 1. No dividas por ningún factor.

Ejercicio C3. La derivada de x^{-1}. Demuestra sobre $\mathbb{R} \setminus \{0\}$ que $D(x^{-1}) = -x^{-2}$ aplicando la proposición del recíproco a la función identidad. Explica por qué el resultado no atribuye derivada en 0.

Ejercicio C4. Recíproco de una composición. Determina el dominio y la derivada de $R(x) = \frac{1}{1+(x^2+1)^3}$.

Justifica primero que el denominador no se anula en ningún real.

Ejercicio C5. Una composición con puntos excluidos. Sean $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$, y $g : \mathbb{R} \setminus \{1\} \to \mathbb{R}$, $g(t) = 1/(t-1)$. Determina el dominio efectivo de $g \circ f$ y su derivada allí, sin reintegrar los puntos excluidos.

Ejercicio C6. Dos pruebas para un mismo cuadrado. Deriva $H(x) = (x^2+1)^2$ de dos maneras: interpretándolo como producto de dos polinomios iguales e interpretándolo como composición $g \circ f$. Comprueba que ambos resultados coinciden para todo real.

Ejercicio C7. No anulación local cuantitativa. Sean $g : A \to \mathbb{R}$ diferenciable en el punto de acumulación $a \in A$, con $g(a) \neq 0$. Demuestra que existe $\delta > 0$ tal que $g(a+h) \neq 0$ si $a+h \in A$ y $|h| < \delta$. Obtén, de hecho, una cota inferior positiva para $|g(a+h)|$.

Nivel D — Inversión y diagnóstico. Ejercicio D1. Una falsa multiplicación de pendientes. Un estudiante escribe $(fg)' = f'g'$ y pretende demostrarlo multiplicando los dos cocientes incrementales. Localiza el error con $f = g = x$ en $a = 1$; después reconstruye la identidad de incrementos que conduce a la fórmula correcta para funciones diferenciables en un dominio común.

Ejercicio D2. Derivar «arriba y abajo» no funciona. Refuta $(f/g)' = f'/g'$ usando $f(x) = x^2$, $g(x) = x$ en $a = 1$. Demuestra que todos los cocientes numéricos que comparas están definidos y escribe la fórmula correcta.

Ejercicio D3. Un incremento interior nulo. Define $f : \mathbb{R} \to [0, \infty)$ por $f(x) = 0$ si $x \leq 0$ y $f(x) = x^2$ si $x > 0$; sea $g : [0, \infty) \to \mathbb{R}$, $g(t) = t^2 + 1$. Explica por qué no es válido dividir siempre por $f(h) - f(0)$ al derivar $g \circ f$ en 0. Repara la demostración con la extensión continua del cociente de g y comprueba la derivada directamente.

Ejercicio D4. Una conversa inexistente. Refuta la afirmación «si $g \circ f$ es diferenciable en a, entonces g lo es en $f(a)$» mediante $f(x) = x^2$, $g(t) = |t|$ declaradas sobre $\mathbb{R}$, con $a = 0$. Distingue este caso de restringir g explícitamente a $[0, \infty)$.

Ejercicio D5. Cancelar no crea un valor perdido. Examina la afirmación «$r(x) = x^2/x$ tiene derivada en 0 porque se simplifica a x». Especifica el dominio natural de r, determina dónde vale la simplificación y construye, si es posible, una extensión distinta que sí tenga derivada en 0.

Ejercicio D6. La intersección puede aislar un punto. Sean $A = \{0\} \cup \{1/n : n \geq 1\}$ y $B = \{0\} \cup \{-1/n : n \geq 1\}$. Define $f : A \to \mathbb{R}$ y $g : B \to \mathbb{R}$ como funciones constantemente nulas. Comprueba que ambas tienen derivada relativa en 0. ¿Tiene derivada relativa en 0 su suma definida sobre $A \cap B$? Justifica sin utilizar un límite vacuo.

Nivel E — Construcción de contraejemplos. Ejercicio E1. Producto regular, factores irregulares. Construye dos funciones no diferenciables en 0 cuyo producto sí lo sea. Usa $f(x) = g(x) = |x|$ y verifica las tres afirmaciones con cocientes incrementales.

Ejercicio E2. El cociente puede ocultar dos esquinas. Construye numerador y denominador no diferenciables en 0 pero con denominador no nulo en toda la recta y cociente diferenciable. Utiliza $f(x) = g(x) = 1 + |x|$ y demuestra cada propiedad.

Ejercicio E3. Una función exterior no diferenciable. Da un ejemplo de funciones completas $f, g : \mathbb{R} \to \mathbb{R}$ tal que $g \circ f$ tenga derivada en 0 pero g no la tenga en $f(0)$. Verifica con $f(x) = x^2$ y $g(t) = |t|$ y explica qué hipótesis de la regla de la cadena falla.

Ejercicio E4. Infinitos incrementos interiores nulos. Define $f : \mathbb{R} \to [0, \infty)$ por $f(x) = 0$ si $x \in \mathbb{Q}$ y $f(x) = x^2$ si $x \notin \mathbb{Q}$. Demuestra que $f'(0) = 0$ pero f no es constante en ningún entorno de 0 y que $f(h) - f(0) = 0$ para infinitos $h \neq 0$ arbitrariamente cercanos a 0. Con $g : [0, \infty) \to \mathbb{R}$, $g(t) = t^2 + 1$, verifica la regla de la cadena en 0 sin dividir por ese incremento interior.

Ejercicio E5. Una función interior no diferenciable. Da una composición diferenciable en 0 cuya función interior no sea diferenciable. Usa $f(x) = |x|$ y una función exterior constante $g(t) = 7$, declarando dominios y verificando la derivada de la composición.

Nivel F — Descubrimiento guiado. Ejercicio F1. Inventar el término intermedio. Para $f, g : A \to \mathbb{R}$ diferenciables en a, punto de acumulación, parte de $f(a+h)g(a+h) - f(a)g(a)$. Añade y sustrae un mismo término para separar el cambio de los factores. Divide por $h \neq 0$, identifica todos los límites necesarios y demuestra la regla del producto sin citarla como premisa.

Ejercicio F2. Construir el recíproco. Sean $g : A \to \mathbb{R}$ diferenciable en a, punto de acumulación, y $g(a) \neq 0$. Obtén un entorno relativo donde g no se anule, transforma exactamente $1/g(a + h) - 1/g(a)$ y deduce la derivada del recíproco desde la definición. Indica en qué paso se requiere la continuidad.

Ejercicio F3. Inducir la regla de las potencias. Usando la derivada de la identidad y la regla del producto, demuestra por inducción que $D(x^n) = nx^{n-1}$ para todo entero $n \geq 1$. Trata por separado el exponente 0 y especifica por qué esta demostración no cubre exponentes fraccionarios.

Ejercicio F4. Reparar la cadena en el centro. Sean $f : A \to B \subseteq \mathbb{R}$ y $g : B \to \mathbb{R}$, con $a \in A$, $b = f(a)$, ambos puntos de acumulación de sus respectivos dominios y ambas derivadas relativas finitas. Define una función $K : B \to \mathbb{R}$ prolongando $(g(t) - g(b))/(t - b)$ en $t = b$; demuestra su continuidad en b y deduce la regla de la cadena mediante una identidad válida incluso cuando $f(a + h) = b$.

Ejercicio F5. La composición desde residuos. En las hipótesis del ejercicio F4, escribe $f(a + h) = b + Lh + r_f(h)$ y $g(b + t) = g(b) + Mt + t\eta(t)$ con $r_f(h) = o(h)$ y $\eta(t) \to 0$, definiendo $\eta(0) = 0$. Para $t = f(a + h) - b$, demuestra directamente que el último término de la segunda linealización es $o(h)$, incluso cuando $t = 0$. Concluye la derivada de la composición sin citar la regla de la cadena.

Nivel G — Desafíos. Ejercicio G1. Ceros en un producto finito. Prueba por inducción la fórmula de la derivada de $P = \prod_{k=1}^{n} f_k$, para $n \geq 2$ y factores diferenciables en a sobre un dominio común. Demuestra que $P'(a) = 0$ si al menos dos factores se anulan en a. ¿Qué fórmula queda si exactamente un factor, f_j, se anula allí? No dividas por ninguno.

Ejercicio G2. Un conjunto denso de incrementos nulos. Construye una función $f : \mathbb{R} \to \mathbb{R}$ diferenciable en 0, con $f'(0) = 0$ e infinitos ceros no nulos arbitrariamente próximos a 0, pero que no sea localmente constante. Usa $f(x) = 0$ en los racionales y $f(x) = x^2$ en los irracionales. Para $g(t) = 3t + 1$ sobre $\mathbb{R}$, comprueba directamente $(g \circ f)'(0) = g'(f(0))f'(0)$ y explica por qué la habitual división por $f(h) - f(0)$ fracasa.

Ejercicio G3. El cociente desde las linealizaciones. Sean $f, g : A \to \mathbb{R}$ diferenciables en un punto de acumulación a, con $g(a) \neq 0$. Escribe sus aproximaciones de primer orden con residuos $o(h)$ y demuestra *directamente* que f/g es diferenciable en a y que

$$\left(\frac{f}{g}\right)'_A(a) = \frac{f'_A(a)g(a) - f(a)g'_A(a)}{g(a)^2}.$$

No uses las reglas del recíproco ni del cociente; identifica el papel de la no anulación local y evita cualquier división por $f(a + h) - f(a)$ o $g(a + h) - g(a)$.

Soluciones desarrolladas.

Soluciones del nivel A. **Solución A1.** Para funciones diferenciables en un punto de acumulación a de un dominio común y escalares fijos, Teorema 7.1 da

$$(\alpha f + \beta g)'_A(a) = \alpha f'_A(a) + \beta g'_A(a).$$

Esta es una propiedad de las *combinaciones lineales*, no de los productos. Si $f(x) = g(x) = x$ en $\mathbb{R}$, entonces $f'(1) = g'(1) = 1$. Pero $(fg)(x) = x^2$ y, para $h \neq 0$, $\frac{(1+h)^2-1}{h} = 2 + h \longrightarrow 2$. Por tanto, $(fg)'(1) = 2 \neq 1 = f'(1)g'(1)$. La regla verdadera exige los valores de los factores: $(fg)' = f'g + fg'$ donde se cumplen sus hipótesis. El ejemplo refuta la multiplicatividad; no contradice la linealidad.

Solución A2. Ambas funciones son diferenciables en a relativamente al mismo dominio, de modo que podemos aplicar linealidad con coeficientes 2 y -3:

$$(2f - 3g)'_A(a) = 2f'_A(a) - 3g'_A(a) = 2 \cdot 3 - 3 \cdot 4 = \boxed{-6}.$$

Los valores $f(a) = 2$ y $g(a) = -1$ no intervienen: la derivada de una combinación lineal depende aquí solo de las derivadas y de los coeficientes constantes. No se está derivando un producto de funciones.

Solución A3. La regla del producto Teorema 7.2, aplicable por diferenciabilidad común, proporciona

$$(fg)'_A(a) = f'_A(a)g(a) + f(a)g'_A(a) = 3(-1) + 2(4) = \boxed{5}.$$

El primer término -3 corresponde a la variación de f, ponderada por el valor central $g(a)$; el segundo, 8, corresponde a la variación de g, ponderada por $f(a)$. Multiplicar solamente las pendientes daría 12, que no es la respuesta.

Solución A4. Se requiere $g(a) \neq 0$. Como g es diferenciable en a, es continua allí por Teorema 6.1. Con $\varepsilon = |g(a)|/2 > 0$ encontramos $\delta > 0$ tal que, para $x \in A$ y $|x - a| < \delta$,

$$|g(x)| \geq |g(a)| - |g(x) - g(a)| > \frac{|g(a)|}{2} > 0.$$

Así a y un entorno relativo suyo pertenecen al dominio natural del cociente; las restricciones tienen las mismas derivadas puntuales por localidad y Teorema 7.3 asegura la derivabilidad de f/g en a. No se necesita $g'_A(a) \neq 0$: si $g = 1$, entonces $g' = 0$ y $f/g = f$ sigue siendo derivable.

Solución A5. El denominador se factoriza como $x^2 - 4 = (x - 2)(x + 2)$. El dominio natural es, por tanto, $A = \mathbb{R} \setminus \{-2, 2\}$.

Los polinomios son diferenciables en toda la recta y el denominador es distinto de cero en A. Por Teorema 7.3,

$$\begin{aligned} r'(x) &= \frac{2x(x^2 - 4) - (x^2 + 1)2x}{(x^2 - 4)^2} \\ &= \boxed{-\frac{10x}{(x^2 - 4)^2}}, \qquad x \in A. \end{aligned}$$

No existe $r(2)$ en la función original; en consecuencia, tampoco existe $r'(2)$. Una fórmula de derivada no puede extender el dominio por sí sola.

Solución A6. Por la definición aplicada a la identidad, $D(x) = 1$. La regla del producto, ya demostrada, da primero

$$D(x^2) = D(x \cdot x) = 1 \cdot x + x \cdot 1 = 2x.$$

Escribimos ahora $x^3 = x^2 \cdot x$ y volvemos a utilizar la misma regla:

$$D(x^3) = D(x^2)\, x + x^2 D(x) = 2x \cdot x + x^2 = \boxed{3x^2}.$$

Todas las funciones se consideran sobre $\mathbb{R}$, así que no hay restricciones de dominio. La fórmula general para las potencias es una conclusión posterior que este cálculo ayuda a reconstruir, no una hipótesis utilizada.

Solución A7. Tomamos $f : \mathbb{R} \to \mathbb{R}$, $f(x) = 3x + 1$, y $g : \mathbb{R} \to \mathbb{R}$, $g(t) = t^4$. Entonces $g \circ f(x) = (3x + 1)^4$. Para un punto a:

$$f(a) = 3a + 1, \quad f'(a) = 3, \quad g'(t) = 4t^3.$$

Ambas funciones son diferenciables en todos los puntos pertinentes. Aplicando Teorema 7.4,

$$(g \circ f)'(a) = g'(f(a))f'(a) = 4(3a + 1)^3 \cdot 3 = \boxed{12(3a + 1)^3}.$$

La derivada exterior se calcula en la *salida* de la función interior, $f(a)$; escribir $g'(a)$ sustituiría injustificadamente una entrada por otra.

Soluciones del nivel B. **Solución B1.** Cada monomio es diferenciable por Proposición 7.2, la constante tiene derivada cero, y Teorema 7.1 permite derivar la combinación finita. Obtenemos, para cualquier $x \in \mathbb{R}$,

$$\begin{aligned} p'(x) &= 3(5x^4) - 2(3x^2) + 7(1) - 0 \\ &= \boxed{15x^4 - 6x^2 + 7}. \end{aligned}$$

El dominio es toda la recta: ninguna de estas operaciones introduce un denominador ni restringe las entradas. No se ha usado una regla de composición.

Solución B2. Definimos $f(x) = x^2 + 1$, $g(x) = x^3 - 2x$ sobre $\mathbb{R}$. Sus derivadas polinómicas son $f'(x) = 2x$, $g'(x) = 3x^2 - 2$. Por la regla del producto,

$$\begin{aligned} p'(x) &= 2x(x^3 - 2x) + (x^2 + 1)(3x^2 - 2) \\ &= 2x^4 - 4x^2 + 3x^4 + x^2 - 2 \\ &= \boxed{5x^4 - 3x^2 - 2}. \end{aligned}$$

De modo independiente, antes de derivar expandimos la función: $p(x) = x^5 - x^3 - 2x$. La derivación polinómica da $5x^4 - 3x^2 - 2$, coincidente con el cálculo anterior. Esta comprobación no presupone la regla del producto para obtener la forma expandida.

Solución B3. El dominio original es $A = \mathbb{R} \setminus \{2\}$. Para $x \in A$, numerador y denominador son polinomios diferenciables y $x - 2 \neq 0$. La regla del cociente da

$$\begin{aligned} r'(x) &= \frac{2x(x-2) - (x^2+1)}{(x-2)^2} \\ &= \boxed{\frac{x^2 - 4x - 1}{(x-2)^2}}. \end{aligned}$$

En $x = 2$ la función r no está definida y no puede tener derivada, con independencia de que una manipulación formal sugiera otra expresión.

Solución B4. En $A = \mathbb{R} \setminus \{0\}$, la potencia cuarta $u(x) = x^4$ no se anula. Su derivada es $u'(x) = 4x^3$ por Proposición 7.2. Como $r = 1/u$, la proposición del recíproco Proposición 7.1 implica

$$r'(x) = -\frac{u'(x)}{u(x)^2} = -\frac{4x^3}{x^8} = \boxed{-4x^{-5}}, \qquad x \neq 0.$$

La simplificación se realiza únicamente donde $x \neq 0$; no se ha supuesto una regla previa para exponentes negativos. En cero, x^{-4} no tiene valor definido.

Solución B5. La función interior es $f(x) = 2x^2 - 1$, la exterior $g(t) = t^5$, ambas de $\mathbb{R}$ en $\mathbb{R}$. Sus derivadas son

$$f'(x) = 4x, \qquad g'(t) = 5t^4.$$

La composición satisface todas las hipótesis de la regla de la cadena. Por ello,

$$\boxed{D((2x^2-1)^5) = 5(2x^2-1)^4\,4x = 20x(2x^2-1)^4}.$$

En particular, el término exterior se evalúa en $t = f(x) = 2x^2 - 1$, no en $t = x$.

Solución B6. Introducimos los valores intermedios

$$u_1 = x^2 + 1, \qquad u_2 = u_1^3, \qquad u_3 = u_2 - 2, \qquad H = u_3^4.$$

Las cuatro funciones elementales de esta cadena son polinómicas, luego sus derivadas existen sobre $\mathbb{R}$. Sucesivas aplicaciones de Teorema 7.4, o Proposición 7.3, entregan

$$\begin{aligned} H'(x) &= 4u_3^3 \cdot 1 \cdot 3u_1^2 \cdot 2x \\ &= \boxed{24x(x^2+1)^2\big((x^2+1)^3 - 2\big)^3}. \end{aligned}$$

La resta de 2 aporta derivada 1 respecto de u_2. Es decisivo sustituir cada valor intermedio solo *después* de calcular la derivada del eslabón correspondiente.

Solución B7. Las funciones $f(x) = x^2$ y $g(x) = x + 1$, restringidas a $A = [0, \infty)$, son diferenciables relativamente a ese dominio en 0. Sus derivadas relativas son $f'_A(0) = 0$, $g'_A(0) = 1$, y sus valores centrales son $f(0) = 0$, $g(0) = 1$. Por Teorema 7.2,

$$p'_A(0) = f'_A(0)g(0) + f(0)g'_A(0) = 0 \cdot 1 + 0 \cdot 1 = \boxed{0}.$$

Para verificarlo desde la definición, los incrementos admisibles no nulos son $h > 0$:

$$\frac{p(h) - p(0)}{h} = \frac{h^2(h+1)}{h} = h(h+1) \longrightarrow 0 \quad (h \downarrow 0).$$

La derivada solicitada es relativa al dominio declarado; no hemos supuesto ninguna extensión bilateral.

Soluciones del nivel C. **Solución C1.** Escribamos $S_n = \sum_{k=1}^{n} c_k f_k$. Para $n = 1$, la multiplicación por un escalar, caso particular de Teorema 7.1, establece la diferenciabilidad de $S_1 = c_1 f_1$ y $S'_{1,A}(a) = c_1 f'_{1,A}(a)$.

Supongamos para cierto $n \geq 1$ que S_n es diferenciable y que $S'_{n,A}(a) = \sum_{k=1}^{n} c_k f'_{k,A}(a)$. Entonces $S_{n+1} = S_n + c_{n+1} f_{n+1}$ es combinación lineal de dos funciones diferenciables en el mismo punto y dominio. La linealidad proporciona

$$\begin{aligned} S'_{n+1,A}(a) &= S'_{n,A}(a) + c_{n+1} f'_{n+1,A}(a) \\ &= \sum_{k=1}^{n+1} c_k f'_{k,A}(a). \end{aligned}$$

El paso está probado y la inducción concluye el resultado para todo $n \geq 1$. La finitud es esencial: aquí no se ha intercambiado una derivada con una serie infinita.

Solución C2. Para $n = 1$, la fórmula tiene un único término $f'_{1,A}(a)$ multiplicado por el producto vacío 1. Supongamos probada la fórmula para $P_n = \prod_{k=1}^{n} f_k$. Es una función diferenciable en a. Como $P_{n+1} = P_n f_{n+1}$, Teorema 7.2 da

$$P'_{n+1,A}(a) = P'_{n,A}(a) f_{n+1}(a) + P_n(a) f'_{n+1,A}(a).$$

Insertando la hipótesis inductiva, el primer sumando contiene los términos indexados por $1 \le j \le n$, cada uno multiplicado por $f_{n+1}(a)$; el segundo es precisamente el término $j = n+1$. Así,

$$\boxed{P'_{n+1,A}(a) = \sum_{j=1}^{n+1} f'_{j,A}(a) \prod_{\substack{1 \le k \le n+1 \\ k \ne j}} f_k(a)}.$$

La inducción vale aun si uno o varios factores se anulan, porque en ningún paso se ha dividido por ellos.

Solución C3. Sea $u : \mathbb{R} \to \mathbb{R}$, $u(x) = x$. Es diferenciable con $u'(a) = 1$ para todo real a. Para $a \ne 0$ podemos tomar su recíproco en un entorno del punto. Proposición 7.1 asegura

$$D\left(\frac{1}{u}\right)(a) = -\frac{u'(a)}{u(a)^2} = -\frac{1}{a^2}.$$

Por tanto, en el dominio $\mathbb{R} \setminus \{0\}$,

$$\boxed{D(x^{-1}) = -x^{-2}}.$$

En $a = 0$ la función recíproca no está definida; tampoco existe una derivada de esa función allí. La fórmula se ha construido desde el recíproco, no desde una regla general para exponentes negativos.

Solución C4. Para todo $x \in \mathbb{R}$, se tiene $x^2+1 \ge 1$, luego $(x^2+1)^3 \ge 1$ y $q(x) = 1+(x^2+1)^3 \ge 2 > 0$.

El dominio natural de $R = 1/q$ es toda $\mathbb{R}$. Por cadena y linealidad,

$$q'(x) = 3(x^2+1)^2 \cdot 2x = 6x(x^2+1)^2.$$

La regla del recíproco, aplicada ahora que q es diferenciable y no nula, da

$$\boxed{R'(x) = -\frac{6x(x^2+1)^2}{\big(1+(x^2+1)^3\big)^2}}, \qquad x \in \mathbb{R}.$$

La demostración de positividad no es una formalidad: justifica que no hay puntos excluidos donde la fórmula resultaría ilegítima.

Solución C5. La composición solo está definida donde $f(x) = x^2$ pertenece al dominio de g, es decir, donde $x^2 \neq 1$. Su dominio efectivo es $A = \mathbb{R} \setminus \{-1, 1\}$.

Para $t \neq 1$, la regla del recíproco aplicada a $t - 1$ da $g'(t) = -1/(t-1)^2$. Como $f'(x) = 2x$, la cadena sobre A permite concluir

$$\boxed{(g \circ f)'_A(x) = g'(x^2)\, 2x = -\frac{2x}{(x^2-1)^2}}, \quad x \in A.$$

En $x = \pm 1$ la función exterior recibe la entrada prohibida $t = 1$. No hay valor de la composición allí y no puede atribuirse una derivada a la función original en esos puntos.

Solución C6. En la ruta del producto, $H = (x^2+1)(x^2+1)$ y ambos factores son diferenciables. Aplicando Teorema 7.2,

$$H'(x) = 2x(x^2+1) + (x^2+1)2x = 4x(x^2+1).$$

En la ruta de composición, $f(x) = x^2 + 1$ y $g(t) = t^2$. Puesto que $g'(t) = 2t$ y $f'(x) = 2x$, Teorema 7.4 da

$$(g \circ f)'(x) = 2f(x)\, f'(x) = 2(x^2+1)2x = 4x(x^2+1).$$

Ambos razonamientos son válidos para todo real y concuerdan. La primera ruta no necesita la cadena; la segunda explica el punto donde se evalúa la derivada exterior.

Solución C7. Como g es diferenciable en a, Teorema 6.1 la hace continua relativamente a A en ese punto. Elegimos $\varepsilon = |g(a)|/2 > 0$. Existe $\delta > 0$ tal que, para $x = a + h \in A$ con $|h| < \delta$, $|g(a+h) - g(a)| < \frac{|g(a)|}{2}$.

Aplicamos la desigualdad triangular inversa:

$$|g(a+h)| \geq |g(a)| - |g(a+h) - g(a)| > \boxed{\frac{|g(a)|}{2}} > 0.$$

Esta cota proporciona simultáneamente la no anulación local y una separación cuantitativa de cero. El dominio del recíproco contiene así un entorno relativo de a, de manera que las derivadas calculadas en él preservan los incrementos suficientemente pequeños del dominio inicial.

Soluciones del nivel D. **Solución D1.** Con $f = g = x$ en $\mathbb{R}$, las pendientes en 1 son 1 y 1; sin embargo, el producto x^2 tiene derivada 2. El error consiste en reemplazar el cociente incremental de fg por el *producto* de los cocientes de f y g: esas expresiones no son iguales. Para encontrar la identidad correcta, añadimos y sustraemos $f(a+h)g(a)$:

$$\begin{aligned} &f(a+h)g(a+h) - f(a)g(a) \\ &= f(a+h)[g(a+h) - g(a)] + g(a)[f(a+h) - f(a)]. \end{aligned}$$

Para $h \neq 0$ admisible dividimos por h. Los cocientes de f y g convergen a sus derivadas; además, $f(a+h) \to f(a)$ porque diferenciabilidad implica continuidad. Por el álgebra de límites,

$$\boxed{(fg)'_A(a) = f(a)g'_A(a) + g(a)f'_A(a)}.$$

La continuidad solo se invoca después de establecer la identidad. En particular, la derivabilidad del producto es una conclusión del argumento.

Solución D2. Tomemos $f(x)=x^2$, $g(x)=x$ y $a=1$. Como $g(1)=1\neq 0$, el cociente $q=f/g$ está definido sobre $\mathbb{R}\setminus\{0\}$ y allí satisface $q(x)=x$. En el punto 1,

$$q'(1)=1, \qquad \frac{f'(1)}{g'(1)}=\frac{2}{1}=2.$$

Todas las cantidades exhibidas existen, pero no coinciden; queda refutada la fórmula. Cuando f y g son diferenciables y $g(a)\neq 0$, la regla demostrada en §7.4 establece en cambio

$$\boxed{\left(\frac{f}{g}\right)'(a)=\frac{f'(a)g(a)-f(a)g'(a)}{g(a)^2}}.$$

Para este ejemplo resulta $(2\cdot 1-1\cdot 1)/1^2=1$. La condición de no anulación se refiere a $g(a)$, no a $g'(a)$.

Solución D3. Aquí $a=b=0$, $f(h)-f(0)=0$ para *todo* $h<0$. Por eso la descomposición que divide por tal diferencia no es una identidad válida para todos los incrementos $h\neq 0$ admisibles. En cambio, la función

$$K(t)=\begin{cases}(g(t)-g(0))/t, & t>0,\\ g'_{[0,\infty)}(0), & t=0\end{cases}=\begin{cases}t, & t>0,\\ 0, & t=0\end{cases}$$

está definida sobre $[0,\infty)$ y es continua en 0. Como $g(t)-g(0)=K(t)t$ incluso para $t=0$, obtenemos para todo $h\neq 0$ la identidad válida $\frac{g(f(h))-g(f(0))}{h}=K(f(h))\frac{f(h)-f(0)}{h}$. La derivada de f en 0 existe y vale 0: su cociente es 0 para $h<0$ y h para $h>0$. También $g'_{[0,\infty)}(0)=0$. Finalmente, $g(f(h))=1$ para $h\leq 0$ y $1+h^4$ para $h>0$; el cociente directo de la composición es 0 o h^3, respectivamente, y converge a $\boxed{0}$. La prueba reparada y la definición concuerdan sin ninguna división por Δf.

Solución D4. La interior $f(x)=x^2$ tiene $f(0)=0$ y derivada $f'(0)=0$. Para la exterior *declarada sobre* $\mathbb{R}$, los cocientes incrementales en 0 son $|t|/t$, que valen -1 si $t<0$ y 1 si $t>0$; por tanto, $g'_{\mathbb{R}}(0)$ no existe. Pero

$$(g\circ f)(x)=|x^2|=x^2, \qquad (g\circ f)'(0)=0.$$

La composición tiene derivada sin que la función exterior original satisfaga las hipótesis suficientes de la cadena. Si definimos *otra función*, $\widetilde{g}=g|_{[0,\infty)}$, entonces $\widetilde{g}'_{[0,\infty)}(0)=1$ por cocientes con $t>0$, y la regla sí se aplica a $\widetilde{g}\circ f$: $1\cdot 0=0$. No es lícito confundir esta derivada relativa con la derivada bilateral inexistente de g.

Solución D5. El cociente original $r(x)=x^2/x$ exige $x\neq 0$, de modo que su dominio natural es $A=\mathbb{R}\setminus\{0\}$. En A podemos cancelar el factor y concluir $r(x)=x$; por consiguiente, $r'_A(a)=1$ para cada $a\in A$. Sin embargo, $r(0)$ no está definido: no existe derivada de *esa función* en 0.

Sí podemos construir una función nueva $\widetilde{r}:\mathbb{R}\to\mathbb{R}$ por $\widetilde{r}(x)=x$, incluida la asignación expresa $\widetilde{r}(0)=0$. Su cociente incremental en 0 es $h/h=1$ para $h\neq 0$, así que $\widetilde{r}'(0)=1$. Este valor pertenece a la extensión, no al cociente original.

Solución D6. Hay elementos $1/n \in A \setminus \{0\}$ que convergen a 0, y elementos $-1/n \in B \setminus \{0\}$ que también lo hacen: 0 es punto de acumulación de ambos dominios. Para sus respectivas funciones nulas, todos los cocientes incrementales admisibles valen 0, luego $f'_A(0) = g'_B(0) = 0$.

La suma natural solo está definida donde ambos dominios coinciden. Como los puntos no nulos de A son positivos y los de B negativos, $A \cap B = \{0\}$. En la intersección, 0 es aislado. La definición de derivada relativa exige un punto de acumulación del dominio, de modo que la suma *no tiene derivada relativa allí en nuestro sentido*. La fórmula de linealidad no es aplicable: no hay incrementos no nulos del dominio común sobre los cuales calcular el límite.

Soluciones del nivel E. **Solución E1.** En $\mathbb{R}$, tanto f como g valen $|x|$. Como $f(0) = g(0) = 0$, para $h \neq 0$ los cocientes incrementales de ambos son

$$\frac{|h|}{h} = \begin{cases} -1, & h < 0, \\ 1, & h > 0. \end{cases}$$

No tienen límite bilateral, luego ninguno de los factores es diferenciable en 0. Sin embargo, $(fg)(x) = |x|^2 = x^2$ en toda la recta y $\frac{(fg)(h)-(fg)(0)}{h} = \frac{h^2}{h} = h \longrightarrow 0$. El producto es diferenciable en 0, con $(fg)'(0) = 0$. La regla del producto solo ofrecía una hipótesis suficiente, no una caracterización necesaria de los productos diferenciables.

Solución E2. Definimos $f = g = 1 + |x|$ sobre $\mathbb{R}$. Para ambos, el cociente incremental en 0 es $\frac{1+|h|-1}{h} = \frac{|h|}{h}$, que tiene límites laterales distintos. Ninguna función es diferenciable en 0. No obstante, $g(x) = 1 + |x| \geq 1 > 0$ para todo x; el cociente original está definido en toda la recta y

$$q(x) = \frac{1 + |x|}{1 + |x|} = 1.$$

Por definición, $(q(h) - q(0))/h = 0$ para $h \neq 0$ y $q'(0) = 0$. La fórmula del cociente no podía aplicarse porque faltan las derivadas de los factores, aunque el resultado sí se deriva directamente. No se ha rellenado ningún punto ausente.

Solución E3. Tomemos $f(x) = x^2$ y $g(t) = |t|$, cada una con dominio $\mathbb{R}$ y codominio $\mathbb{R}$. El punto central es $a = 0$, con imagen $b = f(0) = 0$. La derivada $f'(0) = 0$ existe, pero el cociente de g en b toma los valores -1 a la izquierda y 1 a la derecha, así que $g'_{\mathbb{R}}(0)$ no existe.

Como $f(x) \geq 0$ para todo x,

$$(g \circ f)(x) = |x^2| = x^2, \qquad (g \circ f)'(0) = \lim_{h \to 0} h = 0.$$

Falla la hipótesis de diferenciabilidad de la exterior en b para las funciones completas declaradas. El valor de la derivada compuesta se obtuvo desde su definición, no multiplicando una derivada inexistente por cero.

Solución E4. Puesto que $0 \in \mathbb{Q}$, $f(0) = 0$. Para todo $h \neq 0$ tenemos $0 \leq f(h) \leq h^2$, independientemente de si h es racional o irracional. En consecuencia,

$$\left| \frac{f(h) - f(0)}{h} \right| \leq |h| \longrightarrow 0,$$

por lo que $f'_{\mathbb{R}}(0) = 0$ y, en particular, f es continua en 0. La densidad de los racionales garantiza infinitos racionales no nulos arbitrariamente próximos a 0 con $f(h) = f(0) = 0$; la densidad de los irracionales garantiza, en cada entorno de 0, un irracional no nulo s para el que

$f(s) = s^2 > 0$. Así, f no es localmente constante. La exterior $g : [0,\infty) \to \mathbb{R}$, $g(t) = t^2 + 1$, tiene derivada relativa $g'_{[0,\infty)}(0) = 0$ desde $(g(t) - g(0))/t = t$ para $t > 0$. La composición vale $1 + f(x)^2$ y, para $h \neq 0$,

$$\left|\frac{(g \circ f)(h) - (g \circ f)(0)}{h}\right| = \frac{f(h)^2}{|h|} \leq |h|^3 \longrightarrow 0.$$

Por tanto, $(g \circ f)'(0) = 0 = g'_{[0,\infty)}(0) f'(0)$. La comprobación utiliza cotas y no divide por $f(h) - f(0)$, que es nulo para infinitos incrementos.

Solución E5. Declaramos $f : \mathbb{R} \to [0,\infty)$, $f(x) = |x|$, y $g : [0,\infty) \to \mathbb{R}$, $g(t) = 7$. La derivada de la interior en 0 no existe: $|h|/h$ vale -1 a la izquierda y 1 a la derecha. La exterior es constante y posee derivada relativa cero en todos los puntos de acumulación de su dominio, incluido el extremo 0.

La composición está definida para todo real y satisface $(g \circ f)(x) = 7$. Su cociente incremental en 0 es $\frac{(g\circ f)(h)-(g\circ f)(0)}{h} = \frac{7-7}{h} = 0$, así que $(g \circ f)'(0) = 0$. No podemos *invocar* la regla de la cadena con la interior original, que carece de derivada en 0; el cálculo directo muestra que la composición puede ser diferenciable igualmente.

Soluciones del nivel F. **Solución F1.** La variación que debemos analizar es $f(a+h)g(a+h) - f(a)g(a)$. Añadimos y sustraemos el producto intermedio $f(a+h)g(a)$ y agrupamos:

$$\begin{aligned} &f(a+h)g(a+h) - f(a)g(a) \\ &= f(a+h)[g(a+h) - g(a)] + g(a)[f(a+h) - f(a)]. \end{aligned}$$

La identidad es puramente algebraica y vale incluso con $h = 0$. Para $h \neq 0$ admisible, dividimos por h:

$$\frac{(fg)(a+h) - (fg)(a)}{h} = f(a+h)\frac{g(a+h) - g(a)}{h} + g(a)\frac{f(a+h) - f(a)}{h}.$$

Por diferenciabilidad, los dos cocientes convergen a $g'_A(a)$ y $f'_A(a)$. La diferenciabilidad de f implica además $f(a+h) \to f(a)$ por Teorema 6.1. Aplicamos el álgebra de límites una vez comprobadas estas tres convergencias: el límite existe, es finito y vale $f(a)g'_A(a) + g(a)f'_A(a)$. Esto prueba, desde la definición, la diferenciabilidad de fg y la fórmula correspondiente sin presuponerla.

Solución F2. La continuidad de g en a, consecuencia de su diferenciabilidad, nos permite elegir $\delta > 0$ tal que, si $a + h \in A$ y $|h| < \delta$,

$$|g(a+h) - g(a)| < |g(a)|/2, \qquad |g(a+h)| > |g(a)|/2 > 0.$$

El recíproco se define, por tanto, en un entorno relativo del punto central. Su diferencia exacta es

$$\frac{1}{g(a+h)} - \frac{1}{g(a)} = -\frac{g(a+h) - g(a)}{g(a+h)g(a)}.$$

Para $h \neq 0$ dividimos por h y reordenamos:

$$\frac{1/g(a+h) - 1/g(a)}{h} = -\frac{1}{g(a+h)g(a)}\frac{g(a+h) - g(a)}{h}.$$

La continuidad se usa **dos veces**: primero para probar que los denominadores cercanos no son cero y después para obtener $g(a+h) \to g(a)$. El otro factor tiende a $g'_A(a)$ por definición. El límite existe y da

$$\boxed{(1/g)'(a) = -\frac{g'_A(a)}{g(a)^2}}.$$

No se ha utilizado la regla del cociente que esta prueba ayuda a construir.

Solución F3. Para $n = 1$, la función identidad es diferenciable sobre $\mathbb{R}$ y $D(x) = 1 = 1 \cdot x^0$. Supongamos que, para cierto $n \geq 1$, la potencia x^n es diferenciable y satisface $D(x^n) = nx^{n-1}$. Factorizamos $x^{n+1} = x^n \cdot x$. La regla del producto asegura su diferenciabilidad y

$$\begin{aligned} D(x^{n+1}) &= D(x^n)\, x + x^n D(x) \\ &= nx^{n-1}x + x^n = (n+1)x^n. \end{aligned}$$

Queda probado el paso inductivo y, con la base, $\boxed{D(x^n) = nx^{n-1}}$ para todo entero $n \geq 1$. Para $n = 0$ tratamos $x^0 = 1$ como la función constante sobre $\mathbb{R}$, cuya derivada es cero; así evitamos expresiones ambiguas en $x = 0$ cuando se pretendiera sustituir mecánicamente en $0x^{-1}$. La inducción solo recorre enteros naturales y no demuestra que las potencias de exponentes fraccionarios sean diferenciables: eso exige justificar antes las funciones raíz y sus derivadas.

Solución F4. Definimos en *todo* B la función

$$K(t) = \begin{cases} \dfrac{g(t) - g(b)}{t - b}, & t \in B \setminus \{b\}, \\ g'_B(b), & t = b. \end{cases}$$

El límite definitorio de $g'_B(b)$ asegura $\lim_{t \to b,\, t \in B \setminus \{b\}} K(t) = g'_B(b) = K(b)$. Por tanto, K es continua en b relativamente a B. La identidad $g(t) - g(b) = K(t)(t - b)$ vale para $t \neq b$ por construcción y para $t = b$ porque ambos lados son cero. En particular, para todo $h \neq 0$ admisible en A, incluso cuando $f(a+h) = b$,

$$\frac{(g \circ f)(a+h) - (g \circ f)(a)}{h} = K(f(a+h)) \frac{f(a+h) - f(a)}{h}.$$

La diferenciabilidad de f proporciona tanto $f(a+h) \to b$ como el límite $f'_A(a)$ del segundo factor. La continuidad relativa de K en b proporciona $K(f(a+h)) \to g'_B(b)$. Por las leyes de límites, el cociente de la composición converge al número finito $g'_B(b) f'_A(a)$. Esto demuestra la diferenciabilidad y la regla de la cadena sin cociente interior indefinido.

Solución F5. Sean $L = f'_A(a)$ y $M = g'_B(b)$, con $b = f(a)$. De la diferenciabilidad de f obtenemos, para h admisible,

$$f(a+h) = b + Lh + r_f(h), \qquad \frac{r_f(h)}{h} \longrightarrow 0.$$

Escribamos $t = t(h) = f(a+h) - b = Lh + r_f(h)$; entonces $t(h) \to 0$ y $\frac{t(h)}{h} = L + \frac{r_f(h)}{h} \longrightarrow L$.

En particular, $t(h)/h$ está acotado para $h \neq 0$ suficientemente pequeño. La diferenciabilidad de g permite definir $\eta(t) = \frac{g(b+t)-g(b)}{t} - M$ cuando $t \neq 0$ y $\eta(0) = 0$, restringiendo t a los desplazamientos admisibles de B. Entonces $\eta(t) \to 0$ y la identidad $g(b+t) = g(b) + Mt + t\eta(t)$

es válida también para $t = 0$. Como $\eta(t(h)) \to 0$, se tiene $\frac{t(h)\eta(t(h))}{h} = \frac{t(h)}{h}\eta(t(h)) \longrightarrow 0$. Por ello $t(h)\eta(t(h)) = o(h)$ incluso si $t(h) = 0$ para infinitos incrementos. Sustituyendo la primera linealización en la segunda,

$$\begin{aligned}(g \circ f)(a+h) &= g(b) + M[Lh + r_f(h)] + t(h)\eta(t(h)) \\ &= g(b) + MLh + o(h).\end{aligned}$$

La unicidad del coeficiente lineal, o directamente el cociente incremental, da $(g \circ f)'_A(a) = \boxed{ML = g'_B(f(a))f'_A(a)}$. No se aplicó la regla de la cadena: se reconstruyó su conclusión mediante residuos controlados.

Soluciones del nivel G. **Solución G1.** Para $n = 1$ vale la fórmula con producto vacío igual a 1. Si para $P_n = \prod_{k=1}^{n} f_k$ sabemos que

$$P'_n(a) = \sum_{j=1}^{n} f'_j(a) \prod_{\substack{1 \le k \le n \\ k \ne j}} f_k(a),$$

la regla del producto aplicada a $P_{n+1} = P_n f_{n+1}$ da

$$P'_{n+1}(a) = P'_n(a) f_{n+1}(a) + P_n(a) f'_{n+1}(a).$$

La sustitución de la hipótesis inductiva genera los términos $j = 1, \ldots, n$ con el nuevo factor $f_{n+1}(a)$ y el término adicional $j = n+1$. La fórmula queda demostrada para todo $n \ge 1$, en particular para el $n \ge 2$ pedido, sin dividir por factores. Si $f_p(a) = f_q(a) = 0$ con $p \ne q$, cada sumando de la fórmula omite como máximo *un* factor: contiene entonces al menos uno de los factores nulos. Por tanto,

$$\boxed{P'(a) = 0}.$$

Si exactamente un factor $f_j(a)$ es cero, todos los términos con índice distinto de j incluyen ese valor nulo. Solo sobrevive el término indexado por j:

$$\boxed{P'(a) = f'_j(a) \prod_{k \ne j} f_k(a)}.$$

No se concluye automáticamente que esta última derivada sea distinta de cero, pues también podría anularse $f'_j(a)$.

Solución G2. Definimos $f(0) = 0$ y, para cualquier real x, $f(x) = 0$ si $x \in \mathbb{Q}$ y $f(x) = x^2$ si $x \notin \mathbb{Q}$. Para $h \ne 0$,

$$\left| \frac{f(h) - f(0)}{h} \right| \le |h|,$$

pues $|f(h)| \le h^2$. El criterio del sándwich muestra que $f'(0) = 0$. Por densidad de $\mathbb{Q}$, cada entorno de 0 contiene infinitos racionales $h \ne 0$ tales que $f(h) - f(0) = 0$; por densidad de los irracionales, cada entorno contiene también un irracional $s \ne 0$ con $f(s) = s^2 > 0$. Luego f no es constante en ningún entorno de 0. Tomamos $g : \mathbb{R} \to \mathbb{R}$, $g(t) = 3t + 1$. Como $g'(t) = 3$ para todo t, $g'(f(0))f'(0) = 3 \cdot 0 = 0$. Verificamos la derivada compuesta *sin recurrir a la cadena*: $\frac{(g\circ f)(h) - (g\circ f)(0)}{h} = 3\frac{f(h)}{h} \longrightarrow 0$.

Así, $(g\circ f)'(0)=0=g'(f(0))f'(0)$. La descomposición formal que divide por $f(h)-f(0)$ carece de sentido en cualquiera de los infinitos incrementos racionales citados; la identidad por extensión o el cálculo directo sí son válidos. La exterior es afín y no constante, por lo que la coincidencia no depende de aplanar la composición mediante una función constante.

Solución G3. Escribamos $F=f(a)$, $G=g(a)\neq 0$, $L=f'_A(a)$ y $M=g'_A(a)$. Por diferenciabilidad existen residuos tales que, para incrementos admisibles,

$$f(a+h)=F+Lh+r_f(h),\qquad g(a+h)=G+Mh+r_g(h),$$

con $r_f(h)/h\to 0$ y $r_g(h)/h\to 0$. Como g es continua en a, elegimos $\delta>0$ de modo que $|g(a+h)-G|<|G|/2$ cuando $a+h\in A$ y $|h|<\delta$. En ese entorno, $|g(a+h)|>|G|/2>0$, por lo que el cociente está definido. Para $h\neq 0$ suficientemente pequeño, restamos el valor central *sin usar ninguna regla de derivación de cocientes*:

$$\begin{aligned}\frac{\dfrac{f(a+h)}{g(a+h)}-\dfrac{F}{G}}{h}&=\frac{G[f(a+h)-F]-F[g(a+h)-G]}{h\,G\,g(a+h)}\\&=\frac{G\left(L+\dfrac{r_f(h)}{h}\right)-F\left(M+\dfrac{r_g(h)}{h}\right)}{G\,g(a+h)}.\end{aligned}$$

El numerador tiende a $GL-FM$. Por continuidad, $g(a+h)\to G$; el denominador tiende al número no nulo G^2. Las leyes de límites, aplicadas a esta *identidad exacta*, prueban que el cociente incremental converge al número finito

$$\boxed{\left(\frac{f}{g}\right)'_A(a)=\frac{LG-FM}{G^2}=\frac{f'_A(a)g(a)-f(a)g'_A(a)}{g(a)^2}}.$$

La no anulación local fue necesaria para escribir la fracción inicial y pasar al límite. No se dividió por incrementos de f o g, que podrían ser cero, ni se usaron las reglas del recíproco o del cociente como premisas.

Capítulo 8

Derivación de funciones elementales, inversas e implícitas

Las reglas de derivación permiten operar con funciones cuya diferenciabilidad está establecida. Al estudiar inversas, raíces o ecuaciones entre variables aparece una cuestión previa: debemos identificar la función que existe con los dominios y las propiedades necesarios para aplicar esas reglas. Algunos símbolos esconden decisiones. La escritura f^{-1} exige que f sea invertible con dominio y codominio fijados; una raíz requiere demostrar la existencia del valor elegido y señalar su dominio; una ecuación como $F(x, y) = 0$ describe inicialmente un conjunto de pares, no una función y de x. Además, incluso cuando una función existe y es continua, su diferenciabilidad puede fallar en determinados puntos. No confundir estos problemas será el hilo conductor del capítulo.

Aquí estudiaremos las funciones **algebraicas accesibles con las herramientas disponibles**: polinomios, funciones racionales, raíces, potencias racionales en dominios admisibles y ciertas ramas de ecuaciones polinómicas. Logaritmos y exponenciales se construirán a partir de la integral en el capítulo visible 17. Las funciones trigonométricas y sus inversas requerirán sus propias definiciones y los límites que justifiquen sus derivadas.

Nuestra ruta será

función y dominio
$\longrightarrow$ existencia y continuidad de la inversa
$\longrightarrow$ derivada de la inversa
$\longrightarrow$ raíces y potencias racionales
$\longrightarrow$ ramas y derivación implícita justificada.

Para una relación polinómica probaremos la diferenciabilidad de una **rama continua previamente establecida**, bajo la condición algebraica indicada en el teorema. La prueba no proporcionará una construcción general de ramas a partir de cualquier ecuación; esa cuestión requiere el teorema de la función implícita en varias variables.

8.1. La fórmula no crea su función

Tres preguntas antes de calcular. La familiaridad con una expresión puede hacernos pasar directamente a su derivada. En el cálculo riguroso el orden debe ser otro. Primero preguntamos **qué objeto define la expresión**: una función completa, con dominio y codominio, o solamente una relación entre variables. Después comprobamos **que ese objeto existe con las propiedades necesarias**: por ejemplo, que una función posee inversa o que una ecuación tiene la rama elegida. Solo al final investigamos **si es diferenciable en el punto considerado** y qué fórmula permite hallar su derivada.

Estas comprobaciones responden a problemas distintos. Una función bien definida puede carecer de inversa; una inversa continua puede no tener derivada finita en cierto punto; una

ecuación puede admitir varias ramas. La regla de la cadena se aplica cuando las funciones y sus derivadas pertinentes ya están disponibles.

Una misma expresión no determina una misma función inversa. Consideremos la regla algebraica $x \mapsto x^2$. En el capítulo 2 aprendimos que la regla, aislada, no especifica una función completa: faltan al menos su dominio y su codominio. El problema se hace visible al intentar invertirla.

Ejemplo 8.1. **El cuadrado admite dos inversas después de elegir el dominio, pero no una inversa sobre toda la recta.** La función

$$f : \mathbb{R} \longrightarrow \mathbb{R}, \qquad f(x) = x^2,$$

no es inyectiva, pues $f(1) = f(-1) = 1$. No posee, por tanto, una función inversa f^{-1} sobre su imagen: el valor 1 tendría dos antecedentes. Tampoco la mera escritura $y^2 = t$ permite decidir entre ellos. La situación cambia si restringimos el dominio. Las funciones

$$f_+ : [0, \infty) \longrightarrow [0, \infty), \quad f_+(x) = x^2,$$
$$f_- : (-\infty, 0] \longrightarrow [0, \infty), \quad f_-(x) = x^2$$

son biyectivas. La existencia y unicidad de la raíz cuadrada no negativa, establecidas al construir los reales, dan sus inversas:

$$f_+^{-1}(t) = \sqrt{t}, \qquad f_-^{-1}(t) = -\sqrt{t}, \qquad t \geq 0.$$

Comprobemos las identidades, con sus dominios. Para $t \geq 0$,

$$f_+(\sqrt{t}) = t, \qquad f_-(-\sqrt{t}) = t.$$

Para $x \geq 0$, $\sqrt{x^2} = x$; para $x \leq 0$, $-\sqrt{x^2} = x$. Pero sobre toda la recta la primera igualdad no es cierta: si $x = -2$, entonces $\sqrt{x^2} = 2 \neq -2$. Precisamente por ello,

$$\boxed{\sqrt{x^2} = |x| \quad \text{para todo } x \in \mathbb{R},}$$

no $\sqrt{x^2} = x$ sin restricción. Las dos fórmulas inversas corresponden a **dos funciones de partida diferentes**, obtenidas mediante restricciones explícitas. Ninguna convierte en invertible a $f : \mathbb{R} \to \mathbb{R}$.

El ejemplo muestra una distinción decisiva: **resolver una ecuación no equivale a construir la inversa de una función que no es inyectiva**. La ecuación $y^2 = t$, para $t > 0$, tiene las soluciones $\sqrt{t}$ y $-\sqrt{t}$. Una inversa, en cambio, debe asignar exactamente un antecedente a cada elemento de su dominio. La elección de rama precede a toda fórmula de derivación.

La función inversa no es el recíproco. Los símbolos f^{-1} y $1/f$ pueden parecer próximos, pero expresan operaciones diferentes. Si $f : A \to B$ es una biyección, $f^{-1} : B \to A$ deshace su asignación: satisface $f^{-1}(f(x)) = x$ para $x \in A$. El recíproco de una función real se define, en cambio, por

$$\frac{1}{f}(x) = \frac{1}{f(x)} \qquad \text{cuando } x \in A \text{ y } f(x) \neq 0.$$

El capítulo 7 demostró la derivada del recíproco. Ese resultado **no** demuestra la fórmula de la derivada de una función inversa. En el ejemplo anterior, $f_+^{-1}(t) = \sqrt{t}$ está definida para

$t \geq 0$, mientras que $(1/f_+)(x) = 1/x^2$ solo está definida para $x > 0$. Son funciones distintas, con dominios y valores distintos; conocer cómo derivar la segunda no permite derivar automáticamente la primera.

Hay otra precaución: cuando obtengamos la fórmula de derivación inversa, será necesario indicar en qué puntos se evalúan ambas derivadas. Si $b = f(a)$, la derivada de f^{-1} se estudia en b, mientras que la de f se conoce en a. Una fórmula en la que esos puntos desaparezcan será, como mínimo, una abreviatura que habrá que reconstruir.

Una relación puede contener varias gráficas. Consideremos ahora $x^2 + y^2 = 1$.

La ecuación describe los pares reales que la satisfacen. Para todo $x \in (-1, 1)$, existen dos valores correspondientes de y, a saber, $\sqrt{1 - x^2}$ y $-\sqrt{1 - x^2}$. Así, el conjunto completo de soluciones no es la gráfica de una función y definida sobre $(-1, 1)$: viola la unicidad del valor de salida. Podemos, sin embargo, **seleccionar** una de sus ramas como una función genuina, por ejemplo

$$\varphi_+ : [-1, 1] \longrightarrow \mathbb{R}, \qquad \varphi_+(x) = \sqrt{1 - x^2}.$$

La existencia de la raíz de un número no negativo justifica su definición en ese dominio. Que sea una función y que satisfaga la ecuación no establece todavía la existencia de su derivada en todos sus puntos. Cerca de los extremos, además, los incrementos admisibles son unilaterales. Volveremos a esta distinción cuando abordemos las ramas algebraicas y la derivación implícita.

La frase informal «derivamos $x^2 + y^2 = 1$ respecto de x» oculta una hipótesis: que se ha elegido una función $y = \varphi(x)$ y que esa función admite las derivadas que se pretenden usar. No empezaremos por esa manipulación. Primero identificaremos las ramas y después demostraremos qué condiciones permiten deducir su diferenciabilidad desde la relación.

Dominios, puntos excluidos y condiciones suficientes. Del capítulo anterior conservamos otra enseñanza: una igualdad de expresiones sobre cierto conjunto no autoriza a añadir puntos a la función original. El cociente $(x^2 - 1)/(x - 1)$ coincide con $x + 1$ cuando $x \neq 1$, pero el primer cociente sigue sin estar definido en 1. Algo semejante ocurre al escoger una rama de una raíz: la regla puede ser algebraicamente correcta allí donde existe, pero esa corrección no extiende silenciosamente su dominio.

Del mismo modo, no debemos convertir una condición **suficiente** en una condición **necesaria**. El teorema que obtendremos para la derivada de una inversa exigirá, entre otras hipótesis, que cierta derivada sea no nula; de esa fórmula no se deduce que una función con derivada nula carezca de inversa. Inyectividad, continuidad de la inversa y diferenciabilidad de la inversa son cuestiones diferentes que investigaremos por separado.

Nuestra auditoría de trabajo puede formularse en forma de preguntas: ¿qué función se ha definido y sobre qué dominio?, ¿qué punto del dominio y qué incrementos son admisibles?, ¿la inversa o la rama ha sido construida?, ¿se ha demostrado su continuidad cuando es necesaria?, ¿qué resultado, exactamente, justifica la derivada? Cada respuesta deberá apoyarse en una definición, en un resultado anterior o en una demostración nueva, nunca en el aspecto conocido de la expresión.

Antes de continuar

Sin consultar el ejemplo, explica por qué $x \mapsto x^2$ no tiene inversa sobre $\mathbb{R}$ aunque para cada $t > 0$ pueda resolverse $x^2 = t$. Especifica después las dos restricciones que sí poseen inversa y demuestra las identidades de composición en sus dominios. Finalmente, identifica

el supuesto oculto en la instrucción «derivar implícitamente $x^2 + y^2 = 1$»: ¿qué función y se está derivando y quién ha demostrado que es diferenciable?

El siguiente paso será recuperar, sin rehacer las pruebas cerradas, los resultados sobre existencia y continuidad de funciones inversas de los capítulos 2 y 5. Solo después construiremos el teorema que permite calcular su derivada. **Una fórmula diferencial será la conclusión de ese recorrido, no el sustituto de sus hipótesis.**

8.2. Inversa: existencia y continuidad primero

Deshacer una asignación: tres tareas diferentes. En §8.1 vimos que restringir el cuadrado a los números no negativos o a los no positivos produce dos funciones invertibles, aunque la regla $x \mapsto x^2$ sobre toda la recta no lo sea. La situación sugiere una pregunta más amplia: **¿qué debemos saber de una función para poder invertirla y estudiar después la derivada de su inversa?**

La respuesta tiene tres etapas, que conviene mantener separadas:

1. **Existencia.** Identificamos una función inyectiva $f : A \to \mathbb{R}$ y tomamos su imagen efectiva $J = f(A)$. La aplicación $f : A \to J$ es biyectiva; la teoría del capítulo 2 proporciona una única inversa $g : J \to A$.
2. **Continuidad.** La existencia de g no afirma nada, por sí sola, sobre el comportamiento de sus valores ante entradas próximas. Necesitamos un resultado analítico adicional.
3. **Diferenciabilidad.** Incluso después de demostrar que g es continua, queda pendiente investigar si su cociente incremental converge a un número real. Esta será la tarea de §8.3.

La imagen $J = f(A)$ contiene exactamente los valores que poseen algún antecedente. Si el codominio original incluye otros valores, la función no puede asignarles antecedentes mediante una inversa. Debemos, por tanto, identificar una biyección entre el dominio y la imagen efectiva, o verificar la sobreyectividad respecto del codominio declarado.

Recuperación de un teorema ya demostrado. En el capítulo 5 demostramos Teorema 5.7. Su especial relevancia aquí merece recuperar el enunciado, **sin repetir su demostración**:

Inversa continua sobre un intervalo: resultado disponible
Sean $I \subseteq \mathbb{R}$ un intervalo y $f : I \to \mathbb{R}$ una función continua e inyectiva. Pongamos $J = f(I)$. Entonces J es un intervalo, existe la inversa única $g = f^{-1} : J \longrightarrow I$, y g es continua en cada punto de J, con continuidad relativa al dominio J. Si I contiene al menos dos puntos, f es estrictamente creciente o estrictamente decreciente. El caso de un intervalo unitario se interpreta directamente.

El resultado reúne conclusiones distintas. La **inyectividad** garantiza que una salida no tenga dos antecedentes; al considerar como codominio la **imagen**, garantizamos que tampoco falte ninguno. De ahí procede la existencia de g. La **continuidad de** g, en cambio, se deduce del teorema anterior gracias a que f es continua y su dominio es un *intervalo*. No se obtiene de la mera biyección.

La monotonía estricta que aparece en el teorema tampoco se ha deducido de un signo de f'. Su prueba ya fue construida en el capítulo 5 mediante continuidad, valores intermedios e inyectividad, antes de introducir el teorema del valor medio. Podemos utilizarla aquí porque

pertenece a nuestra teoría previa, pero no sustituir su justificación por un criterio diferencial que todavía no hemos demostrado.

Lee las flechas con su dominio
La inversión intercambia entradas y salidas:

$$f : I \longrightarrow J, \quad x \longmapsto y = f(x),$$
$$g = f^{-1} : J \longrightarrow I, \quad y \longmapsto x = g(y).$$

Sus identidades son $g(f(x)) = x$ para $x \in I$ y $f(g(y)) = y$ para $y \in J$. Para estudiar la continuidad de g en $b = f(a)$, los valores que se aproximan a b deben pertenecer a J; sus antecedentes pertenecen a I. Este cambio de variables será decisivo al derivar.

Una inversa global sin fórmula explícita. Los ejemplos escolares pueden transmitir la impresión de que solo hemos construido una inversa cuando conseguimos despejar una fórmula. No es así. La existencia de una función inversa es una propiedad de la asignación; disponer de una expresión cerrada es una cuestión distinta.

Ejemplo 8.2. **Una inversa existente y continua sin necesidad de despejar.** Sea

$$f : \mathbb{R} \longrightarrow \mathbb{R}, \qquad f(x) = x^3 + x.$$

Inyectividad por álgebra. Si $u < v$, entonces

$$f(v) - f(u) = (v - u)(v^2 + uv + u^2 + 1).$$

El primer factor es positivo. El segundo también lo es, porque

$$v^2 + uv + u^2 = \left(u + \frac{v}{2}\right)^2 + \frac{3v^2}{4} \geq 0.$$

Por tanto $f(v) > f(u)$: la función es estrictamente creciente y, en particular, inyectiva. No hemos empleado la derivada para obtener esta conclusión.

Imagen. Dado $t \in \mathbb{R}$, elijamos $M = |t| + 1$. Se tiene

$$f(-M) = -M^3 - M < t < M^3 + M = f(M).$$

Como f es un polinomio continuo, el teorema del valor intermedio proporciona algún $u \in (-M, M)$ con $f(u) = t$. La inyectividad hace único ese u. Por consiguiente, $f(\mathbb{R}) = \mathbb{R}$ y f es biyectiva.

Inversión y continuidad. Podemos definir sin ambigüedad

$$g : \mathbb{R} \longrightarrow \mathbb{R}, \qquad g(t) = \text{el único real } u \text{ tal que } u^3 + u = t.$$

El resultado Teorema 5.7 asegura que g es continua en toda la recta. No hemos despejado u ni calculado $g'(t)$; ninguna de esas operaciones era necesaria para justificar su existencia y continuidad. Por ejemplo, $g(0) = 0$ y $g(2) = 1$, pues $0^3 + 0 = 0$ y $1^3 + 1 = 2$.

Este ejemplo separa el *conocimiento estructural* del *cálculo simbólico*. Sabemos qué función es g, cuál es su dominio, por qué cada valor tiene una única salida y por qué es continua. Todavía no sabemos, a partir de lo aquí demostrado, cuál es su derivada: esa conclusión exige un argumento nuevo.

El intervalo no es un adorno de las hipótesis. ¿Podríamos eliminar la palabra «intervalo» del teorema y conservar su conclusión para cualquier subconjunto de la recta? El siguiente ejemplo muestra que la continuidad de la función original y su inyectividad no bastan para ese enunciado general.

Ejemplo 8.3. **Una biyección continua cuya inversa es discontinua.** Definamos

$$A = [0,1) \cup \{2\}, \qquad f : A \longrightarrow [0,1],$$

mediante

$$f(x) = \begin{cases} x, & 0 \le x < 1, \\ 1, & x = 2. \end{cases}$$

La función es inyectiva: los valores obtenidos en $[0,1)$ pertenecen a $[0,1)$ y el valor de 2 es 1, distinto de todos ellos. Su imagen es exactamente $[0,1]$, de modo que es biyectiva. Además, f es continua **relativamente a** A. En cada punto $a \in [0,1)$ coincide localmente con la identidad; en 2 la continuidad relativa es inmediata porque 2 es un punto aislado de A: si $|x-2| < 1/2$ y $x \in A$, entonces $x = 2$.

La inversa existe y es

$$g : [0,1] \longrightarrow A, \qquad g(y) = \begin{cases} y, & 0 \le y < 1, \\ 2, & y = 1. \end{cases}$$

Pero g **no es continua en** 1. Basta fijar $\varepsilon = 1/2$. Para cualquier $\delta > 0$, tomemos

$$y = 1 - \min\left\{\frac{\delta}{2}, \frac{1}{2}\right\}.$$

Entonces $y \in [0,1)$ y $|y-1| < \delta$, mientras que

$$|g(y) - g(1)| = |y - 2| > 1 > \varepsilon.$$

Por tanto ningún radio δ garantiza la continuidad de g en 1. La función de partida era continua e inyectiva, pero A no es un intervalo: allí fracasa la conclusión que intentábamos generalizar.

El contraejemplo no dice que una función con dominio no intervalar tenga siempre inversa discontinua. Dice algo más preciso: **sin una hipótesis adicional, la continuidad e inyectividad de f sobre un subconjunto arbitrario no autorizan a afirmar la continuidad de su inversa**. En nuestras aplicaciones al cálculo de una variable, trabajar sobre intervalos nos proporciona exactamente la garantía establecida en el capítulo 5.

Volver a las ramas: restringir también significa cambiar la imagen. Recuperemos brevemente el cuadrado, ahora mediante un diagrama de funciones tipadas. La elección de un intervalo modifica tanto los antecedentes permitidos como la imagen a la que llega la inversa:

Función de partida	Imagen efectiva	Función inversa
$q_+ : [0,2] \to [0,4]$, $q_+(x) = x^2$	$[0,4]$	$q_+^{-1} : [0,4] \to [0,2]$, $q_+^{-1}(t) = \sqrt{t}$
$q_- : [-2,0] \to [0,4]$, $q_-(x) = x^2$	$[0,4]$	$q_-^{-1} : [0,4] \to [-2,0]$, $q_-^{-1}(t) = -\sqrt{t}$

La existencia de las raíces cuadradas no negativas, probada en el capítulo 1, permite verificar las dos imágenes y las dos identidades inversas. Las funciones q_+ y q_- son continuas y estrictamente monótonas en sus respectivos intervalos: para $0 \le u < v$ tenemos $u^2 < v^2$, mientras que para $u < v \le 0$ se cumple $u^2 > v^2$. Por Teorema 5.7, **ambas inversas son continuas sobre** $[0, 4]$, incluidos sus extremos en el sentido relativo.

No debemos reemplazar ese enunciado por «las inversas son derivables también en 0 y 4». Continuidad y derivabilidad son propiedades diferentes; además, una derivada en un extremo, según la definición de nuestro capítulo 6, solo podría considerarse con los incrementos admitidos por el dominio. El estudio de tales cocientes pertenece al paso siguiente.

Comprueba que entendiste las hipótesis
Considera una función continua $f : I \to \mathbb{R}$ sobre un intervalo. ¿Qué información adicional garantiza que tenga inversa sobre $f(I)$? ¿Qué teorema garantiza la continuidad de esa inversa? Explica por qué ninguna de estas respuestas exige conocer una fórmula para f^{-1} ni el signo de f'. Por último, identifica en el ejemplo del dominio $[0, 1) \cup \{2\}$ la hipótesis del teorema que dejó de cumplirse.

Hemos establecido el punto de partida para la derivación de funciones inversas: disponemos de una biyección $f : I \to J$ y de una función $g = f^{-1} : J \to I$ cuya continuidad ya está justificada. En §8.3 fijaremos $a \in I$, escribiremos $b = f(a)$ y estudiaremos directamente el cociente

$$\frac{g(y) - g(b)}{y - b}, \qquad y \in J, \quad y \neq b.$$

Solo allí aparecerá una fórmula para su límite, y solo después de verificar las hipótesis que permiten invertir el correspondiente cociente de f.

8.3. Derivada de una función inversa

¿De dónde puede salir el recíproco de una pendiente? La sección anterior terminó con dos funciones verdaderamente construidas: $f : I \to J$, continua e inyectiva sobre un intervalo, y su inversa continua $g : J \to I$, donde $J = f(I)$. Esta información todavía no contiene una derivada de g. Para descubrirla, conviene comenzar por dos puntos distintos de la gráfica, no por una identidad derivada formalmente. Fijemos $a \in I$ y llamemos $b = f(a)$. Tomemos una salida $y \in J$, distinta de b, y recuperemos su antecedente $x = g(y)$. Las identidades inversas dan

$$g(b) = a, \qquad f(x) = y.$$

Como f es inyectiva, $y \neq b$ equivale a $x \neq a$. Por tanto, los incrementos $y - b$ y $x - a$ son ambos no nulos y podemos escribir la igualdad **exacta**

$$\frac{g(y) - g(b)}{y - b} = \frac{x - a}{f(x) - f(a)} = \left(\frac{f(x) - f(a)}{x - a} \right)^{-1}.$$

La razón de cambio de la inversa es, entre los puntos correspondientes, el recíproco de la razón de cambio de la función original. Pero **una identidad entre cocientes no demuestra todavía que sus límites existan**. Para convertirla en un teorema tenemos que controlar dónde se acercan x e y y cuándo se permite pasar al recíproco.

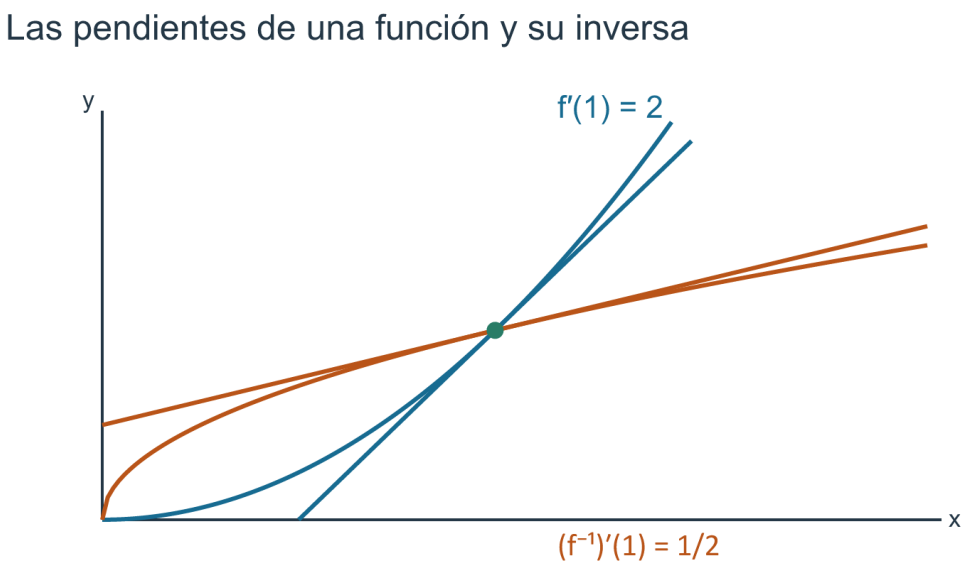


FIGURA 8.1. En el punto $(1,1)$, la tangente a x^2 tiene pendiente 2 y la tangente a su inversa $\sqrt{x}$ tiene pendiente $1/2$. La fórmula exige que la derivada de la función original no sea cero.

El enunciado: dominios y puntos de evaluación visibles.

TEOREMA 8.1. **Teorema de la derivada de una función inversa en un intervalo.** Sea $I \subseteq \mathbb{R}$ un intervalo y sea $f : I \to \mathbb{R}$ continua e inyectiva. Definamos $J = f(I)$ y consideremos la inversa, existente y continua por Teorema 5.7,

$$g = f^{-1} : J \longrightarrow I.$$

Sea $a \in I$ un punto de acumulación de I, y pongamos $b = f(a)$. Supongamos que existe la derivada relativa finita de f en a y que

$$f'_I(a) \neq 0.$$

Entonces b es un punto de acumulación de J, g es diferenciable en b relativamente a J y

$$\boxed{g'_J(b) = \frac{1}{f'_I(a)}.}$$

Equivalente y abreviadamente, siempre que se cumplan las hipótesis en $a = g(b)$,

$$(f^{-1})'(b) = \frac{1}{f'(f^{-1}(b))}.$$

Demostración. Separamos los pasos que suelen quedar ocultos en la fórmula.

1. La inversa existe y es continua antes de derivarla. Por hipótesis, f es inyectiva y J es su imagen; por tanto $f : I \to J$ es biyectiva. El teorema Teorema 5.7 asegura que $g = f^{-1}$ existe y es continua en b, relativamente a J. No hemos supuesto que g sea diferenciable.

2. El punto b permite un límite perforado en J. Como a es punto de acumulación de I, podemos escoger puntos $x \in I \setminus \{a\}$ arbitrariamente próximos a a. La continuidad de f en a garantiza que sus imágenes $f(x)$ se aproximan a $f(a) = b$; la inyectividad garantiza, además, que $f(x) \neq b$. Más explícitamente, dado $\varepsilon > 0$, la continuidad proporciona $\delta > 0$ tal que

$$x \in I, \quad |x - a| < \delta \quad \Longrightarrow \quad |f(x) - b| < \varepsilon.$$

Por acumulación elegimos $x \in I$ con $0 < |x - a| < \delta$. Entonces $y = f(x) \in J$ cumple $0 < |y - b| < \varepsilon$. Esto demuestra que b es punto de acumulación de J, condición necesaria para definir $g'_J(b)$ según Definición 6.1.

3. Los antecedentes se aproximan al punto correcto. Para $y \in J \setminus \{b\}$ escribamos $x = g(y)$. Por la continuidad de g en b,

$$\lim_{\substack{y \to b \\ y \in J}} g(y) = g(b) = a.$$

En consecuencia, cuando $y \to b$ relativamente a J, el antecedente $x = g(y)$ tiende a a relativamente a I; además, permanece distinto de a por inyectividad. **Este es el paso en el que necesitamos la continuidad de la inversa.** La mera igualdad $f(x) = y$ no bastaría para deducirlo.

4. El denominador del recíproco queda separado de cero. Por la definición de derivada relativa,

$$Q(x) := \frac{f(x) - f(a)}{x - a} \longrightarrow L := f'_I(a) \neq 0 \qquad (x \to a,\ x \in I \setminus \{a\}).$$

En particular, para x suficientemente próximo a a dentro del dominio perforado, $|Q(x) - L| < |L|/2$, de modo que $|Q(x)| > |L|/2 > 0$. El cociente $Q(x)$ está bien definido y su recíproco también. Además,

$$\left| \frac{1}{Q(x)} - \frac{1}{L} \right| = \frac{|Q(x) - L|}{|Q(x)|\,|L|} \leq \frac{2}{|L|^2} |Q(x) - L| \longrightarrow 0.$$

Esta estimación muestra directamente por qué es indispensable la condición $L \neq 0$ para aplicar el paso al recíproco.

5. Concluimos mediante el cociente incremental de la inversa. Para $y \in J \setminus \{b\}$ y $x = g(y)$ tenemos

$$\begin{aligned} \frac{g(y) - g(b)}{y - b} &= \frac{x - a}{f(x) - f(a)} \\ &= \frac{1}{Q(x)}. \end{aligned}$$

El paso 3 permite aplicar el límite del paso 4 cuando $y \to b$ dentro de J. Por tanto,

$$\lim_{\substack{y \to b \\ y \in J}} \frac{g(y) - g(b)}{y - b} = \frac{1}{L} = \frac{1}{f'_I(a)}.$$

El límite existe como número real. Según la definición de derivada relativa, hemos demostrado simultáneamente que g es diferenciable en b y que su derivada tiene el valor anunciado. □

Leer la prueba en sentido inverso. El resultado admite una lectura geométrica: las coordenadas de los puntos de las gráficas de f y g se intercambian, y con ellas se intercambian los incrementos horizontal y vertical. De ahí nace el recíproco de las pendientes secantes. La pendiente límite de la inversa solo puede obtenerse de esa observación cuando los antecedentes correspondientes convergen al punto adecuado y la pendiente límite original no es cero. La función f y su inversa se evalúan en **puntos distintos**:

$$\begin{aligned} a \in I, \qquad & b = f(a) \in J, \\ f'_I(a) = L \neq 0, \quad & g'_J(b) = 1/L. \end{aligned}$$

Por eso escribir $(f^{-1})'(x) = 1/f'(x)$ sin explicar qué representa x es, en general, incorrecto. Si la variable de entrada de la inversa se llama y, el antecedente en el que se evalúa f' es $g(y)$: la forma correcta es $g'(y) = 1/f'(g(y))$, exclusivamente donde las hipótesis estén verificadas.

Una justificación circular que debemos evitar
De $g(f(x)) = x$ sería tentador escribir $(g \circ f)'(a) = g'(f(a))f'(a) = 1$ y despejar $g'(f(a))$. **Ese cálculo ya supone que $g'(f(a))$ existe**, precisamente lo que teníamos que demostrar. Ahora sí podemos emplear la regla de la cadena de Teorema 7.4 como comprobación *posterior* del resultado, nunca como su fundamento inicial.

Una inversa cuyo valor conocemos sin disponer de una fórmula cerrada. El ejemplo de §8.2 proporciona una aplicación en la que el teorema evita la necesidad de despejar la función inversa.

Ejemplo 8.4. **Derivar la inversa de $x^3 + x$ en dos puntos.** Sea

$$f : \mathbb{R} \longrightarrow \mathbb{R}, \qquad f(x) = x^3 + x,$$

cuya biyectividad y cuya inversa continua $g : \mathbb{R} \to \mathbb{R}$ quedaron demostradas en Ejemplo 8.2 mediante álgebra y el teorema del valor intermedio. El capítulo 7 permite derivar el polinomio: $f'(x) = 3x^2 + 1$. Esta cantidad es positiva y, en particular, no nula para todo $x \in \mathbb{R}$. **No necesitamos deducir la inyectividad a partir de ese signo:** ya quedó probada independientemente. Sabemos que $f(0) = 0$ y $f(1) = 2$, de modo que $g(0) = 0$ y $g(2) = 1$. Aplicando el teorema en $a = 0$ y $a = 1$, respectivamente, obtenemos

$$\begin{aligned} g'(0) &= \frac{1}{f'(0)} = 1, \\ g'(2) &= \frac{1}{f'(1)} = \frac{1}{4}. \end{aligned}$$

De hecho, para cualquier $t \in \mathbb{R}$ la misma prueba autoriza la identidad $g'(t) = \frac{1}{3[g(t)]^2+1}$. La fórmula expresa la derivada mediante la inversa **ya definida**; no exige encontrar una expresión elemental explícita para $g(t)$.

Cuando la derivada original es cero: un límite que no se puede invertir. El teorema no afirma que una función con $f'(a) = 0$ carezca de inversa. La inyectividad y la diferenciabilidad son propiedades diferentes. Lo que sucede es que el recíproco $1/f'(a)$ deja de ser un número real y nuestro argumento no puede concluir la existencia de una derivada finita para la inversa.

Hay una afirmación más precisa que podemos demostrar: **si ambas funciones inversas son diferenciables con derivadas finitas en los puntos correspondientes, la derivada de la función original no puede ser cero.** En efecto, solo después de suponer ambas derivabilidades es lícito aplicar la regla de la cadena a $g \circ f = \mathrm{id}_I$; obtenemos $1 = (g \circ f)'_I(a) = g'_J(b) f'_I(a)$.

El primer miembro es la derivada relativa de la identidad en el punto de acumulación a. Si $f'_I(a) = 0$, la igualdad sería imposible para cualquier valor real de $g'_J(b)$. Se trata de una **condición necesaria** bajo derivabilidad de ambas, no de una prueba de existencia de inversa a partir de $f'(a) \neq 0$. Veamos un caso concreto que aísla este fenómeno. Definamos

$$h : [-1, 1] \longrightarrow [-1, 1], \qquad h(x) = x^3.$$

Para $u < v$, la factorización

$$v^3 - u^3 = (v - u)(v^2 + uv + u^2) > 0$$

es estrictamente positiva: el primer factor es positivo y el segundo lo es para $u \neq v$, puesto que $v^2 + uv + u^2 = \left(u + \frac{v}{2}\right)^2 + \frac{3v^2}{4}$, cuyos dos sumandos solo se anulan simultáneamente cuando $u = v = 0$, excluido aquí. Por continuidad y el teorema del valor intermedio, $h([-1, 1]) = [-1, 1]$; por tanto tiene inversa continua $k : [-1, 1] \to [-1, 1]$. Se cumple $h'(0) = 0$ y $k(0) = 0$. Para examinar si k tiene derivada finita en el origen, usemos las entradas $y = u^3$, donde $u \in [-1, 1] \setminus \{0\}$. Entonces $k(y) = u$ y $\frac{k(y)-k(0)}{y-0} = \frac{u}{u^3} = \frac{1}{u^2}$.

Cuando $y \to 0$ dentro de $[-1, 1]$, la continuidad de k implica $u \to 0$; en particular, los valores anteriores crecen sin cota al tomar, por ejemplo, $u = 1/n$ con $n \to \infty$. Si el cociente incremental de k tuviera límite real, estaría acotado en un entorno perforado de cero, contradicción. **La inversa existe y es continua, pero no tiene derivada finita en** 0**.** No hemos necesitado introducir aún una notación general de raíces de índice arbitrario.

El mismo cuidado se requiere en los extremos. La inversa del cuadrado restringido $q : [0, \infty) \to [0, \infty)$ es $q^{-1}(t) = \sqrt{t}$, pero $q'_{[0,\infty)}(0) = 0$. Su cociente incremental en el extremo de la imagen satisface, para $t = u^2 > 0$, $\frac{q^{-1}(t)-q^{-1}(0)}{t} = \frac{u}{u^2} = \frac{1}{u}$, que no posee límite real finito cuando $u \to 0^+$. Es una afirmación sobre la derivada **relativa al dominio** $[0, \infty)$, no sobre valores que la raíz ni siquiera admite.

Qué demuestra el teorema y qué no
La hipótesis $f'_I(a) \neq 0$ es suficiente para la fórmula **junto con** continuidad, inyectividad, dominio intervalar y acumulación del punto. No construye una inversa por sí sola. Una derivada no nula en un punto aislado de la discusión global no autoriza a inferir inyectividad en algún entorno: esa cuestión se analizará expresamente en §8.7. Por otro lado, cuando $f'_I(a) = 0$, la inversa puede seguir existiendo, pero no puede tener derivada real finita en b si ambas funciones se relacionan como hemos supuesto y la regla de la cadena es aplicable.

Comprobación activa: reconstruir las hipótesis. Sin mirar la demostración, completa el siguiente argumento. Se sabe que f es continua e inyectiva en un intervalo I; se fija $a \in I$, punto de acumulación, y se define $b = f(a)$ y $g = f^{-1}$. ¿Qué resultado demuestra que g es continua? ¿Cómo justificas que b sea punto de acumulación de $J = f(I)$? ¿Por qué $y \neq b$ implica $g(y) \neq a$? ¿Qué hipótesis permite pasar al límite del recíproco y por qué no es legítimo empezar derivando $g \circ f$?

La respuesta configura la cadena completa:

inyectividad e imagen
$\longrightarrow$ inversa existente y continua
$\longrightarrow$ cocientes recíprocos
$\longrightarrow$ límite real de la inversa.

En §8.4 emplearemos este resultado para estudiar las raíces de cualquier índice, pero solo **después de construir rigurosamente** cada función raíz mediante continuidad, valor intermedio y unicidad algebraica. La fórmula de la derivada nunca reemplazará esa construcción.

8.4. Raíces de cualquier índice: construir antes de derivar

¿Qué asegura el símbolo de una raíz? La construcción de la raíz cuadrada no negativa sugiere extender la pregunta a otros índices naturales. Para cada índice habrá que demostrar existencia y unicidad de la raíz en los valores admisibles de t. Una vez construida la función y fijado su dominio, podremos investigar su derivada mediante el teorema de la inversa. Fijemos un entero $n \geq 1$. Consideraremos inicialmente la función polinómica

$$q_n : [0, \infty) \longrightarrow \mathbb{R}, \qquad q_n(u) = u^n.$$

Su continuidad está disponible desde el capítulo 5. Para invertirla debemos resolver dos cuestiones independientes: demostrar que alcanza cada $t \geq 0$ y que nunca alcanza ese mismo valor en dos puntos distintos del dominio. La primera utilizará el teorema del valor intermedio; la segunda, una identidad algebraica.

Existencia, unicidad y continuidad de la raíz no negativa.

Proposición 8.1. **Construcción y derivación de las raíces de índice natural.** Sea $n \in \mathbb{N}$ con $n \geq 1$.

1. Para cada $t \geq 0$ existe un único $u \geq 0$ tal que $u^n = t$. Denotaremos ese número por $r_n(t)$ o $\sqrt[n]{t}$. La función
$$r_n : [0, \infty) \longrightarrow [0, \infty)$$
es la inversa de $q_n : [0, \infty) \to [0, \infty)$, $q_n(u) = u^n$, y es continua respecto de su dominio.
2. Para cada $t > 0$, r_n posee derivada finita y
$$r_n'(t) = \frac{1}{n\,[r_n(t)]^{n-1}}.$$
3. Si n es impar, existe además una única función raíz real $\widetilde{r}_n : \mathbb{R} \to \mathbb{R}$ caracterizada por $[\widetilde{r}_n(t)]^n = t$. Coincide con r_n para $t \geq 0$, es continua en toda la recta y, para $t \neq 0$,
$$\widetilde{r}_n'(t) = \frac{1}{n\,[\widetilde{r}_n(t)]^{n-1}}.$$
Si n es par, ningún $t < 0$ tiene raíz real de índice n.

Demostración.

1. Existencia para $t \geq 0$. Fijemos $t \geq 0$ y elijamos $M = t + 1 > 0$. Puesto que $M \geq 1$ y $n \geq 1$, tenemos $M^n \geq M > t$. El polinomio q_n es continuo en $[0, M]$, y

$$q_n(0) = 0 \leq t < M^n = q_n(M).$$

Por el teorema del valor intermedio existe $u \in [0, M]$ tal que $u^n = t$. En particular, cuando $t = 0$, la construcción permite $u = 0$: no necesitamos dividir por t ni excluir el extremo.

2. Unicidad sin utilizar derivadas. Si $0 \leq u < v$, la factorización de la diferencia de potencias proporciona

$$v^n - u^n = (v - u)\big(v^{n-1} + v^{n-2}u + \cdots + u^{n-1}\big).$$

Para $n = 1$, esta igualdad se reduce a $v - u > 0$. Para $n \geq 2$, el primer factor es positivo y cada sumando del segundo es no negativo; además $v^{n-1} > 0$, pues $v > u \geq 0$. Por tanto $v^n - u^n > 0$ en todos los casos. La función q_n es estrictamente creciente en $[0, \infty)$ por un argumento puramente algebraico y, en particular, inyectiva. Dos números no negativos distintos no pueden tener la misma potencia n-ésima. Quedan probadas la existencia y la unicidad prometidas. Ahora, y no antes, la asignación $t \mapsto r_n(t)$ está bien definida: $r_n(t)$ es el único $u \geq 0$ hallado. La existencia para todo $t \geq 0$ y la inyectividad prueban que

$$q_n : [0, \infty) \longrightarrow [0, \infty)$$

es biyectiva, con identidades

$$q_n(r_n(t)) = t \quad (t \geq 0), \qquad r_n(q_n(u)) = u \quad (u \geq 0).$$

El teorema Teorema 5.7, aplicado al intervalo $[0, \infty)$, asegura la continuidad de r_n en todo su dominio, **incluido** 0 **en sentido relativo**. Ni la inyectividad ni esta continuidad han sido obtenidas a partir de una derivada.

3. Derivada para $t > 0$. Fijemos $t > 0$ y pongamos $a = r_n(t) > 0$. El capítulo 7 proporciona la derivada de una potencia natural, relativamente a $[0, \infty)$:

$$(q_n)'_{[0,\infty)}(a) = na^{n-1} > 0.$$

Las hipótesis de Teorema 8.1 ya están justificadas: dominio intervalar, continuidad, inyectividad, inversa continua, punto a de acumulación y derivada no nula. Como $q_n(a) = t$, concluimos que

$$r_n'(t) = \frac{1}{(q_n)'(a)} = \frac{1}{na^{n-1}} = \frac{1}{n\,[r_n(t)]^{n-1}}.$$

Para $n = 1$ interpretamos $a^0 = 1$; la fórmula dice simplemente $r_1(t) = t$ y $r_1'(t) = 1$, como corresponde a la identidad.

4. Extensión para índice impar. Supongamos que n es impar. Si $t < 0$, la existencia y unicidad anteriores, aplicadas a $-t > 0$, proporcionan un único $v = r_n(-t) > 0$ con $v^n = -t$. Puesto que n es impar, $u = -v$ satisface

$$u^n = (-v)^n = -v^n = t.$$

Es el único antecedente real: si $w^n = t < 0$, entonces $w < 0$ (las potencias impares conservan el signo); así $-w > 0$ y $(-w)^n = -t$, por lo que la unicidad no negativa da $-w = v$ y $w = u$. Para $t \geq 0$ ya existe exactamente la raíz no negativa; ninguna raíz negativa puede tener potencia positiva o nula distinta de cero. Podemos definir, por tanto,

$$\widetilde{r}_n(t) = \begin{cases} r_n(t), & t \geq 0, \\ -r_n(-t), & t < 0. \end{cases}$$

Acabamos de construir una inversa global de $Q_n : \mathbb{R} \to \mathbb{R}$, $Q_n(u) = u^n$. Esta función es continua por ser polinómica e inyectiva por la unicidad de las raíces para cada signo; su imagen es toda $\mathbb{R}$ por existencia. El teorema Teorema 5.7 aplicado al intervalo $\mathbb{R}$ demuestra que $\widetilde{r}_n$ es continua en toda la recta, incluido el origen: no hace falta presumir continuidad en el empalme de la definición por casos. Si $t < 0$ y $a = \widetilde{r}_n(t) < 0$, se cumple $Q_n'(a) = na^{n-1} > 0$, ya que $n-1$ es par. El teorema Teorema 8.1 da la fórmula anunciada; para $t > 0$ ya la obtuvimos en el paso anterior. El caso $n = 1$ vuelve a ser la identidad, incluida su derivada en 0.

5. Índice par y entrada negativa. Si n es par, $u^n \geq 0$ para cualquier $u \in \mathbb{R}$. Por consiguiente, la igualdad $u^n = t$ es imposible en los reales cuando $t < 0$. Esto no significa que falle un método para hallar la raíz: **no existe el número real buscado**. □

Interpretar la construcción: raíz, dominio y elección de signo. Lo que llamamos $\sqrt[n]{t}$ para $t \geq 0$ es, por definición, **la raíz no negativa**. Si el índice es par y $t > 0$, también $-\sqrt[n]{t}$ satisface la ecuación $u^n = t$; son las dos soluciones reales, no dos valores de una misma función. Para $t = 0$ ambas expresiones coinciden en el único valor 0. Si el índice es impar, la raíz es única en toda la recta: en adelante podremos escribir $\sqrt[n]{t} = \widetilde{r}_n(t)$ también para $t < 0$, sin cambiar de rama. Por ejemplo, de $(-2)^3 = -8$ se deduce $\sqrt[3]{-8} = -2$. En cambio, la escritura $\sqrt[4]{-8}$ no designa un número real. No basta que una manipulación simbólica resulte familiar: es necesario que sus entradas pertenezcan al dominio de la función efectivamente construida.

Hay un detalle de notación que será decisivo en §8.5: **todavía no hemos definido potencias de exponente racional como una operación general**. La igualdad $\sqrt[n]{t} = t^{1/n}$ podrá adoptarse cuando fijemos su significado y sus dominios; no ha intervenido como premisa en la construcción anterior.

¿Qué ocurre en cero? El límite que la fórmula no incluye. El resultado Proposición 8.1 garantiza diferenciabilidad de r_n para $t > 0$, pero no permite sustituir $t = 0$ en un denominador que entonces puede anularse. El caso $n = 1$ es excepcional y sencillo: r_1 es la identidad y su derivada relativa en 0 vale 1. Consideremos $n \geq 2$. Para $h > 0$, pongamos $u = r_n(h) > 0$. Como $u^n = h$ y $r_n(0) = 0$, el cociente incremental derecho es exactamente $\frac{r_n(h)-r_n(0)}{h} = \frac{u}{u^n} = \frac{1}{u^{n-1}}$.

Por la continuidad de r_n en 0, $u = r_n(h) \to 0^+$ cuando $h \to 0^+$. Dado cualquier $K > 0$, basta tomar $u > 0$ suficientemente pequeño para que $u^{n-1} < 1/K$; entonces el cociente es mayor que K. Podemos hacerlo con entradas admisibles $h = u^n \to 0^+$. Por tanto el cociente no tiene límite real finito: r_n **no es derivable en** 0 **respecto de** $[0, \infty)$ **cuando** $n \geq 2$.

Si $n \geq 3$ es impar, la extensión $\widetilde{r}_n$ tampoco tiene derivada finita en cero. Ahora podemos tomar u de cualquiera de los dos signos; para $h = u^n \neq 0$ seguimos teniendo $\frac{\widetilde{r}_n(h)-\widetilde{r}_n(0)}{h} = \frac{1}{u^{n-1}}$, que crece sin cota porque $n-1$ es par. Esto recupera y generaliza el diagnóstico de la inversa de x^3 de §8.3. No estamos asignando el valor $+\infty$ a una derivada real: la definición del capítulo 6 exige un límite **finito**, que aquí no existe.

Un ejemplo que obliga a comprobar todas las etapas.

Ejemplo 8.5. **Cuarta raíz positiva y raíz cúbica negativa.** Sabemos que $3^4 = 81$, con $3 \geq 0$; la unicidad de Proposición 8.1 demuestra $\sqrt[4]{81} = 3$. Para $t = 81 > 0$, la fórmula diferencial sí es aplicable y da

$$\frac{d}{dt}\sqrt[4]{t}\Big|_{t=81} = \frac{1}{4 \cdot 3^3} = \frac{1}{108}.$$

Para el índice impar, $(-2)^3 = -8$ y la raíz construida globalmente satisface $\sqrt[3]{-8} = -2$. Como su antecedente no es cero,

$$\frac{d}{dt}\sqrt[3]{t}\Big|_{t=-8} = \frac{1}{3(-2)^2} = \frac{1}{12}.$$

La comparación importante no es solo numérica: el primer cálculo utiliza una inversa definida sobre $[0,\infty)$; el segundo, una inversa definida sobre toda $\mathbb{R}$. En ninguno hemos deducido existencia, unicidad ni dominio a partir de una supuesta regla de derivación.

Un salto ilegítimo frecuente
«La derivada de u^n es nu^{n-1}; luego $\sqrt[n]{t}$ existe y su derivada es $1/(n\sqrt[n]{t}^{\,n-1})$». La primera premisa no demuestra que una ecuación $u^n = t$ tenga solución ni que esta sea única en el dominio elegido. La fórmula diferencial solo se obtiene **después** de establecer una biyección continua, su inversa continua y una derivada original no nula en el antecedente.

Comprueba las dependencias antes de avanzar. Para $n = 5$ y $t = -32$, justifica la existencia y unicidad de la raíz real, indica su valor y calcula la derivada en ese punto. Repite únicamente las preguntas de existencia y dominio para $n = 4$ y $t = -32$. Finalmente, explica por qué el argumento de $r_n'(t)$ para $t > 0$ no demuestra diferenciabilidad en $t = 0$ si $n \geq 2$, aunque r_n sea continua allí. Nuestra cadena queda cerrada hasta las raíces:

continuidad de u^n
$\longrightarrow$ existencia por valor intermedio
$\longrightarrow$ unicidad algebraica
$\longrightarrow$ inversa continua
$\longrightarrow$ derivada donde procede.

En §8.5 construiremos las potencias de exponente racional y estudiaremos qué fórmulas conservan su significado al combinar raíces, signos y cocientes. No presentaremos identidades de exponentes fuera de sus dominios legítimos.

8.5. Potencias racionales y funciones algebraicas explícitas

¿Qué significa elevar a una fracción? La construcción de §8.4 nos permite afirmar que, para cada $n \geq 1$, existe una raíz n-ésima no negativa de todo número no negativo; además, si n es impar, existe una raíz real única para cualquier argumento. Sin embargo, ninguna de esas afirmaciones define por sí sola la expresión $x^{m/n}$. Falta escoger una representación del exponente, determinar el dominio y decidir qué hacemos con un exponente negativo. Solo después tendremos una función a la que aplicar las reglas de derivación. Fijemos un número racional q. Lo escribiremos en su **forma reducida**

$$q = \frac{m}{n}, \qquad m \in \mathbb{Z}, \quad n \in \mathbb{N}, \quad n \geq 1, \quad \gcd(|m|, n) = 1.$$

Para $q = 0$, la representación reducida es $0/1$. Esta convención evita que las propiedades de una función dependan de una elección accidental de denominador. Por ejemplo, para $x < 0$, escribir una fracción sin reducir puede inducirnos a extraer una raíz de índice par que no existe, pese a que el exponente reducido tenga denominador impar.

Definición con dominio. Si $m \neq 0$ y $x > 0$, ponemos $x^{m/n} := [r_n(x)]^m$, con las potencias enteras interpretadas como en el capítulo 7; el número $r_n(x)$ es estrictamente positivo, de modo

que también están definidos sus exponentes negativos. Si n es impar, utilizamos exactamente la misma expresión para $x < 0$, pero entendiendo $r_n(x)$ como la extensión real impar $\widetilde{r}_n(x)$ de Proposición 8.1. Para $x = 0$, definimos $0^{m/n} = 0$ **si** $m > 0$; si $m < 0$, la expresión queda indefinida porque exigiría el recíproco de cero. Para $q = 0$, adoptamos la función potencia constante $x^0 = 1$ en toda $\mathbb{R}$, en consonancia con el tratamiento funcional del capítulo 7: no estamos asignando aquí un valor a un símbolo aislado 0^0.

Así, para $q \neq 0$ reducido, el dominio real natural que hemos fijado es

$$D_q = \begin{cases} \mathbb{R}, & n \text{ impar y } m > 0, \\ \mathbb{R} \setminus \{0\}, & n \text{ impar y } m < 0, \\ [0, \infty), & n \text{ par y } m > 0, \\ (0, \infty), & n \text{ par y } m < 0. \end{cases}$$

Si $n = 1$, la definición coincide con las potencias enteras ya construidas. Si n es par, una base negativa está excluida: la potencia racional **no es una instrucción para fabricar raíces reales inexistentes**. El convenio $x^0 = 1$ corresponde a una función constante, y se trata por separado de los casos $m \neq 0$.

Una representación reducida no se puede sustituir por una operación ilegítima. Para $x > 0$, incluso si inicialmente escribimos $q = m/n = p/s$ con $n, s \geq 1$, el valor obtenido por raíces y potencias positivas no depende de la representación. En efecto, sean $a = r_n(x) > 0$ y $b = r_s(x) > 0$. Usando únicamente las leyes de exponentes **enteros** para bases positivas,

$$(a^m)^{ns} = (a^n)^{ms} = x^{ms}, \qquad (b^p)^{ns} = (b^s)^{pn} = x^{pn}.$$

Como $ms = pn$ y la potencia de índice ns es inyectiva en $(0, \infty)$ por §8.4, resulta $a^m = b^p$. Esta comprobación no requiere definir potencias reales arbitrarias. El razonamiento **no** autoriza repetir las mismas operaciones cuando una base negativa y un denominador par hacen inexistente alguna raíz intermedia. Nuestra definición real sobre bases negativas utiliza desde el comienzo la fracción reducida y solo se permite cuando su denominador es impar. Así,

$$(-8)^{1/3} = -2, \qquad (-8)^{2/3} = (-2)^2 = 4.$$

La escritura no reducida $(-8)^{2/6}$, si se interpreta mediante nuestra convención, debe reducirse primero a $(-8)^{1/3} = -2$. Por el contrario, calcular $\sqrt[6]{(-8)^2} = 2$ es una **operación diferente**; sustituir una por otra produciría una igualdad falsa. Antes de aplicar una regla de exponentes, hay que comprobar que ambos lados designan funciones reales en el conjunto considerado.

Derivada: primero en una base estrictamente positiva. La derivada de una potencia racional se deduce de las raíces construidas y de la regla de la cadena, no de una tabla que todavía no hemos justificado.

Proposición 8.2. **Potencias racionales: derivación en su dominio no nulo.** Sea $q = m/n \in \mathbb{Q}$ su representación reducida, con $n \geq 1$.

1. Si $x > 0$, la función $x \mapsto x^q$ es diferenciable y
$$\frac{d}{dx} x^q = q\, x^{q-1}.$$
2. Si n es impar, la misma conclusión vale para todo $x < 0$, con ambas potencias racionales entendidas mediante sus exponentes reducidos y sus raíces reales. Para $q = 0$, la

función constante $x^0 = 1$ tiene derivada 0 en toda la recta. El enunciado no incluye automáticamente la derivabilidad en $x = 0$ cuando $q \neq 0$.

Demostración. El caso $q = 0$ está resuelto por la derivada de la función constante. Supongamos $m \neq 0$. En un punto $x > 0$, pongamos $u = r_n(x) > 0$. Por Proposición 8.1, r_n es diferenciable en x y $r_n'(x) = 1/(nu^{n-1})$. La potencia entera $v \mapsto v^m$ es diferenciable en u: para $m > 0$ por la regla de potencias naturales y para $m < 0$ por la regla de potencias enteras negativas del capítulo 7. La regla de la cadena da

$$\begin{aligned}\frac{d}{dx}[r_n(x)]^m &= mu^{m-1}\frac{1}{nu^{n-1}}\\ &= \frac{m}{n}\,u^{m-n}.\end{aligned}$$

El exponente $(m-n)/n$ también está reducido, porque $\gcd(|m-n|, n) = \gcd(|m|, n) = 1$. Por la definición anterior, $u^{m-n} = x^{(m-n)/n} = x^{q-1}$; se obtiene así la fórmula anunciada. Si n es impar y $x < 0$, elegimos $u = \widetilde{r}_n(x) < 0$. En particular, $u \neq 0$, por lo que la potencia entera u^m sigue siendo diferenciable y su derivada es mu^{m-1}. La extensión impar posee, por Proposición 8.1, derivada $1/(nu^{n-1})$ en x, ya que $x \neq 0$. Aplicando la misma cadena se obtiene $(m/n)u^{m-n}$. La fracción de $q-1$ continúa reducida y su denominador sigue siendo impar, de modo que $u^{m-n} = x^{q-1}$ también tiene significado real. En ningún paso se ha exigido que $u > 0$ para la regla de potencias **enteras**; solo se ha exigido $u \neq 0$ cuando el exponente es negativo. □

El caso $n = 1$ recupera las derivadas enteras previas. Para $q \neq 0$, la fórmula se establece en los puntos **no nulos** del dominio correspondiente. Cuando la base es negativa con denominador impar, la potencia x^{q-1} del resultado está igualmente definida, pero su valor no tiene por qué ser positivo: los signos han de obtenerse de la raíz real, no de una interpretación mediante raíces pares.

La frontera del origen se decide por cocientes. Que una potencia esté definida en cero no significa que la fórmula anterior sea válida allí. Si $m > 0$, la función $x^{m/n}$ vale 0 en ese punto. Para incrementos $h > 0$, escribamos $u = r_n(h) > 0$. El cociente es exactamente $\frac{h^{m/n}-0}{h} = \frac{u^m}{u^n} = u^{m-n}$.

Si $m > n$, ese cociente tiende a 0; si $m = n$, vale 1; y si $0 < m < n$, crece sin cota y no define una derivada real finita. Cuando n es par este es todo el estudio relativo admisible en cero. Si n es impar, h también puede aproximarse por la izquierda: escribiendo $h = u^n$ con $u \to 0$ de ambos signos, el mismo cociente u^{m-n} demuestra las mismas tres conclusiones respecto de $\mathbb{R}$. El caso $m = n$ en una fracción reducida solo corresponde a $q = 1$. Para $m < 0$, cero no pertenece al dominio, de modo que allí no hay derivada de la función definida. Para $m = 0$, la derivada en cero es 0, por constancia.

El exponente por sí solo no establece el dominio
Las expresiones $x^{2/3}$ y $x^{1/2}$ están definidas en 0, pero ninguna posee allí derivada finita; $x^{3/2}$ sí posee derivada **relativa derecha** igual a 0 en ese extremo; $x^{-1/3}$ ni siquiera está definida en cero. Es necesario distinguir punto excluido, extremo incluido y punto interior antes de calcular.

Los cocientes muestran ya que una fórmula válida para $x \neq 0$ no determina la derivada en cero. El valor absoluto y las funciones por tramos exigirán examinar esos puntos directamente desde la definición.

Componer exige transportar el dominio. Sea $g : A \to \mathbb{R}$ una función diferenciable en los puntos que nos interesan, y sea $q \in \mathbb{Q}$. La expresión $[g(x)]^q$ define una función **solo** en

$$B = \{x \in A : g(x) \in D_q\} = g^{-1}(D_q).$$

La continuidad y la derivabilidad relativas de g respecto de A no agregan puntos a B. Si a es un punto interior de A y $g(a) \neq 0$ pertenece al dominio real de la potencia racional, la regla de la cadena, aplicada donde las composiciones están definidas, proporciona

$$\frac{d}{dx}[g(x)]^q\Big|_{x=a} = q\,[g(a)]^{q-1} g'(a).$$

Para evitar introducir hipótesis topológicas encubiertas, esta fórmula se afirma aquí con a interior de A: la continuidad de g preserva localmente la condición $g(x) > 0$, si la necesitamos, o $g(x) < 0$ cuando el denominador reducido es impar. Si a es un extremo o un punto de acumulación de un dominio restringido, habrá que formular la versión relativa con su dominio efectivo y verificar separadamente sus incrementos; esa tarea pertenece también a §8.6. Si g es un cociente P/Q, se conserva la exclusión $Q(x) = 0$ aunque una simplificación algebraica produzca otra expresión sin denominador.

Ejemplo 8.6. **Una composición global, una potencia con signo y un punto prohibido.** Comencemos con

$$F(x) = (x^2 + 1)^{3/2}, \qquad x \in \mathbb{R}.$$

La función interior siempre es positiva. Como $3/2$ está reducido, Proposición 8.2 y la regla de la cadena permiten derivar en toda la recta:

$$F'(x) = \frac{3}{2}(x^2 + 1)^{1/2} \cdot 2x = 3x\sqrt{x^2 + 1}.$$

Para $G(x) = (2x - 1)^{2/3}$, el denominador reducido es impar y el exponente es positivo: G existe en toda $\mathbb{R}$, incluso cuando $2x - 1 < 0$. Si $x \neq 1/2$, el argumento no se anula y

$$G'(x) = \frac{2}{3}(2x - 1)^{-1/3} \cdot 2 = \frac{4}{3\sqrt[3]{2x - 1}}.$$

En $x = 1/2$ la función está definida, pero el cociente incremental es, para $h \neq 0$, $\frac{G(1/2+h)-G(1/2)}{h} = \frac{(2h)^{2/3}}{h}$. Su valor absoluto crece sin cota cuando $h \to 0$, pues $(2h)^{2/3} = [\sqrt[3]{2h}]^2 = |2h|^{2/3}$; por tanto G no tiene derivada real finita en ese punto.

Finalmente, definamos de manera literal

$$H(x) = \left(\frac{x^2 - 1}{x - 1}\right)^{-1/2}.$$

El cociente interior exige $x \neq 1$ y coincide con $x + 1$ solamente allí. El exponente $-1/2$ exige además una base **estrictamente positiva**. El dominio real de H es, por tanto, $(-1, \infty) \setminus \{1\}$; en él, $H'(x) = -\frac{1}{2(x+1)^{3/2}}$. Aunque esta expresión derivada tenga sentido al sustituir $x = 1$, H **sigue sin estar definida allí**. Podemos construir otra función prolongando la expresión simplificada, pero no alterar retroactivamente el dominio de H. El valor $x = -1$ también queda excluido: el exponente negativo requiere dividir por una raíz no nula.

Control de lectura: antes de derivar
Para cada una de las expresiones $(-8)^{2/3}$, $(-8)^{2/6}$, $(-8)^{1/6}$, $x^{3/2}$ en $x=0$ y $((x^2-1)/(x-1))^{-1/2}$ en $x=1$, responde por este orden: ¿cuál es el exponente reducido?, ¿existe la raíz real pertinente?, ¿está permitido el recíproco?, ¿pertenece el punto al dominio de la función original?, ¿se puede aplicar la derivada racional o hay que volver al cociente incremental? Explica asimismo por qué $\sqrt[6]{(-8)^2}$ no puede sustituir a $(-8)^{2/6}$ bajo la convención adoptada.

Hemos ampliado el repertorio sin ampliar arbitrariamente los dominios: las raíces existen por §8.4; las potencias racionales se definen en términos de ellas y de potencias enteras; sus derivadas proceden de Proposición 8.2 y de la cadena. La §8.6 estudiará con mayor detalle los puntos donde fallan las hipótesis de esas reglas, las funciones con valor absoluto y los empalmes por tramos.

8.6. Puntos excepcionales y funciones por tramos

La derivada de una expresión no resuelve sus excepciones. La fórmula obtenida en §8.5 para una potencia racional compuesta requiere que el argumento no se anule en el punto de aplicación. Al encontrar una raíz en cero, un cambio de signo o el punto donde se unen dos tramos, debemos recuperar la definición de derivada. No basta con que cada fórmula sea diferenciable **fuera** del punto excepcional: la derivabilidad se decide comparando los valores de la función original en ese punto y en los puntos que se le aproximan **dentro de su dominio**.

En $f(x)=|x|$, los tramos $-x$ y x tienen derivadas -1 y 1, pero la derivada en el punto de unión exige calcular allí el cociente incremental respecto del valor central. Para $|x^3|$ también cambia la expresión en cero y, sin embargo, los cocientes sí tienen un límite común. La existencia de dos reglas laterales obliga a comprobar ese límite, sin anticipar su resultado.

Empalmes: primero el valor, después las pendientes. Supongamos que a es un punto interior de un intervalo I y que una función $f: I \to \mathbb{R}$ está dada a la izquierda y a la derecha de a por dos reglas. La elección de $f(a)=c$ forma parte de la definición. La continuidad en a exige que los límites de ambos tramos sean precisamente c; la igualdad de esos dos límites entre sí no basta si el valor asignado en a es otro. Cuando se ha comprobado lo anterior, la derivada exige todavía algo distinto: las expresiones

$$\frac{f(a+h)-c}{h}, \qquad h<0 \quad \text{y} \quad h>0,$$

han de converger al **mismo número real finito**. Es exactamente el criterio de derivadas laterales demostrado en Proposición 6.3, recuperado aquí para analizar empalmes. Si una de las dos derivadas laterales no existe como número real, o si existen pero difieren, no hay derivada bilateral. Si coinciden, su valor común es la derivada. La continuidad, por sí sola, tampoco garantiza esa igualdad.

Podemos comprobarlo directamente: todo incremento no nulo suficientemente pequeño queda a uno de los dos lados de a. Si los dos cocientes laterales convergen a L, dado $\varepsilon>0$ cada lado proporciona una cota δ_- o δ_+ para que su cociente diste de L menos de ε. Con $\delta=\min\{\delta_-,\delta_+\}$ la estimación vale para ambos signos de h y, por tanto, para el límite bilateral. Recíprocamente, un límite bilateral restringido a cualquiera de los dos lados conserva su valor. No se ha usado ningún teorema del valor medio.

Dos pruebas diferentes en un punto de unión
Continuidad: $\lim_{x\to a^-} f(x) = f(a) = \lim_{x\to a^+} f(x)$. **Derivabilidad:** además, deben coincidir y ser finitos los límites de $[f(a+h)-f(a)]/h$ para $h \to 0^-$ y $h \to 0^+$. Si uno de los lados no pertenece al dominio, ninguna de estas condiciones debe exigirlo: se trabaja con la continuidad o derivada **relativa al lado disponible**.

La última precisión importa. La raíz cuadrada $r_2 : [0,\infty) \to \mathbb{R}$ solo admite incrementos $h > 0$ en cero. La derivada relativa no exige un límite inexistente desde valores negativos: exige que el cociente $\sqrt{h}/h = 1/\sqrt{h}$ converja a un número real cuando $h \downarrow 0$, cosa que no ocurre. En cambio, $x \mapsto x^2$ restringida a $[0,\infty)$ sí tiene derivada relativa 0 en cero porque $h^2/h = h \to 0$ por el único lado permitido. Una derivada relativa en un extremo no se convierte, por ello, en una derivada bilateral de una extensión no especificada.

El valor absoluto de una función: el cero manda. Sea $u : I \to \mathbb{R}$ diferenciable en un punto interior a de I. Si $u(a) > 0$, su continuidad asegura que $u(x) > 0$ para todos los x de un entorno suficientemente pequeño de a. Allí $|u(x)| = u(x)$; por localidad de la derivada, $D|u|(a) = u'(a)$. Si $u(a) < 0$, análogamente $D|u|(a) = -u'(a)$. En ambos casos la derivada se obtiene **después** de probar que el signo permanece fijo, no por dividir formalmente por $|u(a)|$ sin comprobarlo.

¿Y si $u(a) = 0$? Para todo incremento admisible $h \neq 0$ tenemos la igualdad exacta

$$\frac{|u(a+h)| - |u(a)|}{h} = \frac{|h|}{h}\left|\frac{u(a+h)-u(a)}{h}\right|.$$

Si $u'(a) = 0$, el valor absoluto de este cociente tiende a cero. Por tanto, $|u|$ es diferenciable en a y $(|u|)'(a) = 0$. Si $u'(a) \neq 0$, su valor absoluto tiende a $|u'(a)| > 0$, pero el factor $|h|/h$ vale -1 por la izquierda y 1 por la derecha: los cocientes laterales convergen a $-|u'(a)|$ y $|u'(a)|$, respectivamente. No pueden coincidir. Así hemos probado, **en un punto interior donde $u(a) = 0$ y u es diferenciable**, el criterio

$$\boxed{|u| \text{ es diferenciable en } a \iff u'(a) = 0; \qquad (|u|)'(a) = 0 \text{ en tal caso.}}$$

Esta conclusión no afirma nada sobre $|u|$ cuando u no es diferenciable en el punto, ni reemplaza el análisis relativo en los extremos del dominio. Por ejemplo, para $u(x) = x$ restringida a $[0,\infty)$, $|u(x)| = x$ y su derivada relativa en cero vale 1; el criterio bilateral anterior no es aplicable porque cero no es interior de ese dominio.

La comparación entre $|x|$ y $|x^3|$ se vuelve ahora precisa: en cero las derivadas interiores de $u(x) = x$ y $v(x) = x^3$ son 1 y 0. El primer valor absoluto no es diferenciable; el segundo sí, con derivada nula. En efecto, de manera independiente, $\frac{|h^3|-|0|}{h} = \frac{|h|^3}{h} \longrightarrow 0$, lo que confirma el resultado sin recurrir a una regla que fuera ilegítima en el punto de anulación.

Radicales en cero: la cadena puede fallar o quedar restituida. Recordemos de Proposición 8.1 que r_n no tiene derivada relativa real finita en cero cuando $n \geq 2$. Por ello, no podemos aplicar la regla de la cadena a $r_n \circ u$ en un punto a con $u(a) = 0$ alegando que conocemos $u'(a)$: falta una de las derivadas requeridas por el teorema. **Esto no prueba que la composición carezca de derivada.** Para decidirlo hay que volver a su cociente, teniendo en cuenta el dominio efectivo $\{x : u(x) \in D_{r_n}\}$.

Concretamente, $x \mapsto \sqrt{x^2}$ está definida en toda la recta e iguala $|x|$; su cociente en cero vale $|h|/h$, cuyos límites laterales son -1 y 1. No hay derivada. En cambio, $x \mapsto \sqrt{x^4}$ está igualmente

definida en toda la recta, pero la raíz no negativa de x^4 es x^2: su cociente en cero es $h^2/h = h$, y su derivada es 0. Tampoco sería válido concluir que $\sqrt[3]{x^3}$ no es diferenciable en cero porque la raíz cúbica aislada no lo es allí: esta composición coincide con la identidad x y su derivada es 1.

Estos ejemplos separan dos proposiciones lógicamente distintas: «no se cumplen las hipótesis de una regla de derivación» y «no existe la derivada». La primera solo impide aplicar esa regla. La segunda exige una demostración adicional a partir de la función concreta. Además, la igualdad $\sqrt{x^4} = x^2$ se ha establecido en toda la recta, mientras que simplificar una fracción racional puede cambiar su dominio: tampoco una composición aparentemente reparada autoriza a llenar un punto excluido.

Ejemplo 8.7. **Un radical con esquina, un empalme suave y un hueco que no desaparece.** Analicemos tres funciones, declarando antes sus dominios.

Primera: $F : \mathbb{R} \to \mathbb{R}$ dada por $F(x) = \sqrt{x^2 + x^4}$. El radicando es no negativo para cualquier x. Como la raíz cuadrada es no negativa y $1 + x^2 > 0$,

$$F(x) = |x|\sqrt{1 + x^2} \qquad (x \in \mathbb{R}).$$

En $x \neq 0$ el argumento de la raíz original es positivo, y las reglas ya demostradas dan

$$F'(x) = \frac{2x + 4x^3}{2\sqrt{x^2 + x^4}} = \frac{x + 2x^3}{\sqrt{x^2 + x^4}}.$$

No evaluamos esta expresión en cero. Puesto que $F(0) = 0$, para $h \neq 0$ el cociente exacto es $\frac{F(h)-F(0)}{h} = \frac{|h|}{h}\sqrt{1 + h^2}$. Por la derecha tiende a 1 y por la izquierda a -1. Así, F es continua en cero —producto de dos funciones continuas—, pero **no** diferenciable allí.

Segunda: definimos $G : \mathbb{R} \to \mathbb{R}$ por

$$G(x) = \begin{cases} x^2, & x \leq 0, \\ x^2 + x^3, & x > 0. \end{cases}$$

Ambas expresiones se aproximan a $0 = G(0)$ cuando $x \to 0$ desde sus respectivos lados: el empalme es continuo. Para $h < 0$, su cociente en cero es $G(h)/h = h$; para $h > 0$, es $h + h^2$. Los dos tienden a cero. El criterio lateral prueba $G'(0) = 0$, además de las derivadas ordinarias $2x$ a la izquierda y $2x + 3x^2$ a la derecha de cero. La presencia de dos reglas no crea necesariamente una esquina.

Tercera: la función literal

$$H : \mathbb{R} \setminus \{1\} \longrightarrow \mathbb{R}, \qquad H(x) = \frac{x^2 - 1}{x - 1},$$

coincide con $x + 1$ **solo en su dominio**. No tiene valor, ni continuidad puntual, ni derivada en $x = 1$, porque ese punto no pertenece al dominio. Si definimos otra función $\widetilde{H} : \mathbb{R} \to \mathbb{R}$ con $\widetilde{H}(x) = x + 1$, entonces $\widetilde{H}(1) = 2$ y $\widetilde{H}'(1) = 1$. Es una extensión diferenciable de H, no una propiedad que el cociente original poseyera ya en su punto excluido.

Método de inspección en cuatro preguntas. Para una expresión con raíces, signos o tramos, seguiremos este orden: **(i)** fijar el dominio natural o el dominio explícitamente declarado, sin restaurar puntos cancelados; **(ii)** determinar si el punto pertenece al dominio y si es interior, extremo o punto de acumulación con varios lados disponibles; **(iii)** comprobar la continuidad relativa y calcular los cocientes incrementales de cada lado efectivo; **(iv)** aplicar una regla de

derivación solamente donde sus hipótesis estén probadas. Si hay extensión, se la denominará como una función nueva.

Laboratorio breve: clasifica antes de calcular
Sin consultar el ejemplo, estudia $A(x) = \sqrt{x^4}$ y $B(x) = \sqrt{x^2}$ en cero: demuestra sus identidades con potencias enteras y calcula ambos cocientes incrementales. Después define una función por tramos que valga x^2 para $x < 0$, 7 en $x = 0$ y x^2 para $x > 0$. ¿Coinciden los límites laterales? ¿Es continua? ¿Tiene derivada? Finalmente, explica por qué la función $\sqrt{x}$ definida solo en $[0, \infty)$ no necesita un cociente desde la izquierda en cero, aunque tampoco posea allí derivada relativa finita.

Antes de derivar una inversa local debemos justificar que existe. Una derivada no nula en un único punto controla la comparación con el valor central, pero no compara automáticamente las imágenes de cualquier par de puntos próximos.

8.7. Una inversa no se deduce de una derivada no nula aislada

Lo que afirma una derivada y lo que exige una inversa. En §8.3 demostramos Teorema 8.1: si una función continua e inyectiva está definida en un intervalo, su inversa existe y es continua sobre la imagen; si además la derivada de la función original en el punto considerado es finita y no nula, entonces podemos derivar la inversa. El orden lógico era importante: **la inyectividad se estableció antes de utilizar la derivada**. ¿Podríamos invertirlo y fabricar la inyectividad únicamente a partir de $f'(a) \neq 0$?

Conviene separar tres afirmaciones. Que $f'(a) \neq 0$ significa que converge a un número no nulo el cociente de incrementos **con un extremo fijo en** a. Que f sea inyectiva en un intervalo I significa que $f(u) \neq f(v)$ siempre que $u, v \in I$ y $u \neq v$: aquí **ambos extremos pueden variar**. Finalmente, disponer de una inversa diferenciable en $f(a)$ requiere primero una inversa bien definida y continua, además de las condiciones de Teorema 8.1. Ninguna de estas afirmaciones es un mero cambio de notación de otra. La derivada no nula sí da una información local precisa. Sea a un punto interior del dominio de f, y supongamos que existe $c = f'(a) \neq 0$. Por definición de límite, existe $\delta > 0$ tal que

$$0 < |x - a| < \delta \quad \Longrightarrow \quad \left| \frac{f(x) - f(a)}{x - a} - c \right| < \frac{|c|}{2}.$$

Por la desigualdad triangular inversa, el cociente tiene módulo mayor que $|c|/2$ y, de hecho, el mismo signo que c. En particular,

$$f(x) \neq f(a) \qquad (0 < |x - a| < \delta).$$

Por consiguiente, a es el único antecedente de **su propio valor** dentro de ese entorno. Para $c > 0$, además, los valores inmediatamente a la izquierda de a son menores que $f(a)$ y los de la derecha son mayores; si $c < 0$, se invierten esas desigualdades. Nada de ello compara $f(u)$ y $f(v)$ cuando u y v son dos puntos próximos **distintos de** a. Ahí se encuentra la brecha que debemos examinar.

Un contraejemplo: pliegues que se aproximan al origen. Construiremos una función continua en toda la recta con derivada 1 en cero, pero cuya restricción a ningún intervalo abierto que contenga cero es inyectiva. El mecanismo será elemental: añadiremos a la identidad pequeñas elevaciones triangulares, cada vez más cercanas a cero. Su altura será insignificante respecto de

su distancia al origen, aunque su pendiente descendente sea lo bastante pronunciada para que la función repita valores.

Para cada entero $n \geq 2$, definamos

$$a_n = 2^{-n}, \qquad w_n = \frac{a_n^3}{100}, \qquad h_n = \frac{a_n^2}{100}, \qquad I_n = [a_n, a_n + 2w_n].$$

Los intervalos I_n son disjuntos. En efecto, $a_{n+1} = a_n/2$ y

$$a_{n+1} + 2w_{n+1} = a_{n+1}\left(1 + \frac{2a_{n+1}^2}{100}\right) < 2a_{n+1} = a_n.$$

Como $a_n \to 0$, estos intervalos se acumulan exclusivamente en cero. Sobre cada uno definamos una tienda triangular que sube desde 0 hasta una altura h_n y vuelve a bajar a 0:

$$p_n(x) = \begin{cases} \dfrac{h_n}{w_n}(x - a_n), & a_n \leq x \leq a_n + w_n, \\ \dfrac{h_n}{w_n}(a_n + 2w_n - x), & a_n + w_n \leq x \leq a_n + 2w_n. \end{cases}$$

Las fórmulas coinciden en la cima, pues ambas dan h_n. Ahora definamos una sola función $p : \mathbb{R} \to \mathbb{R}$ poniendo $p(x) = p_n(x)$ si $x \in I_n$ para algún $n \geq 2$, y $p(x) = 0$ fuera de la unión de los intervalos. No hay conflicto de definiciones porque estos son disjuntos; en sus extremos, la tienda vale cero y empalma con la función exterior. Finalmente, sea

$$f : \mathbb{R} \longrightarrow \mathbb{R}, \qquad f(x) = x + p(x).$$

Primera verificación: continuidad. En el interior de cada semibase p es afín y, en las cimas y extremos de las tiendas, los valores laterales coinciden. En cualquier punto distinto de cero hay un entorno que corta, a lo sumo, un número finito de tiendas; de hecho, todas salvo un número finito quedan cerca de cero. Por eso f es continua fuera del origen. En el origen tenemos $p(0) = 0$. Si $x \in I_n$, entonces $0 \leq p(x) \leq h_n$, y fuera de los intervalos $p(x) = 0$. Cuando $x \to 0$ a través de estas tiendas, sus índices necesariamente tienden a infinito y $h_n \to 0$; por tanto, $p(x) \to 0$. Concluimos que f es continua también en cero.

Segunda verificación: la derivada en el origen. Si $x \in I_n$, se cumple $x \geq a_n > 0$ y, por tanto,

$$0 \leq \frac{p(x)}{x} \leq \frac{h_n}{a_n} = \frac{a_n}{100} \longrightarrow 0 \qquad (n \to \infty).$$

Fuera de las tiendas, $p(x)/x = 0$ para $x \neq 0$. Estas dos observaciones, junto con que las tiendas se acumulan solo en cero, dan $\lim_{x\to 0} p(x)/x = 0$. Como $f(0) = 0$,

$$\boxed{f'(0) = \lim_{x\to 0} \frac{f(x) - f(0)}{x} = \lim_{x\to 0}\left(1 + \frac{p(x)}{x}\right) = 1.}$$

Se trata de una derivada bilateral: para $x < 0$ la función coincide exactamente con x, y el mismo límite vale por ambos lados.

Tercera verificación: ninguna restricción alrededor de cero es inyectiva. En la mitad ascendente de I_n, f tiene pendiente algebraica $1 + \frac{h_n}{w_n} = 1 + \frac{1}{a_n} > 0$, mientras que, en la mitad descendente, su pendiente es $1 - \frac{h_n}{w_n} = 1 - \frac{1}{a_n} < 0$.

No utilizamos aquí un teorema diferencial de monotonía: estas pendientes son simplemente los coeficientes de dos **expresiones afines**. Los valores en los extremos y en la cima son

$$\begin{aligned} f(a_n) &= a_n, \\ f(a_n + w_n) &= a_n + w_n + h_n, \\ f(a_n + 2w_n) &= a_n + 2w_n. \end{aligned}$$

Como $h_n > w_n$ (pues $0 < a_n < 1$), el valor en la cima supera al del extremo derecho. Elijamos un número c_n estrictamente comprendido entre estos dos valores. La mitad ascendente recorre desde a_n hasta el valor máximo, y la mitad descendente recorre desde el máximo hasta $a_n + 2w_n$. Por continuidad —o despejando en las expresiones afines— existen dos puntos distintos, uno en cada mitad, con imagen c_n. Dado cualquier $\delta > 0$, podemos tomar n tan grande que $I_n \subset (0, \delta)$. Esos dos puntos pertenecen entonces a $(-\delta, \delta)$ y tienen la misma imagen. **No existe ningún entorno intervalar de cero en el que f sea inyectiva**, a pesar de ser continua en toda $\mathbb{R}$ y satisfacer $f'(0) = 1$.

Lo que demuestra exactamente el contraejemplo
La función construida es continua en toda la recta y diferenciable en 0 con derivada no nula; no afirmamos que sea diferenciable en las cimas y extremos de las tiendas. El ejemplo refuta la inferencia basada en la derivada **en un solo punto**. No contradice un eventual teorema que impusiera hipótesis adicionales en todo un entorno: ese es otro enunciado y pertenece al desarrollo posterior de la teoría diferencial.

El caso opuesto: una derivada nula no elimina una inversa. Tampoco es legítima la afirmación recíproca «si $f'(a) = 0$, entonces f no tiene inversa». Para la función

$$q : \mathbb{R} \longrightarrow \mathbb{R}, \qquad q(x) = x^3,$$

si $u < v$, la identidad

$$v^3 - u^3 = (v - u)(v^2 + uv + u^2) > 0$$

prueba su crecimiento estricto: el primer factor es positivo y el segundo también, porque $v^2 + uv + u^2 = (u + v/2)^2 + 3v^2/4$ solo se anula cuando $u = v = 0$, imposible si $u < v$. La continuidad polinómica, el teorema de los valores intermedios y los valores de q en enteros positivos y negativos muestran que su imagen es toda $\mathbb{R}$. Por tanto, q es biyectiva y su inversa $q^{-1}(y) = \sqrt[3]{y}$ existe y es continua por Teorema 5.7. Pero $q'(0) = 0$. La inversa no tiene derivada real finita en el origen: para $y \neq 0$,

$$\frac{q^{-1}(y) - q^{-1}(0)}{y} = \frac{\sqrt[3]{y}}{y} = \frac{1}{(\sqrt[3]{y})^2} \longrightarrow +\infty \qquad (y \to 0).$$

Aquí la inversa **existe y es continua**, pero la condición $q'(0) \neq 0$ de Teorema 8.1 no se cumple y su conclusión de diferenciabilidad tampoco. Hay una distinción más precisa: **si una función y su inversa son ambas derivables con derivadas finitas en puntos correspondientes, la derivada de la primera no puede anularse**. En efecto, aplicada *después* de haber demostrado ambas derivabilidades, la regla de la cadena a $q^{-1} \circ q = \mathrm{id}$ da $1 = (q^{-1})'(q(a))q'(a)$. Esta observación no prueba la derivabilidad de la inversa: solo comprueba una condición necesaria cuando ya se la conoce.

Ejemplo 8.8. **Tres diagnósticos que no deben confundirse.** Compara las siguientes situaciones antes de intentar utilizar una fórmula de derivación inversa.

(a) Derivada no nula, sin inversa local. La función $f = x + p$ construida arriba cumple $f'(0) = 1$, pero sus tiendas repiten valores en cualquier entorno del origen. No tiene inversa como función sobre la imagen de ninguna restricción a un intervalo abierto que contenga cero.

(b) Inversa existente, derivada original nula. La función $q(x) = x^3$ es una biyección continua de $\mathbb{R}$ sobre $\mathbb{R}$ y $q'(0) = 0$. La inversa existe y es continua, pero no es derivable con derivada finita en cero.

(c) Inversa existente y fórmula justificada. La función $r(x) = x^3 + x$ es continua; para $u < v$,

$$r(v) - r(u) = (v - u)(v^2 + uv + u^2 + 1) > 0.$$

La inyectividad proviene de esta desigualdad **algebraica**. La imagen es toda $\mathbb{R}$ por continuidad y los valores $r(N) > M$ y $r(-N) < -M$ para enteros N suficientemente grandes y cualquier $M > 0$. Su inversa s existe y es continua; dado que $r'(0) = 1$, Teorema 8.1 autoriza $s'(0) = 1/r'(0) = 1$. Advertencia: haber calculado únicamente $r'(0) = 1$ no habría demostrado la inyectividad necesaria.

Leer el teorema en la dirección correcta. Podemos condensar lo aprendido sin atribuir a una condición más de lo que afirma:

$$\boxed{\begin{array}{l} f : I \to \mathbb{R} \text{ continua e inyectiva}, \quad J = f(I), \quad a \in I \text{ de acumulación}, \quad f_I'(a) \neq 0 \\ \qquad\qquad \Longrightarrow \quad (f^{-1})_J'(f(a)) = \dfrac{1}{f_I'(a)}. \end{array}}$$

La primera línea no se obtiene de la tercera hipótesis; debe acreditarse por su cuenta. La segunda condición sobre la derivada es **suficiente para derivar una inversa ya existente**, no necesaria para que la inversa exista. Si solo conocemos $f'(a) \neq 0$, podemos separar $f(a)$ de los valores vecinos, pero no excluir coincidencias $f(u) = f(v)$ con $u \neq v$ cerca de a.

Laboratorio de transferencia: prueba o contraejemplo
Explica qué cuantificadores cambian entre «$f(x) \neq f(a)$ para $x \neq a$ suficientemente próximo» e «inyectividad en un entorno». Para $q(x) = x^3$, identifica cuál de las tres tareas —existencia de inversa, continuidad de inversa, derivabilidad de inversa en cero— fracasa. Finalmente, sustituye en la construcción triangular h_n por $w_n/2$: ¿cuál sería entonces la pendiente afín en el tramo descendente? ¿Seguiría funcionando *esta* prueba de la falta de inyectividad? Justifica sin invocar el teorema del valor medio.

La necesidad de **seleccionar y demostrar una función** se volverá todavía más visible al abordar ecuaciones polinómicas con dos variables. Una relación puede contener varias gráficas y una misma ecuación puede admitir ramas que se cruzan. En §8.8 distinguiremos el conjunto de soluciones de una rama particular y prepararemos un argumento algebraico que permita estudiar después su derivada sin suponerla de antemano.

8.8. Relaciones polinómicas y ramas

La ecuación describe pares; la rama asigna valores. Hasta ahora hemos invertido funciones *ya definidas*: antes de calcular la derivada de una inversa debíamos justificar la

inyectividad de la función original. Una ecuación en dos letras plantea un problema anterior todavía más elemental. Si escribimos $F(x, y) = 0$,

¿qué significa «derivar y»? La ecuación, por sí sola, no nos ha dado una función de x. Primero fija un conjunto de pares; puede no contener ninguno sobre cierto valor de x, contener varios o permitir distintas elecciones de un valor para cada entrada. Llamaremos aquí **polinomio en dos variables** a una expresión finita de la forma

$$F(x, y) = \sum_{(i,j)\in E} c_{ij} x^i y^j, \qquad c_{ij} \in \mathbb{R},$$

con E un conjunto finito de pares de enteros no negativos. Como es habitual en un polinomio, la potencia de exponente cero representa el factor constante 1; esto no introduce una operación aislada 0^0. El polinomio determina una función $F : \mathbb{R}^2 \to \mathbb{R}$, pero su ecuación de ceros todavía no asigna una salida única a cada primera coordenada.

Definición 8.1. **Conjunto de ceros y rama de una relación polinómica.** Sea F un polinomio real en dos variables. Su **conjunto de ceros** es

$$Z_F = \{(x, y) \in \mathbb{R}^2 : F(x, y) = 0\}.$$

Una **rama de la relación sobre un intervalo** $I \subseteq \mathbb{R}$ es una función especificada $\varphi : I \to \mathbb{R}$ cuya gráfica está contenida en Z_F; es decir,

$$F(x, \varphi(x)) = 0 \qquad \text{para todo } x \in I.$$

Diremos que la rama *pasa por* (a, b) si $a \in I$ y $\varphi(a) = b$. La llamaremos **rama continua en** a únicamente cuando se haya demostrado, además, que $\varphi(x) \to b$ al tender x a a relativamente a I (o la condición habitual de continuidad si a es aislado). Ni la definición de rama ni la inclusión de su gráfica en Z_F presuponen diferenciabilidad.

El objeto Z_F y la función φ tienen tipos diferentes: el primero es un subconjunto de $\mathbb{R}^2$; la segunda tiene un dominio I, un valor preciso para cada $x \in I$ y una gráfica que es *parte* de aquel conjunto. Una ecuación no selecciona su rama; incluso conocer un punto $(a, b) \in Z_F$ puede dejar abiertas varias elecciones de rama por ese mismo punto. Para hablar de $\varphi'(a)$ habrá que acreditar, además, que a es un punto de acumulación del dominio y que converge el cociente incremental correspondiente.

Un cruce: varias ramas, diferentes derivadas. La relación $y^2 = x^2$ se factoriza en $(y - x)(y + x) = 0$. Por la propiedad del producto nulo, su conjunto de ceros es la unión de dos rectas. La factorización describe las posibilidades, pero no impone una sola función sobre toda la recta.

Ejemplo 8.9. **Tres elecciones sobre la misma ecuación.** Consideremos

$$F(x, y) = y^2 - x^2, \qquad Z_F = \{(x, x) : x \in \mathbb{R}\} \cup \{(x, -x) : x \in \mathbb{R}\}.$$

Las aplicaciones

$$\varphi_+(x) = x, \qquad \varphi_-(x) = -x, \qquad \psi(x) = |x|, \qquad x \in \mathbb{R},$$

son tres ramas genuinas: para cada una, $F(x, \varphi(x)) = 0$ en todo su dominio. Son continuas y pasan por $(0, 0)$. No obstante, las dos primeras tienen derivadas distintas,

$$\varphi'_+(0) = 1, \qquad \varphi'_-(0) = -1,$$

mientras que la tercera **no tiene derivada en cero**, pues para $h \neq 0$

$$\frac{\psi(h) - \psi(0)}{h} = \frac{|h|}{h} = \begin{cases} 1, & h > 0, \\ -1, & h < 0. \end{cases}$$

La misma ecuación y el mismo punto permiten, por tanto, ramas continuas con derivadas diferentes e incluso una rama continua no diferenciable. El defecto no está en la ecuación: falta seleccionar la función sobre la que hacemos una afirmación diferencial. La elección ψ combina por tramos las dos rectas y muestra por qué tampoco basta pedir «una rama continua» para obtener automáticamente una derivada.

Es instructivo intentar «derivar $y^2 = x^2$» como si y fuese ya una función diferenciable. Aparecería $2y\,y' = 2x$, pero en $(0,0)$ esta igualdad se reduce a $0 = 0$ y no determina y'. Tampoco demuestra que y' exista para la rama $|x|$. La manipulación solo podría ser una **comprobación posterior**, una vez elegida una rama y establecida su diferenciabilidad por un argumento independiente. Vamos a preparar precisamente ese argumento.

Una identidad algebraica exacta, sin derivadas implícitas. Necesitamos separar el cambio de F en dos contribuciones: una producida por variar la primera coordenada y otra por variar la segunda. No recurriremos a un teorema de varias variables ni a derivadas parciales. La factorización de las diferencias de potencias basta para obtener una identidad **exacta**, válida para todo par (x, y).

Lema 8.1. **Descomposición polinómica por telescopado.** Sean F un polinomio real en dos variables y $(a, b) \in \mathbb{R}^2$. Existen polinomios $A_{a,b}$ y $B_{a,b}$ tales que, para todos $x, y \in \mathbb{R}$,

$$\boxed{F(x,y) - F(a,b) = (x-a)A_{a,b}(x,y) + (y-b)B_{a,b}(x,y).}$$

Aunque la pareja completa de polinomios de una descomposición no tiene por qué ser única, los dos números $A_{a,b}(a,b)$ y $B_{a,b}(a,b)$ **son independientes de la descomposición**. Se calculan algebraicamente a partir de los monomios de F.

Demostración. Para cada monomio $x^i y^j$, sumamos y restamos $a^i y^j$:

$$\begin{aligned} x^i y^j - a^i b^j &= (x^i - a^i) y^j + a^i (y^j - b^j) \\ &= (x-a) \left(\sum_{k=0}^{i-1} x^{i-1-k} a^k \right) y^j + (y-b) a^i \left(\sum_{\ell=0}^{j-1} y^{j-1-\ell} b^\ell \right). \end{aligned}$$

Si $i = 0$ o $j = 0$, la suma correspondiente se interpreta como vacía y vale cero. Multiplicamos cada igualdad por c_{ij} y sumamos el número finito de monomios de F. Obtenemos, por ejemplo,

$$A_{a,b}(x,y) = \sum_{\substack{(i,j) \in E \\ i \geq 1}} c_{ij} \left(\sum_{k=0}^{i-1} x^{i-1-k} a^k \right) y^j,$$

$$B_{a,b}(x,y) = \sum_{\substack{(i,j) \in E \\ j \geq 1}} c_{ij} a^i \left(\sum_{\ell=0}^{j-1} y^{j-1-\ell} b^\ell \right),$$

que son polinomios y satisfacen la identidad requerida. En particular, al evaluarlos en (a,b), cada suma interior consta de tantas copias del mismo monomio como indica su índice:

$$A_{a,b}(a,b) = \sum_{\substack{(i,j)\in E\\ i\geq 1}} i\, c_{ij} a^{i-1} b^{j},$$
$$B_{a,b}(a,b) = \sum_{\substack{(i,j)\in E\\ j\geq 1}} j\, c_{ij} a^{i} b^{j-1}.$$

Para comprobar que los *valores* no dependen de esta elección, supongamos que otra pareja de polinomios $\widetilde{A}, \widetilde{B}$ cumple la misma identidad. Fijando $y = b$ y tomando $x \neq a$ resulta

$$\widetilde{A}(x,b) = \frac{F(x,b) - F(a,b)}{x-a}.$$

El lado derecho es igual a $A_{a,b}(x,b)$ por nuestra identidad. Como ambos son polinomios continuos en x, al pasar al límite $x \to a$ obtenemos $\widetilde{A}(a,b) = A_{a,b}(a,b)$. Del mismo modo, fijando $x = a$ y haciendo $y \to b$ se obtiene $\widetilde{B}(a,b) = B_{a,b}(a,b)$. Esto prueba la independencia anunciada. De hecho, si H es cualquier polinomio, reemplazar $A_{a,b}$ por $A_{a,b} + (y-b)H$ y $B_{a,b}$ por $B_{a,b} - (x-a)H$ conserva la identidad, pero no cambia sus valores en (a,b). □

Lectura de la demostración. La suma y resta de $a^i y^j$ separa el efecto de modificar x del de modificar y; la factorización de potencias extrae los factores $x-a$ e $y-b$. Así aparece una identidad válida **antes** de decidir si hay una rama, si es continua o si admite derivada. Los números $A_{a,b}(a,b)$ y $B_{a,b}(a,b)$ tienen una expresión formal semejante a la derivación de un polinomio con la otra letra fija, pero en esta sección solo usamos álgebra de polinomios y continuidad de funciones de una variable.

Qué información queda preparada para una rama continua. Consideremos primero la circunferencia ya presentada en §8.1, $F(x,y) = x^2 + y^2 - 1$. En torno a $(a,b) = (0,1)$, su telescopado exacto es

$$F(x,y) - F(0,1) = x(x+0) + (y-1)(y+1),$$

por lo que podemos escoger $A_{0,1}(x,y) = x$ y $B_{0,1}(x,y) = y+1$. De aquí se obtienen los números **verificados**, no supuestos,

$$A_{0,1}(0,1) = 0, \qquad B_{0,1}(0,1) = 2 \neq 0.$$

La rama superior $\varphi(x) = \sqrt{1-x^2}$, definida en $I = (-1,1)$, está construida con una raíz existente; es continua en 0 y satisface $\varphi(0) = 1$ y $F(x,\varphi(x)) = 0$. Su continuidad, junto a la de los polinomios de una variable, permitirá controlar los coeficientes del telescopado cuando $x \to 0$. **Aún no hemos supuesto que φ sea diferenciable**: esa será la conclusión por demostrar en §8.9. En el cruce del ejemplo, en cambio, $F(x,y) = y^2 - x^2$ admite la descomposición

$$F(x,y) - F(a,b) = -(x-a)(x+a) + (y-b)(y+b),$$

así que $A_{0,0}(0,0) = B_{0,0}(0,0) = 0$. No hay un coeficiente no nulo ante el incremento de y que permita despejar un cociente incremental. La diferencia entre ambos casos no autoriza a declarar que las ramas del cruce carecen de derivada: dos sí la tienen, y otra no, como acabamos de probar.

Antes de derivar una relación
Para $F(x,y) = y^2 - x^2$, escribe el conjunto Z_F y las gráficas de φ_+ y ψ como conjuntos de pares: identifica qué objetos son iguales y cuáles solo están contenidos unos en otros. Después reproduce sin consultar el lema el telescopado de $F(x,y) = x^2 + xy + y^2 - 1$ alrededor de (a,b): comprueba la identidad sustituyendo y determina $A_{a,b}(a,b)$ y $B_{a,b}(a,b)$. Por último, explica por qué conocer $F(a,b) = 0$ no demuestra la existencia de una rama continua que pase por (a,b).

Hemos establecido el orden lógico que faltaba: **relación → elección y existencia de la rama → continuidad de la rama → identidad de telescopado**. En §8.9 añadiremos una condición algebraica explícita, $B_{a,b}(a,b) \neq 0$, y demostraremos directamente mediante el cociente incremental que esa rama continua es diferenciable en a. No obtendremos de esa condición la existencia general de ramas; tal afirmación pertenece a una teoría posterior.

8.9. Derivación implícita polinómica sin circularidad

La pendiente debe ser una conclusión. En §8.8 construimos el conjunto de ceros de un polinomio, distinguimos sus ramas como funciones y demostramos una identidad de telescopado. Ahora queremos saber si una rama continua admite derivada. Para la circunferencia, la rama superior ya está definida por $\varphi(x) = \sqrt{1-x^2}$ cerca de 0, pero no vamos a introducir su derivada en la hipótesis ni a escribir de entrada $2x + 2\varphi(x)\varphi'(x) = 0$: esa igualdad, obtenida por la regla de la cadena, solo sería legítima después de saber que φ' existe.

La pregunta correcta es: **¿cuándo permite la ecuación demostrar la convergencia del cociente incremental de una rama continua?** La respuesta requiere que el coeficiente algebraico situado ante el cambio de la segunda coordenada no se anule en el punto. La continuidad garantizará que podamos dividir por él cerca del punto, y no antes.

Teorema 8.2. **Derivación implícita de una rama polinómica continua.** Sean $F \in \mathbb{R}[X,Y]$, $I \subseteq \mathbb{R}$ un intervalo, $a \in I$ un punto de acumulación de I y $b \in \mathbb{R}$ con $F(a,b) = 0$. Supongamos que **ya se ha especificado** una función $\varphi : I \to \mathbb{R}$ tal que

$$\varphi(a) = b, \qquad F(x, \varphi(x)) = 0 \quad (x \in I),$$

y que φ es continua en a, con respecto al dominio I. Sean $A_{a,b}$ y $B_{a,b}$ dos polinomios que satisfacen la descomposición de Lema 8.1:

$$F(x,y) - F(a,b) = (x-a)A_{a,b}(x,y) + (y-b)B_{a,b}(x,y).$$

Si

$$B_{a,b}(a,b) \neq 0,$$

entonces φ tiene derivada relativa **finita** en a, y

$$\boxed{\varphi'_I(a) = -\frac{A_{a,b}(a,b)}{B_{a,b}(a,b)}.}$$

La fórmula no depende de la descomposición escogida. El teorema no afirma la existencia de la rama a partir de $F(a,b) = 0$.

Demostración. Por Lema 8.1, para cada $x \in I$ podemos sustituir $y = \varphi(x)$ en la identidad exacta. Como $F(x,\varphi(x)) = F(a,b) = 0$ y $\varphi(a) = b$, obtenemos

$$0 = (x-a)A_{a,b}(x,\varphi(x)) + (\varphi(x)-b)B_{a,b}(x,\varphi(x)).$$

Pongamos $\beta = B_{a,b}(a,b) \neq 0$. Las funciones de una variable $x \mapsto A_{a,b}(x,\varphi(x))$ y $x \mapsto B_{a,b}(x,\varphi(x))$ son sumas finitas de productos de funciones continuas en a; esto se deduce de las reglas de continuidad ya demostradas, sin invocar teoría multivariable. En particular,

$$B_{a,b}(x,\varphi(x)) \longrightarrow \beta \qquad (x \to a,\ x \in I).$$

Por la definición de límite, existe $\delta > 0$ tal que, si $x \in I$ y $|x-a| < \delta$, entonces

$$|B_{a,b}(x,\varphi(x)) - \beta| < \frac{|\beta|}{2}.$$

La desigualdad triangular inversa implica la cota concreta

$$|B_{a,b}(x,\varphi(x))| > \frac{|\beta|}{2} > 0.$$

Ahora, y solamente ahora, podemos despejar para $x \in I$ con $0 < |x-a| < \delta$:

$$\frac{\varphi(x)-\varphi(a)}{x-a} = -\frac{A_{a,b}(x,\varphi(x))}{B_{a,b}(x,\varphi(x))}.$$

La continuidad de φ y de $A_{a,b}$ da $A_{a,b}(x,\varphi(x)) \to A_{a,b}(a,b)$; la del denominador, junto con $\beta \neq 0$, permite aplicar las leyes de límites de cocientes. Así,

$$\lim_{\substack{x \to a \\ x \in I,\ x \neq a}} \frac{\varphi(x)-\varphi(a)}{x-a} = -\frac{A_{a,b}(a,b)}{B_{a,b}(a,b)} \in \mathbb{R}.$$

El punto de acumulación garantiza que el límite corresponde a incrementos admisibles no vacíos; por definición, es la derivada relativa de φ en a. El último enunciado sobre independencia procede también de Lema 8.1: aunque puedan variar los polinomios A, B, sus valores en (a,b) no varían. □

Lectura de la prueba: dónde entra cada hipótesis. No hemos derivado una igualdad que contuviera una función de regularidad desconocida. La identidad del lema ya era válida para todos los pares (x,y); la **existencia de la rama** permitió sustituir $y = \varphi(x)$; su **continuidad** hizo converger los dos coeficientes a sus valores en (a,b); la **no anulación** de $B_{a,b}(a,b)$ aportó un denominador legítimo; y solo el límite del cociente produjo la **derivada**, como conclusión. El argumento funciona también en un extremo de I, pero allí la derivada es relativa a I, no una derivada bilateral sobre un dominio inexistente.

En el volumen I de *A Course in Analysis*, Jacob y Evans presentan, para la derivada de una función **inversa**, un cálculo motivador por regla de la cadena, advierten que requiere cautelas y aplazan su prueba completa [[**5**], cap. 7, teorema 7.5, pp. 111–112]. Ese pasaje sirve aquí como **control metodológico**, no como fuente de un teorema implícito que el libro no haya establecido: una identidad diferencial formal puede orientar una fórmula, pero no demostrar que existe la derivada que aparece en ella. Nuestra deducción usa directamente el cociente incremental.

Circunferencia: una rama construida y una derivada demostrada.

EJEMPLO 8.10. **La rama superior cerca de** $(0,1)$**.** Consideremos

$$F(x,y)=x^2+y^2-1, \qquad I=(-1,1), \qquad \varphi(x)=\sqrt{1-x^2}.$$

El radicando es positivo en I, de modo que la raíz no negativa existe por los resultados previos; $\varphi: I\to\mathbb{R}$ es una función continua, satisface $F(x,\varphi(x))=0$ y $\varphi(0)=1$. En §8.8 se obtuvo la identidad

$$F(x,y)-F(a,b)=(x-a)(x+a)+(y-b)(y+b),$$

por lo que $A_{a,b}(x,y)=x+a$ y $B_{a,b}(x,y)=y+b$. Para cualquier $a\in I$, pongamos $b=\varphi(a)>0$. Entonces $B_{a,b}(a,b)=2b>0$; el teorema garantiza la derivabilidad en a, sin haberla supuesto, y proporciona

$$\boxed{\varphi'(a)=-\frac{2a}{2b}=-\frac{a}{\sqrt{1-a^2}} \qquad (-1<a<1).}$$

En particular, $\varphi'(0)=0$. En $a=1/2$, donde $b=\sqrt{3}/2$, obtenemos $\varphi'(1/2)=-1/\sqrt{3}$.

Control independiente del valor en cero. Racionalizando el cociente, para $0\neq h\in(-1,1)$ se tiene

$$\frac{\sqrt{1-h^2}-1}{h}=\frac{-h}{\sqrt{1-h^2}+1}\longrightarrow 0.$$

Esta verificación no reemplaza el teorema: comprueba un punto particular de la fórmula obtenida para todos los puntos interiores de la rama.

La ecuación de la circunferencia **no** produce, en torno a $(1,0)$, una rama real $y(x)$ definida sobre un intervalo abierto que contenga $x=1$: cualquier intervalo de ese tipo contiene $x>1$, para el que $y^2=1-x^2<0$ no tiene solución real. La rama superior sí está definida sobre $[-1,1]$, pero $B_{1,0}(1,0)=0$, de modo que el teorema no se aplica en el extremo. Su cociente relativo muestra además que allí no tiene derivada finita:

$$\frac{\sqrt{1-x^2}-0}{x-1}=-\sqrt{\frac{1+x}{1-x}}\longrightarrow-\infty \qquad (x\uparrow 1).$$

No aplicación y **ausencia de derivada** son, sin embargo, afirmaciones distintas: aquí la segunda se ha probado aparte.

Condición suficiente, condición necesaria y conversas falsas. Supongamos ahora que una rama φ ya es diferenciable en a y satisface $F(x,\varphi(x))=0$. Las fórmulas polinómicas que obtuvimos en §8.8 para $A_{a,b}(a,b)$ y $B_{a,b}(a,b)$ permiten, mediante las reglas de derivación de los capítulos 6 y 7, diferenciar **a posteriori** la composición polinómica: el resultado necesario es $A_{a,b}(a,b)+B_{a,b}(a,b)\varphi_I'(a)=0$. Se puede obtener sin introducir derivadas parciales: para cada monomio $x^i\varphi(x)^j$ se aplican las reglas del producto y de la cadena ya demostradas, y se suman los términos finitos. Cuando $B_{a,b}(a,b)\neq 0$, la ecuación determina el mismo valor del teorema; cuando $B_{a,b}(a,b)=0$, **no permite despejar** y no decide por sí sola si la derivada existe. Esta identidad es una *consecuencia necesaria bajo diferenciabilidad previa*, no nuestra demostración de tal diferenciabilidad.

La no degeneración tampoco es necesaria para que una rama particular tenga derivada. En la relación $F(x,y) = y^2 - x^3$, la función ya construida $\psi : [0,\infty) \to \mathbb{R}$, $\psi(x) = x^{3/2}$, es una rama continua por $(0,0)$ y satisface $\psi(x)^2 = x^3$. El telescopado da $B_{0,0}(0,0) = 0$, pero, para $x > 0$, $\frac{\psi(x)-\psi(0)}{x-0} = \sqrt{x} \longrightarrow 0$, por lo que $\psi'_{[0,\infty)}(0) = 0$. En el cruce $y^2 = x^2$ de §8.8, el mismo coeficiente se anula y hay ramas derivables de pendientes distintas y una rama continua no derivable. Por tanto, $B(a,b) = 0$ no demuestra ninguna de esas alternativas; exige estudiar la rama elegida por otros medios.

Lo que el teorema demuestra — y lo que no
De **rama existente + continuidad en** a + $B(a,b) \neq 0$ se deducen diferenciabilidad en a y fórmula de la derivada. Ni $F(a,b) = 0$ ni la desigualdad $B(a,b) \neq 0$, tomadas como únicas hipótesis de este resultado, constituyen aquí una demostración de existencia local de rama. Tampoco convertimos la condición suficiente $B(a,b) \neq 0$ en una caracterización de todas las ramas diferenciables. La cuestión general de existencia local y unicidad pertenece al futuro teorema implícito del Tomo III.

Transferencia: reconstruye la prueba, no la fórmula
Para $F(x,y) = x^2 + xy + y^2 - 1$ y un punto (a,b) de su conjunto de ceros, utiliza las fórmulas del lema para obtener $A_{a,b}(a,b) = 2a + b$ y $B_{a,b}(a,b) = a + 2b$. Formula con precisión qué **dato sobre la rama** debe aportarse antes de aplicar el teorema y qué condición algebraica adicional permite concluir $\varphi'_I(a) = -(2a+b)/(a+2b)$. Explica después por qué esta expresión no tiene significado como «la pendiente de la ecuación» si no se ha identificado una función φ. Como segunda comprobación, explica por qué en la cúspide $y^2 = x^3$ no es lícito concluir «no existe derivada» a partir de $B(0,0) = 0$.

La derivación implícita queda así fundada en un límite verificable. En §8.10 organizaremos las mismas distinciones en un laboratorio de curvas: circunferencia, cruce, punto singular y parametrización restringida, sin ampliar las hipótesis del teorema ni importar resultados multivariables.

8.10. Laboratorio P3: leer una curva antes de derivarla

Una ecuación no indica por sí sola qué función estamos estudiando. Una expresión como $x^2 + y^2 = 1$ puede reconocerse geométricamente antes de decidir qué función representa. Pero un conjunto de puntos y una función real no son el mismo objeto. En este laboratorio recorreremos ocho estaciones, usando exclusivamente las construcciones y demostraciones de §§8.1–8.9. La meta no es acumular fórmulas: es **reconstruir las decisiones que hacen legítimo cada cociente incremental**. Las tres relaciones de referencia serán

$$C = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\},$$
$$X = \{(x,y) \in \mathbb{R}^2 : y^2 = x^2\},$$
$$K = \{(x,y) \in \mathbb{R}^2 : y^2 = x^3\}.$$

La letra C nombra un conjunto de puntos de una circunferencia; X, un cruce de rectas; K, una curva con dos ramas que se encuentran en el origen. Sus descripciones geométricas serán consecuencias de cálculos, no premisas que autoricen derivar funciones inexistentes.

Estación 1. Identificar el objeto: relación, gráfica o rama. Pregunta inicial. ¿Podemos escribir sin más «la función y definida por $x^2 + y^2 = 1$»?

No: para $-1 < x < 1$ la ecuación permite simultáneamente

$$y = \sqrt{1 - x^2} \qquad \text{e} \qquad y = -\sqrt{1 - x^2}.$$

Por tanto, C no es la gráfica de una sola función $y(x)$ definida sobre $(-1, 1)$. La proyección de C sobre el eje horizontal es $[-1, 1]$, pero la unicidad vertical falla en cada punto interior. En cambio, para $x = \pm 1$ ambas expresiones producen el mismo valor 0. Recordemos la definición Definición 8.1: solo después de fijar una función $\varphi : I \to \mathbb{R}$ cuyo grafo esté contenido en el conjunto de ceros podemos hablar de *su* derivada.

Error de lectura
Resolver una ecuación para y puede ofrecer varias expresiones. Anotar y' antes de escoger dominio y rama no especifica qué objeto se pretende derivar.

Estación 2. Verificar existencia y continuidad sin recurrir a la derivada. En la circunferencia, los resultados previos sobre raíces y composición permiten construir, con dominio explícito,

$$\varphi_{\pm} : [-1, 1] \longrightarrow \mathbb{R}, \qquad \varphi_{\pm}(x) = \pm\sqrt{1 - x^2}.$$

Ambas funciones están bien definidas, son continuas en su dominio relativo y cumplen la relación. No hemos aplicado todavía ninguna regla de derivación. En la curva K la restricción $x \geq 0$ es obligatoria, pues $x^3 = y^2 \geq 0$. Sobre $[0, \infty)$ existen las funciones continuas

$$\psi_{\pm}(x) = \pm x^{3/2}, \qquad \psi_{\pm}(x)^2 = x^3.$$

La potencia racional ya fue construida en §8.5; no necesitamos postular una solución implícita. Para apreciar por qué la condición $F(a, b) = 0$ **no es, por sí sola, una construcción de rama**, tomemos $F(x, y) = x^2 + y^2$. Su único cero real es $(0, 0)$: los cuadrados son no negativos y su suma solo puede ser cero cuando ambos se anulan. No existe una rama $\varphi : I \to \mathbb{R}$ por ese punto si exigimos que 0 sea punto de acumulación de I y $F(x, \varphi(x)) = 0$ para todo $x \in I$, porque entonces $I \subseteq \{0\}$, contradiciendo la acumulación. Esto no contradice Teorema 8.2: aquí no se ha dado una rama continua a la cual aplicarlo.

Estación 3. Precisar el dominio y los incrementos admisibles. La rama superior de C existe sobre $[-1, 1]$, pero un cociente incremental centrado en $a = 1$ solo puede usar $x \leq 1$. Para $-1 < x < 1$ y $x \neq 1$,

$$\frac{\varphi_+(x) - \varphi_+(1)}{x - 1} = \frac{\sqrt{1 - x^2}}{x - 1} = -\sqrt{\frac{1 + x}{1 - x}}.$$

El cociente tiende a $-\infty$ cuando $x \uparrow 1$; por ello **no hay derivada relativa real finita** en 1, aunque la rama sea continua allí. Tampoco existe una rama $y(x)$ real definida sobre un intervalo *abierto* alrededor de $x = 1$, pues habría valores $x > 1$ y la ecuación exigiría $y^2 = 1 - x^2 < 0$. Para K sucede algo diferente: el dominio $[0, \infty)$ de ψ_+ solo admite incrementos por la derecha en el origen, pero

$$\frac{\psi_+(x) - \psi_+(0)}{x - 0} = \sqrt{x} \longrightarrow 0 \qquad (x \downarrow 0).$$

Existe, pues, la derivada *relativa* $\psi'_{+,[0,\infty)}(0) = 0$. Un extremo de dominio no implica, por sí solo, ni existencia ni inexistencia de derivada: se debe calcular el límite sobre los incrementos disponibles.

Estación 4. Seleccionar la rama en un punto de cruce. El conjunto X se factoriza exactamente: $y^2 - x^2 = (y - x)(y + x) = 0$. Las funciones $\chi_1(x) = x$ y $\chi_2(x) = -x$, ambas con dominio $\mathbb{R}$, satisfacen la ecuación; sus derivadas en 0 son, respectivamente, 1 y -1. Existe también la selección continua

$$\chi_3(x) = |x| = \begin{cases} -x, & x < 0, \\ x, & x \geq 0, \end{cases}$$

que pertenece al mismo conjunto de ceros, pero tiene cocientes laterales -1 y 1 y no es diferenciable en 0. La misma relación contiene, así, dos ramas lineales de pendientes distintas y otra rama continua sin pendiente finita en su unión. Una relación no elige la rama ni determina automáticamente una única tangente a una gráfica funcional.

Comprueba la selección
¿Qué operación cambió al pasar de χ_1 a χ_3? No cambió el polinomio de la ecuación; cambió la **función elegida**. Reconstruye las dos definiciones por tramos antes de comparar sus cocientes en el origen.

Estación 5. Calcular los coeficientes algebraicos correctos. Para aplicar Teorema 8.2, el punto (a, b) debe pertenecer a la relación, la rama ya debe estar construida y hay que disponer de la identidad exacta de Lema 8.1:

$$F(x, y) - F(a, b) = (x - a)A_{a,b}(x, y) + (y - b)B_{a,b}(x, y).$$

La siguiente tabla recoge identidades obtenidas por diferencias de potencias. Sus letras A, B son coeficientes del telescopado, **no derivadas parciales introducidas por anticipado**.

Relación $F(x, y) = 0$	$A_{a,b}(x, y)$	$B_{a,b}(x, y)$	Valores en (a, b)
$x^2 + y^2 - 1 = 0$	$x + a$	$y + b$	$A(a, b) = 2a$, $B(a, b) = 2b$
$y^2 - x^2 = 0$	$-(x + a)$	$y + b$	$A(a, b) = -2a$, $B(a, b) = 2b$
$y^2 - x^3 = 0$	$-(x^2 + xa + a^2)$	$y + b$	$A(a, b) = -3a^2$, $B(a, b) = 2b$

La identidad de la tercera fila, por ejemplo, se verifica sin tomar límites:

$$y^2 - b^2 - (x^3 - a^3) = (y - b)(y + b) - (x - a)(x^2 + xa + a^2).$$

En $(0, 1) \in C$, $B(0, 1) = 2 \neq 0$. En los orígenes de X y K, tanto A como B se anulan: el cociente $-A/B$ daría una expresión $0/0$ **sin valor definido**. Que una fórmula no proporcione respuesta no autoriza a completar la respuesta por conjetura geométrica.

Estación 6. Probar el límite: de los coeficientes a las pendientes. Tomemos en C el punto $(a,b) = (3/5, 4/5)$ y la rama positiva construida. Se verifica $a^2 + b^2 = 1$ y $b > 0$. Como $B(a,b) = 8/5$, el teorema garantiza que la rama es diferenciable en a y da

$$\varphi_+'\left(\frac{3}{5}\right) = -\frac{A(a,b)}{B(a,b)} = -\frac{6/5}{8/5} = -\frac{3}{4}.$$

También podemos reconstruir el cálculo directamente a partir de la identidad de dos cuadrados, sin presuponer una derivada:

$$\frac{\varphi_+(x) - b}{x - a} = -\frac{x + a}{\varphi_+(x) + b} \longrightarrow -\frac{2a}{2b} = -\frac{3}{4}.$$

El denominador converge a $2b = 8/5 \neq 0$, por lo que la división es válida para x suficientemente cercano a a. Esta igualdad traduce la continuidad en una justificación precisa del límite. El contraste con K muestra el alcance de la condición: en $(0,0)$, $B = 0$ impide usar *este* teorema, pero el cociente de la estación 3 prueba independientemente que ψ_+ tiene derivada relativa 0.

Estación 7. Refutar las inferencias falsas con un contraejemplo específico. No basta advertir que una conclusión «puede fallar»; debe identificarse exactamente la hipótesis omitida y construirse el ejemplo pertinente.

Afirmación apresurada	Diagnóstico verificable
«Si $F(a,b) = 0$, existe una rama sobre un intervalo con a de acumulación».	$F = x^2 + y^2$ y $(a,b) = (0,0)$: el único cero es el origen; no hay tal rama.
«Si $B(a,b) = 0$, ninguna rama puede tener derivada».	$y^2 = x^3$ y $\psi_+(x) = x^{3/2}$ sobre $[0,\infty)$: $B(0,0) = 0$ y derivada relativa 0.
«Una rama continua tiene necesariamente derivada en el cruce».	$y^2 = x^2$ y $\chi_3(x) = \lvert x\rvert$: cocientes laterales incompatibles.
«No se aplica el teorema; por tanto, no existe derivada».	Es la misma confusión lógica anterior: la rama ψ_+ no satisface $B \neq 0$ y, sin embargo, tiene derivada.
«La ecuación fija una sola pendiente en el cruce».	$\chi_1'(0) = 1$ y $\chi_2'(0) = -1$ para la misma ecuación.

Una condición suficiente como $B(a,b) \neq 0$ sirve para **probar** una conclusión cuando concurren *todas* las hipótesis de Teorema 8.2. Su incumplimiento no demuestra la negación de la conclusión. En cambio, la ausencia de rama sí impide formular una derivada de esa rama: no es un defecto de cálculo, sino una ausencia del objeto que se pretende estudiar.

Estación 8. Parametrizar con restricción y recuperar una gráfica local. Una parametrización ofrece pares $(x(t), y(t))$, pero no garantiza que su conjunto de puntos sea la gráfica de una función y de x. Consideremos, sobre el intervalo $T = [0,2]$,

$$x(t) = t^3 + t, \qquad y(t) = t^2.$$

Para $s < t$ en T,

$$x(t) - x(s) = (t - s)(t^2 + ts + s^2 + 1) > 0.$$

Por tanto, $x : T \to J = x(T) = [0, 10]$ es continua, estrictamente creciente y biyectiva sobre su imagen. Teorema 5.7 garantiza que su inversa $g : J \to T$ es continua: ahora sí existe la función

$$\Phi : J \longrightarrow \mathbb{R}, \qquad \Phi(u) = y(g(u)) = g(u)^2.$$

En $t_0 = 1$ obtenemos $u_0 = x(1) = 2$ y $\Phi(2) = 1$. Como $x'(1) = 4 \neq 0$ y $y'(1) = 2$, Teorema 8.1 y la regla de la cadena, aplicados **después de construir la inversa**, producen

$$\Phi'(2) = y'(1)g'(2) = \frac{y'(1)}{x'(1)} = \frac{2}{4} = \frac{1}{2}.$$

La notación abreviada $dy/dx = (dy/dt)/(dx/dt)$ solo recoge aquí esa cadena de argumentos; no es una licencia para dividir derivadas cuando x no admite la inversión requerida. Comprobemos el valor sin recurrir a esa abreviatura. Para $t \neq 1$, la factorización exacta da

$$\frac{y(t) - y(1)}{x(t) - x(1)} = \frac{t^2 - 1}{t^3 + t - 2} = \frac{t + 1}{t^2 + t + 2} \longrightarrow \frac{1}{2}.$$

La continuidad de g asegura que, cuando $u = x(t) \to 2$, el parámetro correspondiente satisface $t \to 1$; además, la inyectividad de x permite cancelar $t - 1$ sin identificar puntos diferentes de la curva. Como control algebraico opcional, eliminar el parámetro proporciona $F(u, v) = u^2 - v(v+1)^2 = 0$, porque $u = t(t^2+1)$ y $v = t^2$. En $(u, v) = (2, 1)$ el telescopado tiene valores $A = 2u = 4$ y $B = -(3v^2 + 4v + 1) = -8$; Teorema 8.2 vuelve a dar $-A/B = 1/2$ para la rama continua Φ ya construida. El acuerdo entre los dos métodos verifica un mismo objeto, no la existencia automática de una rama a partir de la ecuación.

Traducción final: qué significa la pendiente y qué no significa. En una gráfica funcional derivable en un punto interior, el número $\varphi'(a)$ determina la pendiente de su recta tangente en el sentido de la aproximación lineal demostrada en el capítulo 6. Así, la rama superior de C en $(3/5, 4/5)$ tiene pendiente $-3/4$; cada rama lineal de X en el origen posee su propia pendiente; $\chi_3 = |x|$ no posee una sola pendiente allí; las ramas de K poseen derivada **relativa** horizontal en el origen, pero solo están definidas para $x \geq 0$. El punto $(1, 0)$ de C es distinto: el cociente de la rama superior diverge y no hay pendiente real finita para su gráfica en ese extremo. No necesitamos introducir todavía la teoría general de tangentes a curvas parametrizadas ni llamar «derivada infinita» a un número real inexistente.

Mapa de decisión para futuras curvas
Ante una relación, identifica primero el conjunto de ceros; demuestra después que la rama o la inversa elegida **existe** con dominio explícito; establece su continuidad cuando el resultado la exige; verifica el punto de acumulación y el tipo de incremento; comprueba $B(a, b) \neq 0$ antes de usar el teorema implícito; si falla esa hipótesis, vuelve al cociente propio de la rama. Solo entonces interpreta geométricamente la derivada obtenida. La falta de alguno de estos pasos nunca se repara escribiendo una fórmula familiar.

Los cuarenta ejercicios y sus soluciones razonadas combinan estas distinciones: construir una función, justificar una rama, probar diferenciabilidad y calcular la derivada bajo las hipótesis correspondientes. La comparación entre métodos deberá referirse siempre al mismo objeto y al mismo dominio.

8.11. Ejercicios y soluciones

Este banco pone a prueba una secuencia de decisiones antes que la memorización de fórmulas: identificar la función y su dominio, justificar la existencia de una inversa o una rama, establecer la continuidad requerida, comprobar el tipo de derivada en el punto y solo entonces calcular. Una relación polinómica no se tratará como función sin escoger una rama; la no anulación de un coeficiente algebraico proporciona un **criterio suficiente condicionado**, no una conclusión sobre cualquier curva no construida.

Los cuarenta ejercicios originales se distribuyen en los niveles $7A{+}7B{+}7C{+}6D{+}5E{+}5F{+}3G = 40$. Los identificadores `CPM-T1-EXR-0316--0355` y sus homólogos `CPM-T1-SOL-0316--0355` se emparejan al final de esta sección. Se permiten únicamente los resultados de los capítulos anteriores y de las §§8.1–8.10. No se invocarán Rolle, el teorema del valor medio, criterios diferenciales posteriores de monotonía, logaritmos o funciones trigonométricas no construidas, ni el teorema implícito multivariable.

Nivel A — Reconocimiento y comprensión. Ejercicio A1. El cuadrado y sus dos restricciones. Considera $q : \mathbb{R} \to \mathbb{R}$, $q(x) = x^2$. Determina su imagen y explica por qué no admite inversa. Construye dos restricciones a intervalos que sí sean biyecciones sobre su imagen, y tipa sus inversas.

Ejercicio A2. ¿En qué punto se evalúa la derivada? Sean I un intervalo, $f : I \to \mathbb{R}$ continua e inyectiva, $J = f(I)$ y $g = f^{-1}$. Si $a \in I$ es punto de acumulación y $f'_I(a) \neq 0$, escribe correctamente $g'_J(b)$ para $b = f(a)$. Explica por qué la cadena de símbolos «$(f^{-1})' = 1/f'$» no es por sí sola una igualdad de funciones bien tipada.

Ejercicio A3. Un dato puntual no produce inversión. Examina la afirmación «$f'(a) \neq 0$ implica que f es inyectiva en algún intervalo alrededor de a». Determina qué información proporciona realmente el cociente incremental y por qué no compara automáticamente dos puntos móviles distintos de a. Señala el resultado anterior que exige inyectividad como hipótesis.

Ejercicio A4. Una igualdad de expresiones y dos funciones. Compara $u(x) = \sqrt{x^2}$ y $v(x) = x$, ambas definidas en $\mathbb{R}$. Decide en qué subconjunto coinciden, deriva cada función allí donde sea posible y analiza el punto 0 por cocientes.

Ejercicio A5. Dos gráficas en una relación. Para $F(x, y) = y^2 - x^2$, construye dos ramas sobre $I = (-1, 1)$ que pasen por $(0, 0)$ y calcula sus derivadas en el origen. ¿Selecciona la ecuación una única pendiente?

Ejercicio A6. Frontera del repertorio. Decide cuáles de $\ln x$, $\arcsin x$ y $x^{3/5}$ han recibido construcción y regla de derivación suficientes en este capítulo. Para la expresión admitida, especifica su dominio real bajo la convención de fracción reducida con denominador impar y señala dónde es válida su fórmula derivada.

Ejercicio A7. El extremo de la raíz cuadrada. Sea $r_2 : [0, \infty) \to \mathbb{R}$, $r_2(t) = \sqrt{t}$. Escribe la definición de $r'_{2,[0,\infty)}(0)$ y decide, mediante el cociente exacto y los incrementos permitidos, si existe como número real finito.

Nivel B — Aplicación directa. Ejercicio B1. Raíz de una función afín. Determina el dominio natural de $f(x) = \sqrt{5x+4}$ y calcula $f'(x)$ en los puntos donde la regla de la raíz y la cadena son aplicables. Examina por separado el extremo del dominio.

Ejercicio B2. Potencia racional de base positiva. Para $f(x) = (x^2+1)^{3/2}$, prueba que está definida y es diferenciable sobre $\mathbb{R}$, y calcula $f'(x)$.

Ejercicio B3. La raíz cúbica compuesta. Estudia $f(x) = \sqrt[3]{2x-1}$ en toda la recta: da la fórmula de su derivada en los puntos admisibles y determina mediante el cociente incremental qué sucede en $x = 1/2$.

Ejercicio B4. Raíz de un cociente. Halla el dominio de $f(x) = \sqrt{(x-2)/(x+3)}$, su interior y la derivada en este último. ¿Por qué no puedes evaluar la función en -3 ni aplicar sin más la misma fórmula en 2?

Ejercicio B5. La inversa negativa del cuadrado. Define $q : (-\infty, 0] \to [0, \infty)$ por $q(u) = u^2$. Demuestra que es biyectiva, identifica su inversa g y calcula $g'(y)$ para $y > 0$; distingue el resultado de la rama positiva.

Ejercicio B6. Una pendiente en la circunferencia de radio cinco. Construye la rama superior de $x^2 + y^2 = 25$ sobre $[-5, 5]$, comprueba que pasa por $(3, 4)$ y calcula su derivada en $x = 3$ mediante los coeficientes algebraicos del telescopado.

Ejercicio B7. Potencia racional sobre bases negativas. Bajo la convención de raíces impares, considera $f(x) = x^{2/3}$ para $x \in \mathbb{R}$. Calcula $f'(x)$ para $x < 0$ y determina mediante límites laterales si existe una derivada real finita en 0.

Nivel C — Combinación estructural. Ejercicio C1. Construir una raíz antes de derivarla. Para $n \geq 1$, define $r_n(t)$ como el único real no negativo cuya potencia n-ésima es $t \geq 0$. Justifica su existencia, unicidad y continuidad, y prueba para $t > 0$ que $r_n'(t) = 1/(n\, r_n(t)^{n-1})$ sin recurrir al teorema del valor medio.

Ejercicio C2. Una regla con tres condiciones de dominio. Sean $p, q \in \mathbb{R}[X]$ y $r = m/n \in \mathbb{Q}$ una fracción reducida con $n \geq 1$. Sea U un intervalo abierto tal que $q(x) \neq 0$ y $p(x)/q(x) > 0$ para todo $x \in U$. Demuestra que $H(x) = [p(x)/q(x)]^{m/n}$ es diferenciable en U y obtiene una fórmula explícita sin usar logaritmos.

Ejercicio C3. El radical oculta un valor absoluto. Determina si $f(x) = \sqrt{x^2 + x^4}$ tiene derivada en 0. Explica por qué una sustitución automática de $x = 0$ en la fórmula general de la raíz no resuelve el problema.

Ejercicio C4. Un módulo que sí admite derivada en el origen. Estudia $f(x) = |x^3|$ para $x \in \mathbb{R}$: prueba directamente su derivabilidad en 0 y calcula la derivada cuando $x \neq 0$.

Ejercicio C5. Pendiente implícita bajo hipótesis explícitas. Sea $F(x, y) = x^3 + xy + y^3 - 1$. Supón **dada** una rama $\varphi : I \to \mathbb{R}$ continua en 0, donde I es un intervalo con 0 como punto de acumulación, $\varphi(0) = 1$ y $F(x, \varphi(x)) = 0$ en I. Construye el telescopado de F, verifica sus coeficientes en $(0, 1)$ y prueba el valor de $\varphi_I'(0)$. ¿Qué no demuestra por sí mismo este razonamiento?

Ejercicio C6. Parámetro y gráfica. Para $t \in [0,2]$, considera $x(t) = t^3 + t$ y $y(t) = t^2$. Justifica la existencia de una función Y de la coordenada x sobre su imagen y calcula $Y'(2)$ por inversa y regla de la cadena. Comprueba el resultado mediante un cociente de diferencias factorizado.

Ejercicio C7. Raíz cúbica de un polinomio. Sea $f(x) = \sqrt[3]{x^3+1}$. Calcula $f'(x)$ donde la regla de inversión lo autorice, estudia expresamente $x = 0$ y decide por cociente incremental si hay derivada real finita en $x = -1$.

Nivel D — Inversión y diagnóstico. Ejercicio D1. La cadena no construye la inversa diferenciable. Alguien escribe $f(g(y)) = y$ y, derivando, obtiene $f'(g(y))g'(y) = 1$. Identifica la hipótesis que da por supuesta; reconstruye el orden no circular de Teorema 8.1.

Ejercicio D2. ¿La derivada no nula es necesaria para la existencia de la inversa? Contrasta las propiedades de $f(x) = x^3$ y $g = f^{-1}$ en el origen. Precisa qué requisito sirve para construir g y qué falla al tratar de derivarla con Teorema 8.1.

Ejercicio D3. Una conclusión que no figura en el teorema. Si $F(a,b) = 0$ y $B_{a,b}(a,b) \neq 0$, ¿puede invocarse **únicamente** Teorema 8.2 para declarar que existe una rama continua alrededor de a? Identifica con precisión la hipótesis que falta en ese teorema y distingue «no demostrado por este resultado» de «matemáticamente falso».

Ejercicio D4. La degeneración no excluye una derivada. Para $F(x,y) = y^2 - x^3$ y $\varphi : [0,\infty) \to \mathbb{R}$, $\varphi(x) = x^{3/2}$, comprueba que es rama continua por el origen, calcula $B_{0,0}(0,0)$ y halla su derivada relativa en 0.

Ejercicio D5. El módulo no se cancela. Compara en $x = 0$ las derivadas de $f(x) = \sqrt{x^2}$ y $g(x) = x$ sobre $\mathbb{R}$. Explica exactamente por qué la coincidencia sobre $[0,\infty)$ no decide la derivada bilateral de f.

Ejercicio D6. Un extremo de pendiente no finita. Sea $\varphi : [-1,1] \to \mathbb{R}$, $\varphi(x) = \sqrt{1-x^2}$. Demuestra mediante incrementos desde la izquierda que $\varphi'_{[-1,1]}(1)$ no existe como número real finito. No utilices la fórmula interior en el extremo.

Nivel E — Construcción de contraejemplos. Ejercicio E1. Inyectividad con derivada nula. Construye una función continua e inyectiva en un intervalo cuya derivada se anule en un punto de acumulación. Explica qué ocurre con la derivada finita de la función inversa en la imagen de ese punto.

Ejercicio E2. Conjunto de ceros no funcional. Construye una relación polinómica cuyo conjunto de ceros completo no sea la gráfica de una función y de x, pero contenga dos ramas diferenciables sobre un mismo intervalo. Verifica ambas afirmaciones.

Ejercicio E3. Una derivada con $B = 0$. Exhibe una relación polinómica y una rama continua y diferenciable en un punto de acumulación en el que el coeficiente B del telescopado se anule. Aclara si la derivada es bilateral o relativa.

Ejercicio E4. Una raíz par no admite entradas negativas reales. Demuestra que ninguna función real que cumpla $r(t)^2 = t$ puede estar definida en todos los números negativos. Identifica exactamente dónde se define la raíz cuadrada real construida en este libro.

Ejercicio E5. Dos fórmulas derivables, un empalme no derivable. Construye una función continua sobre $\mathbb{R}$ que esté dada por polinomios diferenciables a cada lado de 0, pero no sea diferenciable en 0. Justifica continuidad y fracaso de la derivada.

Nivel F — Descubrimiento guiado. Ejercicio F1. Reconstruir la derivada de la inversa. Sean I intervalo, $f : I \to \mathbb{R}$ continua e inyectiva, $J = f(I)$, $g = f^{-1}$ y $a \in I$ punto de acumulación con $f'_I(a) = L \neq 0$. (a) Justifica que g es continua y que $b = f(a)$ es punto de acumulación de J. (b) Relaciona exactamente los cocientes de f y g. (c) Prueba que el denominador permanece no nulo cerca de a y deduce $g'_J(b)$ sin citar la fórmula del teorema inverso.

Ejercicio F2. Raíces impares de números negativos. Sea $n \geq 1$ impar. (a) Prueba algebraicamente que $u \mapsto u^n$ es estrictamente creciente sobre $\mathbb{R}$. (b) Demuestra existencia de una inversa continua sobre toda la recta. (c) Deduce la fórmula de su derivada en entradas $t < 0$ y explica por qué la no anulación necesaria está asegurada.

Ejercicio F3. Coeficientes independientes del telescopado. Supón que un polinomio F admite dos identidades $F(x, y) - F(a, b) = (x - a)A + (y - b)B = (x - a)\widetilde{A} + (y - b)\widetilde{B}$, con los cuatro coeficientes polinómicos. Demuestra que $A(a, b) = \widetilde{A}(a, b)$ y $B(a, b) = \widetilde{B}(a, b)$, aun cuando los polinomios completos puedan diferir.

Ejercicio F4. Reconstrucción polinómica controlada. Considera $F(x, y) = x^2+xy+y^2-7$ y el punto $(1, 2)$. Supón dada una rama $\varphi : I \to \mathbb{R}$ continua en 1, definida en un intervalo donde 1 es punto de acumulación, con $\varphi(1) = 2$ y $F(x, \varphi(x)) = 0$. (a) Construye $A_{1,2}, B_{1,2}$. (b) Prueba la no anulación local adecuada. (c) Obtén $\varphi'_I(1)$ por cocientes.

Ejercicio F5. Una relación sí determina una inversa global. Para cada $x \in \mathbb{R}$ considera $y^3 + y = x$. Construye rigurosamente la única función $\varphi : \mathbb{R} \to \mathbb{R}$ que satisface esta relación; demuestra su continuidad y su derivabilidad en todos los puntos, y expresa $\varphi'(x)$ en función de $\varphi(x)$ sin teorema del valor medio.

Nivel G — Desafíos. Ejercicio G1. Una derivada positiva aislada no garantiza inyectividad local. Para $n \geq 2$ pon $a_n = 2^{-n}$, $w_n = a_n^3/100$ y $h_n = a_n^2/100$. En $[a_n, a_n + 2w_n]$ define una tienda triangular p_n de altura h_n y vértice en $a_n + w_n$, con valor cero en ambos extremos. Fuera de la unión de las bases toma $p = 0$ y define $f(x) = x + p(x)$ para todo $x \in \mathbb{R}$. Prueba que las bases son disjuntas, f es continua, $f'(0) = 1$ y f no es inyectiva en ningún entorno de 0. Da dos puntos distintos de igual imagen en cada base.

Ejercicio G2. Exponentes racionales y derivada en el origen. Sea $r = m/n > 0$ una fracción reducida, $m, n \in \mathbb{N}$. Considera $f_r(x) = x^{m/n}$ con el dominio real de la convención de §8.5: si n es par, $x \geq 0$; si n es impar, $x \in \mathbb{R}$ con raíz impar real. Clasifica exhaustivamente la existencia y el valor de la derivada **relativa al dominio** en 0 para $r > 1$, $r = 1$ y $0 < r < 1$. Cuando falte la derivada, describe el comportamiento de los cocientes permitidos a ambos lados.

Ejercicio G3. Dos tangentes en un cruce deformado. Para $F(x,y) = y^2 - x^2(1+x)$ en $(0,0)$, construye dos ramas continuas sobre algún intervalo abierto que contenga 0, demuestra directamente que son diferenciables allí y que sus pendientes en el origen son opuestas. Calcula $A(0,0)$ y $B(0,0)$ e interpreta por qué el teorema implícito de §8.9 no selecciona una pendiente única.

Soluciones desarrolladas.

Soluciones del nivel A. **Solución A1.** La imagen de q es $[0,\infty)$: todo cuadrado es no negativo y, para cada $t \geq 0$, la raíz cuadrada construida satisface $(\sqrt{t})^2 = t$. La igualdad $q(1) = q(-1)$ muestra que q no es inyectiva sobre $\mathbb{R}$; reducir solamente el codominio a la imagen no repara ese defecto. Las restricciones

$$q_+ : [0,\infty) \to [0,\infty), \quad q_+(x) = x^2, \qquad q_- : (-\infty, 0] \to [0,\infty), \quad q_-(x) = x^2$$

son biyectivas: cada $t \geq 0$ tiene exactamente una preimagen no negativa y una no positiva, respectivamente. Sus inversas son $q_+^{-1}(t) = \sqrt{t}$ y $q_-^{-1}(t) = -\sqrt{t}$, ambas de $[0,\infty)$ al intervalo de partida correspondiente. Las identidades $q_\pm(q_\pm^{-1}(t)) = t$ y $q_\pm^{-1}(q_\pm(x)) = x$ se comprueban respetando los signos de x.

Solución A2. Por Teorema 5.7, $g : J \to I$ existe y es continua. Las hipótesis de Teorema 8.1 permiten afirmar, con derivadas relativas,

$$g'_J(b) = \frac{1}{f'_I(a)}, \qquad b = f(a), \quad a = g(b).$$

Si la condición se verifica en todos los puntos apropiados de J, puede escribirse punto a punto $g'(y) = 1/f'(g(y))$. El miembro izquierdo se evalúa en una **salida** y de f; el derecho, en el antecedente $g(y)$. Escribir $1/f'(y)$ sin cambiar el argumento confunde los dominios y, aun con argumentos correctos, debe conservarse la hipótesis de no anulación.

Solución A3. Si $f'(a) = L \neq 0$, el cociente $(f(x) - f(a))/(x - a)$ tiene el signo de L cerca de a para $x \neq a$. De ello se deduce que $f(x) \neq f(a)$ suficientemente cerca de a, esto es, el valor central queda aislado entre los valores vecinos. Pero la inyectividad exige $f(u) \neq f(v)$ para **cualesquiera** $u \neq v$ del mismo intervalo, no solo cuando uno sea a. El contraejemplo construido en §8.7 (y reconstruido en G1) presenta $f'(0) = 1$ con valores repetidos en cada entorno de 0. Teorema 5.7 supone continuidad e inyectividad en el intervalo; no deriva esta última de una sola pendiente no nula.

Solución A4. La raíz no negativa satisface $\sqrt{x^2} = |x|$. En $[0,\infty)$, $u = v = x$; si $x < 0$, $u(x) = -x \neq x = v(x)$, salvo en el origen ya incluido. Por tanto, $u'(x) = 1$ para $x > 0$ y $u'(x) = -1$ para $x < 0$, mientras que $v'(x) = 1$ en toda la recta. En 0,

$$\frac{u(h) - u(0)}{h} = \frac{|h|}{h} = \begin{cases} 1, & h > 0, \\ -1, & h < 0. \end{cases}$$

Los límites laterales difieren y $u'(0)$ no existe; en cambio $(v(h) - v(0))/h = 1$. Coincidir en medio dominio no implica igualdad de derivadas bilaterales en su frontera.

Solución A5. Las funciones $\varphi_1(x) = x$ y $\varphi_2(x) = -x$ sobre $I = (-1, 1)$ son continuas, pasan por el origen y verifican $\varphi_i(x)^2 - x^2 = 0$ para todo $x \in I$. Sus cocientes en 0 son $\varphi_1(h)/h = 1$ y $\varphi_2(h)/h = -1$: las pendientes son 1 y -1. Además, $|x|$ es una tercera rama continua de la misma relación que no es diferenciable en 0. La ecuación determina el conjunto de ceros, no una única elección de función ni una pendiente singular del conjunto.

Solución A6. Las expresiones $\ln x$ y $\arcsin x$ pertenecen a teorías de funciones aún no construidas en este curso; no se les puede atribuir aquí una fórmula derivada. Sí está construida $x^{3/5} = (\sqrt[5]{x})^3$. Como la raíz quinta es real para bases negativas, su dominio es $\mathbb{R}$. Para $x \neq 0$, la regla de §8.5 da $\frac{d}{dx}x^{3/5} = \frac{3}{5}x^{-2/5} = \frac{3}{5(\sqrt[5]{x})^2}$.

En 0, el cociente $h^{3/5}/h = |h|^{-2/5}$ tiende a $+\infty$ por ambos lados; no existe derivada real finita. Que una expresión pueda escribirse no sustituye la construcción previa de su función.

Solución A7. La derivada relativa exigiría el límite real

$$\lim_{\substack{h\to 0\\ h>0}} \frac{r_2(h) - r_2(0)}{h} = \lim_{h\downarrow 0} \frac{\sqrt{h}}{h} = \lim_{h\downarrow 0} \frac{1}{\sqrt{h}}.$$

Para cualquier $M > 0$, basta tomar $0 < h < 1/M^2$ para obtener $1/\sqrt{h} > M$. El cociente no converge a un número real: no existe derivada relativa finita en el extremo. La fórmula $1/(2\sqrt{t})$ demostrada para $t > 0$ no se prolonga mediante sustitución de $t = 0$.

Soluciones del nivel B. **Solución B1.** La condición para una raíz cuadrada real es $5x + 4 \geq 0$, por lo que el dominio es $[-4/5, \infty)$. Si $x > -4/5$, el argumento es positivo y la cadena, junto con la derivada de la raíz ya demostrada, proporciona $f'(x) = \frac{5}{2\sqrt{5x+4}}$.

En $a = -4/5$, para $h > 0$ el cociente relativo vale $\sqrt{5h}/h = \sqrt{5}/\sqrt{h}$, no finito al tender $h \downarrow 0$. La función está definida y es continua allí, pero no posee derivada relativa real finita.

Solución B2. Para cualquier $x \in \mathbb{R}$ se tiene $x^2 + 1 \geq 1 > 0$. En consecuencia, la potencia racional $u^{3/2}$ está definida y es diferenciable en la base $u = x^2 + 1$ en todos los puntos. La regla de potencia de §8.5 y la cadena de C08 dan

$$f'(x) = \frac{3}{2}(x^2 + 1)^{1/2}(2x) = 3x\sqrt{x^2 + 1}.$$

No hay extremos, ceros de la base ni divisiones excluidas que estudiar por separado.

Solución B3. La raíz cúbica está definida en todo $\mathbb{R}$. Si $x \neq 1/2$, su argumento $2x - 1$ no se anula; la regla de la inversa y la cadena permiten

$$f'(x) = \frac{2}{3\left(\sqrt[3]{2x - 1}\right)^2}.$$

En $a = 1/2$, $f(a) = 0$ y, para $h \neq 0$,

$$\frac{f(a + h) - f(a)}{h} = \frac{\sqrt[3]{2h}}{h} = \frac{\sqrt[3]{2}}{|h|^{2/3}}.$$

El cociente crece sin cota por ambos lados; no existe derivada real finita en ese punto. El denominador nulo de la fórmula exterior advierte una excepción, pero es el cociente el que la resuelve.

Solución B4. El cociente $u(x) = (x-2)/(x+3)$ es no negativo exactamente si $x < -3$ o $x \geq 2$; $x = -3$ se excluye porque anula el denominador. El dominio es $(-\infty, -3) \cup [2, \infty)$ y su interior es $(-\infty, -3) \cup (2, \infty)$. En él $u > 0$, y la regla del cociente da $u'(x) = 5/(x+3)^2$. Por cadena,

$$f'(x) = \frac{5}{2(x+3)^2\sqrt{(x-2)/(x+3)}}.$$

El punto -3 no pertenece a la función. En 2, el cociente derecho es $f(2+h)/h = 1/(\sqrt{h}\sqrt{5+h})$ para $h > 0$ y diverge, por lo que tampoco existe derivada relativa finita allí.

Solución B5. Si $u < v \leq 0$, entonces $u^2 > v^2$, de modo que q es inyectiva en el semieje negativo. Toda salida $y \geq 0$ tiene una única preimagen allí, $u = -\sqrt{y}$; la función es biyectiva y $g(y) = -\sqrt{y}$. La continuidad de la inversa también se obtiene por Teorema 5.7. Para $y > 0$, $g'(y) = -\frac{1}{2\sqrt{y}}$, por la regla de la raíz o Teorema 8.1, pues $q'(-\sqrt{y}) = -2\sqrt{y} \neq 0$. La rama inversa positiva del cuadrado posee el signo opuesto, $+1/(2\sqrt{y})$; ambas requieren indicar qué función se está invirtiendo.

Solución B6. Definimos $\varphi : [-5, 5] \to \mathbb{R}$ por $\varphi(x) = \sqrt{25 - x^2}$, continua por composición de polinomio y raíz sobre valores no negativos. Se cumple $3^2 + \varphi(3)^2 = 9 + 16 = 25$ y $\varphi(3) = 4$. Para $F(x, y) = x^2 + y^2 - 25$, la identidad exacta es

$$F(x, y) - F(a, b) = (x - a)(x + a) + (y - b)(y + b).$$

En $(a, b) = (3, 4)$, $A(a, b) = 6$ y $B(a, b) = 8 \neq 0$. Por Teorema 8.2, la rama continua es diferenciable en 3 y $\varphi'(3) = -6/8 = \boxed{-3/4}$. La existencia y continuidad de la rama precedieron a la aplicación del criterio.

Solución B7. Como el denominador reducido 3 es impar, $x^{2/3} = (\sqrt[3]{x})^2$ existe para todo real. Para $x < 0$, la derivada de la raíz impar en argumento no nulo y la cadena dan $f'(x) = 2\sqrt[3]{x} \cdot \frac{1}{3(\sqrt[3]{x})^2} = \frac{2}{3\sqrt[3]{x}}$.

En 0 debemos usar el cociente: $f(0) = 0$ y $f(h)/h = |h|^{2/3}/h$. Si $h > 0$ vale $h^{-1/3} \to +\infty$; si $h < 0$ vale $-|h|^{-1/3} \to -\infty$. No hay derivada finita ni un límite bilateral en los reales.

Soluciones del nivel C. **Solución C1.** El polinomio $P(u) = u^n$ es continuo en $[0, \infty)$. Para $t \geq 0$ elegimos $M > \max\{1, t\}$: $P(0) = 0 \leq t < M^n$, por lo que el teorema del valor intermedio garantiza $u \in [0, M]$ con $u^n = t$. Si $0 \leq u < v$, entonces

$$v^n - u^n = (v - u) \sum_{k=0}^{n-1} v^{n-1-k} u^k > 0,$$

así que la solución es única. La restricción $P : [0, \infty) \to [0, \infty)$ es una biyección continua en un intervalo, cuya inversa r_n es continua por Teorema 5.7. En $a = r_n(t) > 0$, $P'(a) = na^{n-1} > 0$; Teorema 8.1 da $r_n'(t) = 1/(na^{n-1})$. Para $n = 1$, la misma fórmula es 1. Ni la existencia ni la inyectividad se obtuvieron del signo de una derivada.

Solución C2. Por hipótesis, $u = p/q$ está bien definido en U y toma valores estrictamente positivos. Como p y q son polinomios diferenciables y $q \neq 0$, la regla del cociente de C08 proporciona $u' = (p'q - pq')/q^2$. La función $v \mapsto v^{m/n}$ ha sido construida y derivada para $v > 0$, incluso si $m < 0$. Por composición y cadena,

$$H'(x) = \frac{m}{n}\left(\frac{p(x)}{q(x)}\right)^{m/n-1} \frac{p'(x)q(x) - p(x)q'(x)}{q(x)^2}, \qquad x \in U.$$

La condición de base positiva evita los casos de signo y las posibles singularidades en cero; $q \neq 0$ evita un cociente indefinido. Si $m = 0$, la fracción reducida representa la constante 1 en U, la expresión anterior es 0 y no requiere interpretar 0^0.

Solución C3. La identidad exacta $x^2 + x^4 = x^2(1 + x^2)$ y la elección de la raíz no negativa dan $f(x) = |x|\sqrt{1 + x^2}$.

Por tanto $f(0) = 0$ y, para $h \neq 0$, $(f(h) - f(0))/h = (|h|/h)\sqrt{1 + h^2}$. El límite por la derecha es 1 y por la izquierda es -1. No existe $f'(0)$. La derivada exterior de $\sqrt{u}$ se demostró en $u > 0$, mientras que aquí el argumento interior vale 0 en el punto investigado: aplicarla formalmente sería usar una premisa ausente.

Solución C4. Se verifica $|x^3| = |x|^3$. En el origen,

$$\left|\frac{f(h) - f(0)}{h}\right| = \frac{|h|^3}{|h|} = |h|^2 \longrightarrow 0,$$

así que $f'(0) = 0$. Si $x > 0$, $f(x) = x^3$ y $f'(x) = 3x^2$; si $x < 0$, $f(x) = -x^3$ y $f'(x) = -3x^2$. La presencia de un valor absoluto no implica por sí sola un ángulo: el factor cúbico hace desaparecer la diferencia de pendientes al acercarse a 0.

Solución C5. En (a, b) el telescopado de monomios da

$$\begin{aligned} F(x,y) - F(a,b) &= (x - a)(x^2 + xa + a^2 + y) \\ &\quad + (y - b)\big(a + y^2 + yb + b^2\big), \end{aligned}$$

pues primero variamos x manteniendo y fijo y luego y manteniendo $x = a$. Por tanto $A_{a,b}(a, b) = 3a^2 + b$ y $B_{a,b}(a, b) = a + 3b^2$; en $(0, 1)$ valen 1 y 3, respectivamente. Como la rama ha sido **dada** y es continua en 0, Teorema 8.2 asegura que $\varphi_I'(0) = -1/3$. El enunciado no ha construido una rama para esta ecuación ni demuestra, por sí solo, un teorema general de existencia local; la diferenciabilidad es una conclusión condicionada a la rama continua especificada.

Solución C6. Para $s < t$ en $[0, 2]$,

$$x(t) - x(s) = (t - s)(t^2 + ts + s^2 + 1) > 0.$$

La función $x : [0, 2] \to [0, 10]$ es continua, estrictamente creciente y sobreyectiva por el valor intermedio; por Teorema 5.7 posee inversa continua $g : [0, 10] \to [0, 2]$. Así existe la función $Y(u) = g(u)^2$ y $Y(2) = 1$. En $t_0 = 1$, $x'(1) = 4 \neq 0$ y $y'(1) = 2$; la regla inversa seguida de cadena proporciona $Y'(2) = 2/4 = 1/2$. Independientemente, si $t \neq 1$,

$$\frac{y(t) - y(1)}{x(t) - x(1)} = \frac{t^2 - 1}{t^3 + t - 2} = \frac{t + 1}{t^2 + t + 2} \longrightarrow \frac{2}{4}.$$

La continuidad de g justifica que $u = x(t) \to 2$ corresponda a $t \to 1$; no se ha dividido por una derivada antes de construir la función de u.

Solución C7. La raíz cúbica existe para todos los argumentos reales, así que f tiene dominio $\mathbb{R}$. Para $x \neq -1$ se cumple $x^3 + 1 \neq 0$ y la regla de inversión con la cadena da

$$f'(x) = \frac{3x^2}{3\big(\sqrt[3]{x^3+1}\big)^2} = \frac{x^2}{\big(\sqrt[3]{x^3+1}\big)^2}.$$

En particular, en $x = 0$ la base es 1 y $f'(0) = 0$. En -1, $f(-1) = 0$ y, para $h \neq 0$,

$$\frac{f(-1+h) - f(-1)}{h} = \frac{\sqrt[3]{3h - 3h^2 + h^3}}{h} = \frac{\sqrt[3]{3 - 3h + h^2}}{|h|^{2/3}}.$$

El numerador tiende a $\sqrt[3]{3} > 0$, mientras el denominador tiende a 0 por valores positivos; el cociente diverge a $+\infty$. No existe derivada real finita en -1.

Soluciones del nivel D. **Solución D1.** La cadena $D(f \circ g)(y) = f'(g(y))g'(y)$ solo puede invocarse si **ya sabemos que g es diferenciable**. Pero esa es precisamente la conclusión buscada. El orden correcto es: (i) demostrar que $f : I \to J = f(I)$ es biyectiva a partir de la inyectividad y de su imagen; (ii) obtener la continuidad de $g : J \to I$ por Teorema 5.7; (iii) fijar $b = f(a)$ y escribir, para $y \neq b$, con $x = g(y)$,

$$\frac{g(y) - g(b)}{y - b} = \left(\frac{f(x) - f(a)}{x - a}\right)^{-1}.$$

La continuidad de g lleva $x \to a$ y la hipótesis $f_I'(a) \neq 0$ permite invertir el límite. Solo **después** resulta legítimo usar la cadena para comprobar que $f'(a)g'(b) = 1$.

Solución D2. La función $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^3$, es continua y estrictamente creciente: si $u < v$, entonces $u^3 < v^3$ (por signos y factorización de potencias). Tiene imagen $\mathbb{R}$ y una inversa continua $g(y) = \sqrt[3]{y}$. Sin embargo, $f'(0) = 0$. En el origen de la imagen,

$$\frac{g(h) - g(0)}{h} = \frac{\sqrt[3]{h}}{h} = \frac{1}{|h|^{2/3}} \longrightarrow +\infty,$$

por lo que g no tiene derivada real finita allí. La inyectividad y la correspondencia con la imagen crean la inversa; la condición $f'(a) \neq 0$ es una hipótesis **suficiente para la fórmula diferencial** de Teorema 8.1, no un requisito para que exista una inversa como función.

Solución D3. Teorema 8.2 tiene la forma: «si hay un intervalo I, un punto a de acumulación y una rama $\varphi : I \to \mathbb{R}$ **ya dada y continua** que satisface la relación, entonces, bajo $B(a, b) \neq 0$, esa rama es diferenciable». Por tanto, no puede invocarse el propio teorema para obtener la existencia de φ; sería afirmar una premisa a partir de su consecuencia condicionada. La frase «no está demostrado por este teorema» **no** implica que el enunciado de existencia con $B \neq 0$ sea falso en matemáticas. La cuestión general de existencia local requiere una demostración adicional que no forma parte del resultado aquí utilizado y se reserva para el teorema implícito posterior. Sin condición sobre B, incluso $F(x, y) = x^2 + y^2$ en $(0, 0)$ muestra que una ecuación satisfecha en un punto puede carecer de ramas sobre intervalos no triviales.

Solución D4. Para $x \geq 0$, $\varphi(x) = x^{3/2} \geq 0$, es continua, vale 0 en el origen y satisface $\varphi(x)^2 = x^3$. En el telescopado $F(x,y) - F(0,0) = -x\,x^2 + y\,y$, puede tomarse $B_{0,0}(x,y) = y$, de modo que $B_{0,0}(0,0) = 0$. Pero, para $h > 0$, $\frac{\varphi(h)-\varphi(0)}{h} = \frac{h^{3/2}}{h} = \sqrt{h} \longrightarrow 0$.

La derivada **relativa a** $[0,\infty)$ sí existe y vale 0. No es una derivada bilateral de una función definida a ambos lados del origen; este ejemplo invalida la pretensión de convertir $B \neq 0$ en condición necesaria.

Solución D5. $f(x) = \sqrt{x^2} = |x|$. Para $h > 0$, $f(h)/h = 1$; para $h < 0$, $f(h)/h = -1$. Por tanto no existe $f'(0)$ como límite bilateral, mientras que $g(h)/h = 1$ y $g'(0) = 1$. Las funciones coinciden para $x \geq 0$ y tienen allí la misma derivada relativa desde la derecha; pero los incrementos negativos también pertenecen al dominio de f y no pueden desecharse al estudiar su derivada en $\mathbb{R}$.

Solución D6. En el extremo $a = 1$ los incrementos permitidos son $h < 0$. Pongamos $h = -s$ con $s > 0$ y $s \downarrow 0$. Como $\varphi(1) = 0$,

$$\frac{\varphi(1-s) - \varphi(1)}{-s} = -\frac{\sqrt{1-(1-s)^2}}{s} = -\frac{\sqrt{2-s}}{\sqrt{s}}.$$

Para $0 < s < 1$, $\sqrt{2-s} \geq 1$, así que el cociente es a lo sumo $-1/\sqrt{s}$ y tiende a $-\infty$ en el sentido de barreras. Ningún límite real finito define la derivada relativa. La fórmula interior $-x/\sqrt{1-x^2}$ no incluye el punto $x = 1$ y no puede evaluarse allí.

Soluciones del nivel E. **Solución E1.** Tomemos $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^3$. Es polinómica, luego continua. Si $u < v$, el cubo preserva estrictamente el orden, de modo que f es inyectiva; $f'(0) = 3 \cdot 0^2 = 0$. La inversa existe y es continua por Teorema 5.7, pero $g(y) = \sqrt[3]{y}$ tiene en 0 el cociente $g(h)/h = |h|^{-2/3}$, que no posee límite real finito. Se ha construido una inversa a pesar de una derivada original nula; no se ha afirmado que todas las inversas presenten ese comportamiento.

Solución E2. Sirve $F(x,y) = y^2 - x^2$. Para todo $x \in (0,1)$, los pares (x,x) y $(x,-x)$ pertenecen al conjunto de ceros y tienen la **misma primera coordenada con segundas coordenadas distintas**; por el criterio vertical, el conjunto completo no es gráfica de una función $y(x)$. Sin embargo, $\varphi_+(x) = x$ y $\varphi_-(x) = -x$ son dos ramas sobre $(-1,1)$, ambas polinómicas y derivables en todos sus puntos, con pendientes 1 y -1.

Solución E3. Sea $F(x,y) = y^2 - x^3$ y elijamos $\varphi : [0,\infty) \to \mathbb{R}$, $\varphi(x) = x^{3/2}$. Esta función está construida mediante raíz no negativa, es continua y satisface $F(x,\varphi(x)) = 0$. Para $(a,b) = (0,0)$ el telescopado $F(x,y) = -x\,x^2 + y\,y$ tiene $B(0,0) = 0$. La diferencia $\varphi(h)/h = \sqrt{h}$ para $h > 0$ tiende a 0. Por ello la rama es derivable en 0 **relativamente al semieje**. El ejemplo demuestra que el requisito $B \neq 0$ del teorema de §8.9 no es necesario para la diferenciabilidad de una rama particular.

Solución E4. Para cualquier real v se tiene $v^2 \geq 0$ por la propiedad de orden. Si un número $t < 0$ tuviera raíz cuadrada real $r(t)$, la ecuación $r(t)^2 = t$ implicaría simultáneamente $r(t)^2 \geq 0$ y $r(t)^2 < 0$, contradicción. La raíz cuadrada construida en el capítulo 1 es $r_2 : [0,\infty) \to [0,\infty)$; no es una regla real sobre entradas negativas. El signo $\pm$ al resolver $v^2 = t > 0$ designa dos soluciones, no una raíz no negativa con dos valores.

Solución E5. Definamos $f(x) = -x$ si $x < 0$ y $f(x) = x$ si $x \geq 0$; es $f(x) = |x|$. Ambas expresiones parciales son polinomios y se derivan en los interiores respectivos. Como $|f(x)| = |x| \to 0 = f(0)$, el empalme es continuo. Pero el cociente $(f(h) - f(0))/h$ vale -1 para $h < 0$ y 1 para $h > 0$. No hay límite común y f no es diferenciable en 0. La diferenciabilidad de las piezas sin verificar sus límites laterales en el empalme no basta.

Soluciones del nivel F. **Solución F1.** (a) $f : I \to J = f(I)$ es biyectiva por inyectividad y definición de imagen. Teorema 5.7 proporciona su inversa continua g. Dado $\varepsilon > 0$, la continuidad de f y la acumulación de a permiten encontrar $x \neq a$ con $|x-a|$ pequeño y $0 < |f(x) - f(a)| < \varepsilon$; la segunda desigualdad estricta usa inyectividad. Luego b es punto de acumulación de J.

(b) Para $y \in J \setminus \{b\}$ sea $x = g(y) \neq a$. Se cumple exactamente

$$\frac{g(y) - g(b)}{y - b} = \frac{x - a}{f(x) - f(a)} = \left(\frac{f(x) - f(a)}{x - a} \right)^{-1}.$$

(c) El cociente interior tiende a $L \neq 0$ cuando $x \to a$. Existe $\delta > 0$ tal que $|Q(x) - L| < |L|/2$ para $0 < |x - a| < \delta$ en I; entonces $|Q(x)| > |L|/2$. Por continuidad de g, $y \to b$ implica $x = g(y) \to a$, y

$$\left| \frac{1}{Q(x)} - \frac{1}{L} \right| \leq \frac{2}{|L|^2} |Q(x) - L| \longrightarrow 0.$$

Así $g'_J(b) = 1/L$. La demostración construye la derivada a partir de límites de cocientes y no presupone g' en una cadena.

Solución F2. (a) Sean $u < v$. Si ambos son no negativos, $u^n < v^n$ por la factorización de $v^n - u^n$ con factores no negativos. Si ambos son negativos, $0 < -v < -u$ y, como n es impar, $u^n = -(-u)^n < -(-v)^n = v^n$. Si $u < 0 \leq v$, también $u^n < 0 \leq v^n$. La potencia impar es estrictamente creciente en toda $\mathbb{R}$.

(b) El polinomio $P(u) = u^n$ es continuo. Para cualquier t elegimos $M > |t| + 1$, de modo que $P(-M) < t < P(M)$; el valor intermedio da una solución y la estrictez la hace única. Por Teorema 5.7 la inversa $r_n : \mathbb{R} \to \mathbb{R}$ es continua. Equivale a la raíz impar construida en §8.4.

(c) Para $t < 0$, $a = r_n(t) < 0$, y $P'(a) = na^{n-1} > 0$ porque $n - 1$ es par (para $n = 1$, $P' = 1$). El teorema inverso da

$$r_n'(t) = \frac{1}{n\,[r_n(t)]^{n-1}}, \qquad t < 0.$$

La entrada distinta de cero evita la singularidad del caso $n \geq 3$ en el origen.

Solución F3. Restando las dos descomposiciones y fijando $y = b$, obtenemos $(x-a)\big[A(x,b) - \widetilde{A}(x,b)\big] = 0$.

Para $x \neq a$ se puede cancelar $x - a$, y entonces $A(x,b) = \widetilde{A}(x,b)$. Como ambos miembros son polinomios continuos, al hacer $x \to a$ resulta $A(a,b) = \widetilde{A}(a,b)$. Análogamente, fijando $x = a$ y cancelando $y - b$ para $y \neq b$, se prueba $B(a,b) = \widetilde{B}(a,b)$. La unicidad se refiere a **los valores en el punto**, no a los polinomios enteros: si H es un polinomio, reemplazar A por $A + (y-b)H$ y B por $B - (x-a)H$ conserva la identidad y puede cambiar ambos coeficientes como funciones.

Solución F4. En (a,b), telescopar primero x y después y da la identidad exacta

$$F(x,y)-F(a,b)=(x-a)(x+a+y)+(y-b)(a+y+b).$$

Al fijar $(a,b)=(1,2)$, se obtiene $A_{1,2}(x,y)=x+1+y$ y $B_{1,2}(x,y)=1+y+2=y+3$. Se comprueba $F(1,2)=1+2+4-7=0$, $A_{1,2}(1,2)=4$ y $B_{1,2}(1,2)=5$. La continuidad de φ implica $B_{1,2}(x,\varphi(x))\to 5$; en un entorno relativo de 1 su valor absoluto es, por ejemplo, mayor que $5/2$. De la ecuación, para $x\neq 1$ en ese entorno,

$$\frac{\varphi(x)-2}{x-1}=-\frac{x+1+\varphi(x)}{\varphi(x)+3}\longrightarrow -\frac{4}{5}.$$

Así $\varphi_I'(1)=-4/5$. La prueba utiliza continuidad previa de la rama y no genera por sí misma su existencia.

Solución F5. Definamos $P(u)=u^3+u$. Si $v>u$,

$$P(v)-P(u)=(v-u)(v^2+uv+u^2+1)>0,$$

pues $v^2+uv+u^2\geq 0$. Así P es inyectiva; para cualquier $x\in\mathbb{R}$, eligiendo $M>|x|+1$ se obtiene $P(-M)<x<P(M)$ y el valor intermedio garantiza un único antecedente. Podemos definir $\varphi=P^{-1}:\mathbb{R}\to\mathbb{R}$; es continua por Teorema 5.7. Como $P'(u)=3u^2+1\geq 1$, Teorema 8.1 demuestra que la inversa es diferenciable en cada salida x y

$$\boxed{\varphi'(x)=\frac{1}{3[\varphi(x)]^2+1}}.$$

Se ha construido una rama global mediante inversión justificada; no se ha usado la regla posterior «derivada positiva implica monotonía» ni el teorema del valor medio.

Soluciones del nivel G. **Solución G1.** Denotemos $I_n=[a_n,a_n+2w_n]$. Dado que $a_{n+1}=a_n/2$ y $2w_{n+1}=a_{n+1}^3/50<a_{n+1}$, se tiene $a_{n+1}+2w_{n+1}<2a_{n+1}=a_n$: las bases son disjuntas y se acumulan únicamente en 0. En I_n, escribiendo $s=x-a_n$, definimos explícitamente

$$p_n(x)=\begin{cases}(h_n/w_n)s, & 0\leq s\leq w_n,\\ (h_n/w_n)(2w_n-s), & w_n\leq s\leq 2w_n.\end{cases}$$

La función p coincide con estas tiendas dentro de sus bases y vale 0 fuera. Cada tienda es continua y se anula en los extremos, por lo que los empalmes son continuos. En el único punto de acumulación adicional, 0, si $x\in I_n$ entonces $0\leq p(x)\leq h_n\to 0$; para x fuera de las bases $p(x)=0$. Esto demuestra continuidad en 0 y, por tanto, de $f=x+p$ en toda la recta.

Para $x\neq 0$ perteneciente a I_n,

$$0\leq\left|\frac{f(x)-f(0)}{x}-1\right|=\frac{p(x)}{x}\leq\frac{h_n}{a_n}=\frac{a_n}{100}\longrightarrow 0$$

cuando $x\to 0$; fuera de las bases la diferencia es exactamente cero. Luego $f'(0)=1$.

En cada I_n, $h_n/w_n=1/a_n>1$ y

$$f(a_n)=a_n,\quad f(a_n+w_n)=a_n+w_n+h_n>a_n+2w_n=f(a_n+2w_n).$$

En el tramo ascendente, tomemos $s_n = 2w_n/(1 + h_n/w_n)$. Se cumple $0 < s_n < w_n$ y $f(a_n + s_n) = a_n + 2w_n = f(a_n + 2w_n)$, con dos puntos distintos. Como $a_n + 2w_n \to 0$, cada entorno de 0 contiene una base íntegra y, por ende, dos puntos de igual imagen. La pendiente puntual positiva en 0 no produce inyectividad local.

Solución G2. Como $r = m/n > 0$, en todos los dominios admitidos $f_r(0) = 0$. La derivada relativa, si existiera, sería el límite finito de $f_r(h)/h$ para los incrementos h que pertenecen al dominio.

Incrementos positivos. Para $h > 0$, $\frac{f_r(h)}{h} = h^{r-1}$.

Si $r > 1$, este cociente tiende a 0; si $r = 1$, es constantemente 1; si $0 < r < 1$, crece sin cota hacia $+\infty$.

Incrementos negativos, cuando n es impar. Para $h < 0$, la raíz impar es negativa y $(\sqrt[n]{h})^m = (-1)^m|h|^r$, por lo que $\frac{f_r(h)}{h} = (-1)^{m+1}|h|^{r-1}$.

Si $r > 1$, también tiende a 0 y concuerda con el límite derecho. Si $r = 1$, la fracción reducida exige $m = n = 1$ y el cociente es 1 a ambos lados. Si $0 < r < 1$, no hay límite real finito: por la izquierda crece hacia $+\infty$ si m es impar y decrece hacia $-\infty$ si m es par; por la derecha siempre tiende a $+\infty$.

Clasificación. Para $r > 1$, $f'_{r,D}(0) = 0$; para $r = 1$, $f'_{r,D}(0) = 1$; para $0 < r < 1$, no existe derivada relativa real finita. Cuando n es par solo hay incrementos desde $h > 0$, de modo que los mismos tres casos se entienden unilateralmente. La clasificación procede del dominio y de los cocientes exactos, no de sustituir $x = 0$ en rx^{r-1}.

Solución G3. Para x en el intervalo abierto $I = (-1/2, 1/2)$ se cumple $1 + x > 0$, así que la raíz $s(x) = \sqrt{1 + x}$ está construida y es continua y diferenciable. Definamos

$$\varphi_+(x) = x\sqrt{1 + x}, \qquad \varphi_-(x) = -x\sqrt{1 + x}.$$

Ambas satisfacen $\varphi_\pm(x)^2 = x^2(1 + x)$ y pasan por $(0, 0)$; la relación tiene, por ejemplo para $x > 0$, dos salidas distintas y no selecciona rama. Por la regla del producto y la raíz positiva,

$$\varphi'_\pm(x) = \pm\left(\sqrt{1 + x} + \frac{x}{2\sqrt{1 + x}}\right),$$

por lo que $\varphi'_+(0) = 1$ y $\varphi'_-(0) = -1$. Se puede comprobar directamente en 0 con $\varphi_\pm(h)/h = \pm\sqrt{1 + h} \to \pm 1$, sin presumir diferenciabilidad de una función implícita. En el telescopado respecto de $(0, 0)$ podemos tomar $A_{0,0}(x, y) = -x(1+x)$ y $B_{0,0}(x, y) = y$, pues $F(x, y) = -x\,[x(1+x)]+y\,y$. Ambos valores en $(0, 0)$ son 0. El criterio suficiente de Teorema 8.2 no se aplica ($B = 0$), pero **no contradice** que cada rama construida posea su propia derivada. Las dos rectas de pendiente 1 y -1 son tangentes a dos gráficas distintas, no una pendiente escogida por la ecuación completa.

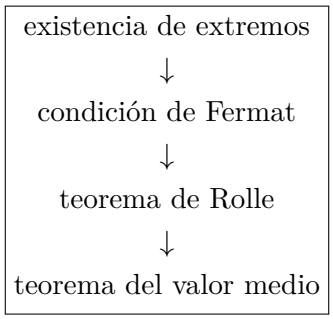
El banco cierra la progresión del capítulo: una función debe existir antes de que podamos derivarla; la diferenciabilidad requiere límites con el dominio correcto; una condición suficiente nunca puede invertirse sin argumento adicional.

Capítulo 9

Teoremas de Rolle y del valor medio

En los capítulos anteriores construimos la derivada como un límite local y demostramos las reglas necesarias para calcularla. Sabemos derivar polinomios y numerosas funciones algebraicas, y hemos aprendido a no aplicar una fórmula antes de comprobar la existencia de la función y la legitimidad de sus operaciones. Pero todavía no sabemos responder una pregunta que parece natural cuando observamos una función sobre **todo un intervalo**: ¿qué información acerca de sus derivadas puede deducirse de los valores que toma en dos puntos distintos?

La dificultad consiste en relacionar dos escalas. El cociente de diferencias compara los extremos de un intervalo; la derivada examina un solo punto. No basta con escribir una fórmula para cada uno: tendremos que demostrar que, bajo hipótesis precisas, existe un punto interior en el que ambas cantidades coinciden. La construcción seguirá una cadena en la que cada resultado autorizará el siguiente:



La existencia de extremos permitirá construir la cadena de argumentos. Seguiremos el uso de la continuidad en el intervalo cerrado y de la diferenciabilidad en su interior para precisar las conclusiones de Fermat, Rolle y el teorema del valor medio.

9.1. De una pendiente local a una afirmación global

Dos medidas de cambio que no debemos confundir. Sea $a < b$ y consideremos una función $f\colon [a,b] \longrightarrow \mathbb{R}$.

Los valores de la función en los extremos del dominio proporcionan dos puntos de la gráfica:

$$P = (a, f(a)), \qquad Q = (b, f(b)).$$

La pendiente de la secante determinada por ellos es el número $m_{a,b} := \frac{f(b)-f(a)}{b-a}$. El denominador es distinto de cero porque $a < b$. Además, esta cantidad solo necesita conocer $f(a)$ y $f(b)$: no requiere que hayamos examinado el comportamiento de la función entre ambos puntos. Resume la variación de salida por unidad de incremento de entrada a lo largo de **todo el tramo**. La derivada, en cambio, se construyó mediante un límite de cocientes de incrementos. Si $x \in (a,b)$ y f es diferenciable en x, podemos escribir

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

entendiendo que $h \neq 0$ y que $x + h \in [a, b]$. Como x es interior, están disponibles incrementos suficientemente pequeños de ambos signos. Esta fórmula solo estudia lo que ocurre **arbitrariamente cerca de** x. La notación $f'(x)$ abrevia aquí la derivada relativa al dominio $[a, b]$ introducida en el capítulo 6. La diferencia entre los datos puede disponerse así:

Cantidad	Qué compara	Qué información necesita
$m_{a,b}$	Los valores en a y b	Dos valores de una función definida en ambos extremos
$f'(x)$	Variaciones cada vez más próximas a x	La existencia del límite diferencial en ese punto interior

No existe una identidad algebraica que obligue por sí sola a que $f'(x)$ sea igual a $m_{a,b}$ para algún x. Lo que estamos buscando es una **afirmación de existencia**: necesitamos condiciones que enlacen el comportamiento local de la función con el intervalo completo.

La pregunta exacta. Una vez que la función y el intervalo están fijados, el problema que guiará este capítulo puede escribirse con un cuantificador:

$$\boxed{\text{¿Existe } c \in (a, b) \text{ tal que } f'(c) = \frac{f(b) - f(a)}{b - a}?}$$

Hay tres aspectos que no debemos perder al leer esta pregunta. Primero, el punto buscado pertenece al **intervalo abierto**: ni a ni b sirven como sustitutos automáticos. Segundo, preguntamos por la existencia de *algún* punto, no por su unicidad. Tercero, aunque el miembro derecho esté siempre definido para $a < b$, la expresión $f'(c)$ solo tiene sentido si la función es diferenciable en el punto elegido. La existencia que queremos justificar no se obtiene de las reglas algebraicas de derivación ya demostradas. Esas reglas calculan derivadas cuando se satisfacen ciertas hipótesis; no garantizan, por sí mismas, que una derivada tome un valor fijado de antemano en algún punto de un intervalo.

Un caso en el que podemos encontrar el punto directamente.

Ejemplo 9.1. **La secante de una parábola encuentra una tangente paralela.** Sean $a < b$ y

$$f\colon [a, b] \longrightarrow \mathbb{R}, \qquad f(x) = x^2.$$

La pendiente de la secante entre los extremos es

$$\begin{aligned} m_{a,b} &= \frac{b^2 - a^2}{b - a} \\ &= \frac{(b - a)(a + b)}{b - a} \\ &= a + b. \end{aligned}$$

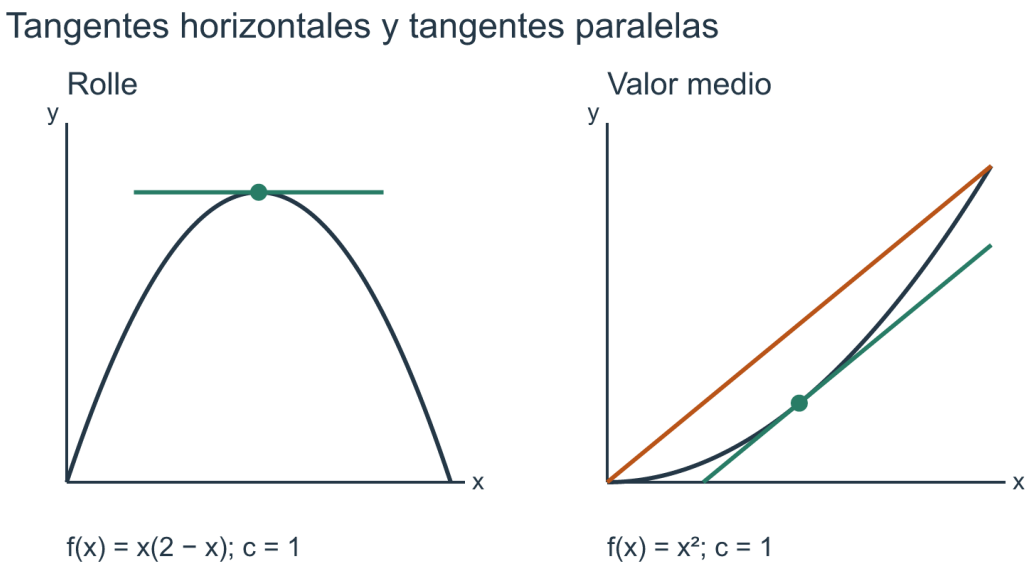


FIGURA 9.1. En $[0,2]$, $x(2-x)$ ilustra Rolle con tangente horizontal en $c=1$; x^2 ilustra el valor medio con tangente paralela a la secante entre los extremos, también en $c=1$.

La cancelación es legítima porque $b-a \neq 0$. Por las reglas de derivación de los capítulos anteriores, para todo $x \in (a,b)$, $f'(x) = 2x$. Por tanto, buscar un punto en el que ambas pendientes coincidan equivale a resolver $2c = a+b$. La única solución de esta ecuación es $c = \frac{a+b}{2}$.

Y, dado que $a < b$, $a < \frac{a+b}{2} < b$.

En este ejemplo sí hemos encontrado un punto **admisible** y comprobado la igualdad requerida:

$$f'\left(\frac{a+b}{2}\right) = a+b = \frac{f(b)-f(a)}{b-a}.$$

En el caso cuadrático, la fórmula explícita de la derivada permite despejar el punto buscado. Para una función sin una expresión manejable necesitamos otro argumento: demostrar que existe algún punto apropiado aunque no podamos calcularlo explícitamente.

Una verificación particular no es un teorema universal
El ejemplo demuestra la existencia del punto para $f(x) = x^2$ y para los intervalos no degenerados considerados. No demuestra que el punto buscado sea siempre el punto medio ni que cualquier función definida en $[a,b]$ satisfaga la misma conclusión. Esas afirmaciones requerirían argumentos independientes; no pueden deducirse de un caso favorable.

Qué tendremos que demostrar antes de responder. El primer obstáculo aparece cuando los valores de la función coinciden en los extremos. En ese caso,

$$f(a) = f(b) \quad \Longrightarrow \quad m_{a,b} = 0,$$

y la pregunta se reduce a saber si debe existir algún punto interior con derivada nula. La igualdad de los valores extremos no basta por sí sola: todavía debemos investigar qué ocurre con la continuidad en el intervalo cerrado y con la diferenciabilidad en su interior. Una función continua en un intervalo cerrado y acotado alcanza máximo y mínimo (Teorema 5.5). Si uno de esos extremos se alcanza en el interior, la diferenciabilidad impone una condición a los cocientes incrementales de ambos lados. Esa condición dará el teorema de Fermat sin recurrir a criterios de monotonía.

Con ese resultado podremos investigar la situación de los extremos iguales y demostrar el teorema de Rolle. Solo después transformaremos el problema de una secante de pendiente arbitraria en otro de pendiente cero mediante una función auxiliar. Esa transformación será la clave del teorema del valor medio.

Antes de continuar
En el ejemplo de la parábola, distingue los datos que proceden únicamente de los extremos de los datos que proceden de la derivada. Identifica después las dos afirmaciones diferentes que hemos probado: que la ecuación de pendientes tiene una solución y que esa solución pertenece al intervalo abierto. Finalmente, explica por qué el cálculo particular no autoriza una conclusión sobre todas las funciones.

La posición del extremo determina qué incrementos están disponibles: en el interior podremos comparar ambos signos, mientras que en el borde solo se dispone de los incrementos admitidos por el intervalo.

9.2. Extremos y puntos interiores

La palabra *extremo* puede referirse a objetos distintos. Los números a y b son los **extremos del intervalo** $[a, b]$; un máximo o un mínimo es un **valor de la función**; y el punto c en el que se obtiene ese valor es el lugar donde el extremo **se alcanza**. Esta distinción será esencial: la derivada se calcula en puntos del dominio, no en valores de la imagen.

Alcanzar un máximo o un mínimo en todo el dominio.

Definición 9.1. **Máximos y mínimos absolutos.** Sean $A \subseteq \mathbb{R}$ un conjunto no vacío, $f\colon A \to \mathbb{R}$ y $c \in A$.

- Diremos que f **alcanza un máximo absoluto** (o *global*) en c si

$$\forall x \in A, \qquad f(x) \leq f(c).$$

- Diremos que f **alcanza un mínimo absoluto** (o *global*) en c si

$$\forall x \in A, \qquad f(c) \leq f(x).$$

En el primer caso, el número $f(c)$ es el valor máximo de f en A; en el segundo, es su valor mínimo. Un mismo valor extremo puede alcanzarse en varios puntos distintos.

El punto donde se alcanza el extremo debe pertenecer al dominio. La identidad en $(0, 1)$ tiene supremo 1 e ínfimo 0, pero ninguno se alcanza: no existe un argumento admisible cuya imagen sea uno de esos valores. Esta diferencia será necesaria al elegir el punto al que aplicamos el resultado diferencial. La definición tampoco exige continuidad ni diferenciabilidad. Estas propiedades serán hipótesis de ciertos **teoremas de existencia**, no ingredientes de la noción de máximo o mínimo.

Comparar solamente con los puntos cercanos. Un extremo absoluto se caracteriza por una desigualdad válida en todo A. Para reconocer un extremo *local*, en cambio, bastará con comparar los valores de la función en una parte suficientemente pequeña del dominio.

Definición 9.2. **Máximos y mínimos locales relativos al dominio.** Sean $A \subseteq \mathbb{R}$ no vacío, $f\colon A \to \mathbb{R}$ y $c \in A$. Diremos que f alcanza un **máximo local en c relativo a A** si existe $\delta > 0$ tal que

$$\forall x \in A: \quad |x - c| < \delta \implies f(x) \leq f(c).$$

Sustituyendo la última desigualdad por $f(c) \leq f(x)$ obtenemos la definición de **mínimo local relativo a A**. El extremo local es **estricto** cuando la desigualdad correspondiente es estricta para todos los puntos del dominio que satisfacen $0 < |x - c| < \delta$. Así, para un máximo local estricto exigimos

$$\forall x \in A: \quad 0 < |x - c| < \delta \implies f(x) < f(c).$$

Para un mínimo local estricto se invierte esa desigualdad.

El conjunto de comparación es siempre $A \cap (c - \delta, c + \delta)$. No se exige que la función esté definida fuera de A. Un punto de máximo absoluto es también un punto de máximo local: como la desigualdad vale en **todo** el dominio, vale en particular en cualquier entorno relativo. Análogamente para los mínimos. La implicación contraria no es válida; veremos un ejemplo antes de terminar la sección.

Cómo leer los cuantificadores
En un máximo absoluto, la desigualdad debe cumplirse para **todo** $x \in A$. En un máximo local basta **encontrar un radio** $\delta > 0$ para el que se cumpla con todos los puntos del dominio que caen dentro de ese radio. El radio puede depender del punto c y de la función: no hay una exigencia uniforme sobre todo A.

Interior y borde: la geometría del dominio cambia la prueba. A partir de ahora fijemos $a < b$ y $A = [a, b]$. Un punto $c \in A$ es **interior al intervalo** cuando $a < c < b$. Entonces podemos escoger, por ejemplo, $\delta_0 := \frac{1}{2}\min\{c - a, b - c\} > 0$, de modo que $(c - \delta_0, c + \delta_0) \subseteq (a, b)$. Existen por tanto incrementos admisibles de los dos signos: si $0 < |h| < \delta_0$, tanto $c + h$ como $c - h$ pertenecen al dominio. Podemos comparar valores situados a la izquierda y a la derecha de c.

Los puntos a y b, en cambio, son los **extremos del intervalo**. Para estudiar el comportamiento local en a solamente disponemos de puntos de $[a, b]$ a su derecha; para hacerlo en b, de puntos a su izquierda. Un máximo o mínimo local relativo puede encontrarse perfectamente en el borde. Pero allí no podremos razonar como si la función tuviese argumentos admisibles a ambos lados del punto.

Ejemplo 9.2. **Un extremo de borde con derivada relativa no nula.** Sea

$$f\colon [0, 1] \to \mathbb{R}, \qquad f(x) = x.$$

Para todo $x \in [0, 1]$ tenemos

$$f(0) = 0 \leq f(x) \leq 1 = f(1).$$

Por tanto, 0 es punto de mínimo absoluto y 1 es punto de máximo absoluto. Ambos son también extremos locales relativos y, de hecho, estrictos: para $x \in (0,1]$ se tiene $f(x) > f(0)$, mientras que para $x \in [0,1)$ se tiene $f(x) < f(1)$. Calculemos ahora la derivada **relativa al dominio** usando exclusivamente incrementos admisibles. Para $c = 0$ o $c = 1$, $\frac{f(c+h)-f(c)}{h} = \frac{(c+h)-c}{h} = 1$ siempre que $h \neq 0$ y $c + h \in [0,1]$. En consecuencia, $f'(0) = f'(1) = 1$ como derivadas relativas: la primera utiliza incrementos positivos y la segunda, negativos. No hay contradicción alguna. La existencia de un extremo en el borde no proporciona una condición bilateral sobre la derivada.

Este ejemplo sirve también para descartar una confusión de vocabulario: que 1 sea el extremo derecho del *intervalo* no significa que la función alcance allí su *mínimo*; la naturaleza del extremo funcional depende de los valores de f, no de la posición horizontal.

Ejemplo 9.3. **Un extremo interior y dos extremos de borde.** Consideremos

$$g\colon [-1,1] \to \mathbb{R}, \qquad g(x) = x^2.$$

Como $0 \leq x^2 \leq 1$ para todo $x \in [-1,1]$,

$$g(0) = 0 = \min g([-1,1]), \qquad g(-1) = g(1) = 1 = \max g([-1,1]).$$

El mínimo absoluto se alcanza en el punto interior 0 y es también un mínimo local estricto, pues $x^2 > 0$ si $x \neq 0$. El máximo absoluto se alcanza en **dos puntos**, ambos situados en el borde. La definición de máximo no afirma que haya un único punto que lo realice. Las reglas de derivación ya demostradas dan, en el interior, $g'(x) = 2x$. Incluso en los bordes, el cálculo relativo por cocientes produce

$$g'(-1) = -2, \qquad g'(1) = 2.$$

Así, los dos puntos de máximo de borde presentan derivadas relativas diferentes entre sí y no nulas. Su condición de máximo no depende de que esas derivadas se anulen.

La misma gráfica permite observar los tres hechos sin mezclarlos: un extremo funcional puede estar en el interior; otro puede alcanzarse en el borde; y un mismo valor máximo puede corresponder a más de un argumento. Ninguno de esos hechos se obtiene por una regla de derivación: se comprueba primero mediante desigualdades.

Un extremo local no tiene por qué ser absoluto. La distinción entre comparar en todo el dominio y comparar solamente cerca de un punto es real, no una diferencia terminológica.

Ejemplo 9.4. **Un máximo local que no es máximo absoluto.** Sea

$$p\colon [-2,2] \to \mathbb{R}, \qquad p(x) = (1 - x^2)^2.$$

Se tiene $p(0) = 1$. Si $0 < |x| < 1$, entonces $0 < 1 - x^2 < 1$ y, por consiguiente,

$$p(x) = (1 - x^2)^2 < 1 = p(0).$$

Por tanto, 0 es un punto interior de **máximo local estricto**, tomando $\delta = 1$ en Definición 9.2. Sin embargo,

$$p(-2) = p(2) = 9 > 1 = p(0).$$

No es un punto de máximo absoluto. De hecho, para todo $x \in [-2,2]$ tenemos $0 \leq x^2 \leq 4$, de donde $-3 \leq 1 - x^2 \leq 1$ y $p(x) \leq 9$: el máximo absoluto es 9 y se alcanza en los dos bordes.

No invertir una implicación
De «máximo absoluto» se sigue «máximo local» porque la desigualdad global sigue siendo verdadera después de restringir los puntos de comparación. El ejemplo anterior demuestra que **máximo local no implica máximo absoluto**. La misma distinción rige para los mínimos.

Lo que aporta el teorema de los valores extremos. El teorema de los valores extremos del capítulo 5 (Teorema 5.5) afirma lo siguiente: si $f\colon [a,b] \to \mathbb{R}$ es continua y $a < b$, existen $x_m, x_M \in [a,b]$ tales que

$$f(x_m) \leq f(x) \leq f(x_M) \qquad \text{para todo } x \in [a,b].$$

Existencia es aquí la palabra central. El resultado garantiza que se alcanzan ambos valores extremos, pero no dice que x_m o x_M sean interiores, ni que sean únicos, ni que f sea diferenciable. Por ejemplo, para una función constante $f(x) = k$ todo punto del intervalo es simultáneamente de máximo y mínimo absolutos: se cumple $f(x) = k$ en todo el dominio. En esa situación tampoco hay extremos locales estrictos cuando $a < b$.

La continuidad es suficiente para la conclusión de existencia del teorema, no una condición necesaria para que una función particular alcance un máximo o mínimo. Y, aunque un extremo absoluto sea siempre local, **todavía no sabemos qué puede imponer la derivabilidad en un punto interior donde se alcanza**. Esa será exactamente la pregunta de la sección siguiente.

Laboratorio de lectura antes de Fermat
Para cada ejemplo, identifica primero el dominio y distingue los extremos del intervalo de los valores extremos de la función. Localiza después cada punto de máximo o mínimo y decide si es interior o de borde. Finalmente, comprueba qué desigualdad demuestra el extremo y si los cocientes de derivación pueden utilizar incrementos de ambos signos. No infieras que una derivada debe anularse antes de contar con un teorema que lo justifique.

El siguiente paso consistirá en estudiar un **extremo local alcanzado en un punto interior y con derivada existente**. Solo al combinar esas condiciones podremos obtener una conclusión general sobre el valor de la derivada.

9.3. Fermat: condición necesaria de extremo interior

Un extremo local se reconoce comparando el valor central con los de puntos próximos del dominio. Si además el centro es interior y la función es diferenciable allí, los cocientes incrementales de ambos signos deben ser compatibles con esa comparación. Esa compatibilidad determinará el valor de la derivada.

Enunciado: una condición necesaria, no un criterio de búsqueda.

Teorema 9.1. **Teorema de Fermat (condición necesaria para un extremo interior).** Sean $A \subseteq \mathbb{R}$ un conjunto no vacío, $f\colon A \to \mathbb{R}$ y c un punto interior de A. Supongamos que f es diferenciable en c respecto de A y que alcanza allí un máximo local o un mínimo local relativo a A. Entonces

$$\boxed{f'_A(c) = 0.}$$

No se exige que el extremo sea estricto, que sea absoluto ni que f sea diferenciable en otros puntos del dominio.

Demostración. Supongamos primero que c es un punto de **máximo local**. Por Definición 9.2, existe $\delta_1 > 0$ tal que

$$f(c+h) - f(c) \leq 0$$

siempre que $c + h \in A$ y $|h| < \delta_1$. Como c es interior de A, existe $\delta_0 > 0$ con

$$(c - \delta_0, c + \delta_0) \subseteq A.$$

Tomemos $\rho := \min\{\delta_0, \delta_1\} > 0$. Para todo h con $0 < |h| < \rho$ está definido el cociente

$$Q(h) := \frac{f(c+h) - f(c)}{h}.$$

Al **dividir por un incremento positivo**, conservamos el sentido de la desigualdad:

$$0 < h < \rho \quad \Longrightarrow \quad Q(h) \leq 0.$$

Al **dividir por un incremento negativo**, debemos invertirlo:

$$-\rho < h < 0 \quad \Longrightarrow \quad Q(h) \geq 0.$$

La diferenciabilidad en c afirma que existe un número real $L = f'_A(c)$ tal que $Q(h) \to L$ cuando $h \to 0$ por los incrementos admisibles. En particular, las restricciones de Q a los incrementos positivos y negativos tienen **el mismo límite** L; ambos lados están disponibles porque c es interior. Aplicando la preservación del orden bajo límites (Proposición 4.4) a cada lado por separado, obtenemos

$$\underbrace{L \leq 0}_{h \to 0^+}, \qquad \underbrace{L \geq 0}_{h \to 0^-}.$$

Por antisimetría del orden, $L = 0$, y, por tanto, $f'_A(c) = 0$. Si c es un punto de **mínimo local**, la desigualdad inicial se invierte: $f(c+h) - f(c) \geq 0$. Entonces $Q(h) \geq 0$ para $h > 0$ y $Q(h) \leq 0$ para $h < 0$. El mismo argumento de límites da $L \geq 0$ y $L \leq 0$, luego nuevamente $L = 0$. □

La demostración no dice que un máximo o un mínimo «tenga una tangente horizontal» por inspección de la gráfica: **obtiene la nulidad de la derivada desde el orden de los cocientes y la igualdad de sus dos límites laterales**. La interpretación geométrica viene después del argumento, no lo sustituye.

Dónde trabaja cada hipótesis.

Dato	Uso exacto en la demostración
Extremo **local**	Proporciona el signo de $f(c+h) - f(c)$ para incrementos suficientemente pequeños.
Punto **interior**	Garantiza incrementos admisibles positivos y negativos.
Diferenciabilidad en c	Garantiza un único límite real finito común a ambos lados.

No hemos utilizado la continuidad en todo un intervalo cerrado ni el teorema de los valores extremos. Este último será necesario **después**, cuando queramos garantizar que existe algún punto extremo; Fermat parte de un punto extremo ya dado y obtiene información sobre su derivada. Si un extremo **absoluto** se alcanza en un punto interior diferenciable, el teorema se aplica porque todo extremo absoluto es local (Definición 9.1 y Definición 9.2).

La recíproca es falsa: una derivada nula sin extremo.

Ejemplo 9.5. **La función cúbica en el origen.** Consideremos

$$f\colon [-1,1] \to \mathbb{R}, \qquad f(x) = x^3.$$

Por las reglas de derivación ya demostradas, $f'(x) = 3x^2$ en el interior y, en particular, $f'(0) = 0$. Sin embargo, dado cualquier $\delta > 0$, podemos elegir t con $0 < t < \min\{\delta, 1\}$; entonces $-t, t \in [-1,1]$ y

$$f(-t) = -t^3 < 0 = f(0) < t^3 = f(t).$$

Por pequeños que sean los entornos del origen, contienen valores estrictamente inferiores y superiores a $f(0)$. Por tanto, 0 no es punto de máximo ni de mínimo local.

En lenguaje lógico, Fermat demuestra

$$\text{extremo local interior} + \text{derivabilidad} \quad \Longrightarrow \quad f'(c) = 0,$$

pero **no** demuestra la implicación inversa. Una derivada nula identifica un punto que puede merecer estudio; no clasifica por sí sola su comportamiento. Los procedimientos sistemáticos de clasificación se reservarán al capítulo siguiente.

Si la derivada no existe, puede haber un extremo.

Ejemplo 9.6. **Una cúspide con máximo.** Sea

$$g\colon [-1,1] \to \mathbb{R}, \qquad g(x) = -|x|.$$

Para todo $x \in [-1,1]$, $g(x) \le 0 = g(0)$, con desigualdad estricta cuando $x \neq 0$. Así, el punto interior 0 es un máximo absoluto y local estricto. Pero, si $h \neq 0$,

$$\frac{g(h) - g(0)}{h} = -\frac{|h|}{h} = \begin{cases} -1, & h > 0, \\ 1, & h < 0. \end{cases}$$

Los límites laterales no coinciden y, por el criterio Proposición 6.3, $g'(0)$ no existe como número real. El extremo subsiste: lo que falla es **una de las hipótesis de Fermat**, no su conclusión aplicada ilegítimamente.

También debemos mantener la condición de **interioridad**. La identidad sobre $[0,1]$ del Ejemplo 9.2 alcanza extremos en ambos bordes y tiene allí derivadas relativas iguales a 1. En esos puntos solo es posible acercarse desde un lado, de modo que no aparecen las dos desigualdades opuestas que obligaron a concluir $L = 0$.

Tres afirmaciones diferentes
Que $f'(c) = 0$ **no garantiza** un extremo local (función cúbica). Que exista un extremo local **no garantiza** diferenciabilidad (cúspide). Y un extremo de borde, incluso con derivada relativa existente, **no obliga** a que esa derivada sea cero (función identidad restringida). Ninguna de estas observaciones contradice Teorema 9.1: en cada caso falta una condición necesaria para aplicarlo o se intenta invertir su implicación.

Laboratorio de demostración: reconstruir el argumento. Antes de pasar a Rolle, reconstruye el razonamiento sin consultar la prueba: (i) escribe la desigualdad que caracteriza un máximo local; (ii) explica por qué el denominador cambia el sentido de esa desigualdad al pasar de $h > 0$ a $h < 0$; (iii) identifica qué hipótesis asegura que los dos límites corresponden al **mismo número**; (iv) repite el argumento para un mínimo; (v) señala en cuál de esos pasos fracasa cada uno de los ejemplos anteriores. Si no puedes justificar una flecha, todavía falta una hipótesis o una propiedad del límite.

La siguiente sección combinará dos resultados que ya están disponibles: la **existencia** de máximos y mínimos para funciones continuas en un intervalo cerrado (Teorema 5.5) y la **nulidad de la derivada** cuando uno de esos extremos se alcanza en un punto interior diferenciable (Teorema 9.1). La igualdad de los valores en los extremos del intervalo permitirá enlazar ambos hechos y obtener el teorema de Rolle.

9.4. El teorema de Rolle

Fermat nos dice qué ocurre **si ya conocemos** un punto interior en el que se alcanza un extremo diferenciable. Para pasar de esa información local a un resultado sobre un intervalo completo falta justificar la existencia de un punto semejante. La continuidad sobre un intervalo cerrado garantizará extremos absolutos; la igualdad de los valores en los bordes permitirá localizar al menos uno en el interior, salvo en un caso que resolveremos por separado.

Enunciado y alcance.

Teorema 9.2. **Teorema de Rolle.** Sean $a, b \in \mathbb{R}$ con $a < b$ y sea $f\colon [a, b] \to \mathbb{R}$. Supongamos que:

1. f es continua en $[a, b]$ (relativamente a su dominio);
2. f es diferenciable en cada punto de (a, b);
3. $f(a) = f(b)$.

Entonces **existe** $c \in (a, b)$ tal que

$$\boxed{f'(c) = 0.}$$

No se afirma que ese punto sea único ni que podamos determinarlo mediante una fórmula.

Demostración. Llamemos k al valor común de los bordes:

$$k := f(a) = f(b).$$

Primer paso: existencia de extremos absolutos. Como f es continua en el intervalo cerrado y acotado $[a, b]$, el teorema de los valores extremos (Teorema 5.5) proporciona puntos $u, v \in [a, b]$ y números $m := f(u)$ y $M := f(v)$ tales que

$$m \le f(x) \le M \qquad (x \in [a, b]).$$

En particular, evaluando en cualquiera de los dos bordes, obtenemos

$$m \leq k \leq M.$$

Segundo paso: distinguir si la función es constante. Si $m = M$, las desigualdades anteriores fuerzan $f(x) = m$ para todo $x \in [a, b]$. Tomemos

$$c := \frac{a+b}{2} \in (a, b).$$

Como f es constante, para todo incremento admisible no nulo su cociente incremental en c es 0; por tanto, $f'(c) = 0$. Este caso queda resuelto. Supongamos ahora que $m < M$. **Al menos una** de las desigualdades $m \leq k \leq M$ debe ser estricta: si ambas fuesen igualdades, tendríamos $m = M = k$, en contradicción con $m < M$.

- Si $M > k$, el punto v donde se alcanza el máximo absoluto **no puede ser** a **ni** b, porque $f(a) = f(b) = k < M = f(v)$. De $v \in [a, b]$ deducimos $v \in (a, b)$.
- Si no ocurre $M > k$, entonces $M = k$ y, puesto que $m < M$, se tiene $m < k$. El punto u donde se alcanza el mínimo absoluto tampoco puede ser a ni b, ya que $f(u) = m < k = f(a) = f(b)$. En consecuencia, $u \in (a, b)$.

En cualquiera de los dos casos existe un punto $c \in (a, b)$ donde f alcanza un extremo absoluto, máximo o mínimo. Por Definición 9.1 y Definición 9.2, ese extremo es también local. La segunda hipótesis asegura que f es diferenciable precisamente en ese c. Podemos aplicar ahora el teorema de Fermat (Teorema 9.1) y concluir

$$f'(c) = 0.$$

Hemos establecido la existencia requerida tanto si f es constante como si no lo es. □

La prueba no comienza buscando una solución de $f'(x) = 0$. **Construye la existencia** de un punto adecuado mediante un teorema de extremos y solo después utiliza Fermat. Por eso funciona sin disponer de una fórmula para f o para su derivada. Tampoco exige derivadas en los bordes: el punto donde se aplica Fermat quedó localizado en (a, b).

La igualdad en los bordes no obliga a que ambos extremos sean interiores. Conviene identificar la inferencia exacta del segundo paso. El teorema de los valores extremos solo entrega $u, v \in [a, b]$; por sí mismo no sitúa ninguno en (a, b). Tampoco basta escribir $f(a) = f(b)$ y afirmar sin más que «el máximo está en el interior»: una función no constante puede alcanzar su máximo exclusivamente en los bordes. Lo que sí sabemos es que, al ser no constante, **algún valor de la función difiere del valor común** k; si un valor lo supera, un máximo absoluto debe alcanzarse adentro; si ninguno lo supera, algún valor es inferior a k y un mínimo absoluto debe alcanzarse adentro. El argumento cubre ambas posibilidades.

Mapa de dependencias de Rolle
Continuidad en $[a, b] \longrightarrow$ extremos absolutos existentes (Teorema 5.5). Igualdad $f(a) = f(b)$ y no constancia $\longrightarrow$ uno de esos extremos tiene un punto de realización interior. Diferenciabilidad en $(a, b) \longrightarrow$ la derivada existe en ese punto. Extremo interior y derivabilidad $\longrightarrow$ derivada nula (Teorema 9.1). Si la función es constante, basta calcular directamente el cociente incremental en cualquier punto interior.

Un ejemplo que permite encontrar el punto, pero no lo presupone.

Ejemplo 9.7. **El punto de Rolle no tiene por qué ser el punto medio.** Sea

$$f\colon [0,2]\to\mathbb{R}, \qquad f(x)=x^2(x-2).$$

Por ser polinómica, f es continua en $[0,2]$ y diferenciable en $(0,2)$. Además, $f(0)=0=f(2)$. Las tres hipótesis de Teorema 9.2 están verificadas **antes de derivar**. Rolle garantiza, por tanto, al menos un $c\in(0,2)$ con $f'(c)=0$. La derivada, obtenida mediante las reglas ya demostradas, es $f'(x)=3x^2-4x=x(3x-4)$. Los únicos ceros reales de esa expresión son $x=0$ y $x=4/3$. El primero no pertenece al intervalo abierto; el segundo satisface $0<4/3<2$. Así, el único punto que cumple la conclusión en el intervalo requerido es

$$c=\frac{4}{3}, \qquad f'\!\left(\frac{4}{3}\right)=0.$$

El punto medio es 1 y $f'(1)=-1\neq 0$. La igualdad de los valores en los bordes no obliga a que la derivada se anule en el punto medio.

Existencia no significa unicidad
Para la función constante $f(x)=k$ en cualquier $[a,b]$ con $a<b$, **todos** los puntos de (a,b) satisfacen $f'(c)=0$. En Ejemplo 9.7 encontramos exactamente uno. Ambos comportamientos cumplen Rolle: su conclusión tiene la forma «existe al menos un punto».

Laboratorio de lectura: localizar cada salto lógico. Reconstruye la demostración tapando el texto: identifica primero el teorema que garantiza la existencia de m y M; explica por qué $m\le k\le M$; separa el caso $m=M$; justifica que $m<M$ obliga a tener $M>k$ o $m<k$; demuestra en cada alternativa que el punto extremo elegido no puede ser un borde; comprueba que la diferenciabilidad se utiliza **después** de localizar ese punto. Finalmente, indica por qué el razonamiento no invoca el teorema del valor medio, que todavía no se ha demostrado. Hemos obtenido una afirmación global a partir de dos resultados anteriores, pero cada hipótesis cumple una función distinta. En la sección siguiente examinaremos qué puede fallar cuando retiramos alguna de ellas, antes de utilizar Rolle para construir el teorema del valor medio.

9.5. Auditoría de las hipótesis de Rolle

La demostración de Teorema 9.2 utiliza tres datos distintos. La continuidad en el intervalo cerrado asegura la **existencia de extremos absolutos**; la igualdad de los valores de borde permite encontrar un punto extremo interior cuando la función no es constante; y la diferenciabilidad en el interior autoriza aplicar Fermat en ese punto. ¿Qué sucede si conservamos dos de esas condiciones y retiramos la tercera? Para responder no basta con señalar que un paso de la demostración deja de estar justificado: construiremos funciones en las que también **fracasa la conclusión**.

Sin continuidad en el intervalo cerrado.

Ejemplo 9.8. **Una discontinuidad en el borde elimina la garantía.** Definamos la función

$$f\colon [0,1]\longrightarrow\mathbb{R}, \qquad f(x)=\begin{cases} x, & 0\le x<1,\\ 0, & x=1.\end{cases}$$

Los extremos del dominio tienen la misma imagen: $f(0) = 0 = f(1)$. Además, para todo $x \in (0,1)$ existe un entorno de x contenido en $(0,1)$ en el que $f(t) = t$. Por la definición de derivada, f es diferenciable en **cada punto interior** y

$$f'(x) = 1 \qquad (0 < x < 1).$$

Por tanto, no existe $c \in (0,1)$ para el que $f'(c) = 0$. La condición que falta es la continuidad en el intervalo cerrado: al aproximarnos a 1 mediante puntos del dominio menores que 1,

$$\lim_{x \to 1^-} f(x) = 1 \neq 0 = f(1).$$

En efecto, $f([0,1]) = [0,1)$: la función tiene supremo 1, pero no alcanza un máximo absoluto. Ha fallado precisamente la garantía de existencia de extremos utilizada al comenzar la prueba de Rolle.

El ejemplo no permite reemplazar «continua en $[a,b]$» por «continua y diferenciable en (a,b)». La diferenciabilidad interior ya implica continuidad en el interior (Teorema 6.1), pero **no controla la continuidad relativa en los dos bordes**. La falla de este ejemplo se concentra en $b = 1$.

Sin diferenciabilidad en un punto interior.

Ejemplo 9.9. **Una esquina puede concentrar el extremo sin producir derivada nula.** Consideremos

$$g\colon [-1,1] \longrightarrow \mathbb{R}, \qquad g(x) = |x|.$$

La desigualdad $||x| - |y|| \leq |x - y|$ prueba que g es continua en todo su dominio. Sus valores de borde coinciden: $g(-1) = 1 = g(1)$. En los puntos interiores distintos de cero, el cálculo por las dos ramas de la función da

$$g'(x) = \begin{cases} -1, & -1 < x < 0, \\ 1, & 0 < x < 1. \end{cases}$$

En 0, en cambio, los cocientes incrementales satisfacen

$$\frac{g(h) - g(0)}{h} = \frac{|h|}{h} = \begin{cases} -1, & h < 0, \\ 1, & h > 0. \end{cases}$$

Como sus límites laterales difieren, $g'(0)$ no existe (Proposición 6.3). No hay **ningún punto interior con derivada existente e igual a cero**: fuera de 0 las derivadas son -1 o 1, y en 0 la derivada no está definida. Precisamente donde se alcanza el mínimo absoluto, $g(0) = 0$, no podemos aplicar Fermat.

Continuidad y valores de borde iguales sí bastan para encontrar un extremo interior en este caso; **no bastan para convertirlo en un cero de la derivada**. Un extremo angular no contradice Fermat: le falta la diferenciabilidad exigida por Teorema 9.1. Compárese con la cúspide de Ejemplo 9.6, donde también existe un extremo interior sin derivada, aunque el signo de la función es opuesto.

Sin igualdad de los valores en los bordes.

Ejemplo 9.10. **La identidad no tiene derivada nula.** Recuperemos la función de Ejemplo 9.2, ahora para comprobar una hipótesis diferente:

$$h\colon [0,1] \longrightarrow \mathbb{R}, \qquad h(x) = x.$$

Es continua en todo $[0,1]$ y diferenciable en cada punto interior. Sin embargo,

$$h(0) = 0 \neq 1 = h(1), \qquad h'(x) = 1 \quad (0 < x < 1).$$

La conclusión de Rolle es falsa: no existe $c \in (0,1)$ con $h'(c) = 0$. El máximo y el mínimo absolutos se alcanzan en bordes distintos y no hay un valor común de borde que fuerce un extremo interior. Aquí faltó exclusivamente la tercera hipótesis.

Este contraejemplo no significa que valores de borde diferentes impidan siempre una derivada nula. Por ejemplo, $q\colon [-1,2] \to \mathbb{R}$, $q(x) = x^2$, satisface $q(-1) = 1 \neq 4 = q(2)$, pero $q'(0) = 0$ y $0 \in (-1,2)$. Se distingue así una hipótesis **suficiente para una garantía universal** de una condición que fuese necesaria para que una función particular presentase la conclusión.

Matriz de control: fallar una sola hipótesis. En cada fila se conserva el mismo tipo de dominio, un intervalo cerrado no degenerado, y se mantienen las **otras dos hipótesis**.

Función	Continua en el cerrado	Diferenciable en el abierto	Valores de borde iguales	¿Existe c interior con $f'(c) = 0$?
f de Ejemplo 9.8	No	Sí	Sí	No
g de Ejemplo 9.9	Sí	No	Sí	No
h de Ejemplo 9.10	Sí	Sí	No	No

Las tres filas prueban algo preciso: **ninguna de las tres hipótesis puede suprimirse sin más del enunciado de Rolle, conservando la misma conclusión para todas las funciones**. No prueban que las hipótesis sean individualmente necesarias para que *alguna función concreta* tenga una derivada nula; el ejemplo q anterior ya descarta esa interpretación.

No confundir el fracaso de una prueba con el de una conclusión
Si falta una hipótesis, Rolle deja de ser aplicable. De ahí solo se deduce que **el teorema no garantiza la conclusión**: para afirmar que esta es falsa debemos examinar la derivada en *todo* el intervalo abierto. Eso hicimos en los tres contraejemplos. Tampoco se exige que f sea diferenciable en a o en b: añadir esa condición cambiaría el teorema sin necesidad.

Laboratorio de reparación de enunciados. Considera las siguientes frases incompletas y repara cada una aportando la condición ausente, no simplemente una gráfica: «si $f(a) = f(b)$ y f es diferenciable en (a, b), existe un cero de f'»; «si f es continua en $[a, b]$ y $f(a) = f(b)$, existe un cero de f'»; «si f es continua en $[a, b]$ y diferenciable en (a, b), existe un cero de f'». Para cada reparación, identifica primero cuál de los tres ejemplos la refuta y explica después en qué línea de la demostración original se necesita la hipótesis incorporada. Finalmente, contrasta «no puedo aplicar Rolle» con «sé que no hay derivada nula»: son afirmaciones lógicamente diferentes.

Para tratar valores de borde diferentes, restaremos a la función la recta secante. La nueva función tendrá valores iguales en ambos extremos y conservará las condiciones de continuidad y diferenciabilidad necesarias para aplicar Rolle.

9.6. El teorema del valor medio

El teorema de Rolle resuelve el caso en que los valores de la función coinciden en los bordes: bajo continuidad y diferenciabilidad, garantiza una derivada interior igual a cero. Para responder la pregunta de §9.1 no podemos exigir esa igualdad; la pendiente de la secante puede ser cualquier número real. En lugar de modificar Rolle, **transformaremos la función** de manera que las hipótesis de ese teorema queden satisfechas.

La construcción: restar la recta secante. Sean $a < b$ y $f\colon [a, b] \to \mathbb{R}$. La pendiente de la secante entre $(a, f(a))$ y $(b, f(b))$ es $m := \frac{f(b)-f(a)}{b-a}$. El cociente está definido porque $b - a > 0$. La función afín

$$\ell\colon [a, b] \to \mathbb{R}, \qquad \ell(x) := f(a) + m(x - a)$$

satisface $\ell(a) = f(a)$ y $\ell(b) = f(a) + m(b - a) = f(b)$. Es, por tanto, la recta que une los dos puntos de la gráfica especificados por los bordes. Si definimos

$$g\colon [a, b] \to \mathbb{R}, \qquad \boxed{g(x) := f(x) - \ell(x) = f(x) - f(a) - m(x - a),}$$

la diferencia tiene valores de borde iguales, y de hecho nulos:

$$g(a) = f(a) - f(a) - m(a - a) = 0,$$
$$g(b) = f(b) - f(a) - m(b - a) = 0.$$

Esta igualdad se obtiene por construcción, sin suponer que $f(a) = f(b)$. Si conservamos la continuidad y la diferenciabilidad de f, la resta de una función afín no las destruye. Ya contamos con los ingredientes para aplicar Rolle a g, no necesariamente a f.

Enunciado y demostración.

Teorema 9.3. **Teorema del valor medio de Lagrange.** Sean $a, b \in \mathbb{R}$ con $a < b$ y sea $f\colon [a, b] \to \mathbb{R}$. Supongamos que f es continua en $[a, b]$ respecto de ese dominio y diferenciable en cada punto de (a, b). Entonces existe $c \in (a, b)$ tal que

$$\boxed{f'(c) = \frac{f(b) - f(a)}{b - a}.}$$

No se requieren derivadas en a ni en b; la conclusión afirma la existencia de al menos un punto interior, no su localización mediante una fórmula.

Demostración. Definamos m y g como en la construcción anterior:

$$m := \frac{f(b)-f(a)}{b-a}, \qquad g(x) := f(x) - f(a) - m(x-a) \quad (x \in [a,b]).$$

Verifiquemos, una por una, las hipótesis de Rolle.

Continuidad en el cerrado. Por hipótesis, f es continua en $[a,b]$. La función $x \mapsto f(a) + m(x-a)$ es afín y continua en ese mismo dominio. Las reglas de continuidad para diferencias aseguran que g es continua en $[a,b]$, incluidos los dos bordes en el sentido relativo.

Diferenciabilidad en el abierto. Para cada $x \in (a,b)$, la derivada $f'(x)$ existe. La función afín es diferenciable y tiene derivada m; por la linealidad de la derivada (Teorema 7.1),

$$g'(x) = f'(x) - m \qquad (x \in (a,b)).$$

Igualdad de los valores de borde. La elección de m da explícitamente

$$g(a) = 0 = g(b).$$

El intervalo es no degenerado porque $a < b$. Todas las hipótesis de Teorema 9.2 están ahora comprobadas **para** g. Por Rolle existe $c \in (a,b)$ tal que $g'(c) = 0$. Utilizando la fórmula para la derivada de g obtenemos

$$0 = g'(c) = f'(c) - m,$$

de donde

$$f'(c) = m = \frac{f(b)-f(a)}{b-a}.$$

Ese es precisamente el punto cuya existencia debíamos demostrar. □

La prueba no presupone que la derivada alcance la pendiente de la secante. **Produce** la igualdad al aplicar un resultado ya demostrado a una función nueva: la nulidad de g' equivale, por una identidad obtenida antes de invocar Rolle, a la igualdad entre las dos pendientes de f. No hemos utilizado el teorema del valor medio dentro de su propia demostración.

Mapa de la transformación
La pendiente m depende solamente de los datos de borde. Restar $\ell(x) = f(a) + m(x-a)$ produce $g(a) = g(b) = 0$ y mantiene continuidad y diferenciabilidad. Rolle proporciona $g'(c) = 0$ para **algún** c interior. La identidad $g'(c) = f'(c) - m$ devuelve la conclusión para la función original. Cada flecha está justificada antes de pasar a la siguiente.

Un cálculo particular después del teorema.

Ejemplo 9.11. **La secante de una cúbica y su pendiente interior.** Consideremos

$$f\colon [0,2] \to \mathbb{R}, \qquad f(x) = x^3.$$

La función es continua en $[0,2]$ y diferenciable en $(0,2)$ por ser polinómica. La pendiente de su secante es

$$m = \frac{f(2)-f(0)}{2-0} = \frac{8}{2} = 4.$$

El teorema garantiza algún $c \in (0,2)$ con $f'(c) = 4$. Como $f'(x) = 3x^2$, resolver la ecuación **después de haber justificado la existencia** da

$$3c^2 = 4 \quad \Longleftrightarrow \quad c = \pm\frac{2}{\sqrt{3}}.$$

Solo la raíz positiva pertenece a $(0,2)$: $0 < 2/\sqrt{3} < 2$, porque $1 < \sqrt{3} < 2$. En consecuencia,

$$c = \frac{2}{\sqrt{3}}, \qquad f'(c) = 4.$$

La función auxiliar correspondiente es $g(x) = x^3 - 4x$, con $g(0) = g(2) = 0$ y $g'(x) = 3x^2 - 4$. Rolle aplicado a ella conduce al mismo punto. Obsérvese, además, que el punto medio del intervalo es 1 y $f'(1) = 3 \neq 4$: no debemos sustituir el punto de existencia por el punto medio sin una comprobación independiente.

Qué hemos demostrado y qué queda por examinar. Si $f(a) = f(b)$, la pendiente m vale cero y el valor medio recupera exactamente la conclusión de Rolle para la función original; la construcción auxiliar puede escribirse simplemente $g = f - f(a)$. Si los valores de borde son distintos, Rolle no se aplica necesariamente a f, pero sí a la diferencia g una vez verificadas sus hipótesis. **No hay una hipótesis de igualdad de bordes en el teorema del valor medio.**

La formulación también puede escribirse como una igualdad entre incrementos:

$$f(b) - f(a) = f'(c)(b - a) \qquad \text{para algún } c \in (a,b).$$

Es la misma afirmación porque $b-a \neq 0$; no introduce un nuevo teorema. No hemos demostrado que el punto c sea único, que sea siempre el punto medio ni que podamos encontrarlo resolviendo una ecuación explícita para cualquier función. Estas distinciones, junto con el sentido exacto de la igualdad entre secante y derivada, serán el objeto de la sección siguiente.

Laboratorio de reconstrucción: diseñar la función auxiliar
Sin consultar la demostración, intenta restar a f una función afín desconocida $p(x) = u + vx$. Impón $f(a) - p(a) = f(b) - p(b) = 0$ y determina v y u mediante estas dos ecuaciones; comprueba por qué $a < b$ garantiza que la pendiente queda determinada. Verifica después continuidad y diferenciabilidad de $f - p$, aplica Rolle solamente tras comprobar sus valores de borde y traduce la igualdad $(f-p)'(c) = 0$ a una igualdad sobre f'. Finalmente, identifica en qué lugar fracasa el argumento si se omite la continuidad en uno de los bordes o la diferenciabilidad en un punto interior.

Hemos alcanzado el resultado central de este capítulo. Antes de utilizarlo para obtener estimaciones y estudiar funciones con la misma derivada, precisaremos en qué sentido la pendiente global coincide con una pendiente local y cuáles son los límites lógicos de esa coincidencia.

9.7. Qué afirma realmente el valor medio

El teorema del valor medio (Teorema 9.3) ha cerrado la pregunta planteada en §9.1, pero su fórmula admite lecturas incorrectas si confundimos **igualdad de pendientes**, **coincidencia de rectas** y **determinación de un punto**. En esta sección no añadiremos hipótesis ni demostraremos un teorema nuevo: examinaremos exactamente la conclusión que ya está disponible.

La pendiente común y las dos rectas. Sean $a < b$ y $f\colon [a,b] \to \mathbb{R}$ continua en $[a,b]$ y diferenciable en (a,b). Escribamos

$$m := \frac{f(b)-f(a)}{b-a}, \qquad \ell(x) := f(a) + m(x-a).$$

La recta $y = \ell(x)$ pasa por $(a, f(a))$ y $(b, f(b))$: es la **secante** determinada por los datos de borde. El teorema asegura la existencia de un $c \in (a,b)$ con $f'(c) = m$. La recta tangente a la gráfica en el punto $(c, f(c))$ tiene, en ese caso, la ecuación

$$T_c(x) := f(c) + f'(c)(x-c) = f(c) + m(x-c).$$

Ambas rectas tienen pendiente m. Por ello son **paralelas o coincidentes**. Para decidir si coinciden es necesario un dato adicional: al evaluar las dos expresiones en $x = c$ obtenemos

$$T_c(c) = f(c), \qquad \ell(c) = f(a) + m(c-a).$$

En consecuencia, bajo la igualdad de pendientes,

$$\boxed{T_c = \ell \quad \Longleftrightarrow \quad f(c) = \ell(c).}$$

La igualdad $f'(c) = m$ **no contiene** la segunda condición. Dicho de otro modo, Lagrange garantiza que alguna tangente tiene la dirección de la secante, pero no que el punto de tangencia pertenezca a esa secante. La interpretación geométrica ayuda a leer la fórmula; la prueba de existencia continúa siendo la de §9.6, basada en Rolle, no una inspección del dibujo.

El cuantificador: al menos uno, no exactamente uno. La conclusión de Teorema 9.3 puede expresarse mediante el conjunto

$$S_{f,[a,b]} := \left\{ x \in (a,b) : f'(x) = \frac{f(b)-f(a)}{b-a} \right\}.$$

El teorema afirma $S_{f,[a,b]} \neq \varnothing$. No determina cuántos elementos tiene ese conjunto. Tampoco permite reemplazar «existe c» por «para todo c», elegir de antemano $c = (a+b)/2$ o decidir la posición del punto a partir de los valores de borde. Encontrar todas las soluciones de la ecuación diferencial concreta es una tarea adicional, posible en algunos ejemplos y no incluida en la afirmación universal.

Ejemplo 9.12. **Dos puntos de valor medio y dos tangentes distintas de la secante.** Tomemos

$$f\colon [-1,1] \longrightarrow \mathbb{R}, \qquad f(x) = x^3.$$

Por ser polinómica, la función es continua en $[-1,1]$ y diferenciable en $(-1,1)$. Los valores de borde son $f(-1) = -1$ y $f(1) = 1$; por tanto,

$$m = \frac{1-(-1)}{1-(-1)} = 1, \qquad \ell(x) = x.$$

Lagrange garantiza algún punto interior con derivada igual a 1. Para determinar **todos** esos puntos usamos $f'(x) = 3x^2$:

$$3c^2 = 1 \quad \Longleftrightarrow \quad c = \frac{1}{\sqrt{3}} \text{ o } c = -\frac{1}{\sqrt{3}}.$$

Ambos pertenecen a $(-1,1)$, ya que $\sqrt{3}>1$. Son las únicas soluciones reales de la ecuación y, por tanto,

$$S_{f,[-1,1]}=\left\{-\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}}\right\}.$$

El punto medio del intervalo es 0, pero $f'(0)=0\neq 1$; **ninguno** de los dos puntos prometidos es el punto medio. Si c es cualquiera de los puntos hallados, $3c^2=1$ implica $c^3=c/3$. La recta tangente en c es entonces $T_c(x)=c^3+1\cdot(x-c)=x-\frac{2c}{3}$. Tiene pendiente 1, igual que la secante $y=x$, pero su término independiente es distinto de cero porque $c\neq 0$. Las dos tangentes son paralelas a la secante y **ninguna coincide con ella**. En particular, igualdad de pendientes no equivale a que la tangente pase por alguno de los extremos del tramo.

El ejemplo separa tres operaciones: verificar las hipótesis para **garantizar** que $S_{f,[a,b]}$ no está vacío; resolver $f'(x)=m$ para **localizar** sus elementos; y comparar T_c con ℓ para decidir si las rectas coinciden. La segunda y la tercera no forman parte de la demostración del teorema.

Los tres comportamientos de la cantidad de puntos. Disponemos ya de ejemplos que muestran comportamientos diferentes, todos compatibles con el mismo enunciado.

Función e intervalo	Pendiente de la secante	Puntos interiores que la reproducen
x^3 en $[0,2]$ (Ejemplo 9.11)	4	Exactamente uno: $2/\sqrt{3}$
x^3 en $[-1,1]$ (Ejemplo 9.12)	1	Exactamente dos: $\pm 1/\sqrt{3}$
$p(x)=2x+1$ en $[0,2]$	2	Todos los puntos de $(0,2)$, pues $p'(x)=2$

En el último caso, la secante y todas las tangentes son la misma recta $y=2x+1$; en el segundo, las tangentes y la secante son rectas distintas. Por tanto, ni la cantidad de puntos ni la coincidencia de las rectas pueden añadirse gratuitamente a la conclusión de Lagrange. La única cota de cantidad garantizada es que **hay al menos uno**.

No leer más de lo que dice la igualdad
La existencia de c con $f'(c)=m$ no implica que c sea el punto medio, que sea único, que $f(c)$ esté sobre la secante, ni que $f'(x)=m$ para todos los x interiores. Tampoco exige continuidad de la **función derivada**: las hipótesis son continuidad de f en $[a,b]$ y diferenciabilidad de f en (a,b). No añadamos requisitos sobre f' que el teorema no contiene.

Laboratorio de interpretación y lectura lógica. Repite el ejemplo de la cúbica sin consultar los cálculos anteriores: (i) verifica primero las hipótesis de Lagrange; (ii) halla la secante a partir de los datos de borde; (iii) resuelve la ecuación $f'(c)=m$ y descarta cualquier solución exterior al abierto; (iv) construye la tangente en cada solución y contrástala con la secante; (v) determina si el punto medio satisface la ecuación. Después examina la función afín de la tabla y explica por qué **todos** sus puntos interiores cumplen el teorema. En cada paso separa lo que garantiza el resultado general de lo que procede de los cálculos particulares.

La continuidad en el intervalo cerrado y la diferenciabilidad en el abierto cumplen funciones distintas. Los contraejemplos siguientes mostrarán cómo puede fallar la conclusión al retirar cada una de ellas.

9.8. Auditoría de las hipótesis del valor medio

El teorema del valor medio (Teorema 9.3) parte de una función $f\colon [a,b] \to \mathbb{R}$ con $a < b$ y exige **continuidad en todo** $[a,b]$ y **diferenciabilidad en todo** (a,b). Entonces garantiza un punto interior c donde la derivada reproduce la pendiente de la secante. Para auditar las hipótesis construiremos dos funciones: en cada una conservará su validez una de las condiciones, fallará la otra y **la igualdad prometida no tendrá ninguna solución admisible**. Esto es más exigente que señalar una simple imposibilidad de aplicar el teorema.

Retirar la continuidad en un borde, manteniendo la diferenciabilidad interior.

Ejemplo 9.13. **Una pendiente de secante que no aparece en ninguna derivada.** Definamos

$$f\colon [0,1] \longrightarrow \mathbb{R}, \qquad f(x) = \begin{cases} x, & 0 \le x < 1, \\ 2, & x = 1. \end{cases}$$

En todo punto $x \in (0,1)$, la función coincide con la identidad en un entorno de x. Por tanto, es diferenciable **en cada punto del abierto** y

$$f'(x) = 1 \qquad (0 < x < 1).$$

Sin embargo, en el extremo derecho la continuidad relativa falla:

$$\lim_{x \to 1^-} f(x) = 1 \neq 2 = f(1).$$

La pendiente de la secante entre los puntos de borde es

$$m = \frac{f(1) - f(0)}{1 - 0} = \frac{2 - 0}{1} = 2.$$

Para todo $c \in (0,1)$ tenemos $f'(c) = 1 \neq 2 = m$. Por consiguiente,

$$\boxed{\nexists c \in (0,1) : \quad f'(c) = \frac{f(1) - f(0)}{1 - 0}.}$$

La diferenciabilidad interior se ha conservado íntegramente: la única hipótesis analítica incumplida es la continuidad en $b = 1$. El cambio del valor en ese borde modifica la pendiente de la secante **sin modificar ninguna derivada interior**.

Podemos localizar el fracaso en la prueba de §9.6. La recta secante del ejemplo es $\ell(x) = 2x$ y la función auxiliar $g = f - \ell$ satisface $g(0) = g(1) = 0$; no obstante,

$$g(x) = -x \quad (0 \le x < 1), \qquad \lim_{x \to 1^-} g(x) = -1 \neq 0 = g(1).$$

Aunque $g'(x) = -1$ en todo $(0,1)$ y sus valores en los dos bordes coinciden, **no es continua en** $[0,1]$. Rolle (Teorema 9.2) no puede utilizarse, y aquí comprobamos además que su conclusión, $g'(c) = 0$, es efectivamente falsa. Compárese con Ejemplo 9.8: modificar el dato de borde permite refutar tanto Rolle como Lagrange, pero en este ejemplo la secante tiene pendiente 2, no pendiente cero.

Retirar la diferenciabilidad en el interior, manteniendo la continuidad global.

Ejemplo 9.14. **Una esquina omite la pendiente promedio.** Consideremos

$$h\colon [-1,2] \longrightarrow \mathbb{R}, \qquad h(x) = |x|.$$

La desigualdad $||x| - |y|| \leq |x - y|$ demuestra que h es continua en todo $[-1,2]$, incluidos los bordes. Sus valores en los extremos del dominio son $h(-1) = 1$ y $h(2) = 2$, de modo que

$$m = \frac{h(2) - h(-1)}{2 - (-1)} = \frac{2-1}{3} = \frac{1}{3}.$$

En los puntos interiores distintos de cero, las ramas de la función proporcionan

$$h'(x) = \begin{cases} -1, & -1 < x < 0, \\ 1, & 0 < x < 2. \end{cases}$$

En 0 el cociente incremental es

$$\frac{h(0+t) - h(0)}{t} = \frac{|t|}{t} = \begin{cases} -1, & t < 0, \\ 1, & t > 0. \end{cases}$$

Los límites laterales no coinciden, por lo que $h'(0)$ no existe (Proposición 6.3). Hemos examinado **todos** los puntos interiores: donde existe la derivada vale -1 o 1, nunca $1/3$; en el único punto restante ni siquiera está definida. Por tanto,

$$\boxed{\nexists c \in (-1,2) : \quad h'(c) = \frac{h(2) - h(-1)}{2 - (-1)}.}$$

La continuidad en todo el cerrado sigue siendo verdadera. La hipótesis que ha fallado, y solo ella, es la diferenciabilidad en el punto interior 0.

Restar la recta secante tampoco repara una esquina. Si $\ell(x) = 1 + (x+1)/3$ y $g = h - \ell$, se cumple $g(-1) = g(2) = 0$ y g es continua en todo $[-1,2]$. Sin embargo, en los dos tramos interiores diferenciables tenemos

$$g'(x) = \begin{cases} -\dfrac{4}{3}, & -1 < x < 0, \\ \dfrac{2}{3}, & 0 < x < 2, \end{cases}$$

mientras que en 0 su derivada no existe: restar una función afín diferenciable no puede hacer coincidir los dos límites laterales diferentes del cociente de h. De nuevo, Rolle deja de ser aplicable a g y ninguna derivada existente de g se anula.

Qué demuestran exactamente los dos ensayos.

Función e intervalo	Continuidad en el cerrado	Diferenciabilidad en el abierto	Pendiente de la secante	Valores de la derivada interior existente	¿Se alcanza la pendiente?
f de Ejemplo 9.13, $[0,1]$	No: falla en 1	Sí	2	Siempre 1	No
h de Ejemplo 9.14, $[-1,2]$	Sí	No: falla en 0	$1/3$	-1 o 1	No

Las dos filas prueban que **ninguna de las dos hipótesis analíticas puede simplemente omitirse de la garantía universal de Lagrange**. La primera muestra que la diferenciabilidad interior no controla los valores fijados en los bordes; la segunda, que la continuidad global no obliga a que exista una derivada en el lugar donde sería necesaria la pendiente intermedia. No existe en Lagrange una hipótesis adicional de igualdad de los valores extremos: esa condición pertenece al caso particular de Rolle y no debe incorporarse por error al enunciado de Teorema 9.3.

Ausencia de garantía no equivale a imposibilidad en cada función
De que falle una hipótesis no se deduce automáticamente que la igualdad del valor medio sea falsa; hay que comprobarlo, como hicimos. Por ejemplo, si $q\colon [0,1] \to \mathbb{R}$ está dada por $q(x) = x^2$ para $0 \le x < 1$ y $q(1) = 1/2$, falla la continuidad en 1; sin embargo, la secante tiene pendiente $1/2$ y $q'(1/4) = 1/2$. De forma similar, $r(x) = x^2 - |x|$ en $[-1,1]$ es continua y no diferenciable en 0, pero sus bordes valen 0 y $r'(-1/2) = r'(1/2) = 0$. Las hipótesis de Lagrange son condiciones **suficientes para asegurar** la existencia en toda función que las cumpla; estos ejemplos favorables no permiten suprimirlas del teorema general.

Laboratorio de prueba: diagnosticar sin confundir enunciado y conclusión. Para cada uno de los ejemplos principales, reconstruye el argumento en este orden: (i) determina los puntos exactos donde la continuidad y la diferenciabilidad se cumplen o fallan; (ii) calcula la pendiente de la secante usando los valores **efectivos** de borde; (iii) determina el recorrido de los valores de la derivada en todos los puntos interiores donde existe; (iv) comprueba por separado que el valor de la secante no pertenece a ese recorrido; (v) resta la recta secante y señala exactamente qué hipótesis de Rolle deja de cumplirse. Finalmente, contrasta cada contraejemplo con uno de los casos favorables de la advertencia anterior y explica la diferencia lógica entre «el teorema no se aplica» y «la conclusión es falsa».

La auditoría completa el control de las hipótesis antes de utilizar el valor medio como herramienta cuantitativa. En §9.9 emplearemos la igualdad **ya demostrada**, bajo todas sus condiciones, para convertir cotas de la derivada en cotas de los incrementos de la función, sin introducir todavía criterios sistemáticos de monotonía.

9.9. Estimaciones mediante la derivada

El teorema del valor medio no se limita a asegurar la existencia de una tangente paralela a una secante. También permite obtener información sobre la **variación total** de una función

sin calcularla exactamente. Supongamos que conocemos cotas válidas para todas sus derivadas interiores: ¿qué límites imponen a la diferencia entre dos valores de la función? Esta vez no buscaremos el punto c que promete Lagrange; utilizaremos únicamente la certeza de que existe.

De las cotas puntuales a una cota del incremento.

PROPOSICIÓN 9.1. **Proposición (estimación de incrementos por cotas de la derivada).** Sean $a, b \in \mathbb{R}$, $a < b$, y $f\colon [a,b] \to \mathbb{R}$ una función continua en $[a,b]$ y diferenciable en cada punto de (a,b). Si existen números reales m, M tales que

$$m \leq f'(t) \leq M \qquad \text{para todo } t \in (a,b),$$

entonces, para cualesquiera $x, y \in [a,b]$ con $x < y$,

$$\boxed{m(y-x) \leq f(y) - f(x) \leq M(y-x).}$$

En particular, tomando $x = a$ e $y = b$,

$$m(b-a) \leq f(b) - f(a) \leq M(b-a).$$

Demostración. Fijemos arbitrariamente $x, y \in [a,b]$ con $x < y$. Como $[x,y] \subseteq [a,b]$, la restricción $f|_{[x,y]}$ es continua en $[x,y]$ y diferenciable en (x,y). Estas propiedades se heredan de las dos hipótesis sobre f; no estamos suponiendo una derivada en x ni en y.

Por el teorema del valor medio (Teorema 9.3), existe un punto $c \in (x,y)$ para el que

$$\frac{f(y) - f(x)}{y - x} = f'(c).$$

La hipótesis de cotas se aplica a este c porque $(x,y) \subseteq (a,b)$. Por tanto,

$$m \leq \frac{f(y) - f(x)}{y - x} \leq M.$$

Finalmente, $y - x > 0$: al multiplicar por ese número **no cambian de sentido** las desigualdades. Obtenemos

$$m(y-x) \leq f(y) - f(x) \leq M(y-x).$$

Como $x < y$ fueron arbitrarios, la estimación rige para todos los pares ordenados indicados. □

El orden de la prueba es importante. Primero se comprueban las hipótesis de Lagrange en el **subintervalo elegido**; después se obtiene un punto interior c; solo entonces se aplica la cota a $f'(c)$. No se requiere que f' sea continua, que tenga máximo o mínimo, ni que conozcamos el valor exacto de c. Los números m y M son cotas, no necesariamente valores que alcance la derivada. La desigualdad es una estimación cuantitativa, no una fórmula exacta para $f(y) - f(x)$. Cuando $m = M$, las dos cotas coinciden; la consecuencia estructural de tener una misma derivada en todos los puntos se examinará sistemáticamente en §9.10.

Controlar los cambios sin atender a su signo. Supongamos, con las mismas hipótesis sobre f, que existe $K \in \mathbb{R}$ con $K \geq 0$ y

$$|f'(t)| \leq K \qquad (t \in (a,b)).$$

Esta condición equivale a $-K \leq f'(t) \leq K$. Aplicando Proposición 9.1 con $m = -K$ y $M = K$, para $x < y$ obtenemos

$$-K(y-x) \leq f(y) - f(x) \leq K(y-x).$$

Por la caracterización del valor absoluto,

$$|f(y) - f(x)| \leq K(y-x) = K|y-x|.$$

Si $x > y$, se aplica la misma desigualdad al par intercambiado: $|f(y) - f(x)| = |f(x) - f(y)|$. Si $x = y$, ambos miembros valen cero. En consecuencia,

$$\boxed{\forall x, y \in [a,b], \qquad |f(y) - f(x)| \leq K|y-x|.}$$

El paso desde $f'(c)$ hasta esta desigualdad **no utiliza** ningún criterio previo de monotonía. Lo que controla es la magnitud del cambio de la función respecto de la distancia entre argumentos, incluso si la derivada cambia de signo a lo largo del intervalo.

La condición de Lipschitz.

Definición 9.3. **Función Lipschitz en su dominio.** Sean $A \subseteq \mathbb{R}$ no vacío y $f\colon A \to \mathbb{R}$. Diremos que f es **Lipschitz en** A si existe una constante real $K \geq 0$ tal que

$$\forall x, y \in A, \qquad |f(y) - f(x)| \leq K|y-x|.$$

Cualquier K que satisfaga esta desigualdad se llama una **constante de Lipschitz** para f en A. La definición no presupone que f sea diferenciable ni exige que exista una constante óptima.

Hemos probado una consecuencia precisa: **si f es continua en un intervalo cerrado y diferenciable en su interior, y su derivada satisface la cota uniforme $|f'| \leq K$ allí, entonces f es Lipschitz en todo el cerrado con constante K.** La palabra *uniforme* significa que se emplea el mismo número K para todos los puntos interiores, no una cota distinta para cada uno. La definición proporciona también continuidad, sin recurrir a ninguna derivada. Dado $x_0 \in A$ y $\varepsilon > 0$, si $K > 0$ elegimos $\delta = \varepsilon/K$; para $x \in A$ con $|x - x_0| < \delta$, $|f(x) - f(x_0)| \leq K|x - x_0| < \varepsilon$.

Si $K = 0$, la propia desigualdad Lipschitz fuerza $f(x) = f(x_0)$ para todos los argumentos del dominio y la continuidad es inmediata. Por tanto, toda función Lipschitz es continua **relativamente a su dominio**, incluidos los puntos de borde que pertenezcan a él.

No invertir el resultado sin comprobar sus hipótesis
De «derivada existente y uniformemente acotada en el interior, junto con continuidad en el cerrado» se deduce «Lipschitz». La recíproca no exige diferenciabilidad: $u\colon [-1,1] \to \mathbb{R}$, $u(x) = |x|$, satisface $||x| - |y|| \leq |x - y|$, pero no es diferenciable en 0 (Ejemplo 9.9). Si una función Lipschitz es diferenciable en un punto interior, entonces allí sí se obtiene $|f'(c)| \leq K$ pasando al límite en $|(f(c+h) - f(c))/h| \leq K$; esto **no crea** una derivada donde no existe.

Ejemplo: una cota global y su precisión.

Ejemplo 9.15. **La función cuadrática en un intervalo cerrado.** Sea

$$f\colon [0,2] \to \mathbb{R}, \qquad f(x) = x^2.$$

La función es continua en $[0,2]$ y diferenciable en $(0,2)$. En todo punto interior, $f'(t) = 2t$ y, por tanto, $0 \le f'(t) \le 4$. La proposición Proposición 9.1 proporciona, para $0 \le x < y \le 2$,

$$\boxed{0 \le y^2 - x^2 \le 4(y-x).}$$

En particular, $|f(y) - f(x)| \le 4|y - x|$ para todos los puntos del dominio: 4 es una constante de Lipschitz. Podemos verificarlo además mediante una identidad puramente algebraica,

$$|y^2 - x^2| = |y - x|\,|x + y| \le 4|y - x|,$$

porque $0 \le x + y \le 4$. Son dos vías de comprobación diferentes: una invoca el valor medio; la otra aprovecha la fórmula particular del cuadrado.

¿Podemos sustituir 4 por una constante menor que sea válida para **todos** los pares? Para $0 < h < 2$, tomemos $x = 2 - h$ e $y = 2$. Su cociente de incrementos es $\frac{f(2)-f(2-h)}{h} = \frac{4-(2-h)^2}{h} = 4 - h$. Si $K < 4$, basta escoger $0 < h < \min\{2, 4 - K\}$ para obtener $4 - h > K$ y refutar $|f(y) - f(x)| \le K|y - x|$. De modo que 4 no es solamente una cota válida: es la **menor constante de Lipschitz** posible para esta función en $[0,2]$.

El ejemplo separa dos cuestiones: encontrar *alguna* cota válida para la derivada basta para estimar los incrementos, pero determinar la **mejor** constante requiere un argumento adicional. La derivada no necesita alcanzar el valor 4 en el abierto: se aproxima a él cuando el argumento tiende a 2 desde la izquierda, y los cocientes de incrementos también se aproximan a 4.

Ejemplo: la continuidad no basta para la cota Lipschitz.

Ejemplo 9.16. **Una raíz cuadrada con variación demasiado grande cerca del borde.** Definamos

$$g\colon [0,1] \to \mathbb{R}, \qquad g(x) = \sqrt{x}.$$

Esta función es continua en $[0,1]$ y diferenciable en $(0,1)$, donde $g'(t) = \frac{1}{2\sqrt{t}}$. No existe un K finito que acote $|g'(t)|$ para todo $t \in (0,1)$: el denominador se aproxima a cero cuando $t \to 0^+$. Pero **la ausencia de una cota obtenida por este método no bastaría por sí sola para negar Lipschitz**. Comprobemos la definición directamente: para $0 < h \le 1$,

$$\frac{|g(h) - g(0)|}{|h - 0|} = \frac{\sqrt{h}}{h} = \frac{1}{\sqrt{h}}.$$

Dado cualquier $K \ge 0$, podemos elegir $h \in (0,1]$ suficientemente pequeño para que $1/\sqrt{h} > K$. Así, ninguna constante finita satisface la desigualdad de Definición 9.3 en todos los pares de $[0,1]$. La función es continua en el cerrado, pero **no es Lipschitz allí**.

Una frontera que importa
En cualquier subintervalo $[\varepsilon, 1]$ con $0 < \varepsilon < 1$, sí se tiene $|g'(t)| \le 1/(2\sqrt{\varepsilon})$ para los puntos interiores, de modo que g es Lipschitz en ese subintervalo. Una familia de constantes que depende de ε y crece sin cota cuando $\varepsilon \to 0^+$ no proporciona una única constante válida para todo $[0, 1]$. Por el contrario, Ejemplo 9.14 muestra que una función puede ser Lipschitz aunque no sea diferenciable en un punto interior.

Laboratorio: reconstruir una estimación sin resolver la ecuación de Lagrange. Sean $f\colon [a, b] \to \mathbb{R}$ continua en $[a, b]$ y diferenciable en (a, b), con $a < b$. Reconstruye la prueba de Proposición 9.1 eligiendo primero $x < y$, justificando todas las hipótesis sobre la restricción $f|_{[x,y]}$, señalando en qué momento aparece $c \in (x, y)$ y explicando por qué es legítimo multiplicar las desigualdades por $y - x$. Repite la deducción cuando solo conoces $|f'(t)| \le K$; comprueba por separado $x < y$, $x > y$ y $x = y$. Para detectar una reparación necesaria, intenta aplicar la estimación con $M = 1$ a Ejemplo 9.13: $f'(t) = 1$ para todo punto interior, pero $f(1) - f(0) = 2 > 1$. Identifica exactamente la hipótesis omitida.

Ya disponemos de un procedimiento para controlar incrementos y diferencias sin conocer el punto de valor medio. La siguiente sección examinará una consecuencia estructural especial: qué ocurre cuando **dos funciones tienen la misma derivada** en todo un intervalo y cómo reconstruir correctamente las demostraciones correspondientes.

9.10. Derivadas iguales y laboratorio de demostraciones

En §9.9 convertimos una cota sobre la derivada en una cota sobre los incrementos. Si la cota superior e inferior coinciden en cero, el margen de variación desaparece. Esto permite pasar de una afirmación **puntual** —la derivada vale cero en cada punto interior— a otra **global** —todos los valores de la función son iguales—. El puente será nuevamente el teorema del valor medio; no necesitaremos anticipar el estudio sistemático de la monotonía.

Una derivada idénticamente nula.

Proposición 9.2. **Proposición (derivada nula y constancia).** Sean $a, b \in \mathbb{R}$ con $a < b$ y sea $f\colon [a, b] \to \mathbb{R}$ continua en $[a, b]$ y diferenciable en todos los puntos de (a, b). Si

$$f'(t) = 0 \qquad \text{para todo } t \in (a, b),$$

entonces f es constante en $[a, b]$: existe $k \in \mathbb{R}$ tal que $f(x) = k$ para todo $x \in [a, b]$.

Demostración. Fijemos un punto $x_0 \in [a, b]$ y pongamos $k := f(x_0)$. Para un $x \in [a, b]$ arbitrario, si $x = x_0$, ya tenemos $f(x) = k$. Si $x \neq x_0$, definamos

$$u := \min\{x, x_0\}, \qquad v := \max\{x, x_0\}.$$

Entonces $a \le u < v \le b$. La restricción $f|_{[u,v]}$ es continua en $[u, v]$ y diferenciable en (u, v). Por el teorema del valor medio (Teorema 9.3), existe $c \in (u, v) \subseteq (a, b)$ con

$$\frac{f(v) - f(u)}{v - u} = f'(c) = 0.$$

Puesto que $v - u > 0$, concluimos $f(v) = f(u)$. El par u, v está formado exactamente por x y x_0; por tanto, $f(x) = f(x_0) = k$. Como x fue arbitrario, f es constante en todo $[a, b]$. □

También puede recuperarse la conclusión de Proposición 9.1 tomando $m = M = 0$, pues sus dos desigualdades obligan a $f(y) - f(x) = 0$ para $x < y$. La prueba anterior expone directamente **dónde** aparece el punto de Lagrange y por qué podemos comparar cualquier par de argumentos. No hemos requerido derivadas en los bordes ni continuidad de la función derivada.

La implicación recíproca, con el mismo dominio, es inmediata por la definición: si $f(x) = k$ para todos los argumentos, cada cociente incremental interior es $(k - k)/h = 0$. Así, bajo las hipótesis de la proposición, la constancia equivale a que la derivada sea idénticamente nula **en todo el abierto**, no simplemente en uno o varios puntos.

Dos funciones con la misma derivada. La igualdad de derivadas no obliga a la igualdad de los valores de las funciones. Lo que queda sin detectar al derivar es una constante aditiva. Para precisar esa afirmación necesitamos comprobar que las dos funciones se definen sobre el **mismo intervalo** y que podemos derivar su diferencia.

Proposición 9.3. **Proposición (igualdad de derivadas y diferencia constante).** Sean $a, b \in \mathbb{R}$, $a < b$, y dos funciones $f, g\colon [a, b] \to \mathbb{R}$, ambas continuas en $[a, b]$ y diferenciables en (a, b). Si

$$f'(t) = g'(t) \qquad \text{para todo } t \in (a, b),$$

entonces existe una constante $C \in \mathbb{R}$ tal que

$$\boxed{f(x) = g(x) + C \qquad \text{para todo } x \in [a, b].}$$

Más precisamente, fijado cualquier $x_0 \in [a, b]$, la única constante posible es $C = f(x_0) - g(x_0)$.

Demostración. Definamos la diferencia sobre el dominio común,

$$h\colon [a, b] \to \mathbb{R}, \qquad h(x) := f(x) - g(x).$$

Por las reglas de continuidad y la linealidad de la derivada (Teorema 7.1), h es continua en $[a, b]$, diferenciable en (a, b) y satisface

$$h'(t) = f'(t) - g'(t) = 0 \qquad (t \in (a, b)).$$

La proposición Proposición 9.2 se aplica a h y proporciona una constante C con $h(x) = C$ para todo $x \in [a, b]$; esto equivale a $f(x) = g(x) + C$. Evaluar en x_0 da forzosamente $C = f(x_0) - g(x_0)$, lo que demuestra además su unicidad. □

Por consiguiente, si además de $f' = g'$ sabemos que **coinciden en un punto** x_0, entonces $C = 0$ y $f = g$ sobre todo el intervalo. El dato de un solo valor elimina la ambigüedad aditiva; no se necesita conocer ambos valores en todos los argumentos. Recíprocamente, sumar una constante no cambia las derivadas interiores: si $f = g + C$, la linealidad proporciona $f' = g'$. Esta afirmación se refiere a las funciones bajo sus hipótesis de diferenciabilidad, y no construye derivadas en puntos donde faltan.

Ejemplo 9.17. **Una derivada compartida y el papel de un valor inicial.** Consideremos, sobre $[-2, 2]$,

$$f(x) = x^3 + 7, \qquad g(x) = x^3 - 2.$$

Son funciones polinómicas, continuas en el cerrado y diferenciables en el abierto, y

$$f'(t) = 3t^2 = g'(t) \qquad (-2 < t < 2).$$

Sin embargo, $f(0) = 7 \neq -2 = g(0)$. La diferencia vale

$$f(x) - g(x) = (x^3 + 7) - (x^3 - 2) = 9$$

para todo $x \in [-2, 2]$: la constante de Proposición 9.3 es $C = f(0) - g(0) = 9$. Si, en cambio, se prescribe para otra función q la misma derivada $q'(t) = 3t^2$ y el valor $q(0) = 7$, la proposición fuerza $q(x) = x^3 + 7$ en todo el intervalo, siempre que q satisfaga las hipótesis de continuidad y diferenciabilidad exigidas.

Por qué el intervalo y la continuidad no son detalles. En ambas proposiciones, lo decisivo es que para **cada par** $x < y$ del dominio podemos utilizar la restricción al segmento completo $[x, y]$. En realidad, el mismo razonamiento sirve sobre cualquier intervalo real I que contenga al menos dos puntos: si la función es continua relativamente a I y diferenciable en los puntos interiores de I, para $x < y$ pertenecientes a I se aplican las pruebas sobre $[x, y] \subseteq I$. El punto de valor medio está en (x, y) y, por ello, es interior de I. Esta observación abarca intervalos abiertos, cerrados, semiabiertos o no acotados, sin afirmar que deban existir derivadas en sus bordes. No podemos sustituir *intervalo* por *conjunto arbitrario* manteniendo la misma conclusión.

Ejemplo 9.18. **Una derivada nula en cada componente no determina una única constante.** Sea $A := [-2, -1] \cup [1, 2]$ y definamos $u\colon A \to \mathbb{R}$ mediante

$$u(x) := \begin{cases} 0, & -2 \leq x \leq -1, \\ 1, & 1 \leq x \leq 2. \end{cases}$$

La función es continua relativamente a A: cada punto admite un entorno relativo contenido en una de las dos partes, donde u es constante. Además, en todo punto interior de A, es diferenciable y tiene derivada cero; lo mismo ocurre con las derivadas relativas en los cuatro bordes. Pese a ello, $u(-1) = 0 \neq 1 = u(1)$. La razón es verificable: $[-1, 1]$ **no** está contenido en A, por lo que no podemos aplicar Lagrange a una restricción sobre ese segmento. Tenemos constancia en cada componente, pero no una sola constante válida para todo el dominio.

Para la versión de las derivadas iguales, tomemos $p(x) = x^2$ y $q(x) = x^2 + u(x)$ en A. Ambas funciones son continuas relativamente a A, sus derivadas coinciden en los puntos interiores, pero $q - p = u$ vale 0 en una componente y 1 en la otra. Tampoco existe una única constante global que exprese esa diferencia.

Una segunda frontera ya está disponible en Ejemplo 9.13. Allí $f(x) = x$ para $0 \leq x < 1$ y $f(1) = 2$. La función $h(x) := f(x) - x$ cumple $h'(t) = 0$ en todo $(0, 1)$, pero $h(x) = 0$ para $x < 1$ y $h(1) = 1$. **La continuidad relativa en el borde derecho falla**, de modo que Proposición 9.2 no se aplica. Del mismo modo, la igualdad de derivadas en *un solo punto* no controla todo un intervalo: $p(x) = x^2$ y $q(x) = 0$ sobre $[-1, 1]$ cumplen $p'(0) = q'(0)$, pero $p - q = x^2$ no es constante.

Tres límites lógicos de estas conclusiones
«La derivada vale cero en algún punto» no equivale a «vale cero en todos los puntos interiores». «Las derivadas coinciden en un punto» no equivale a «coinciden en todo el abierto». Y «la derivada es nula en todas las partes de un dominio desconectado» no obliga a una sola constante entre partes: el teorema del valor medio exige disponer del segmento completo entre los argumentos comparados.

Laboratorio de demostraciones: verificar, reconstruir y reparar. Trabaja con un intervalo $[a, b]$ no degenerado antes de generalizar. **Estación A:** en Proposición 9.2, fija x_0, elige x arbitrario, construye $[u, v]$ y marca el paso que utiliza cada una de las dos hipótesis analíticas. Explica por qué basta la existencia de algún c sin calcularlo. **Estación B:** vuelve a demostrar la constancia desde Proposición 9.1 con $m = M = 0$ y comprueba que ambas rutas tienen las mismas hipótesis. **Estación C:** en Proposición 9.3, escribe el dominio de $h = f - g$, justifica por separado su continuidad y su diferenciabilidad, y localiza el uso de la linealidad. **Estación D:** identifica qué aporta conocer $f(x_0) = g(x_0)$ y por qué una igualdad de derivadas aislada no lo reemplaza.

Estación E: intenta aplicar la primera proposición a $h = f - \mathrm{id}$ construida desde Ejemplo 9.13; señala la hipótesis exacta que falta y verifica que la conclusión es falsa. **Estación F:** intenta comparar -1 y 1 en Ejemplo 9.18 y escribe la inclusión de conjuntos que necesitarías para usar Lagrange; comprueba que esa inclusión es falsa. **Estación G:** formula con cuantificadores las dos proposiciones y señala el cambio ilícito cuando se reemplaza «para todo t interior» por «existe t interior». **Estación H:** reconstruye la extensión a un intervalo general y explica por qué la continuidad relativa en los extremos de cada subintervalo cerrado se conserva.

Los ejercicios reúnen la existencia de extremos, los argumentos de Fermat, Rolle y Lagrange, y sus consecuencias para estimaciones y constancia. Para aplicar cada resultado habrá que verificar las hipótesis en el intervalo concreto de comparación.

9.11. Ejercicios y soluciones

La cuestión rectora de este banco no es solamente encontrar un número c, sino decidir **qué teorema permite afirmar que existe**, verificar cada hipótesis en el dominio indicado y distinguir una garantía universal de una conclusión accidentalmente verdadera. Los cuarenta ejercicios, originales para este libro, forman siete niveles: $6A+6B+7C+7D+6E+5F+3G = 40$. Sus soluciones razonadas siguen a todos los enunciados y conservan la correspondencia estricta `CPM-T1-EXR-0356--0395` / `CPM-T1-SOL-0356--0395`. Se admiten solamente los resultados anteriores y los establecidos en §§9.1–9.10; no se utilizarán integración, Taylor, convexidad ni criterios sistemáticos de monotonía del capítulo siguiente.

Nivel A — Extremos y Fermat. Ejercicio A1. Localización antes de derivación. Para $f : [0, 3] \to \mathbb{R}$, $f(x) = (x - 1)^2$, encuentra sus valores máximo y mínimo absolutos y todos sus puntos de realización. Calcula la derivada en el punto extremo interior y las derivadas relativas en los bordes. Explica en cuáles de ellos puede invocarse Fermat.

Ejercicio A2. El signo de un cociente. Sean $A \subseteq \mathbb{R}$, c interior de A y $f : A \to \mathbb{R}$ diferenciable en c. Supón que c es un mínimo local relativo. Demuestra directamente, separando incrementos positivos y negativos, que $f'_A(c) = 0$, sin invocar Fermat como caja negra.

Ejercicio A3. Un cero no es una clasificación. Sea $p : [-1, 1] \to \mathbb{R}$, $p(x) = x^3 + 2$. Comprueba que $p'(0) = 0$ y demuestra mediante valores arbitrariamente próximos que 0 no es un extremo local. Formula la implicación de Fermat que este ejemplo impide invertir.

Ejercicio A4. Mínimo angular desplazado. Estudia $q : [1, 3] \to \mathbb{R}$, $q(x) = |x - 2|$. Determina el mínimo absoluto y prueba por cocientes laterales que $q'(2)$ no existe. ¿Contradice esto el teorema de Fermat?

Ejercicio A5. Una parábola con parámetros. Dados $a < b$, sea $r : [a, b] \to \mathbb{R}$, $r(x) = (x - a)(x - b)$. Prueba mediante una identidad cuadrática dónde alcanza su mínimo y su máximo absolutos. Verifica la condición de Fermat en el punto extremo interior; no presupongas que un extremo debe ser único.

Ejercicio A6. Derivadas no nulas en los bordes. Para $s : [0, 2] \to \mathbb{R}$, $s(x) = 3x + 1$, identifica los extremos absolutos y calcula las dos derivadas relativas. Explica por qué no se deduce de Fermat una derivada nula en ningún borde.

Nivel B — Rolle. Ejercicio B1. Verificar y localizar. Comprueba todas las hipótesis de Rolle para $f : [-2, 2] \to \mathbb{R}$, $f(x) = x^2 - 4$, y encuentra exactamente los puntos interiores donde se anula su derivada.

Ejercicio B2. Más de un punto de Rolle. Para $p : [0, 3] \to \mathbb{R}$, $p(x) = x(x - 1)(x - 3)$, comprueba los valores de borde y determina todos los puntos de Rolle. Justifica que cada solución está estrictamente entre 0 y 3.

Ejercicio B3. Construcción simétrica. Para $a < b$, aplica Rolle a $f : [a, b] \to \mathbb{R}$, $f(x) = (x - a)^2(x - b)^2$. Demuestra que el punto medio es el único cero de f' en (a, b) e identifica qué otros ceros de la derivada quedan fuera del intervalo abierto.

Ejercicio B4. Denominador controlado. Sea $r : [1, 3] \to \mathbb{R}$, $r(x) = \frac{(x-1)(x-3)}{x+1}$. Justifica primero la continuidad y la diferenciabilidad requeridas; aplica Rolle y localiza el punto interior con derivada cero.

Ejercicio B5. Derivadas en los bordes innecesarias. Sea $u : [-1, 1] \to \mathbb{R}$, $u(x) = \sqrt{1 - x^2}$. Demuestra que Rolle es aplicable, encuentra el punto prometido y examina por cocientes por qué no se requieren derivadas reales finitas en -1 ni en 1.

Ejercicio B6. Tres ceros interiores. Para $v : [-2, 2] \to \mathbb{R}$, $v(x) = x^4 - 5x^2 + 4$, verifica Rolle y determina cuántos puntos interiores satisfacen $v'(c) = 0$. Comprueba la pertenencia de todos ellos al intervalo abierto.

Nivel C — Auditoría de hipótesis y contraejemplos. Ejercicio C1. Rolle sin continuidad de borde. Define $f : [0, 1] \to \mathbb{R}$ por $f(x) = x^2$ si $0 \le x < 1$ y $f(1) = 0$. Para cada hipótesis de Rolle, escribe «se cumple» o «no se cumple», justificándolo. Decide mediante el estudio completo de f' si la conclusión es verdadera.

Ejercicio C2. Rolle y una esquina. Para $g : [-2, 2] \to \mathbb{R}$, $g(x) = |x|$, audita las tres hipótesis de Rolle y determina si existe un punto interior donde la derivada exista y valga cero.

Ejercicio C3. Valores de borde distintos. Para $h : [0, 1] \to \mathbb{R}$, $h(x) = x^2 + x$, comprueba continuidad y diferenciabilidad, calcula los valores extremos de borde y prueba que ningún punto interior satisface la conclusión de Rolle.

Ejercicio C4. La secante detecta un salto. Sea $f : [0, 2] \to \mathbb{R}$ dada por $f(x) = x$ para $0 \le x < 2$, $f(2) = 5$. Audita las dos hipótesis del teorema del valor medio. Calcula su secante y todas sus derivadas interiores; decide si la conclusión falla efectivamente.

Ejercicio C5. Una secante entre dos pendientes laterales. Para $g : [-1, 2] \to \mathbb{R}$, $g(x) = |x|$, encuentra la pendiente de secante, las derivadas existentes y el punto de no diferenciabilidad. Demuestra que no hay un punto de valor medio, sin confundir derivada inexistente con derivada nula.

Ejercicio C6. Falta una hipótesis y sí hay testigos. Sea $k : [-1, 1] \to \mathbb{R}$, $k(x) = |x| - x^2$. Comprueba que el teorema de Lagrange no puede invocarse en todo el intervalo. Sin embargo, calcula **todos** los puntos interiores donde su derivada existente coincide con la pendiente secante. Explica la distinción lógica.

Ejercicio C7. Un salto que no destruye la coincidencia. Sea $q : [0, 1] \to \mathbb{R}$, $q(x) = x^2$ para $0 \leq x < 1$ y $q(1) = \frac{1}{2}$. Comprueba cuál hipótesis de Lagrange falta, calcula la pendiente de secante y encuentra un punto interior cuya derivada coincide con ella. ¿Qué conclusión general queda excluida por el contraejemplo C4?

Nivel D — El valor medio y su interpretación exacta. Ejercicio D1. Secante de una cuadrática trasladada. Para $f : [-1, 3] \to \mathbb{R}$, $f(x) = x^2 + 3x + 1$, verifica Lagrange y encuentra el conjunto completo de puntos que satisfacen su conclusión.

Ejercicio D2. Dos puntos en un intervalo simétrico. Para $p : [-2, 2] \to \mathbb{R}$, $p(x) = x^3$, encuentra todos los puntos de valor medio y demuestra por qué no es admisible el punto medio del intervalo.

Ejercicio D3. Solo una raíz admisible. Para $r : [1, 3] \to \mathbb{R}$, $r(x) = x^3$, halla la pendiente de secante y resuelve la ecuación de Lagrange sobre los reales. Conserva exclusivamente las soluciones del intervalo abierto y justifica la exclusión de las restantes.

Ejercicio D4. Raíz cuadrada y derivada de borde. Sea $s : [0, 4] \to \mathbb{R}$, $s(x) = \sqrt{x}$. Verifica las hipótesis del valor medio, determina su punto interior y explica por qué el cociente incremental en 0 no es un obstáculo.

Ejercicio D5. Una raíz cúbica bien localizada. Para $t : [-1, 2] \to \mathbb{R}$, $t(x) = x^4$, determina exactamente los puntos interiores de valor medio, usando desigualdades para justificar el lugar del punto obtenido.

Ejercicio D6. Paralela no significa coincidente. Para $f : [0, 2] \to \mathbb{R}$, $f(x) = x^2$, calcula la recta secante entre los extremos y la recta tangente en el punto prometido por Lagrange. Comprueba si son la misma recta y explica qué igualdad adicional exigiría su coincidencia.

Ejercicio D7. ¿Cuándo hay infinitos puntos? Sea $P : [a, b] \to \mathbb{R}$, $P(x) = Ax^2 + Bx + C$, con $a < b$ y parámetros reales A, B, C. Determina el conjunto completo de puntos de valor medio cuando $A \neq 0$ y cuando $A = 0$. Identifica cuál de las hipótesis de Lagrange utiliza la afirmación de existencia.

Nivel E — Estimaciones y Lipschitz. Ejercicio E1. Cotas sin calcular la función. Sea $f : [-1, 3] \to \mathbb{R}$ continua en el cerrado y diferenciable en el abierto, con $2 \leq f'(t) \leq 5$ para todo $t \in (-1, 3)$. Da las mejores cotas universales deducibles para $f(3) - f(-1)$ y exhibe una función admisible que alcance cada extremo de las cotas.

Ejercicio E2. Un valor inicial limita todos los demás. Sea $f : [-2, 2] \to \mathbb{R}$ continua, diferenciable en el abierto, con $|f'(t)| \leq 3$ allí y $f(0) = 4$. Prueba una cota dependiente de x para $|f(x) - 4|$ y deduce una cota uniforme de todos los valores $f(x)$.

Ejercicio E3. Constante óptima de una cúbica. Prueba que $x \mapsto x^3$ es Lipschitz en $[-2, 2]$ con constante 12; demuestra que ninguna constante menor sirve, sin presuponer que la derivada alcanza su cota en un punto interior.

Ejercicio E4. Una función recíproca. Para $f : [1, 4] \to \mathbb{R}$, $f(x) = 1/x$, halla cotas bilaterales para $f(y) - f(x)$ cuando $1 \leq x < y \leq 4$, deduce una constante Lipschitz y prueba que es óptima.

Ejercicio E5. Reparar una cota mediante restricción. Demuestra que $\sqrt{x}$ no es Lipschitz en $[0, 1]$, pero sí lo es en $[\frac{1}{4}, 1]$. Determina la constante óptima en este segundo intervalo.

Ejercicio E6. Separar tres propiedades. Estudia $u : [0, 2] \to \mathbb{R}$, $u(x) = |x - 1|$: demuestra directamente que es Lipschitz con constante óptima 1 y que no es diferenciable en 1. Usa después una función continua del capítulo para refutar «continua implica Lipschitz».

Nivel F — Lectura, reconstrucción y reparación de pruebas. Ejercicio F1. Un signo omitido. Al probar Fermat para un máximo local, alguien escribe «$f(c + h) - f(c) \leq 0$; luego $(f(c + h) - f(c))/h \leq 0$ para ambos signos de h; por tanto $f'(c) \leq 0$». Señala el paso inválido y reescribe una prueba completa para un máximo interior diferenciable.

Ejercicio F2. El extremo que falta. Una prueba de Rolle dice «la función continua alcanza su máximo; como los bordes coinciden, el máximo es interior y Fermat resuelve». Identifica el error, cubre el caso constante y reconstruye el argumento cuando la función no es constante.

Ejercicio F3. Una función auxiliar mal normalizada. Se toma $m = (f(b) - f(a))/(b - a)$ y $g(x) = f(x) - m(x - a)$; se afirma $g(a) = g(b) = 0$ y se aplica Rolle. Determina los valores reales de borde, repara la igualdad falsa y termina correctamente la demostración de Lagrange.

Ejercicio F4. Un punto no es todo el intervalo. Examina la frase «si una función continua en $[-1, 1]$ tiene derivada en 0, entonces Rolle se aplica cuando los bordes coinciden». Usa $f(x) = |x - \frac{1}{2}|$ modificada, si es necesario, por un término afín, para producir bordes iguales, derivabilidad en 0, un punto interior no diferenciable y ausencia de derivada nula donde exista. Explica por qué la cuantificación correcta es universal en el intervalo abierto.

Ejercicio F5. La continuidad de borde es una hipótesis. Sea $h : [0, 1] \to \mathbb{R}$ con $h(x) = 0$ para $x < 1$ y $h(1) = 1$. Verifica que $h' = 0$ en $(0, 1)$, refuta la conclusión de constancia y repara el enunciado de «derivada cero implica constante» para que sea válido.

Nivel G — Desafíos de síntesis. Ejercicio G1. La geometría del dominio. Sea $I \subseteq \mathbb{R}$ un intervalo con al menos dos puntos y $f : I \to \mathbb{R}$ continua relativamente a I y diferenciable en su interior. Supón $|f'(t)| \leq K$ allí para algún $K \geq 0$. Prueba que f es K-Lipschitz sobre todo I, aunque sea abierto o no acotado. Construye luego un dominio desconectado y una función de derivada nula en todos sus puntos interiores que no sea 0-Lipschitz globalmente.

Ejercicio G2. Tres puntos de valor medio y dos rectas tangentes distintas. Para $f : [-1, 1] \to \mathbb{R}$, $f(x) = x^4 - x^2$, determina exhaustivamente los puntos interiores de Lagrange y las rectas tangentes en ellos. ¿Cuáles coinciden con la secante de extremos y cuáles solo son paralelas? Justifica que no faltan otros puntos.

Ejercicio G3. La constante óptima leída desde la derivada. Sea $f : [a, b] \to \mathbb{R}$ continua en el cerrado y diferenciable en (a, b), con $a < b$. Prueba que si es K-Lipschitz entonces $|f'(t)| \leq K$ en cada punto interior. Deduce que f es Lipschitz si y solo si su derivada interior está acotada; cuando lo está, identifica la menor constante admisible como $\sup_{t\in(a,b)} |f'(t)|$. Explica por qué no se necesita continuidad de f' y por qué $|x|$ sigue siendo Lipschitz aunque no sea diferenciable en cero.

Soluciones desarrolladas.

Soluciones del nivel A. **Solución A1.** La identidad $(x - 1)^2 \geq 0$ muestra que el mínimo es 0, alcanzado únicamente en $x = 1 \in (0, 3)$. Como $-1 \leq x - 1 \leq 2$, tenemos $(x - 1)^2 \leq 4$, con igualdad exclusivamente en $x = 3$; el máximo es 4 en ese borde. La derivada interior es $f'(x) = 2(x - 1)$, de modo que $f'(1) = 0$, como exige Teorema 9.1 en el mínimo interior. Por cocientes relativos, las derivadas de borde son $f'(0) = -2$ y $f'(3) = 4$, no nulas. Fermat no se aplica en esos puntos porque ninguno pertenece al interior del dominio. Tampoco debemos suponer que el máximo de borde debe proporcionar una derivada cero.

Solución A2. Por la interioridad de c existe $\rho_0 > 0$ con $(c - \rho_0, c + \rho_0) \subseteq A$. El mínimo local proporciona $\rho_1 > 0$ tal que $f(c + h) - f(c) \geq 0$ para incrementos admisibles con $|h| < \rho_1$. Con $\rho = \min\{\rho_0, \rho_1\}$, para $0 < h < \rho$ el cociente $Q(h) = [f(c + h) - f(c)]/h$ es **no negativo**; para $-\rho < h < 0$ es **no positivo**, pues dividimos por un número negativo. La diferenciabilidad afirma que ambos cocientes laterales convergen al mismo real $L = f'_A(c)$. Por preservación del orden en los límites, el primero da $L \geq 0$ y el segundo $L \leq 0$. Por antisimetría, $L = 0$. No usamos continuidad en todo un intervalo, sino el carácter interior, el extremo local y la existencia de la derivada en c.

Solución A3. $p'(x) = 3x^2$ en $(-1, 1)$, así que $p'(0) = 0$. Dado cualquier $\delta > 0$, elegimos t con $0 < t < \min\{1, \delta\}$. Entonces $p(-t) = 2 - t^3 < 2 = p(0) < 2 + t^3 = p(t)$. Todo entorno relativo del origen contiene un valor inferior y otro superior; por Definición 9.2, no hay extremo local en 0. Fermat establece «extremo local interior y derivabilidad ⇒ derivada cero». El ejemplo invalida la **recíproca** «derivada cero ⇒ extremo local»; no invalida el teorema en su dirección demostrada.

Solución A4. Para todo $x \in [1, 3]$, $q(x) = |x - 2| \geq 0 = q(2)$; la igualdad exige $x = 2$, por lo que es el único punto de mínimo absoluto, interior y estricto. Para h no nulo suficientemente pequeño, $[q(2 + h) - q(2)]/h = |h|/h$: vale -1 cuando $h < 0$ y 1 cuando $h > 0$. Los límites laterales difieren; $q'(2)$ no existe. La conclusión de Fermat no puede reclamarse porque **falta diferenciabilidad en el punto extremo**. Encontrar un extremo no demuestra por sí solo que allí se anule una derivada.

Solución A5. Sea $m = (a + b)/2$. Completando el cuadrado, $r(x) = (x - m)^2 - \frac{(b-a)^2}{4}$.

El primer término es no negativo, y $m \in (a, b)$, de modo que el mínimo absoluto es $-(b-a)^2/4$ y se alcanza únicamente en m. Por otra parte, $x - a \geq 0$ y $x - b \leq 0$ en $[a, b]$, luego $r(x) \leq 0$, con igualdad exactamente en a y b: ambos bordes realizan el máximo absoluto 0. Las reglas de derivación dan $r'(x) = 2x - a - b$ y, en consecuencia, $r'(m) = 0$, coherentemente con Fermat. El valor máximo tiene **dos** puntos de realización, aunque el mínimo tenga uno.

Solución A6. En $[0,2]$ se cumple $1 \le 3x+1 \le 7$, con igualdad inferior solo en 0 y superior solo en 2. Así, el mínimo absoluto vale 1 y el máximo 7, ambos en bordes. Para cualquier incremento admisible no nulo, $[s(c+h)-s(c)]/h = 3$; las derivadas relativas en 0 y 2 son, pues, ambas 3. La hipótesis «c interior» de Teorema 9.1 es falsa en los dos puntos extremos. Allí únicamente existen incrementos admisibles de un signo y el argumento bilateral de Fermat no obliga a obtener cero.

Soluciones del nivel B. **Solución B1.** f es polinómica, por lo que es continua en $[-2,2]$ y diferenciable en $(-2,2)$. Además, $f(-2) = 0 = f(2)$: están verificadas las tres hipótesis de Teorema 9.2. La derivada es $f'(x) = 2x$, cuyo único cero real es $x = 0$. Como $-2 < 0 < 2$, ese es el único punto admisible. Rolle asegura al menos uno; el cálculo adicional permite demostrar aquí la unicidad, que no forma parte de su enunciado general.

Solución B2. Los factores muestran $p(0) = p(3) = 0$. El polinomio es continuo en $[0,3]$ y diferenciable en $(0,3)$, así que Rolle se aplica. Al expandir, $p(x) = x^3 - 4x^2 + 3x$, y

$$p'(x) = 3x^2 - 8x + 3 = 0 \quad \Longleftrightarrow \quad x = \frac{4 \pm \sqrt{7}}{3}.$$

Puesto que $2 < \sqrt{7} < 3$, se obtiene $1/3 < (4-\sqrt{7})/3 < 2/3$ y $2 < (4+\sqrt{7})/3 < 7/3 < 3$. Ambos pertenecen al intervalo abierto y la ecuación cuadrática no tiene otras raíces. Por tanto, hay exactamente **dos** puntos de Rolle: la existencia no implica unicidad.

Solución B3. Los dos factores cuadrados definen un polinomio, continuo y diferenciable sobre los dominios requeridos, y $f(a) = f(b) = 0$. Por Rolle hay un cero interior de f'. Derivando el producto,

$$f'(x) = 2(x-a)(x-b)\big((x-b)+(x-a)\big) = 2(x-a)(x-b)(2x-a-b).$$

Sus ceros reales son a, b y $m = (a+b)/2$. Solo m cumple $a < m < b$, por lo que es el único punto de Rolle. Los otros dos ceros pertenecen al dominio, pero quedan excluidos del conjunto de puntos **interiores** exigidos por el teorema.

Solución B4. En $[1,3]$ se cumple $x + 1 \ge 2 > 0$: el cociente de polinomios es continuo en el cerrado y diferenciable en el abierto. El numerador se anula en ambos bordes, luego $r(1) = r(3) = 0$ y Rolle garantiza al menos un punto. Usando la regla del cociente,

$$r'(x) = \frac{(2x-4)(x+1) - (x^2-4x+3)}{(x+1)^2} = \frac{x^2+2x-7}{(x+1)^2}.$$

El denominador no se anula; las raíces del numerador son $-1 \pm 2\sqrt{2}$. La negativa queda fuera de $[1,3]$; para la positiva, $1 < \sqrt{2} < 2$ implica $1 < -1 + 2\sqrt{2} < 3$. Es el único punto interior con derivada cero.

Solución B5. Como $1 - x^2 \ge 0$ en $[-1,1]$, la raíz no negativa define una función continua en todo el dominio. En $(-1,1)$ la expresión bajo la raíz es positiva, de modo que es diferenciable y $u'(x) = -x/\sqrt{1-x^2}$. Los valores de borde son $u(-1) = u(1) = 0$; Rolle se aplica y la única raíz interior de la derivada es $c = 0$. Para precisar los bordes, si $h > 0$,

$$\frac{u(-1+h) - u(-1)}{h} = \frac{\sqrt{2h - h^2}}{h} = \sqrt{\frac{2-h}{h}} \longrightarrow +\infty.$$

En el borde derecho, con $h < 0$, $[u(1+h) - u(1)]/h = \sqrt{-2h - h^2}/h \to -\infty$. No existen derivadas relativas **finitas** en esos bordes; Rolle no las necesita y su conclusión se refiere solo a un punto del abierto.

Solución B6. La función es polinómica, luego continua en el cerrado y diferenciable en el abierto. En ambos extremos, $v(\pm 2) = 16 - 20 + 4 = 0$; las hipótesis de Rolle están completas. Factorizamos $v'(x) = 4x^3 - 10x = 2x(2x^2 - 5)$.

Los ceros son 0 y $\pm\sqrt{5/2}$. Como $0 < 5/2 < 4$, resulta $0 < \sqrt{5/2} < 2$, y los tres números pertenecen a $(-2, 2)$. No hay otros ceros porque hemos factorizado íntegramente la expresión cúbica. Rolle garantiza existencia, no el número **tres** que hemos establecido aquí por cálculo.

Soluciones del nivel C. **Solución C1.** Se cumplen los bordes iguales: $f(0) = 0 = f(1)$. En todo $x \in (0, 1)$, localmente $f(t) = t^2$, de donde $f'(x) = 2x > 0$; la diferenciabilidad interior también se cumple. Pero $\lim_{x\to 1^-} f(x) = 1 \neq f(1) = 0$, así que falta la continuidad relativa en el borde derecho. La conclusión de Rolle es **falsa**, no simplemente indemostrable con ese teorema: hemos comprobado que $f'(x)$ es estrictamente positiva en todos los puntos interiores. La falla de continuidad invalida la garantía general.

Solución C2. g es continua en $[-2, 2]$ por la desigualdad $||x| - |y|| \leq |x - y|$; además $g(-2) = 2 = g(2)$. En $(-2, 0)$ su derivada es -1 y en $(0, 2)$ es 1. En 0, $[g(h) - g(0)]/h = |h|/h$ tiene límites laterales -1 y 1: la diferenciabilidad en todo el abierto falla **solo allí**. No existe ningún punto interior con derivada real igual a cero: donde existe vale -1 o 1, y en el origen no existe. Por ello tampoco se cumple la conclusión de Rolle.

Solución C3. El polinomio es continuo en $[0, 1]$ y diferenciable en $(0, 1)$. No obstante, $h(0) = 0$ y $h(1) = 2$, luego falta únicamente la igualdad de los valores de borde. Para $x \in (0, 1)$, $h'(x) = 2x + 1 > 0$, por lo que nunca es cero. La conclusión de Rolle es efectivamente falsa. La función no contradice Lagrange: para este último la igualdad de bordes **no** es una hipótesis, y la pendiente secante vale 2, alcanzada por la derivada en $c = 1/2$.

Solución C4. Para todo $x \in (0, 2)$ hay un entorno contenido en $(0, 2)$ en el que $f(t) = t$; f es diferenciable en todos los puntos interiores, con $f'(x) = 1$. En el borde 2, $\lim_{x\to 2^-} f(x) = 2 \neq 5 = f(2)$, de modo que no hay continuidad en el cerrado. La pendiente de los extremos es $m = [f(2) - f(0)]/(2 - 0) = 5/2$. Como $f'(x) = 1 \neq 5/2$ para **cada** $x \in (0, 2)$, la conclusión del valor medio falla. La ausencia de continuidad se ha identificado independientemente de la comprobación de esa falsedad.

Solución C5. La continuidad de $g = |x|$ vale en todo $[-1, 2]$. Sus valores de borde son $g(-1) = 1$ y $g(2) = 2$, y su secante tiene pendiente $(2 - 1)/(2 - (-1)) = 1/3$. La derivada existe y vale -1 en $(-1, 0)$ y 1 en $(0, 2)$. En cero los cocientes incrementales $|h|/h$ tienden a -1 y 1 por los lados respectivos; $g'(0)$ no es un número real. Ninguna derivada **existente** del abierto vale $1/3$. Ha fallado solamente la hipótesis de diferenciabilidad interior y, de manera comprobada, también la conclusión del valor medio.

Solución C6. Los extremos satisfacen $k(-1) = 1 - 1 = 0 = k(1)$, por lo que la pendiente secante es 0. k es continua en $[-1, 1]$, pero en 0 sus cocientes laterales son -1 por la izquierda y 1 por la derecha: allí no es diferenciable, y Lagrange no es aplicable al intervalo completo. En cambio,

$$k'(x) = \begin{cases} -1 - 2x, & -1 < x < 0, \\ 1 - 2x, & 0 < x < 1. \end{cases}$$

La primera rama se anula exactamente en $c = -1/2$ y la segunda en $c = 1/2$; ambos son interiores. No hay otros puntos, porque cada expresión es afín y el origen carece de derivada. Conclusión: **faltarle una hipótesis al teorema no impide que alguna función particular satisfaga su conclusión**.

Solución C7. La función es diferenciable en cada $x \in (0, 1)$, donde $q'(x) = 2x$, pero $\lim_{x\to 1^-} q(x) = 1 \neq 1/2 = q(1)$; falta continuidad en el cerrado. La pendiente secante es $m = [q(1) - q(0)]/(1 - 0) = 1/2$. Resolver $2c = 1/2$ produce $c = 1/4 \in (0, 1)$, que satisface efectivamente la conclusión. C4 demuestra que **no se puede eliminar la continuidad de borde de una garantía universal de Lagrange**; este ejemplo demuestra que la falta de esa hipótesis tampoco obliga, para cada función, al fracaso de la coincidencia de pendientes. Son proposiciones lógicamente distintas.

Soluciones del nivel D. **Solución D1.** La cuadrática es continua en $[-1, 3]$ y diferenciable en $(-1, 3)$. Sus valores son $f(-1) = -1$ y $f(3) = 19$, de modo que la secante tiene pendiente $m = (19 - (-1))/4 = 5$. Como $f'(x) = 2x + 3$, la ecuación $f'(c) = 5$ equivale a $2c = 2$, y su única solución real es $c = 1$. Verificamos $-1 < 1 < 3$. Por tanto, el conjunto completo de testigos es $\{1\}$; la existencia se justifica por Lagrange y la unicidad por haber resuelto la ecuación diferencial concreta.

Solución D2. El polinomio satisface las dos hipótesis analíticas en el intervalo elegido. $p(2) = 8$ y $p(-2) = -8$, de donde $m = (8 + 8)/4 = 4$. La ecuación $p'(c) = 3c^2 = 4$ tiene dos soluciones reales, $c = \pm 2/\sqrt{3}$. Son admisibles porque $2/\sqrt{3} < 2$ y ambas tienen valor absoluto positivo. El punto medio es 0, pero $p'(0) = 0 \neq 4$: no sirve. Hemos encontrado exactamente dos puntos de Lagrange, lo que no contradice un teorema que afirma solamente que existe **al menos uno**.

Solución D3. La función cúbica es continua y diferenciable en los conjuntos exigidos. La pendiente secante es $m = (27 - 1)/(3 - 1) = 13$. Resolver $3c^2 = 13$ sobre $\mathbb{R}$ da $c = \pm\sqrt{13/3}$. Como $3 < 13 < 27$, se cumple $1 < \sqrt{13/3} < 3$: la raíz positiva pertenece a $(1, 3)$. La negativa es menor que cero y no pertenece al intervalo, aunque satisfaga la **ecuación numérica**. El conjunto de puntos de valor medio es exactamente $\{\sqrt{13/3}\}$; localizar el dominio es parte esencial de la resolución.

Solución D4. La raíz cuadrada no negativa es continua en $[0, 4]$ y diferenciable en todos los puntos de $(0, 4)$, donde $s'(x) = 1/(2\sqrt{x})$. La secante tiene pendiente $(s(4) - s(0))/4 = 2/4 = 1/2$. Igualar $1/(2\sqrt{c}) = 1/2$ produce $\sqrt{c} = 1$ y $c = 1 \in (0, 4)$. En el extremo izquierdo, para $h > 0$, $[s(h) - s(0)]/h = 1/\sqrt{h}$ no converge a un real finito. Esto **no** impide aplicar Lagrange: solo se requiere continuidad en los bordes, no diferenciabilidad allí.

Solución D5. t es polinómica, continua en $[-1, 2]$ y diferenciable en $(-1, 2)$. Se tiene $t(-1) = 1$, $t(2) = 16$ y $m = (16 - 1)/3 = 5$. La derivada es $t'(x) = 4x^3$, de modo que $4c^3 = 5$ equivale a $c = \sqrt[3]{5/4}$. Puesto que $1 < 5/4 < 8$ y la función cúbica conserva estrictamente el orden (por $v^3 - u^3 = (v - u)(u^2 + uv + v^2) > 0$ cuando $0 < u < v$), resulta $1 < c < 2$. La raíz cúbica real es única; por tanto, no existe otro punto admisible. No hemos recurrido a un criterio diferencial de monotonía.

Solución D6. La pendiente de la secante entre $(0,0)$ y $(2,4)$ es $m=2$, y su recta es $\ell(x)=2x$. Lagrange es aplicable porque la parábola es continua y diferenciable. Como $f'(x)=2x$, el único punto de valor medio es $c=1$. La recta tangente allí tiene pendiente 2 y pasa por $(1,f(1))=(1,1)$: $T_1(x)=1+2(x-1)=2x-1$.

Las rectas son paralelas, pero $T_1(0)=-1\neq 0=\ell(0)$; no coinciden. Coincidir requeriría, además de $f'(c)=m$, que **el punto de tangencia perteneciera a la secante**, es decir, $f(c)=\ell(c)$. Aquí $f(1)=1\neq 2=\ell(1)$.

Solución D7. Todo polinomio es continuo en el intervalo cerrado y diferenciable en el abierto, así que la existencia de un punto está garantizada por Teorema 9.3, con independencia de A. La pendiente secante se simplifica usando $b^2-a^2=(b-a)(a+b)$: $m=A(a+b)+B$.

La derivada es $P'(x)=2Ax+B$. Si $A\neq 0$, $2Ac+B=m$ equivale a $c=(a+b)/2$, el único valor real y además interior porque $a<b$. Si $A=0$, $P'(x)=B=m$ en **todos** los puntos interiores, de modo que el conjunto buscado es (a,b) completo. La hipótesis de existencia utiliza continuidad en $[a,b]$ y diferenciabilidad en (a,b); la clasificación del conjunto de testigos utiliza además la estructura particular del polinomio.

Soluciones del nivel E. **Solución E1.** El intervalo tiene longitud $3-(-1)=4$. Aplicando Proposición 9.1 a $x=-1$, $y=3$ y a las cotas dadas,

$$2\cdot 4\le f(3)-f(-1)\le 5\cdot 4,\qquad \text{es decir,}\qquad 8\le f(3)-f(-1)\le 20.$$

Ninguna de las cotas puede mejorarse para **todas** las funciones admitidas: la función afín $f_-(x)=2x$ es continua, diferenciable, tiene derivada idéntica a 2 y diferencia de bordes 8; análogamente $f_+(x)=5x$ da diferencia 20. Ambas satisfacen las hipótesis. No hemos supuesto una fórmula previa para la función f original.

Solución E2. Por la cota de derivada y Proposición 9.1, f es 3-Lipschitz en $[-2,2]$. Comparar x con el argumento conocido 0 da, para cualquier x del dominio,

$$|f(x)-4|=|f(x)-f(0)|\le 3|x-0|=3|x|.$$

Como $|x|\le 2$, se deduce $|f(x)-4|\le 6$, equivalente a $-2\le f(x)\le 10$ para todos los argumentos. La primera desigualdad es más informativa cerca de 0 que el intervalo uniforme. Los valores -2 y 10 son universalmente alcanzables en el dominio, por ejemplo con $f(x)=4+3x$ en los extremos -2 y 2; no se pueden estrechar ambas cotas bajo los datos dados.

Solución E3. La función es polinómica, continua en $[-2,2]$ y diferenciable en $(-2,2)$. Allí $f'(x)=3x^2$, por lo que $|f'(x)|\le 3\cdot 4=12$; Proposición 9.1 proporciona

$$|x^3-y^3|\le 12|x-y|\qquad (x,y\in[-2,2]).$$

Para probar optimalidad, tomemos $0<h<2$ y el par $x=2$, $y=2-h$. Su cociente de diferencias absolutas es

$$\frac{2^3-(2-h)^3}{h}=12-6h+h^2\longrightarrow 12\quad (h\to 0^+).$$

Si $K<12$ fuese otra constante Lipschitz, todos esos cocientes deberían ser $\le K$, y el límite conservaría esa desigualdad, contradiciendo $12>K$. El valor máximo de la **derivada en el abierto** no tiene por qué alcanzarse: basta aproximarse al borde para probar la constante óptima.

Solución E4. El denominador es positivo en $[1,4]$; la función es continua en el cerrado y diferenciable en el abierto, donde

$$f'(t) = -\frac{1}{t^2}, \qquad -1 \le f'(t) \le -\frac{1}{16}.$$

Por Proposición 9.1, para $1 \le x < y \le 4$, $-(y-x) \le f(y) - f(x) \le -\frac{y-x}{16}$. La cota $|f'| \le 1$ da $|f(y) - f(x)| \le |y - x|$, así que $K = 1$ sirve. Para ver que es mínima, comparemos 1 y $1 + h$ con $h > 0$ pequeño:

$$\frac{|f(1+h) - f(1)|}{h} = \frac{1 - 1/(1+h)}{h} = \frac{1}{1+h} \longrightarrow 1.$$

Cualquier constante global debe dominar esos cocientes y, por paso al límite, ser al menos 1.

Solución E5. En $[0,1]$, el par $0 < h \le 1$ y 0 produce el cociente

$$\frac{|\sqrt{h} - \sqrt{0}|}{|h - 0|} = \frac{1}{\sqrt{h}},$$

que no tiene cota superior finita cuando $h \to 0^+$. No existe constante Lipschitz global. En el intervalo restringido $[1/4, 1]$ la función es continua y diferenciable, con

$$0 < f'(x) = \frac{1}{2\sqrt{x}} \le 1 \qquad (1/4 < x < 1).$$

Por Proposición 9.1, $K = 1$ sirve allí. Para su optimalidad, tomando $h > 0$ suficientemente pequeño, racionalizamos:

$$\frac{\sqrt{1/4 + h} - 1/2}{h} = \frac{1}{\sqrt{1/4 + h} + 1/2} \longrightarrow 1.$$

Una constante menor que 1 contradiría el límite de esas pendientes secantes. La restricción elimina la singularidad de borde responsable del fracaso global.

Solución E6. La desigualdad triangular, aplicada a $x - 1$ e $y - 1$, y la cancelación de los términos constantes dan $\big||x-1| - |y-1|\big| \le |x - y|$.

La desigualdad vale para todo par en $[0,2]$, por lo que u es 1-Lipschitz. Si $x = 1$ e $y = 1 + h$ con $0 < h \le 1$, el cociente Lipschitz es exactamente $h/h = 1$; ninguna constante menor sirve. En el punto interior 1, $\frac{u(1+h)-u(1)}{h} = \frac{|h|}{h}$

Este cociente vale -1 para incrementos negativos y 1 para incrementos positivos, por lo que no hay derivada bilateral. Lipschitz **no implica** diferenciabilidad. Por otro lado, $x \mapsto \sqrt{x}$ es continua en $[0,1]$, pero Ejercicio 9.31 demuestra por cocientes sin cota que no es Lipschitz. La continuidad tampoco implica Lipschitz; este último sí implica continuidad, como se demostró en §9.9.

Soluciones del nivel F. **Solución F1.** El paso inválido consiste en dividir por h sin atender a su signo. Sea c interior y supongamos que $f(c+h)-f(c)\le 0$ para incrementos admisibles suficientemente pequeños. Cuando $h>0$ obtenemos $Q(h)=[f(c+h)-f(c)]/h\le 0$; cuando $h<0$, la desigualdad se **invierte**, $Q(h)\ge 0$. La interioridad permite incrementos de ambos signos, y la diferenciabilidad exige que $Q(h)$ tienda al mismo real $L=f'(c)$ desde la izquierda y la derecha. La conservación del orden bajo límites implica $L\le 0$ por el lado positivo y $L\ge 0$ por el negativo; luego $L=0$. Decir solo $L\le 0$ no aprovecha la información bilateral y, además, se había deducido de una división incorrecta.

Solución F2. La igualdad $f(a)=f(b)=k$ no fuerza que el **máximo** esté en el interior: por ejemplo, $f(x)=(x-a)(x-b)$ es no constante, vale cero en ambos bordes y es negativa dentro, de modo que su máximo se alcanza en los bordes. La reparación empieza con el teorema de valores extremos: hay valores $m\le k\le M$ realizados en $[a,b]$. Si $m=M$, f es constante y cualquier $c\in(a,b)$ tiene derivada cero. Si $m<M$, no pueden cumplirse simultáneamente $m=k=M$; por tanto, o bien $M>k$ o bien $m<k$. En el primer caso, cualquier punto que realice M debe ser interior; en el segundo, cualquier punto que realice m debe ser interior. Allí el extremo absoluto es local y la diferenciabilidad permite aplicar Fermat. En ambos casos existe el cero interior requerido por Rolle.

Solución F3. Con la definición propuesta, $g(a)=f(a)$, mientras que

$$g(b)=f(b)-m(b-a)=f(b)-[f(b)-f(a)]=f(a).$$

Los valores de borde son **iguales**, pero no necesariamente cero. Ese g ya sirve para Rolle si se verifican sus propiedades analíticas; lo falso es la afirmación de que ambos son nulos. Una normalización que vuelve verdadera esa afirmación es $\widetilde{g}(x)=f(x)-f(a)-m(x-a)$. Si f es continua en $[a,b]$ y diferenciable en (a,b), también lo es $\widetilde{g}$ por restarle una función afín, y $\widetilde{g}(a)=\widetilde{g}(b)=0$. Rolle proporciona $c\in(a,b)$ con $0=\widetilde{g}'(c)=f'(c)-m$, esto es, $f'(c)=[f(b)-f(a)]/(b-a)$. No se ha supuesto igualdad previa de $f(a)$ y $f(b)$.

Solución F4. La función que permite cumplir todas las restricciones del ejercicio es

$$f(x)=\left|x-\frac{1}{2}\right|+\frac{x}{2},\qquad -1\le x\le 1.$$

Es continua en todo el intervalo. Los valores de borde son $f(-1)=3/2-1/2=1$ y $f(1)=1/2+1/2=1$. Para $x<1/2$ se tiene $f(x)=1/2-x/2$, por lo que $f'(0)=-1/2$ existe. Para $x>1/2$, $f(x)=3x/2-1/2$, cuya derivada vale $3/2$. En el punto interior $x=1/2$ las derivadas laterales son $-1/2$ y $3/2$: la derivada no existe allí. En ningún otro punto interior la derivada vale cero. Rolle exige $\forall t\in(-1,1)$ diferenciabilidad; conocer $f'(0)$ solo verifica un punto. La conclusión es falsa en este ejemplo, lo que muestra que sustituir un cuantificador universal por un dato puntual es ilegítimo.

Solución F5. En cualquier punto interior $x\in(0,1)$ la función coincide localmente con la constante cero. Sus cocientes incrementales son todos cero, por lo que $h'(x)=0$ para **todo** el abierto. Sin embargo, $h(0)=0\neq 1=h(1)$; la función no es constante en su dominio. La razón es que $\lim_{x\to 1^-}h(x)=0\neq h(1)$, de modo que falla continuidad relativa en el extremo derecho. El enunciado correcto es: «si $a<b$ y $f:[a,b]\to\mathbb{R}$ es continua en $[a,b]$, diferenciable en (a,b) y $f'(t)=0$ para todo $t\in(a,b)$, entonces f es constante en **todo** $[a,b]$». La prueba compara cualquier par $x<y$ mediante Lagrange en $[x,y]$: la pendiente secante es $f'(c)=0$, y $f(x)=f(y)$.

Soluciones del nivel G. **Solución G1.** Sean $x, y \in I$. Si son iguales, la desigualdad Lipschitz es $0 \le 0$. Si $x < y$, la propiedad de intervalo garantiza $[x, y] \subseteq I$. La restricción de f es continua en $[x, y]$ porque f es continua relativamente a I, y es diferenciable en (x, y), cuyos puntos son interiores de I. Lagrange proporciona $c \in (x, y)$ tal que $\frac{f(y)-f(x)}{y-x} = f'(c)$.

Tomando valores absolutos y usando $y - x > 0$, $|f(y) - f(x)| \le K(y - x)$. Para el orden inverso intercambiamos los nombres de los puntos; así se prueba la propiedad en todo I. No hemos necesitado que el intervalo tenga extremos, sea acotado ni sea cerrado. Como contraejemplo al reemplazo por un dominio arbitrario, tomemos $A = [-2, -1] \cup [1, 2]$ y $u : A \to \mathbb{R}$ igual a 0 en la primera componente y a 1 en la segunda. Es continua relativamente a A y su derivada es cero en todo punto interior; sin embargo, $|u(1) - u(-1)| = 1 > 0 \cdot |1 - (-1)| = 0$. No es 0-Lipschitz globalmente. Entre -1 y 1 falta el segmento necesario para aplicar Lagrange; la conclusión **por componente** no autoriza una constante única entre componentes.

Solución G2. f es polinómica, continua en $[-1, 1]$ y diferenciable en $(-1, 1)$. Como $f(-1) = 1 - 1 = 0 = f(1)$, la secante es la recta $\ell(x) = 0$ de pendiente $m = 0$. Los puntos interiores de valor medio son exactamente las raíces de

$$f'(x) = 4x^3 - 2x = 2x(2x^2 - 1) = 0, \qquad c \in \left\{ -\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right\}.$$

Los tres valores pertenecen a $(-1, 1)$ porque $0 < 1/\sqrt{2} < 1$, y la factorización muestra que no faltan raíces. Como la pendiente en todos es cero, cada tangente es una recta horizontal de ecuación $T_c(x) = f(c)$. En $c = 0$ vale $f(0) = 0$, de modo que $T_0(x) = 0$ **coincide** con la secante. En los otros dos puntos, $c^2 = 1/2$ y $c^4 = 1/4$, por lo que $f(c) = -1/4$ y ambas tangentes tienen la ecuación $T_c(x) = -1/4$. Estas dos coinciden entre sí, pero son distintas de la secante: comparten su pendiente, no su ordenada.

Solución G3. Supongamos primero que f es K-Lipschitz. Para cualquier punto interior t, todos los incrementos no nulos suficientemente pequeños son admisibles, y

$$\left| \frac{f(t + h) - f(t)}{h} \right| \le K.$$

La diferenciabilidad en t proporciona el límite de estos cocientes, y la continuidad del valor absoluto junto con la preservación del orden da $|f'(t)| \le K$. Por consiguiente, **toda** constante Lipschitz domina el conjunto $\{|f'(t)| : t \in (a, b)\}$. Recíprocamente, si f' está acotada en (a, b), el número $S = \sup_{t \in (a,b)} |f'(t)|$ es finito y no negativo (el abierto es no vacío). En todo punto interior, $|f'(t)| \le S$ por definición de supremo. Aplicando Proposición 9.1 a cualquier par del intervalo concluimos $|f(y) - f(x)| \le S|y - x|$. Así, S es una constante Lipschitz, y la primera parte demuestra que ninguna menor es admisible: **es la mínima**. Si la derivada no está acotada, la primera parte descarta toda constante Lipschitz finita.

No se ha supuesto que f' sea continua ni que alcance su supremo; bastan sus cotas puntuales y el teorema de Lagrange. La equivalencia requiere la **diferenciabilidad interior que figura en el enunciado**: $x \mapsto |x|$ es 1-Lipschitz en $[-1, 1]$ por desigualdad triangular, pero carece de derivada en el origen. No constituye contraejemplo a la equivalencia condicionada, sino a una generalización que suprimiera la diferenciabilidad.

El banco vuelve sobre una misma disciplina de lectura: primero deben existir la función, el dominio y sus propiedades; después se aplican los teoremas con cuantificadores y bordes correctos;

por último se comprueba si una conclusión de existencia puede precisarse a partir de la función particular.

Capítulo 10

Monotonía, extremos, convexidad y forma de las gráficas

El teorema del valor medio nos permitió pasar de un intervalo a un punto: si una función es continua entre dos extremos y diferenciable en su interior, existe un punto donde su derivada coincide con la pendiente de la secante. Pero el resultado también se puede leer en sentido contrario. Si sabemos algo acerca de la derivada **en todos los puntos** de un intervalo, ¿qué podremos afirmar sobre los valores de la función al recorrerlo?

El signo de la derivada permitirá comparar valores de la función, localizar ciertos extremos y estudiar la inclinación de la gráfica. Para precisar la convexidad utilizaremos desigualdades entre valores de la función, de manera que la conclusión pueda justificarse sin depender de la apariencia del dibujo. Una tabla de signos resume información útil cuando cada signo está vinculado al resultado que autoriza la conclusión. El dominio y las hipótesis de ese resultado deben conservarse al interpretar la tabla.

10.1. Del incremento medio al sentido del cambio

La información local puede controlar un intervalo. Consideremos una función f definida y diferenciable en un intervalo abierto I. Elegimos dos puntos cualesquiera $x, y \in I$ con $x < y$. La palabra *intervalo* cumple aquí una función esencial: nos garantiza que el segmento $[x, y]$ está contenido en I. Como la diferenciabilidad implica continuidad, podemos aplicar el teorema del valor medio a la restricción de f a $[x, y]$. Existe, por tanto, un punto $c \in (x, y)$ tal que $\frac{f(y)-f(x)}{y-x} = f'(c)$. Multiplicando por $y-x$, que es positivo, obtenemos $f(y) - f(x) = f'(c)(y-x)$. Como $y - x > 0$, la identidad del valor medio hace coincidir el signo de $f(y) - f(x)$ con el de la derivada en el punto intermedio. Si f' es positiva en todo el intervalo, cualquiera de los puntos que entregue el teorema producirá una diferencia positiva.

La condición sobre la derivada debe valer en todo el intervalo para que el razonamiento funcione con cualquier par de puntos. La continuidad de f' no interviene: el teorema del valor medio se aplica a f, continua en el segmento y diferenciable en su interior.

Una primera consecuencia, con demostración. Supongamos que $f'(t) > 0$ para todo $t \in I$. Para cualesquiera $x < y$ de I, el punto c obtenido arriba satisface $f'(c) > 0$. Puesto que $y - x > 0$, $f(y) - f(x) = f'(c)(y - x) > 0$.

Por consiguiente,

$$x < y \quad \Longrightarrow \quad f(x) < f(y).$$

En palabras: al avanzar hacia la derecha dentro del intervalo, los valores de la función aumentan estrictamente. Hemos establecido una conclusión sobre *todos los pares* $x < y$, no solamente sobre la inclinación de la gráfica cerca de un punto. Si sustituimos la hipótesis $f'(t) > 0$ por $f'(t) \geq 0$, exactamente el mismo razonamiento proporciona $f(x) \leq f(y)$. La distinción entre aumento estricto y no estricto será importante: una derivada puede anularse sin que por ello la función deje de crecer estrictamente.

¿Dónde se empleó cada hipótesis? El carácter intervalar de I permitió tomar el segmento $[x, y]$ sin salir del dominio. La diferenciabilidad proporcionó continuidad y permitió aplicar el teorema del valor medio. La positividad de f' determinó el signo del incremento. Si falta cualquiera de estos elementos, el argumento debe revisarse; una fórmula de derivación, por sí sola, no autoriza la conclusión global.

Dos advertencias antes de convertir esto en una regla. La primera concierne a la **conversa**. Sería tentador sostener que una función que crece estrictamente debe tener derivada estrictamente positiva en cada punto. Consideremos, sin embargo, $f(x) = x^3$. Para comprobar que es estrictamente creciente, tomemos $x < y$. Entonces

$$y^3 - x^3 = (y - x)(x^2 + xy + y^2) > 0,$$

pues $y - x > 0$ y $x^2 + xy + y^2 = \frac{1}{2}[(x + y)^2 + x^2 + y^2] > 0$ cuando $x < y$. Sin embargo, su derivada es

$$f'(x) = 3x^2, \qquad f'(0) = 0.$$

La derivada nula en un punto aislado no ha producido un tramo horizontal. La conclusión correcta para una función creciente y diferenciable será una desigualdad no estricta para la derivada; más adelante demostraremos también esa dirección de la relación. La segunda advertencia concierne al **dominio**. Sea

$$g\colon \mathbb{R} \setminus \{0\} \longrightarrow \mathbb{R}, \qquad g(x) = -\frac{1}{x}.$$

Su derivada es

$$g'(x) = \frac{1}{x^2} > 0 \qquad (x \neq 0).$$

Esto nos permite demostrar que g es estrictamente creciente en cada uno de los intervalos $(-\infty, 0)$ y $(0, \infty)$. Pero no en todo su dominio, considerado como un conjunto ordenado: aunque $-1 < 1$,

$$g(-1) = 1 > -1 = g(1).$$

¿Qué falló? No podemos aplicar el teorema del valor medio al segmento $[-1, 1]$, porque g no está definida en 0. La desconexión del dominio no se arregla diciendo que «la derivada es positiva en todas partes donde existe»: el puente entre los dos valores atraviesa un punto que falta.

Una regla que aprenderemos a leer correctamente. El signo de la derivada informa sobre el sentido del cambio **dentro de cada intervalo en el que se cumplen las hipótesis**. No debemos combinar intervalos separados ni convertir una condición suficiente en una condición necesaria sin demostrarlo.

Hacia los extremos y la convexidad. Una vez que podamos determinar dónde crece y dónde decrece una función, estaremos preparados para investigar qué sucede al pasar de un comportamiento al otro. No bastará con resolver $f'(x) = 0$: también habrá que examinar los puntos donde la derivada no existe y los extremos del dominio. Después haremos una pregunta de orden superior. La derivada nos dice cómo cambia la función; ¿qué significa que la propia derivada aumente o disminuya? La respuesta conducirá a la convexidad, pero comenzaremos por

una definición geométrica precisa basada en cuerdas, para que los criterios diferenciales sean teoremas y no simples convenciones de dibujo.

Antes de seguir. ¿Qué parte de la demostración dejaría de ser válida si supiéramos que $f'(c) > 0$ para *un solo* punto c? ¿Por qué el ejemplo de x^3 impide afirmar que el crecimiento estricto equivale a $f' > 0$ en todas partes?

10.2. Funciones crecientes y decrecientes: definiciones y cuantificadores

En §10.1 dedujimos una desigualdad entre los valores de una función a partir del signo de su derivada. Antes de sistematizar ese procedimiento, debemos precisar **qué significa** que una función crezca o decrezca, sin mencionar derivadas. Esa propiedad ya apareció en el capítulo 2, al estudiar las funciones inversas (Definición 2.8). Aquí recuperaremos su definición para aprender a utilizarla en argumentos sobre intervalos. No introduciremos una segunda definición incompatible con la primera.

Cuatro propiedades que no deben confundirse. Sean $D \subseteq \mathbb{R}$ y $f\colon D \to \mathbb{R}$. La definición Definición 2.8 compara **cualquier par** de puntos distintos del dominio, colocados en orden. Reunamos sus cuatro posibilidades:

Propiedad de f en D	Exigencia para todo $x, y \in D$ con $x < y$
No decreciente	$f(x) \leq f(y)$
Estrictamente creciente	$f(x) < f(y)$
No creciente	$f(x) \geq f(y)$
Estrictamente decreciente	$f(x) > f(y)$

En este libro, **monótona** significa no decreciente o no creciente; cuando sea necesario especificaremos «estrictamente». No emplearemos «creciente» a secas si pudiera confundirse el sentido débil con el estricto. El cuantificador universal merece una lectura cuidadosa. Por ejemplo, «f es estrictamente creciente en D» afirma

$$\forall x, y \in D, \qquad x < y \Longrightarrow f(x) < f(y).$$

No afirma que exista al menos un par favorable, ni que la desigualdad se cumpla únicamente entre puntos consecutivos de una tabla, ni exige que conozcamos una fórmula para f. La definición habla de todos los pares comparables del dominio.

La negación proporciona un método de refutación
Para demostrar que f **no** es estrictamente creciente basta exhibir $x, y \in D$ con $x < y$ y $f(x) \geq f(y)$. En cambio, para negar que sea *no decreciente* hace falta un par con $f(x) > f(y)$. La igualdad entre dos valores distintos refuta el crecimiento estricto, pero no el crecimiento débil. Los cuantificadores determinan exactamente qué contraejemplo debemos buscar.

¿Qué hace la palabra «estrictamente»? Toda función estrictamente creciente es no decreciente: basta observar que $f(x) < f(y)$ implica $f(x) \leq f(y)$. La implicación análoga vale para las funciones decrecientes. Las conversas son falsas, pues una desigualdad no estricta permite que dos entradas diferentes tengan la misma imagen. Consideremos una función constante $k\colon \mathbb{R} \to \mathbb{R}$, $k(x) = 3$. Para todo $x < y$, $k(x) = k(y) = 3$. Por ello, k es a la vez no decreciente y no creciente, aunque no es estrictamente creciente ni estrictamente decreciente. Esto no es una paradoja: las dos condiciones débiles son compatibles con la igualdad.

De hecho, si una función es simultáneamente no decreciente y no creciente en un conjunto D, entonces es constante allí. Dados $x, y \in D$, si $x < y$ las dos condiciones proporcionan $f(x) \leq f(y)$ y $f(x) \geq f(y)$; por antisimetría, $f(x) = f(y)$. Si $y < x$, intercambiamos los papeles; si $x = y$, la igualdad es inmediata. Hemos probado una afirmación válida sin continuidad, derivadas ni hipótesis de intervalo.

Hay otra conexión útil con la teoría anterior. Si f es no decreciente e inyectiva en D, entonces $x < y$ obliga a $f(x) \leq f(y)$, mientras que la inyectividad impide $f(x) = f(y)$; de aquí se sigue $f(x) < f(y)$. Recíprocamente, toda función estrictamente creciente es inyectiva por Teorema 2.2. Por tanto, **no decreciente más inyectiva equivale a estrictamente creciente**; de manera simétrica, no creciente más inyectiva equivale a estrictamente decreciente. No debemos sustituir «inyectiva» por «continua»: la continuidad no excluye los tramos constantes.

Un caso degenerado que aclara los cuantificadores
Si D contiene cero o un elemento, no existen $x, y \in D$ con $x < y$. Las cuatro condiciones de monotonía se satisfacen entonces de manera vacía. En particular, una función sobre un conjunto unitario puede ser estrictamente creciente y estrictamente decreciente a la vez; esa simultaneidad es imposible si el dominio contiene al menos dos puntos. Las afirmaciones sobre *cambio efectivo* necesitan dos entradas distintas.

Estudiar una restricción no equivale a estudiar toda la función. Sea $f\colon D \to \mathbb{R}$ y $I \subseteq D$. Decir que «f es no decreciente en I» significa aplicar la definición a la restricción $f|_I$. Todas las comparaciones deben tener ambos argumentos en I. Si f es monótona en D, conserva ese sentido de monotonía en cualquier subconjunto; pero una propiedad demostrada para una restricción no se extiende automáticamente al dominio completo.

La función

$$p\colon \mathbb{R} \to \mathbb{R}, \qquad p(x) = x^2$$

permite comprobarlo sin cálculo diferencial. Si $x < y \leq 0$, tenemos $y - x > 0$ y $x + y < 0$, por lo que $p(y) - p(x) = (y - x)(x + y) < 0$. Así, p es estrictamente decreciente en $(-\infty, 0]$. En cambio, para $0 \leq x < y$, ambos factores del mismo producto son positivos y $p(y) > p(x)$: es estrictamente creciente en $[0, \infty)$. No es monótona en todo $\mathbb{R}$. El par $-1 < 0$ y la desigualdad $p(-1) > p(0)$ descartan que sea no decreciente; el par $0 < 1$ y $p(0) < p(1)$ descartan que sea no creciente. Este ejemplo muestra una forma correcta de comunicar el resultado: debemos **nombrar cada intervalo** y el sentido de monotonía correspondiente. Decir solamente «x^2 decrece y luego crece» puede servir como orientación, pero no sustituye la identificación de los conjuntos en los que se verifican los cuantificadores.

Componentes separadas: la comparación que suele olvidarse. La definición de monotonía es válida en cualquier subconjunto ordenado D, aunque no sea un intervalo. Lo que exige un intervalo es el razonamiento con el teorema del valor medio utilizado en §10.1: para comparar $f(x)$ con $f(y)$ necesitamos que la función esté definida en todo $[x, y]$ y cumpla allí las hipótesis pertinentes. Volvamos a $g(x) = -1/x$ sobre $D = \mathbb{R} \setminus \{0\}$. En §10.1 verificamos que g es estrictamente creciente en $(-\infty, 0)$ y también en $(0, \infty)$. Sin embargo, si tomamos $x = -1$ y $y = 1$, obtenemos

$$x < y, \qquad g(x) = 1 > -1 = g(y).$$

Ese par cruza las dos componentes y refuta que g sea no decreciente en D. La monotonía **en cada componente** no es lo mismo que la monotonía **en la unión**: esta última añade

todas las comparaciones entre componentes, que todavía deben comprobarse por otros medios. Podemos ver que tampoco es correcto inferir que un dominio desconectado impide por sí mismo la monotonía global. La identidad $h\colon D \to \mathbb{R}$, $h(x) = x$, es estrictamente creciente en ese mismo D: la desigualdad $x < y$ se traduce directamente en $h(x) < h(y)$, incluso si el intervalo que separa los argumentos contiene puntos excluidos del dominio. **La definición no necesita intervalos; el criterio mediante derivadas sí necesita un tramo donde pueda aplicarse el teorema del valor medio.**

Lectura de una tabla: evidencia finita y afirmación universal. Supongamos que conocemos tres valores $f(-1) = 2$, $f(0) = 2$ y $f(1) = 5$. Estos datos son compatibles con una función no decreciente y descartan que sea estrictamente creciente en cualquier dominio que contenga -1 y 0. Pero **no demuestran** que sea no decreciente en todo un intervalo: podrían existir dos puntos no observados cuyo orden de imágenes fuera el contrario.

La diferencia entre prueba y exploración puede expresarse como un procedimiento. Para refutar una propiedad universal, basta un par testigo. Para demostrarla hay que tomar puntos arbitrarios $x < y$ y justificar la desigualdad correspondiente. A veces una identidad algebraica bastará; en la sección siguiente, el teorema del valor medio nos permitirá deducirla de información acerca de la derivada.

Laboratorio breve: leer antes de calcular
Considera $f\colon [-2, 2] \to \mathbb{R}$, $f(x) = |x|$. Decide primero, comparando valores, en cuál de los intervalos $[-2, 0]$ y $[0, 2]$ es estrictamente creciente o decreciente. Para negar la monotonía en todo $[-2, 2]$, proporciona **dos pares diferentes**: uno que refute «no decreciente» y otro que refute «no creciente». Finalmente, explica por qué esos dos testigos no dicen nada, por sí solos, sobre la existencia de derivadas.

Fijado un intervalo en el que queremos comparar valores, podemos usar el signo de f' en cada punto interior para obtener una desigualdad entre las imágenes de cualquier par ordenado. Los criterios siguientes expresan ese paso del control puntual a la comparación sobre el intervalo.

10.3. Criterios de monotonía a partir de la derivada

Las definiciones de §10.2 exigen comparar todos los pares $x < y$ de un conjunto. Cuando conocemos la derivada, el teorema del valor medio ofrece una vía para demostrar esas desigualdades sin efectuar cada comparación mediante álgebra. Sin embargo, debemos fijar primero el intervalo y verificar que la función se comporte bien también en sus extremos, si los tiene.

Cuatro criterios, una misma demostración.

Teorema 10.1. **Teorema (criterios de monotonía por el signo de la derivada).** Sea $I \subseteq \mathbb{R}$ un intervalo que contiene al menos dos puntos. Supongamos que $f\colon I \to \mathbb{R}$ es continua en I (con la continuidad relativa en los extremos que pertenezcan a I) y diferenciable en el interior I°. Entonces se cumplen las cuatro implicaciones siguientes, siempre que la condición sobre f' valga para **todo** $t \in I^\circ$:

1. Si $f'(t) \geq 0$, entonces f es no decreciente en I.
2. Si $f'(t) > 0$, entonces f es estrictamente creciente en I.
3. Si $f'(t) \leq 0$, entonces f es no creciente en I.
4. Si $f'(t) < 0$, entonces f es estrictamente decreciente en I.

Demostración. Tomemos $x, y \in I$ arbitrarios con $x < y$. Puesto que I es un intervalo, $[x, y] \subseteq I$. La restricción de f es continua en $[x, y]$ y diferenciable en (x, y); por el teorema del valor medio de Lagrange (Teorema 9.3), existe $c \in (x, y)$ tal que

$$f(y) - f(x) = f'(c)(y - x).$$

El factor $y - x$ es estrictamente positivo. Si $f' \geq 0$ en el interior, el miembro derecho es no negativo y $f(x) \leq f(y)$: queda probado el primer caso. Si $f' > 0$, el producto es positivo y $f(x) < f(y)$, lo que demuestra el segundo. Para $f' \leq 0$ obtenemos $f(y) - f(x) \leq 0$ y, para $f' < 0$, obtenemos $f(y) - f(x) < 0$: son los casos tercero y cuarto. Como $x < y$ fue arbitrario, las conclusiones valen para todos los pares exigidos por la definición. □

Aplicamos el teorema del valor medio a f, por lo que basta que su derivada exista en cada punto interior y tenga allí el signo requerido. En los extremos de un intervalo cerrado se necesita continuidad relativa de f para admitirlos como puntos de comparación; no se exigen derivadas bilaterales en ellos.

Auditoría de la demostración
La cadena lógica es: $x < y$ arbitrarios $\longrightarrow$ segmento $[x, y]$ contenido en I $\longrightarrow$ hipótesis de Lagrange para la restricción $\longrightarrow$ un punto c interior $\longrightarrow$ signo de $f'(c)(y-x)$ $\longrightarrow$ desigualdad entre los valores. La condición «para todo t» nos permite utilizar el signo de la derivada **en cualquier punto c que proporcione el teorema**.

Qué dicen exactamente las conversas. Las cuatro condiciones anteriores son suficientes para obtener monotonía. Para averiguar cuáles son también necesarias, no sirve invertir las flechas del enunciado: hay que volver a la definición de derivada.

Proposición 10.1. **Proposición (signo necesario para una función monótona).** Sea I un intervalo y sea $f\colon I \to \mathbb{R}$. Si f es no decreciente en I y es diferenciable en un punto interior c, entonces $f'(c) \geq 0$. Si f es no creciente y es diferenciable en c, entonces $f'(c) \leq 0$. Las mismas conclusiones no estrictas valen, respectivamente, para funciones estrictamente crecientes y estrictamente decrecientes.

Demostración. Supongamos primero que f es no decreciente. Como c es interior, para todo $h \neq 0$ suficientemente pequeño, $c + h \in I$. Si $h > 0$, tenemos $f(c + h) \geq f(c)$, y por tanto

$$\frac{f(c+h) - f(c)}{h} \geq 0.$$

Si $h < 0$, la monotonía aplicada al par $c + h < c$ da $f(c + h) \leq f(c)$. Ahora numerador y denominador son no positivos, de modo que el mismo cociente vuelve a ser no negativo. Si existe su límite bilateral, no puede ser negativo; por definición, dicho límite es $f'(c) \geq 0$. Si f es no creciente, las desigualdades de los valores se invierten y ambos cocientes incrementales son no positivos; su límite, cuando existe, satisface $f'(c) \leq 0$. Finalmente, la monotonía estricta implica la correspondiente monotonía débil (§10.2), por lo que acabamos de demostrar también los dos últimos casos. □

Observemos el alcance exacto de la proposición. No exige continuidad de f' ni que f sea diferenciable en todo I: produce una conclusión **en cada punto interior donde la derivada exista**. Si, adicionalmente, f cumple las hipótesis del teorema anterior, podemos combinar ambas direcciones:

$$f \text{ no decreciente en } I \iff f'(t) \geq 0 \quad (t \in I^\circ),$$
$$f \text{ no creciente en } I \iff f'(t) \leq 0 \quad (t \in I^\circ).$$

Estas equivalencias son para las versiones **no estrictas** y conservan las hipótesis de continuidad en I y diferenciabilidad en su interior. En cambio, el crecimiento estricto solo fuerza $f' \geq 0$, no $f' > 0$; el decrecimiento estricto solo fuerza $f' \leq 0$, no $f' < 0$. El ejemplo $f(x) = x^3$ de §10.1 ya refuta una de las falsas conversas: crece estrictamente en $\mathbb{R}$, pero $f'(0) = 0$. Para obtener el caso simétrico basta considerar $-x^3$, estrictamente decreciente con derivada nula en 0. Un valor cero de la derivada no informa, por sí solo, de la existencia de un tramo constante ni de un extremo.

¿Puede una derivada anularse y mantenerse el crecimiento estricto? La respuesta es afirmativa, pero debemos precisar **cuánto** puede anularse. La condición $f' \geq 0$ solo garantiza crecimiento débil; para excluir dos valores iguales necesitamos descartar que la función sea constante entre ellos.

Proposición 10.2. **Proposición (criterio exacto para el crecimiento estricto bajo $f' \geq 0$).** Supongamos las hipótesis de Teorema 10.1 y que $f'(t) \geq 0$ para todo $t \in I^\circ$. Entonces f es estrictamente creciente en I si, y solo si, f' **no es idénticamente nula en ningún intervalo abierto no vacío contenido en** I°.

La afirmación simétrica vale para el decrecimiento estricto si $f' \leq 0$ y tampoco es idénticamente nula en ningún intervalo abierto no vacío del interior.

Demostración. Por Teorema 10.1, f es no decreciente. Supongamos que no es estrictamente creciente. Según §10.2, existen $x < y$ en I con $f(x) \geq f(y)$. La monotonía débil proporciona la desigualdad opuesta y, por tanto, $f(x) = f(y)$. Para cualquier $z \in [x, y]$, se tiene

$$f(x) \leq f(z) \leq f(y) = f(x).$$

Así, f es constante en $[x, y]$, por lo que $f'(t) = 0$ para todo $t \in (x, y)$. Este intervalo abierto es no vacío y está contenido en I°, contradiciendo la condición anunciada. Luego f es estrictamente creciente. Recíprocamente, si f' fuese idénticamente nula en algún intervalo abierto no vacío $J \subseteq I^\circ$, la proposición de constancia por derivada nula (Proposición 9.2), aplicada a cada segmento cerrado contenido en J, mostraría que f toma el mismo valor en dos puntos distintos de J. Eso contradice el crecimiento estricto. El caso decreciente sigue reemplazando f por $-f$. □

Una condición suficiente particularmente cómoda es que $f' \geq 0$ y sus ceros sean **aislados**: entonces ningún intervalo abierto puede estar formado enteramente por ceros. Lo mismo ocurre si f' se anula solo en un conjunto finito. Pero no necesitamos imponer ninguna de esas dos condiciones: lo indispensable para el criterio anterior es que la derivada no se anule *en todo un intervalo abierto*. Es una afirmación global sobre los intervalos, no una clasificación del comportamiento en un punto aislado.

Ejemplo 10.1. **Un polinomio cuyo signo cambia en puntos conocidos.** Consideremos

$$p\colon \mathbb{R} \to \mathbb{R}, \qquad p(x) = x^3 - 3x.$$

Por las reglas de derivación anteriores, $p'(x) = 3x^2 - 3 = 3(x-1)(x+1)$. Si $x < -1$, ambos factores son negativos y su producto es positivo; si $-1 < x < 1$, el producto es negativo; si $x > 1$, ambos factores son positivos. Como p es continua en toda la recta, el criterio se aplica también a los intervalos con los extremos -1 o 1 incluidos. Concluimos que p es estrictamente creciente en $(-\infty, -1]$ y en $[1, \infty)$, y estrictamente decreciente en $[-1, 1]$. Estos tres resultados no permiten declarar que p sea monótona en su dominio completo. Por ejemplo, $p(-2) = -2 < p(-1) = 2$

refuta que sea no creciente, mientras que $p(-1) = 2 > p(0) = 0$ refuta que sea no decreciente. En §10.4 veremos qué información adicional proporciona el paso de un signo al opuesto para estudiar extremos.

Dos hipótesis que no podemos suprimir.

Ejemplo 10.2. **Una derivada positiva sin continuidad en el borde.** Definamos $q\colon [0,1] \to \mathbb{R}$ por

$$q(0) = 2, \qquad q(x) = x \quad (0 < x \le 1).$$

Para todo $t \in (0,1)$ existe $q'(t) = 1 > 0$. Sin embargo, $0 < \frac{1}{2}$ y $q(0) = 2 > \frac{1}{2} = q(\frac{1}{2})$, de modo que q ni siquiera es no decreciente en $[0,1]$. El criterio no falla: falta la continuidad de q en el extremo 0, necesaria para aplicar Lagrange a $[0, \frac{1}{2}]$. El signo de una derivada interior no controla por sí solo un valor de borde asignado discontinuamente.

La otra hipótesis es la geometría del dominio. Como comprobamos en §§10.1–10.2, $g(x) = -1/x$ tiene $g'(x) > 0$ en cada punto de $\mathbb{R} \setminus \{0\}$, y es estrictamente creciente por separado a ambos lados de 0, pero **no** es no decreciente en la unión. El teorema no autoriza atravesar el punto excluido. Si el dominio tiene varias componentes, habrá que estudiarlas y después comprobar, mediante la definición, las comparaciones entre ellas antes de afirmar monotonía global.

Cómo leer una tabla de signos
Primero se fija un intervalo I donde f esté definida y sea continua, y se verifica su diferenciabilidad interior. Después se demuestra el signo de f' en **todos** los puntos interiores, incluidos los casos que no aparecen en una tabla de muestras. Solo entonces se aplica Teorema 10.1. Los ceros aislados de f' no obligan a separar artificialmente el intervalo si el signo débil se mantiene; tampoco autorizan por sí solos a proclamar un extremo.

Laboratorio de demostraciones: reparar una conclusión. Considera la siguiente argumentación: «$f'(a) > 0$; por tanto, f es estrictamente creciente en todo su dominio». El dato se refiere a un solo punto, mientras que el teorema exige un signo en todo el interior de un intervalo y continuidad hasta los extremos incluidos. Para reparar la prueba habría que formular esas hipótesis y precisar *en qué intervalo* se obtiene la conclusión. Ni siquiera conocer el signo de la derivada en muchos puntos de una tabla equivale a demostrarlo universalmente.

La igualdad $f'(c) = 0$ es compatible con crecimiento estricto, como muestra x^3. Para obtener monotonía a partir de $f' \ge 0$ en un intervalo, y decidir si es estricta, tendremos que examinar además si la función puede permanecer constante en algún subintervalo; el criterio correspondiente se expresa mediante la anulación de la derivada en él.

Antes de seguir. ¿Por qué el límite de los cocientes incrementales puede ser cero aunque cada cociente sea positivo? ¿Qué parte del argumento de Teorema 10.1 se rompe si $[x, y]$ sale del dominio? ¿Por qué la continuidad en 0 habría impedido el contraejemplo de Ejemplo 10.2?

10.4. Puntos críticos y cambios de signo

En §10.3 demostramos que el signo de la derivada determina el sentido de la monotonía sobre un intervalo. Ahora podemos utilizar esa información para investigar lo que ocurre cerca de un punto donde la función pasa de aumentar a disminuir, o viceversa. El ejemplo $p(x) = x^3 - 3x$ ya anticipó dos cambios de comportamiento. La cuestión es justificar cuándo esos cambios producen un máximo o un mínimo **local**, y distinguir esa conclusión de la simple igualdad $f'(c) = 0$.

El punto donde debemos detenernos a examinar la función. Recordemos que Definición 9.2 define los extremos locales mediante comparaciones con todos los puntos suficientemente próximos del **dominio**. La definición no exige que exista la derivada. Por otra parte, el teorema de Fermat (Teorema 9.1) establece que, si un extremo local se alcanza en un punto interior donde la función es diferenciable, la derivada vale cero. Esto motiva la siguiente terminología.

Definición 10.1. **Punto crítico interior.** Sean $A \subseteq \mathbb{R}$, $f\colon A \to \mathbb{R}$ y $c \in A^\circ$. Diremos que c es un **punto crítico interior** de f si se cumple al menos una de las condiciones siguientes:

- existe la derivada finita $f'(c)$ y $f'(c) = 0$;
- la derivada $f'(c)$ no existe como número real finito.

La condición se refiere a un **punto del dominio**, no al valor $f(c)$. En esta convención, los puntos de borde no se llaman críticos interiores: se estudiarán por separado en §10.5.

La segunda posibilidad abarca esquinas y cúspides, pero también admite, por sí sola, puntos de discontinuidad. Por ello, *crítico* no significa *extremo*, ni garantiza continuidad. Tampoco significa que la gráfica tenga obligatoriamente una tangente horizontal: esa descripción corresponde solo al primer caso, cuando la derivada es cero.

Proposición 10.3. **Proposición (los extremos locales interiores son puntos críticos).** Sean $A \subseteq \mathbb{R}$, $f\colon A \to \mathbb{R}$ y $c \in A^\circ$. Si f alcanza en c un máximo o un mínimo local relativo a A, entonces c es un punto crítico interior en el sentido de Definición 10.1.

Demostración. Si f no es diferenciable en c, se cumple la segunda cláusula de la definición. Si existe su derivada finita, Fermat (Teorema 9.1) implica $f'(c) = 0$, y se cumple la primera. Hemos agotado las dos posibilidades. □

Esta proposición es una **condición necesaria para un extremo interior**, no un método que clasifique todos los puntos críticos. Una lista de candidatos puede ser útil para organizar una investigación; verificar que un candidato sea efectivamente un máximo o mínimo exige todavía comparar valores o aplicar un criterio adicional.

El primer criterio: demostrar lo que el cambio de signo permite concluir. Fijemos un punto interior c y supongamos que podemos mirar la derivada a ambos lados. Para pasar de la monotonía de dos intervalos a una comparación *con* $f(c)$, necesitamos que la función sea continua en c: una función podría aumentar a la izquierda, disminuir a la derecha y, sin embargo, tener en el punto central un valor aislado que destruya el supuesto máximo.

Teorema 10.2. **Teorema (primer criterio de la derivada para extremos locales).** Sean $A \subseteq \mathbb{R}$, $f\colon A \to \mathbb{R}$ y $c \in A^\circ$. Supongamos que existe $r > 0$ tal que $(c - r, c + r) \subseteq A$, que f es continua en c y que es diferenciable en

$$(c - r, c) \cup (c, c + r).$$

Si cada signo de la tabla vale en **todos** los puntos del intervalo lateral indicado, se obtienen las siguientes conclusiones:

Signo de f' en $(c - r, c)$	Signo de f' en $(c, c + r)$	Conclusión en c
$+$	$-$	Máximo local estricto.
$-$	$+$	Mínimo local estricto.
$+$	$+$	Ni máximo ni mínimo local.
$-$	$-$	Ni máximo ni mínimo local.

El teorema **no exige** que $f'(c)$ exista, ni que f' sea continua en los intervalos laterales.

Demostración. Supongamos primero $f'(t) > 0$ a la izquierda y $f'(t) < 0$ a la derecha. Para cualquier $x \in (c-r, c)$, f es continua en $[x, c]$: lo es en x y en todos los puntos anteriores a c por diferenciabilidad, y lo es en c por hipótesis. También es diferenciable en (x, c). Por Lagrange (Teorema 9.3), existe $u \in (x, c)$ tal que

$$f(c) - f(x) = f'(u)(c - x) > 0.$$

Análogamente, para $y \in (c, c+r)$ existe $v \in (c, y)$ con

$$f(y) - f(c) = f'(v)(y - c) < 0.$$

Así, $f(x) < f(c)$ para todo x situado suficientemente cerca de c y distinto de él, tanto a la izquierda como a la derecha. Esta es exactamente la definición de máximo local estricto de Definición 9.2. Si los signos son $-$ a la izquierda y $+$ a la derecha, las dos desigualdades se invierten y obtenemos un mínimo local estricto. Si ambos signos son positivos, los mismos argumentos dan, para cualesquiera $x < c < y$ suficientemente próximos,

$$f(x) < f(c) < f(y).$$

Los puntos de la izquierda impiden un mínimo local y los de la derecha impiden un máximo local. Cuando ambos signos son negativos, las desigualdades se invierten y tampoco puede haber un extremo local. □

La demostración no afirma que la derivada «cambie de signo en c» como si estuviera obligada a existir en ese punto. Afirma algo más preciso: hay un intervalo perforado a la izquierda y otro a la derecha en los que se conocen, respectivamente, los signos. El punto c pertenece al dominio y su continuidad une las comparaciones laterales. **La información decisiva está en los dos lados y en el valor central.**

La continuidad en el punto central no es una formalidad
Aplicar Lagrange a $[x, c]$ y a $[c, y]$ utiliza continuidad en los extremos de ambos segmentos. La diferenciabilidad se exige únicamente en sus interiores: por eso el criterio puede detectar un máximo o mínimo en una esquina donde $f'(c)$ no exista. Tampoco estamos suponiendo continuidad de f'.

Si solo conocemos signos débiles, hay una variante útil pero de alcance menor: $f' \geq 0$ a la izquierda y $f' \leq 0$ a la derecha garantizan, con las mismas hipótesis, un **máximo local no necesariamente estricto**; la combinación inversa garantiza un mínimo local no necesariamente estricto. Se repite la prueba reemplazando $>$ y $<$ por $\geq$ y $\leq$. Para una función constante se satisfacen ambas combinaciones débiles y todos sus puntos son máximos y mínimos locales no estrictos. Por eso no debemos introducir la palabra «estricto» cuando solo hemos demostrado desigualdades débiles.

Volver a un ejemplo conocido: dos extremos, dos comparaciones.

Ejemplo 10.3. **El polinomio $x^3 - 3x$ revisitado.** En Ejemplo 10.1 estudiamos

$$p(x) = x^3 - 3x, \qquad p'(x) = 3(x-1)(x+1).$$

El signo es positivo en $(-\infty, -1)$, negativo en $(-1, 1)$ y positivo en $(1, \infty)$. La función es continua y diferenciable en toda la recta, por lo que Teorema 10.2 se aplica en los dos puntos

críticos. En $c=-1$ el signo pasa de + a − y $p(-1)=2$: se alcanza un **máximo local estricto**. En $c=1$ pasa de − a + y $p(1)=-2$: se alcanza un **mínimo local estricto**.

No hemos demostrado que esos valores sean extremos *absolutos* sobre $\mathbb{R}$. De hecho, $p(x)$ toma valores arbitrariamente grandes o pequeños, pues $p(x)=x(x^2-3)$ tiene el signo de x y magnitud no acotada cuando $|x|$ crece suficientemente. La comparación local y la comparación sobre todo el dominio son problemas diferentes.

En este ejemplo, resolver $p'(x)=0$ descubrió los dos candidatos -1 y 1, pero fue el análisis **de los signos a cada lado** lo que los clasificó. Resolver una ecuación es una etapa de búsqueda, no la demostración completa.

Puntos críticos sin derivada: ni invisibles ni automáticamente extremos.

Ejemplo 10.4. **Una esquina que da un mínimo y otra que no da extremo.** Para $u(x)=|x|$ en $\mathbb{R}$,

$$u'(x)=-1 \quad (x<0), \qquad u'(x)=1 \quad (x>0).$$

La función es continua en 0, pero los cocientes incrementales allí son $|h|/h$, con límites laterales distintos. Así, 0 es crítico por no diferenciabilidad, y el cambio de − a + prueba que es un mínimo local estricto. Cambiando u por $-u$ obtenemos un máximo local estricto, también sin derivada en el origen. Comparemos este comportamiento con la función continua $v\colon \mathbb{R}\to\mathbb{R}$ dada por

$$v(x)=\begin{cases} x, & x\le 0,\\ 2x, & x>0. \end{cases}$$

En 0 sus derivadas laterales son 1 y 2, por lo que $v'(0)$ no existe. Sin embargo, $v'(x)=1$ a la izquierda y $v'(x)=2$ a la derecha: los signos son + y +. El primer criterio muestra que 0 **no** es un extremo local. Ambos ejemplos contienen un punto crítico por no diferenciabilidad; la clasificación cambia cuando cambian las desigualdades laterales.

Derivada cero: tres resultados incompatibles con una clasificación automática.

Ejemplo 10.5. **El mismo dato puntual, conclusiones diferentes.** Consideremos las tres funciones sobre $\mathbb{R}$:

$$a(x)=x^3, \qquad b(x)=x^4, \qquad d(x)=-x^4.$$

Las tres satisfacen

$$a'(0)=b'(0)=d'(0)=0.$$

Pero $a'(x)=3x^2>0$ para $x\neq 0$: su signo es + a ambos lados, y 0 no es extremo local. Para b, $b'(x)=4x^3$ es negativa a la izquierda y positiva a la derecha, de modo que 0 es mínimo local estricto. Para d, $d'(x)=-4x^3$ es positiva a la izquierda y negativa a la derecha, de modo que 0 es máximo local estricto. Todos estos resultados también se verifican directamente comparando x^3, x^4 y $-x^4$ con sus respectivos valores en 0. Por tanto, la afirmación «$f'(c)=0$ implica un extremo» es falsa; también es falsa la idea de que el valor cero permita distinguir máximo de mínimo.

Dos conversas y una hipótesis que conviene auditar. La proposición Proposición 10.3 dice que todo extremo local **interior** es crítico. Su conversa es falsa por $a(x) = x^3$ en Ejemplo 10.5. Tampoco podemos invertir el primer criterio y afirmar que *todo máximo local estricto obliga a que la derivada sea positiva en todo un intervalo izquierdo y negativa en todo un intervalo derecho.* Un extremo se define mediante valores, y esa desigualdad puede coexistir con oscilaciones de las pendientes.

Ejemplo 10.6. **Máximo estricto sin signos laterales uniformes de la derivada.** Podemos construirlo sin funciones trigonométricas. Sea $B(t) = t^3(1-t)^3$ para $0 \le t \le 1$, y $B(t) = 0$ fuera de ese intervalo. El polinomio, su primera derivada y su segunda derivada se anulan en los extremos: por tanto, esta función por tramos es dos veces diferenciable en toda la recta. Además, $0 \le B \le 1/64$.

Para $n \ge 2$, pongamos $a_n = 2^{-n}$, $d_n = a_n^3$ e $I_n = [a_n, a_n + d_n]$. Son intervalos disjuntos, pues $a_{n+1} + d_{n+1} < a_n$, y se acumulan solamente en cero. Definamos

$$p(x) = \begin{cases} \dfrac{a_n^2}{10} B\Big(\dfrac{x - a_n}{d_n}\Big), & x \in I_n, \\ 0, & x \notin \bigcup_{n\ge 2} I_n, \end{cases} \qquad w(x) = -x^2 + p(x).$$

En cada intervalo I_n, $x \ge a_n$ y $0 \le p(x) \le a_n^2/640 \le x^2/640$. Fuera de ellos, $p = 0$. Así $w(x) < 0 = w(0)$ para todo $x \neq 0$: el origen es un máximo estricto. La misma cota da $|p(x)/x| \le |x|/640 \to 0$, de modo que $w'(0) = 0$. En los extremos de cada intervalo, p y p' empalman con cero; en cualquier punto no nulo solo intervienen finitos intervalos localmente. Por ello w es diferenciable en toda la recta.

En los puntos $r_n = a_n + d_n/4$ y $s_n = a_n + 3d_n/4$, usamos $B'(1/4) = 27/512$ y $B'(3/4) = -27/512$:

$$w'(r_n) = -2r_n + \frac{27}{5120 a_n} > 0 \quad \text{para } n \text{ grande}, \qquad w'(s_n) = -2s_n - \frac{27}{5120 a_n} < 0.$$

Ambas sucesiones tienden a cero por la derecha. No hay ningún intervalo derecho completo en el que $w' < 0$, aunque el máximo sea estricto. La condición de signos del primer criterio es suficiente, pero no necesaria.

Ejemplo 10.7. **Los signos correctos no compensan una discontinuidad central.** Definamos $z\colon \mathbb{R} \to \mathbb{R}$ por

$$z(0) = -2, \qquad z(x) = -|x| \quad (x \neq 0).$$

Para todo $x < 0$, $z'(x) = 1 > 0$; para todo $x > 0$, $z'(x) = -1 < 0$. Los signos parecen los de un máximo, pero z no es continua en 0, ya que $\lim_{x\to 0} z(x) = 0 \neq z(0)$. Para $0 < |x| < 1$ tenemos $z(x) = -|x| > -1 > -2 = z(0)$: en realidad, 0 es un **mínimo local estricto**, no un máximo. Falta exactamente la continuidad en c exigida por Teorema 10.2.

El último contraejemplo no desmiente el teorema: muestra por qué no debe omitirse una de sus hipótesis. Tampoco la ausencia de derivada en un punto autoriza a declarar un extremo. Para resolver un problema hay que separar **identificación de candidatos**, **verificación de hipótesis** y **clasificación por comparaciones**.

Protocolo de lectura para los extremos interiores
Primero determina el dominio y comprueba que el punto estudiado sea interior. Identifica dónde se anula la derivada y dónde no existe, sin desechar estos últimos puntos. Después, si vas a utilizar el primer criterio, verifica continuidad en el punto, diferenciabilidad a ambos lados y un signo válido en *cada intervalo lateral completo*. Si esas condiciones no se conocen, vuelve a la definición de extremo local: compara directamente $f(x)$ con $f(c)$ para todos los puntos próximos permitidos. Los extremos de borde requieren un examen separado.

Laboratorio: una prueba, tres diagnósticos. Considera la frase: «$f'(c) = 0$; entonces f tiene un extremo en c». Indica si el fallo está en el cálculo, en la implicación lógica o en el dominio, y utiliza Ejemplo 10.5 para refutarla. Considera después: «f' es positiva antes de c y negativa después de c; luego hay un máximo». Señala qué significa *antes* y *después* en términos de intervalos, qué hipótesis sobre $f(c)$ falta y cómo Ejemplo 10.7 muestra que no es opcional. Por último, para la función v de Ejemplo 10.4, comprueba la no existencia de la derivada en 0 mediante dos cocientes laterales y demuestra directamente, con puntos a ambos lados, que no hay extremo.

Antes de seguir. ¿Por qué el teorema se aplica a $|x|$ aunque su derivada no exista en el origen? ¿Qué diferencia lógica separa «todo extremo interior es crítico» de «todo punto crítico es extremo»? ¿En qué momento del estudio de un intervalo cerrado debemos añadir sus puntos de borde a la lista de candidatos? Esta última pregunta nos llevará a §10.5.

10.5. Extremos absolutos y extremos del dominio

En §10.4 aprendimos a clasificar ciertos extremos *locales* mediante signos de la derivada. Un máximo local, sin embargo, solo exige comparaciones con puntos próximos: no hay razón para que sea el mayor valor de la función sobre todo su dominio. Volvemos ahora a la definición de máximo y mínimo absolutos (Definición 9.1) y distinguiremos dos preguntas que conviene resolver en orden: **¿existe un valor extremo?** y, si existe, **¿en qué puntos puede alcanzarse?**

Primero la existencia; después la localización. Sea $f\colon [a,b] \to \mathbb{R}$ continua, con $a < b$. El teorema de los valores extremos (Teorema 5.5) afirma que existen $x_m, x_M \in [a,b]$ tales que

$$f(x_m) \le f(x) \le f(x_M) \qquad \text{para todo } x \in [a,b].$$

Por tanto, los valores mínimo y máximo existen y se alcanzan. El teorema, por sí solo, no identifica los puntos x_m, x_M ni afirma que sean únicos. Para localizarlos aprovecharemos la distinción entre puntos interiores y extremos del intervalo: un punto de máximo absoluto interior también es de máximo local, mientras que a y b no tienen por qué satisfacer la conclusión bilateral de Fermat.

Proposición 10.4. **Proposición (candidatos necesarios para los extremos absolutos en un intervalo cerrado).** Sean $a < b$ y $f\colon [a,b] \to \mathbb{R}$ continua. Definamos

$$C_f = \{a,b\} \cup \{c \in (a,b) : f'(c) = 0 \text{ o } f'(c) \text{ no existe como número real finito}\}.$$

Si f alcanza un máximo o un mínimo absoluto en $c \in [a,b]$, entonces $c \in C_f$. En particular, existen $u, v \in C_f$ tales que

$$f(u) \le f(x) \le f(v) \qquad \text{para todo } x \in [a,b].$$

Cuando f es diferenciable en todo (a, b), la segunda parte de la unión se reduce a los ceros de f'.

Demostración. Por Teorema 5.5 existen $u, v \in [a, b]$ en los que f alcanza sus valores mínimo y máximo absolutos, respectivamente. Consideremos cualquiera de esos puntos, al que llamaremos c. Si $c = a$ o $c = b$, pertenece a C_f por definición. Si $a < c < b$, el extremo absoluto es también local relativo a $[a, b]$ por Definición 9.2. La proposición Proposición 10.3 demuestra entonces que c es crítico interior: su derivada finita es cero o no existe. En ambos casos $c \in C_f$. Aplicado por separado a u y v, esto prueba las dos afirmaciones. □

La proposición dice que los puntos donde se alcanzan los extremos **están dentro** de C_f; no dice que todos los puntos de C_f sean extremos. El ejemplo x^3 en Ejemplo 10.5 ya mostró un cero de la derivada que ni siquiera produce un extremo local. Tampoco dice que el conjunto de candidatos sea finito. El hecho de que $u, v \in C_f$ existan proviene de la continuidad sobre el intervalo cerrado; resolver $f' = 0$ no sustituye esa garantía.

La forma lógica del procedimiento
Existencia: continuidad de f en $[a, b]$ $\longrightarrow$ Weierstrass. **Localización necesaria:** cualquier punto de extremo absoluto es un borde o un punto crítico interior $\longrightarrow$ Fermat y su extensión a puntos no diferenciables. **Determinación:** comparar los valores pertinentes o demostrar desigualdades globales. El segundo paso no demuestra el primero ni convierte cada candidato en una solución.

Una lista finita: comparar valores, no etiquetas. Si C_f es finito y podemos determinarlo completamente, la proposición reduce la búsqueda a una comparación finita. Para cada $c \in C_f$ calculamos $f(c)$; el menor de esos números es el mínimo absoluto y el mayor es el máximo absoluto. La justificación no es que un valor en la lista «parezca más extremo»: ya sabemos que *algún* punto de mínimo y *algún* punto de máximo aparecen en ella. Al comparar los valores también debemos conservar todos los argumentos que produzcan el mismo valor extremo.

Ejemplo 10.8. **Los extremos locales no tienen por qué superar a los bordes.** Restrinjamos a $[-2, 2]$ el polinomio estudiado en Ejemplo 10.3:

$$p(x) = x^3 - 3x, \qquad p'(x) = 3(x - 1)(x + 1).$$

Es continuo en el intervalo cerrado y sus únicos puntos críticos interiores son -1 y 1. El conjunto de candidatos es, por tanto, $C_p = \{-2, -1, 1, 2\}$. La evaluación exacta da

Punto candidato c	-2	-1	1	2
$p(c)$	-2	2	-2	2

Concluimos que el máximo absoluto vale 2 y se alcanza **en dos puntos**, -1 y 2; el mínimo absoluto vale -2 y se alcanza en -2 y 1. Un extremo interior y un extremo de borde pueden compartir el mismo valor absoluto. La clasificación local de §10.4, por sí sola, no habría revelado los empates con los bordes.

No hemos calculado derivadas en -2 ni en 2 para incluirlos en C_p. Lo que importa es su pertenencia al dominio: un extremo absoluto puede alcanzarse en cualquiera de ellos. Si $a = b$, caso excluido del enunciado de la proposición para evitar una distinción artificial entre interior y bordes, la función tiene simultáneamente máximo y mínimo absolutos en su único argumento.

Los puntos no diferenciables también pueden decidir el resultado.

Ejemplo 10.9. **Una esquina interior y un máximo en el borde.** Consideremos $q\colon [0,3] \to \mathbb{R}$ definida por $q(x) = |x-1|$. La continuidad sobre $[0,3]$ asegura la existencia de ambos extremos. En $(0,1)$ su derivada es -1; en $(1,3)$ es 1; en 1 no existe derivada bilateral. No hay ceros de la derivada en los demás puntos interiores. Por tanto,

$$C_q = \{0,1,3\}, \qquad q(0) = 1, \quad q(1) = 0, \quad q(3) = 2.$$

El mínimo absoluto es 0 y se alcanza en 1, donde la derivada no existe. El máximo absoluto es 2 y se alcanza en el extremo derecho 3. Si hubiésemos buscado exclusivamente soluciones de $q'(x) = 0$, no habríamos encontrado ninguno de los dos extremos.

No diferenciar en un borde no significa que ese borde sea un punto crítico interior
Nuestra definición Definición 10.1 exige $c \in A^\circ$. Los extremos a y b de un intervalo cerrado se incorporan como candidatos *por separado*, aunque existan derivadas laterales y sean distintas de cero. La lista correcta no es «resolver $f' = 0$ y terminar», ni «llamar críticos a todos los puntos donde falta una derivada bilateral».

Qué ocurre si no tenemos las hipótesis de existencia. El teorema Teorema 5.5 es una garantía para funciones continuas **en el intervalo cerrado completo**. No podemos citarlo en un intervalo abierto, en un conjunto no acotado ni para una función discontinua sin comprobar otras propiedades. Tampoco debemos confundir «el teorema no es aplicable» con «los extremos no existen»: cada caso requiere una comprobación independiente.

Ejemplo 10.10. **Tres maneras de perder una conclusión automática.**

Un borde ausente. Para $f(x) = x$ en $(0,1)$ tenemos $0 < f(x) < 1$. El ínfimo de la imagen es 0 y el supremo es 1, pero ninguno se alcanza: no hay mínimo ni máximo absolutos. Los números 0 y 1 no son argumentos de la función.

Un dominio no acotado. El polinomio $p(x) = x^3 - 3x$ en $\mathbb{R}$ tiene los extremos locales descritos en Ejemplo 10.3, pero no tiene máximo ni mínimo absolutos. En efecto, para $x \geq 2$ se cumple $p(x) = x(x^2-3) \geq x$, de modo que sus valores crecen sin cota superior; para $x \leq -2$, la misma identidad y $x^2 - 3 \geq 1$ dan $p(x) \leq x$, y no existe cota inferior. Los extremos locales no resuelven la pregunta global.

Un borde presente sin continuidad. Sea $h\colon [0,1] \to \mathbb{R}$ dada por $h(x) = x$ si $0 \leq x < 1$ y $h(1) = 0$. Su dominio es cerrado y acotado, pero h es discontinua en 1. Tiene mínimo absoluto 0, alcanzado en 0 y 1, y **no tiene máximo**: si $x \in [0,1)$, el punto $y = (x+1)/2 \in (x,1)$ satisface $h(y) > h(x)$, mientras que para $x = 1$ basta elegir $y = 1/2$. Por tanto, ningún argumento maximiza h, aunque $\sup h([0,1]) = 1$.

Una función sobre un intervalo abierto *puede*, sin embargo, tener extremos: x^2 sobre $(-1,1)$ alcanza un mínimo absoluto en 0, aunque no alcanza máximo. Una función discontinua en un compacto también puede alcanzar los dos extremos. Los ejemplos anteriores no convierten las hipótesis de Weierstrass en condiciones necesarias para cada función particular; delimitan el alcance de su garantía universal.

¿Y si hay infinitos candidatos?

Ejemplo 10.11. **Una meseta y un conjunto infinito de ceros.** Definamos $r\colon [0,2] \to \mathbb{R}$ por

$$r(x) = \begin{cases} 0, & 0 \le x \le 1, \\ (x-1)^2, & 1 < x \le 2. \end{cases}$$

Las dos expresiones concuerdan en 1, así que r es continua en todo el intervalo. En $(0,1)$ su derivada vale 0; para $x \in (1,2)$ vale $2(x-1) > 0$. En $x = 1$ existe la derivada y vale 0, pues el cociente incremental es 0 a la izquierda e igual a h a la derecha. Así, $C_r = [0,1] \cup \{2\}$.

No existe una lista *finita* de todos los candidatos: cada punto de $(0,1)$ tiene derivada nula. Aun así, la comparación global es sencilla, porque $r(x) \ge 0$ para todo $x \in [0,2]$, con igualdad exactamente si $x \in [0,1]$, y $r(x) \le 1 = r(2)$. El mínimo absoluto es 0, alcanzado en todos los puntos de $[0,1]$; el máximo absoluto es 1, alcanzado solo en 2.

Este caso enseña qué hacer cuando la ecuación de puntos críticos produce una familia infinita: describirla como conjunto, estudiar el comportamiento por intervalos y comparar rangos de valores. No hay ningún teorema que obligue a reducir los candidatos a una enumeración finita; tampoco podemos elegir unos cuantos valores muestrales y anunciar que representan a todos los demás.

Un protocolo completo para un intervalo cerrado. Para una función $f\colon [a,b] \to \mathbb{R}$, la búsqueda de extremos absolutos requiere justificar la siguiente cadena de afirmaciones:

1. **Dominio y existencia.** Comprobar $a < b$, que la función está definida en todo $[a,b]$ y es continua allí. Si esto falla, investigar existencia o no existencia por otros argumentos, sin atribuir a Weierstrass una conclusión que no garantiza.
2. **Candidatos.** Incluir a y b; determinar los ceros de f' y los puntos interiores donde la derivada no existe. En particular, no perder esquinas ni puntos excepcionales de una definición por tramos.
3. **Suficiencia de la búsqueda.** Invocar Proposición 10.4 para probar que todos los puntos de extremo absoluto pertenecen al conjunto obtenido. Si no se ha caracterizado por completo ese conjunto, la búsqueda sigue incompleta.
4. **Comparación global.** Si la lista es finita, calcular todos sus valores y compararlos. Si es infinita, usar monotonía, desigualdades u otra información demostrada para comparar todos los candidatos. Registrar posibles empates y distinguir argumentos de valores.
5. **Conclusión tipada.** Indicar el valor máximo y mínimo, los puntos en que se alcanzan y el dominio al que se refiere la afirmación. No sustituir «absoluto» por «local» ni extrapolar el resultado a un dominio mayor.

El borde pertenece a la búsqueda, aunque no a la ecuación de Fermat
Weierstrass asegura la existencia; Fermat controla solo los extremos interiores diferenciables. Los puntos interiores no diferenciables y los extremos a, b completan el conjunto de candidatos necesarios. El criterio de signos de §10.4 puede ayudar a clasificar algunos de ellos localmente, pero la palabra *absoluto* exige una comparación sobre **todo** $[a,b]$.

Laboratorio: corregir tres pruebas incompletas. Una solución afirma: «La derivada de p se anula únicamente en -1 y 1; por eso sus valores son los extremos absolutos en $[-2,2]$». Señala exactamente qué candidatos faltan y usa Ejemplo 10.8 para reconstruir la comparación completa. Otra solución examina $q'(x)$ en $(0,3)$, comprueba que nunca vale cero y concluye que q no tiene extremos: localiza la hipótesis falsa escondida en ese razonamiento y repara la búsqueda mediante Ejemplo 10.9. Finalmente, una tercera solución escribe «f está acotada, luego

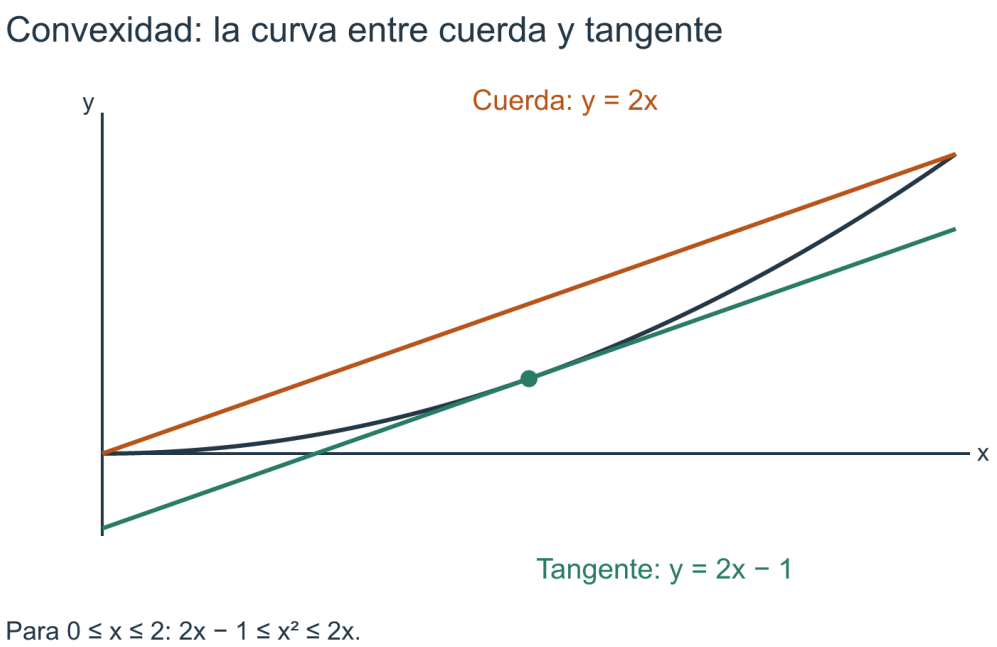


FIGURA 10.1. La parábola convexa x^2 queda por debajo de la cuerda $y = 2x$ en $[0, 2]$ y por encima de la tangente $y = 2x - 1$ en $x = 1$. Las dos comparaciones expresan propiedades distintas.

alcanza su supremo» para la identidad en $(0, 1)$: distingue el dato correcto de la conclusión falsa y proporciona el argumento preciso de no consecución.

Antes de seguir. ¿En qué línea del razonamiento se garantiza la *existencia* del punto de máximo? ¿Por qué la proposición de candidatos es necesaria, pero no clasifica individualmente cada candidato? ¿Cómo modificarías el procedimiento si el dominio fuese una unión de intervalos disjuntos? El paso a la convexidad en §10.6 no eliminará estas obligaciones: también allí comenzaremos por fijar el dominio y las desigualdades que queremos demostrar.

10.6. Convexidad como propiedad de las cuerdas

Hasta aquí hemos comparado valores de una función para averiguar si crece o decrece, o para localizar sus extremos. Ahora compararemos algo distinto: el valor de la función **entre dos puntos** con el valor que tendría allí la recta que une los puntos correspondientes de su gráfica. Esa comparación permitirá definir la convexidad sin mencionar derivadas. Los criterios mediante pendientes y derivadas vendrán después, como consecuencias que habrá que demostrar.

De dos puntos a una cuerda. Sea $I \subseteq \mathbb{R}$ un intervalo y sean $x, y \in I$. Para $0 \le t \le 1$, el número $z = (1 - t)x + ty = x + t(y - x)$ está entre x e y y, por ser I un intervalo, también pertenece a I. Los coeficientes $1 - t$ y t son no negativos y suman 1: decimos que z es una **combinación convexa** de x e y. Cuando $t = 0$ recuperamos x; cuando $t = 1$, recuperamos y; para $0 < t < 1$ y $x \neq y$, obtenemos un punto estrictamente intermedio. Los puntos de la gráfica situados sobre los extremos son $(x, f(x))$ e $(y, f(y))$. El segmento rectilíneo que los une contiene, a la altura de la abscisa z, el punto

$$\big((1-t)x+ty,\,(1-t)f(x)+tf(y)\big).$$

La pregunta geométrica tiene, por tanto, una formulación algebraica: ¿está el punto $(z, f(z))$ a la misma altura, por debajo o por encima de esa cuerda? La respuesta se expresará mediante una desigualdad válida **para todos los extremos y todas las proporciones permitidas**, no mediante la apariencia de un dibujo particular.

Definición 10.2. **Definición (convexidad, convexidad estricta y concavidad).** Sea $I \subseteq \mathbb{R}$ un intervalo y $f\colon I \to \mathbb{R}$.

Diremos que f es **convexa en** I si, para todo $x, y \in I$ y todo $t \in [0,1]$, se cumple

$$f((1-t)x+ty) \le (1-t)f(x)+tf(y).$$

Diremos que f es **estrictamente convexa en** I si la desigualdad es estricta siempre que $x \neq y$ y $0 < t < 1$. La función f es **cóncava** en I si la desigualdad de convexidad se invierte para todos esos x, y, t; es **estrictamente cóncava** si la desigualdad inversa es estricta cuando $x \neq y$ y $0 < t < 1$.

Geométricamente, una función convexa tiene su gráfica a la altura de las cuerdas o por debajo de ellas; para una función cóncava, a la altura o por encima. La expresión «por debajo» incluye la igualdad: una recta puede coincidir enteramente con su cuerda. La convexidad estricta excluye esa igualdad en los puntos estrictamente intermedios de dos abscisas distintas.

Hay tres detalles lógicos que conviene separar. Primero, las condiciones $t = 0$ y $t = 1$ dan igualdad automáticamente, sea cual sea la función: nunca deben exigirse desigualdades estrictas allí. Segundo, si $x = y$, ambos miembros coinciden para cualquier t. Tercero, la definición no presupone que f sea diferenciable ni siquiera continua en los extremos de I; esas propiedades no deben introducirse en secreto como parte de la palabra «convexa».

Leer el cuantificador antes de mirar una figura
Una sola cuerda bien situada no demuestra convexidad, porque la definición exige **todas** las parejas x, y y **todos** los $t \in [0,1]$. En cambio, un único triple admisible (x, y, t) que viole la desigualdad basta para refutarla. Tampoco alcanza con verificar algunos puntos medios de una tabla: la definición se refiere a cualquier proporción intermedia.

Dos formas equivalentes de escribir la misma propiedad. Cuando $x < y$ y elegimos $z \in [x, y]$, existe un único parámetro $t = \frac{z-x}{y-x} \in [0,1]$. Al sustituirlo en la combinación convexa, los coeficientes adquieren una forma simétrica:

$$1-t = \frac{y-z}{y-x}, \qquad t = \frac{z-x}{y-x}.$$

La ecuación de la cuerda entre las abscisas x e y puede escribirse, para $z \in [x, y]$, como $\ell_{x,y}(z) = \frac{y-z}{y-x}f(x) + \frac{z-x}{y-x}f(y)$.

Proposición 10.5. **Proposición (criterio equivalente de las cuerdas).** Sea I un intervalo y $f\colon I \to \mathbb{R}$. La función f es convexa en I si, y solo si, para cualesquiera $x < y$ de I y cualquier $z \in [x, y]$, se cumple

$$f(z) \le \ell_{x,y}(z).$$

La equivalencia también vale para la convexidad estricta si se exige $x < z < y$ y se reemplaza $\leq$ por $<$; para concavidad y concavidad estricta se invierten las desigualdades correspondientes.

Demostración. Supongamos primero que f es convexa. Dados $x < y$ y $z \in [x, y]$, tomemos $t = (z - x)/(y - x)$. Entonces $z = (1 - t)x + ty$, y la desigualdad de la definición da exactamente $f(z) \leq \ell_{x,y}(z)$.

Recíprocamente, supongamos que se cumple la desigualdad con las cuerdas. Dados $x, y \in I$ y $t \in [0, 1]$, si $x = y$ obtenemos igualdad y no hay nada que probar. Si $x < y$, aplicamos la hipótesis a $z = (1 - t)x + ty$. Si $y < x$, intercambiamos los nombres de los extremos y reemplazamos t por $1 - t$; la combinación convexa y su valor interpolado no cambian. Así se obtiene la desigualdad de la definición en todos los casos. Para las versiones estrictas, $x \neq y$ y $0 < t < 1$ equivalen a que z quede estrictamente entre las dos abscisas; el mismo argumento conserva las desigualdades estrictas. Las versiones cóncavas se demuestran invirtiendo los signos. $\square$

El denominador $y - x$ explica por qué en esta formulación exigimos $x < y$: no se puede escribir la ecuación de una cuerda mediante ese cociente cuando las abscisas coinciden. La definición mediante combinaciones convexas, en cambio, contempla sin dificultad el caso $x = y$.

Los ejemplos deben probarse, no adivinarse.

Ejemplo 10.12. **Una función afín coincide con todas sus cuerdas.** Sea $f(x) = mx + b$ en un intervalo I, con $m, b \in \mathbb{R}$. Para todo $x, y \in I$ y $t \in [0, 1]$,

$$\begin{aligned} f((1 - t)x + ty) &= m((1 - t)x + ty) + b \\ &= (1 - t)(mx + b) + t(my + b) \\ &= (1 - t)f(x) + tf(y). \end{aligned}$$

Por tanto, toda función afín es simultáneamente convexa y cóncava. Si I contiene al menos dos puntos, no es estrictamente convexa ni estrictamente cóncava: para dos abscisas distintas y $t = 1/2$ obtenemos igualdad donde se requeriría una desigualdad estricta. El signo de m no interviene; una función puede ser decreciente y convexa a la vez.

Ejemplo 10.13. **El cuadrado: una desigualdad estricta sin derivadas.** Consideremos $q(x) = x^2$ en $\mathbb{R}$. La identidad algebraica

$$(1 - t)x^2 + ty^2 - ((1 - t)x + ty)^2 = t(1 - t)(x - y)^2$$

muestra que la diferencia entre la altura de la cuerda y la altura de la gráfica es no negativa para $t \in [0, 1]$. Si $x \neq y$ y $0 < t < 1$, los tres factores del miembro derecho son positivos. En consecuencia, q es **estrictamente convexa** sobre toda la recta. La función $-q$ es estrictamente cóncava porque multiplica por -1 la desigualdad recién demostrada. La prueba revela algo que una figura no muestra por sí sola: sabemos **cuánto** separa la cuerda de la gráfica, y sabemos exactamente cuándo esa separación es cero. No hemos utilizado el valor de q' ni de q''.

Ejemplo 10.14. **El valor absoluto: convexidad sin diferenciabilidad y sin carácter estricto.** Para $f(x) = |x|$, la desigualdad triangular y la homogeneidad del valor absoluto proporcionan

$$\begin{aligned} |(1 - t)x + ty| &\leq |(1 - t)x| + |ty| \\ &= (1 - t)|x| + t|y|, \end{aligned}$$

pues t y $1-t$ son no negativos. Por lo tanto, $|x|$ es convexa en $\mathbb{R}$, aunque no sea diferenciable en 0, como vimos en Ejemplo 10.4. No es estrictamente convexa: para $x=1$, $y=2$ y $0<t<1$, todos los números implicados son positivos y se obtiene igualdad,

$$|(1-t)\cdot 1 + t\cdot 2| = (1-t)|1| + t|2|.$$

En cambio, para $x=-1$, $y=1$ y $t=1/2$ la desigualdad sí es estricta: $0<1$. Una cuerda estrictamente por encima de la gráfica **no** basta para afirmar convexidad estricta; esa propiedad exige todas las cuerdas con extremos distintos.

Ejemplo 10.15. **Una función que no es ni convexa ni cóncava en toda la recta.** Retomemos $p(x)=x^3$. Para $x=-2$, $y=0$ y $t=1/2$, el punto medio es -1 y

$$p(-1) = -1 > -4 = \frac{p(-2)+p(0)}{2}.$$

Esto refuta la desigualdad de convexidad. Por otra parte, con $x=0$, $y=2$ y $t=1/2$ obtenemos

$$p(1) = 1 < 4 = \frac{p(0)+p(2)}{2},$$

lo que refuta la desigualdad de concavidad. Así, x^3 no es convexa ni cóncava **en** $\mathbb{R}$. No concluimos de ello que sus restricciones a otros intervalos tengan necesariamente el mismo comportamiento: como ocurría con la monotonía, toda afirmación debe indicar el dominio sobre el cual se comparan los puntos.

Convexidad, concavidad y las funciones afines. La definición cóncava no constituye una teoría desconectada de la convexa. Si $g=-f$, la desigualdad

$$f((1-t)x+ty) \geq (1-t)f(x)+tf(y)$$

es equivalente, tras multiplicar ambos lados por -1 e invertir el signo, a la desigualdad de convexidad para g. En particular, f es cóncava si, y solo si, $-f$ es convexa; la misma equivalencia funciona para las versiones estrictas.

Proposición 10.6. **Proposición (convexidad y concavidad simultáneas).** Si I es un intervalo con al menos dos puntos, una función $f\colon I\to\mathbb{R}$ es convexa y cóncava a la vez si, y solo si, es afín en I: existen $m,b\in\mathbb{R}$ tales que $f(x)=mx+b$ para todo $x\in I$.

Demostración. Si f es afín, Ejemplo 10.12 verifica ambas desigualdades con igualdad. Para probar la conversa, supongamos que f es convexa y cóncava. Las desigualdades opuestas fuerzan

$$f((1-t)x+ty) = (1-t)f(x)+tf(y)$$

para todos $x,y\in I$ y $t\in[0,1]$. Escogemos $a<b$ en I y llamamos L a la única función afín que coincide con f en a y b:

$$L(z) = f(a) + \frac{f(b)-f(a)}{b-a}(z-a).$$

Si $a\leq z\leq b$, la igualdad anterior aplicada a los extremos a,b da $f(z)=L(z)$. Si $z>b$ y $z\in I$, el punto b es una combinación convexa estrictamente intermedia de a y z. Aplicando la igualdad a f y a L con los mismos coeficientes y usando $f(a)=L(a)$, $f(b)=L(b)$, obtenemos $f(z)=L(z)$ porque el coeficiente de $f(z)$ es positivo. Si $z<a$, repetimos el argumento expresando a como combinación convexa de z y b; ahora el coeficiente de $f(z)$ vuelve a ser positivo. Por tanto, $f=L$ en todo I. $\square$

Convexa no significa estrictamente convexa
La desigualdad débil admite segmentos rectos e incluso funciones enteramente afines; la estricta prohíbe igualdad en cualquier punto interior de cualquier cuerda con abscisas distintas. Las dos propiedades no deben intercambiarse. Tampoco debe confundirse convexidad con monotonía: x^2 es estrictamente convexa en $\mathbb{R}$, pero decrece en $(-\infty, 0]$ y crece en $[0, \infty)$, como se comprobó en §10.2.

El dominio y los extremos también cuentan. Si una función es convexa en I, su restricción a cualquier subintervalo $J \subseteq I$ sigue siendo convexa: cada triple admisible para J también lo era para I. La misma observación vale para convexidad estricta y para las dos formas de concavidad. La conversa no se obtiene automáticamente al reunir intervalos: faltan todas las cuerdas con extremos situados en piezas diferentes.

En esta exposición exigimos que el **dominio de convexidad sea un intervalo**. Si el dominio original es una unión disjunta de intervalos, podremos estudiar sus restricciones por separado, pero no proclamaremos convexidad en la unión utilizando nuestra definición, que requiere que las combinaciones convexas de los extremos permanezcan en el dominio. La identidad x sobre $\mathbb{R} \setminus \{0\}$ puede satisfacer igualdades de interpolación siempre que los tres puntos estén definidos; sin embargo, el segmento entre -1 y 1 no pertenece al dominio. No confundiremos esa propiedad parcial con convexidad sobre un intervalo. Hay además una advertencia acerca de los extremos incluidos. Definamos $h\colon [0,1] \to \mathbb{R}$ por

$$h(x) = \begin{cases} 1, & x = 0 \text{ o } x = 1, \\ 0, & 0 < x < 1. \end{cases}$$

Esta función es convexa. En efecto, si $z = (1-t)x + ty$ está en $(0,1)$, entonces $h(z) = 0$ y el miembro derecho de la desigualdad es una combinación de números no negativos. Si z es uno de los extremos, la combinación solo puede producirlo con los puntos implicados situados en ese mismo extremo o con un coeficiente igual a 0 o 1; en cualquiera de esos casos hay igualdad. No obstante, h es discontinua en 0 y en 1. La convexidad **en un intervalo cerrado** no implica, por sí sola, continuidad en los extremos. Los resultados diferenciales de las secciones siguientes indicarán explícitamente sus hipótesis adicionales; no las deduciremos de un dibujo.

Laboratorio: tres diagnósticos antes de derivar. Una persona observa que el punto medio de una cuerda de f queda bajo la cuerda y escribe: «f es convexa». Explica por qué la conclusión no se desprende de una sola pareja de extremos y reconstruye los cuantificadores de Definición 10.2. Otra persona presenta $|x|$, advierte que es estricta la desigualdad para los extremos -1 y 1, y concluye «convexidad estricta»: utiliza Ejemplo 10.14 para localizar el error. Una tercera afirma que «toda función convexa tiene derivada segunda positiva»; identifica qué propiedad ni siquiera se exigió en la definición y qué ejemplo de esta sección impide justificar esa frase sin hipótesis nuevas.

Antes de seguir. Si una cuerda une $(x, f(x))$ y $(y, f(y))$, ¿por qué el cociente $(f(y) - f(x))/(y - x)$ resulta natural para estudiar su inclinación? ¿Qué desigualdades entre pendientes de distintas cuerdas podemos deducir de que una gráfica convexa permanezca bajo ellas? §10.7 responderá esas preguntas y permitirá llegar a las derivadas **desde** la definición geométrica, no al revés.

10.7. Pendientes secantes, tangentes y primera derivada

En §10.6 definimos la convexidad mediante una comparación de *alturas*: la gráfica queda bajo la cuerda que une dos de sus puntos. Ahora convertiremos esa comparación en una relación de *pendientes*. Solo después pasaremos al límite que define la derivada. El orden de construcción importa: la convexidad se definió sin pedir diferenciabilidad, y el criterio diferencial debe deducirse de ella, no reemplazarla.

Tres puntos y una cadena de pendientes. Para $u < v$ de un intervalo I, escribamos $S_f(u, v) = \frac{f(v)-f(u)}{v-u}$. Este número es la pendiente de la cuerda entre $(u, f(u))$ y $(v, f(v))$. No es una derivada: está definido para cualquier función real sobre los dos argumentos, con independencia de que existan límites de cocientes incrementales.

Proposición 10.7. **Proposición (criterio de las pendientes secantes).** Sea $I \subseteq \mathbb{R}$ un intervalo que contiene al menos tres puntos y $f\colon I \to \mathbb{R}$. La función f es convexa en I si, y solo si, para toda terna $x < y < z$ de I se cumple

$$S_f(x, y) \le S_f(x, z) \le S_f(y, z).$$

La convexidad estricta equivale a las dos desigualdades estrictas para toda terna $x < y < z$. Para la concavidad se invierten ambos signos, y lo mismo ocurre, con signos estrictos, para la concavidad estricta.

Demostración. Fijemos $x < y < z$. Puesto que

$$y = \frac{z-y}{z-x}x + \frac{y-x}{z-x}z,$$

la desigualdad de convexidad de Definición 10.2 para los extremos x, z dice

$$(z-x)f(y) \le (z-y)f(x) + (y-x)f(z).$$

Restando y agrupando términos, esta desigualdad es equivalente, respectivamente, a

$$\frac{f(y)-f(x)}{y-x} \le \frac{f(z)-f(x)}{z-x}$$

y a

$$\frac{f(z)-f(x)}{z-x} \le \frac{f(z)-f(y)}{z-y}.$$

Los denominadores son positivos; por ello ninguna división invierte los signos. Las dos desigualdades de pendientes equivalen, a su vez, a la misma desigualdad de la cuerda. Si f es convexa, esta se cumple para toda terna y obtenemos la cadena anunciada.

Para la conversa, dados $x < z$ en I y un punto $y \in (x, z)$, la primera desigualdad de pendientes de la hipótesis equivale a $f(y) \le \ell_{x,z}(y)$. En los extremos $y = x, z$ hay igualdad. La proposición Proposición 10.5 permite concluir que f es convexa. Si la convexidad es estricta, la desigualdad de la cuerda es estricta para cada punto intermedio, y los mismos pasos dan las dos desigualdades estrictas; recíprocamente, una cadena estricta implica la desigualdad estricta de la cuerda. Las afirmaciones cóncavas se obtienen aplicando lo probado a $-f$. □

Hay que leer los cuantificadores: no basta encontrar tres puntos cuya cadena sea correcta. Para demostrar convexidad por esta vía debemos tomar *cualquier* terna ordenada del intervalo. Tampoco se ha invocado f'' ni continuidad de f. La pendiente de una cuerda larga queda entre

las pendientes de sus dos partes; esta es una propiedad geométrica que ya sabemos demostrar algebraicamente.

La cadena no compara intervalos desconectados
La condición $x < y < z$ se impone dentro de un mismo intervalo I. Si el dominio original tiene un hueco, la definición de convexidad adoptada en Definición 10.2 no permite llamar convexa a la unión de sus componentes. Las desigualdades válidas por separado no justifican una afirmación sobre cuerdas que atraviesen puntos ausentes.

Acercar los extremos de una cuerda al punto de tangencia. Si c es interior de I y f es diferenciable en c, los cocientes $S_f(c, c+h)$, cuando $h > 0$ y $c + h \in I$, tienden a $f'(c)$ al hacer $h \to 0^+$. Los cocientes $S_f(c+h, c)$ para $h < 0$ tienden al mismo número por la definición bilateral de derivada. La proposición anterior permite situar ese límite entre las pendientes de las cuerdas que llegan a c desde puntos fijos. Para $x < c < y$ en I, la convexidad implica $S_f(x, c) \le f'(c) \le S_f(c, y)$. En efecto, para $c < u < y$ tenemos $S_f(c, u) \le S_f(c, y)$ por Proposición 10.7; hacemos $u \to c^+$ para obtener la segunda desigualdad. Para $x < u < c$, la misma proposición da $S_f(x, c) \le S_f(u, c)$; haciendo $u \to c^-$ obtenemos la primera. Los límites existen precisamente porque supusimos diferenciabilidad en c; no se ha supuesto continuidad de f'.

La desigualdad de la izquierda, multiplicada por $c - x > 0$, y la de la derecha, multiplicada por $y - c > 0$, dicen respectivamente

$$f(x) \ge f(c) + f'(c)(x - c), \qquad f(y) \ge f(c) + f'(c)(y - c).$$

Así aparece una recta especial: la tangente en c es una **recta de apoyo**, porque ninguna otra altura de la gráfica puede quedar por debajo de ella. Conviene demostrar que esta descripción también permite recuperar la definición inicial.

Proposición 10.8. **Proposición (criterio de la tangente de apoyo).** Sea I un intervalo con interior no vacío. Supongamos que $f\colon I \to \mathbb{R}$ es diferenciable en cada punto de I°. Entonces las siguientes condiciones son equivalentes:

1. f es convexa en I.
2. Para todo $c \in I^\circ$ y todo $x \in I$ se cumple

$$f(x) \ge f(c) + f'(c)(x - c).$$

La recta $T_c(x) = f(c) + f'(c)(x - c)$ se llama aquí **tangente de apoyo en** c. No se requiere continuidad de f en los extremos de I para esta equivalencia.

Demostración. Si f es convexa, la desigualdad se acaba de demostrar para $x < c$ y para $x > c$; cuando $x = c$ hay igualdad. La demostración también abarca los extremos de I que pertenezcan al dominio, porque las desigualdades de cuerdas los incluyen.

Recíprocamente, supongamos que todas las tangentes interiores son de apoyo. Elijamos $x < y$ en I y $0 < t < 1$, y definamos $z = (1 - t)x + ty$. Como $x < z < y$, el punto z es interior de I. Aplicamos la hipótesis en $c = z$ a los dos extremos:

$$f(x) \ge f(z) + f'(z)(x - z),$$
$$f(y) \ge f(z) + f'(z)(y - z).$$

Multiplicando la primera desigualdad por $1 - t$ y la segunda por t, y sumando, obtenemos

$$(1-t)f(x)+tf(y)\geq f(z)+f'(z)((1-t)x+ty-z)=f(z).$$

La expresión entre paréntesis es cero por la elección de z. Si $t=0$, $t=1$ o $x=y$, la desigualdad de convexidad es una igualdad inmediata. Por Definición 10.2, f es convexa. □

El alcance es preciso: en un punto donde no existe $f'(c)$ no hemos definido una *tangente mediante la derivada*. La función $|x|$ sigue siendo convexa por Ejemplo 10.14, aunque no posee esa tangente en 0. La proposición afirma una equivalencia **bajo la diferenciabilidad interior**, no que la convexidad implique derivabilidad.

La derivada como criterio y no como definición. Ahora podemos precisar la afirmación anunciada en el diseño del capítulo. Una derivada no decreciente representa una pendiente tangente que no disminuye al avanzar por el intervalo. Para que esa información interior controle también los valores en los extremos incluidos, necesitaremos continuidad allí.

Teorema 10.3. **Teorema (convexidad y monotonía de la primera derivada).** Sea $I\subseteq\mathbb{R}$ un intervalo con al menos dos puntos. Supongamos que $f\colon I\to\mathbb{R}$ es continua en I —con continuidad relativa en sus extremos incluidos— y diferenciable en I°. Entonces

$$f \text{ es convexa en } I \quad\Longleftrightarrow\quad f' \text{ es no decreciente en } I^\circ.$$

Bajo las mismas hipótesis,

$$f \text{ es estrictamente convexa en } I \quad\Longleftrightarrow\quad f' \text{ es estrictamente creciente en } I^\circ.$$

Las afirmaciones correspondientes para concavidad y concavidad estricta sustituyen *no decreciente* y *estrictamente creciente* por *no creciente* y *estrictamente decreciente*, respectivamente.

Demostración. Supongamos primero que f es convexa. Dados $u<v$ interiores, la proposición Proposición 10.8 aplicada en u y en v produce

$$f(v)\geq f(u)+f'(u)(v-u),$$
$$f(u)\geq f(v)+f'(v)(u-v).$$

Al dividir por $v-u>0$ y ordenar ambas desigualdades resulta

$$f'(u)\leq S_f(u,v)\leq f'(v),$$

de donde $f'(u)\leq f'(v)$. Esto prueba que f' es no decreciente en todo el interior. En sentido inverso, supongamos que f' es no decreciente en I°. Para $c\in I^\circ$ y $x>c$ en I, la continuidad de f en $[c,x]$ y su diferenciabilidad en (c,x) permiten aplicar Lagrange (Teorema 9.3). Existe $u\in(c,x)$ tal que

$$f(x)-f(c)=f'(u)(x-c)\geq f'(c)(x-c).$$

Para $x<c$ obtenemos, por Lagrange en $[x,c]$, un $v\in(x,c)$ con

$$f(x)-f(c)=f'(v)(x-c)\geq f'(c)(x-c).$$

En la última desigualdad hemos utilizado $f'(v)\leq f'(c)$ y el signo negativo de $x-c$. Para $x=c$ hay igualdad. Queda demostrada la condición de todas las tangentes de apoyo; Proposición 10.8 implica convexidad en I. Para la parte estricta, supongamos que f es estrictamente convexa. Ya sabemos que f' es no decreciente. Si existieran $u<v$ interiores con $f'(u)=f'(v)$, la

monotonía obligaría a que f' fuese constante en $[u, v]$. Por la proposición de derivada nula y constancia (Proposición 9.2) aplicada a $f(x) - f'(u)x$ sobre $[u, v]$, f sería afín en ese segmento. Esto contradice la desigualdad estricta de Definición 10.2 para sus dos extremos y su punto medio. Por consiguiente, $f'(u) < f'(v)$ siempre que $u < v$.

Finalmente, supongamos que f' es estrictamente creciente. En particular es no decreciente, luego f es convexa por la equivalencia ya demostrada. Si la desigualdad de convexidad fuese igualdad para algún $x < z < y$ de I, la convexidad situaría f **a lo sumo** a la altura de la cuerda de x, y y las desigualdades de apoyo en z la situarían **por lo menos** a la altura de su tangente. La igualdad en z, junto con la prueba de Proposición 10.8, fuerza igualdad en las dos desigualdades de apoyo correspondientes a x e y: los dos términos no negativos de su suma ponderada deben anularse. Así la recta tangente en z coincide con la cuerda en ambos extremos. Para cualquier $w \in [x, y]$, la tangente de apoyo da $f(w) \geq T_z(w)$ y la desigualdad de la cuerda da $f(w) \leq \ell_{x,y}(w) = T_z(w)$. Se sigue que f es afín en $[x, y]$ y su derivada es constante en (x, y), contradiciendo el crecimiento estricto de f' en el interior. No puede darse esa igualdad: f es estrictamente convexa. Aplicar el resultado a $-f$ demuestra todas las variantes cóncavas. □

La continuidad en los bordes es una hipótesis, no una consecuencia
Una función convexa sobre un intervalo cerrado puede ser discontinua en sus extremos, como vimos al final de §10.6. La monotonía de f' en el interior, por sí sola, tampoco controla valores arbitrarios asignados en esos extremos. La equivalencia de Teorema 10.3 **incluye** la continuidad en todo I; para un intervalo abierto, esta queda garantizada por la diferenciabilidad en cada uno de sus puntos. No exigimos nunca que la función derivada f' sea continua.

Dos ejemplos: verificar el criterio y probar su frontera.

Ejemplo 10.16. **El cuadrado y la distancia exacta a su tangente.** Para $q(x) = x^2$ en $\mathbb{R}$, $q'(x) = 2x$ es estrictamente creciente. El teorema anterior recupera su convexidad estricta, que en Ejemplo 10.13 probamos sin derivadas. En cualquier $c \in \mathbb{R}$, su tangente de apoyo es $T_c(x) = c^2 + 2c(x - c)$, y la diferencia se factoriza exactamente: $q(x) - T_c(x) = (x - c)^2 \geq 0$.

La desigualdad es estricta para $x \neq c$. Así relacionamos tres descripciones de la misma propiedad: separación respecto de todas las cuerdas, crecimiento de la primera derivada y posición sobre cada tangente. Esta tercera prueba utiliza diferenciabilidad; la prueba inicial por cuerdas no la necesitaba.

Ejemplo 10.17. **Derivada interior creciente, pero función no convexa por un borde discontinuo.** Definamos $r\colon [0, 1] \to \mathbb{R}$ mediante

$$r(0) = -1, \qquad r(x) = x^2 \quad (0 < x \leq 1).$$

Para todo $x \in (0, 1)$ existe $r'(x) = 2x$, estrictamente creciente. Pero r no es continua en 0. La cuerda entre 0 y 1 tiene altura 0 en $1/2$, mientras que

$$r(1/2) = \frac{1}{4} > 0 = \frac{r(0) + r(1)}{2}.$$

Hemos exhibido una violación efectiva de la convexidad, no solo la ausencia de una hipótesis. El criterio no es aplicable porque falta continuidad relativa en 0. De hecho, la tangente en $c = 1/2$ tendría altura $-1/4$ en $x = 0$, mayor que $r(0) = -1$: tampoco es recta de apoyo de toda la gráfica.

La función $|x|$ proporciona una frontera distinta: su convexidad está demostrada, pero la derivada bilateral en 0 no existe. No podemos utilizar Teorema 10.3 sobre toda $\mathbb{R}$ porque allí falla su hipótesis de diferenciabilidad; sí podemos recuperar la convexidad mediante la definición de §10.6, sin forzar una derivada inexistente.

Laboratorio: qué prueba cada criterio. Una primera argumentación afirma: «f' es creciente en $(0,1)$; por tanto, f es convexa en $[0,1]$». Explica qué continuidad de borde falta y refuta el razonamiento con Ejemplo 10.17. Una segunda argumentación dice: «f es convexa; por tanto, tiene tangente en cada punto». Identifica la hipótesis no garantizada y contrasta Ejemplo 10.14. Una tercera persona conoce que f' es no decreciente y concluye que f es *estrictamente* convexa: prueba por qué las funciones afines de Ejemplo 10.12 contradicen ese salto y formula la condición estricta correcta de Teorema 10.3.

Antes de seguir. ¿En qué paso usamos que $y - x > 0$ al comparar pendientes? ¿Por qué la desigualdad de la tangente debe valer para *todo* argumento del intervalo y no solo cerca de c? ¿Por qué ninguna de las pruebas anteriores requiere continuidad de f'? En §10.8 estudiaremos qué cambia cuando f' es, a su vez, diferenciable: el signo de f'' proporcionará nuevos criterios, pero no sustituirá el fundamento por cuerdas establecido aquí.

10.8. Segunda derivada y pruebas de curvatura

En §10.7 demostramos que, bajo continuidad en el intervalo y diferenciabilidad interior, una función es convexa precisamente cuando su primera derivada es no decreciente. Esto sugiere una pregunta nueva: si también podemos derivar f', ¿podemos reconocer esa monotonía mediante el signo de su derivada? La respuesta será afirmativa, pero no introduciremos una definición nueva de convexidad: el punto de partida sigue siendo la desigualdad de las cuerdas de Definición 10.2.

Qué significa derivar por segunda vez. Sea c un punto interior del dominio de f. Para hablar de $f''(c)$ necesitamos que f' esté definida en algún intervalo abierto alrededor de c y que exista, como número real finito, el límite

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

La segunda derivada mide la variación local de la *pendiente tangente*. Su existencia en c implica que f' es continua en c, pero no que f'' sea continua allí, ni siquiera que esté definida en los puntos vecinos. Por eso distinguiremos cuidadosamente dos clases de enunciados: los criterios de convexidad sobre **todo un intervalo**, que requieren información sobre f'' en cada punto interior, y las pruebas de extremos en **un solo punto**, para las cuales bastará la existencia de f'' en ese punto junto con la diferenciabilidad de f en una vecindad.

Derivar dos veces no exige una derivada segunda continua
La escritura $f''(c)$ designa un límite de cocientes incrementales de f'. No contiene la hipótesis adicional «f'' es continua en c». Si necesitamos continuidad de alguna función, debemos indicarla y justificarla; para los resultados siguientes bastará la continuidad de f en el intervalo considerado y la diferenciabilidad que se especifica en cada enunciado.

El signo de f'' como consecuencia de la teoría de las cuerdas.

Teorema 10.4. **Teorema (criterio de convexidad por la segunda derivada).** Sea $I \subseteq \mathbb{R}$ un intervalo con al menos dos puntos. Supongamos que $f\colon I \to \mathbb{R}$ es continua en todo I y que f' y f'' existen como números reales finitos en cada punto de I°. Entonces

$$f \text{ es convexa en } I \iff f''(x) \geq 0 \quad \text{para todo } x \in I^\circ.$$

Además, la condición $f''(x) > 0$ para todo $x \in I^\circ$ implica que f es estrictamente convexa en I. Análogamente, f es cóncava en I si, y solo si, $f'' \leq 0$ en el interior; la condición $f'' < 0$ en todo el interior garantiza concavidad estricta.

Demostración. Como f' es diferenciable en cada punto de I°, es continua allí. Si f es convexa, Teorema 10.3 asegura que f' es no decreciente en I°. Apliquemos la proposición de signo necesario para funciones monótonas (Proposición 10.1) a la función $g = f'$: para cualquier $c \in I^\circ$ se obtiene

$$f''(c) = g'(c) \geq 0.$$

Recíprocamente, supongamos $f''(c) \geq 0$ para todo $c \in I^\circ$. La función $g = f'$ es continua y diferenciable en el intervalo abierto I°, y satisface $g' = f'' \geq 0$. El criterio de monotonía Teorema 10.1, aplicado a g en ese intervalo, prueba que f' es no decreciente. Ahora Teorema 10.3 devuelve la convexidad de f en **todo** I, incluidos sus extremos cuando pertenecen al dominio: para este último paso empleamos explícitamente la continuidad de f en I.

Si $f'' > 0$ en cada punto interior, la versión estricta de Teorema 10.1 muestra que f' es estrictamente creciente. La equivalencia estricta de Teorema 10.3 implica entonces que f es estrictamente convexa. Todas las afirmaciones sobre concavidad se obtienen aplicando los resultados ya probados a $-f$, cuya segunda derivada es $-f''$. □

El teorema muestra exactamente dónde intervienen las hipótesis. La existencia de f'' en todo el interior nos permite aplicar un criterio de monotonía a f'; la continuidad de f en los extremos incluidos permite trasladar esa conclusión a las cuerdas que llegan hasta ellos. No exigimos que f'' sea continua. Tampoco deducimos convexidad sobre una unión de intervalos separados: cada aplicación necesita un único intervalo en el que las cuerdas estén definidas.

Una segunda derivada positiva no repara un borde discontinuo
En Ejemplo 10.17, la función $r(0) = -1$ y $r(x) = x^2$ para $0 < x \leq 1$ satisface $r''(x) = 2 > 0$ para todos los puntos interiores de $[0, 1]$, pero no es convexa sobre el cerrado: la cuerda entre los extremos incumple la desigualdad. La continuidad de r en 0 falta. No basta calcular f'' en el interior y olvidar los valores de borde.

Positividad estricta: suficiente, pero no necesaria. Conviene distinguir dos proposiciones que se parecen solo en su formulación verbal:

$$f'' > 0 \text{ en } I^\circ \implies f \text{ estrictamente convexa en } I,$$

pero la flecha inversa es falsa. La condición exacta resulta de combinar el criterio anterior con la caracterización del crecimiento estricto que ya demostramos en §10.3.

Proposición 10.9. **Proposición (convexidad estricta cuando f'' puede anularse).** Bajo todas las hipótesis de Teorema 10.4, la función f es estrictamente convexa en I si, y solo si, se cumplen simultáneamente las dos condiciones siguientes:

1. $f''(x) \geq 0$ para todo $x \in I^\circ$;
2. no existe ningún intervalo abierto no vacío $J \subseteq I^\circ$ en el cual f'' sea idénticamente nula.

La formulación simétrica caracteriza la concavidad estricta mediante $f'' \leq 0$ y la ausencia de intervalos abiertos donde f'' sea idénticamente cero.

Demostración. Supongamos que f es estrictamente convexa. Por Teorema 10.3, f' es estrictamente creciente; en particular, es no decreciente. Aplicando Proposición 10.1 a f' obtenemos $f'' \geq 0$. Si f'' fuese idénticamente nula en un intervalo abierto no vacío J del interior, Proposición 9.2 aplicada a f' en los segmentos cerrados contenidos en J mostraría que f' es constante en J. Esto contradice su crecimiento estricto.

Para la conversa, supongamos las dos condiciones. Como f' es continua y diferenciable en el intervalo abierto I°, podemos aplicarle Proposición 10.2: su derivada, que es f'', es no negativa y no se anula idénticamente en ningún intervalo abierto no vacío. Por tanto, f' es estrictamente creciente en I°. La equivalencia estricta de Teorema 10.3 demuestra que f es estrictamente convexa en I. Sustituir f por $-f$ proporciona el caso cóncavo. □

La proposición no afirma que los ceros de f'' tengan que ser aislados. Esa sería una hipótesis adicional y más fuerte de lo necesario. Lo que descarta es **un tramo entero** de segunda derivada nula: allí f' sería constante y f sería afín, lo cual impediría la convexidad estricta.

Ejemplo 10.18. **Una función estrictamente convexa cuya segunda derivada se anula.** Sea $q\colon \mathbb{R} \to \mathbb{R}$, $q(x) = x^4$. Sus derivadas son

$$q'(x) = 4x^3, \qquad q''(x) = 12x^2 \geq 0.$$

El único cero de q'' es 0; en consecuencia, q'' no es idénticamente nula en ningún intervalo abierto no vacío. Por Proposición 10.9, q es estrictamente convexa en toda $\mathbb{R}$, a pesar de que $q''(0) = 0$. También podemos comprobar directamente que $q'(x) = 4x^3$ es estrictamente creciente, lo que permite aplicar Teorema 10.3. Este ejemplo prueba que $f'' > 0$ en todo el intervalo es una **condición suficiente**, no necesaria, para la convexidad estricta. La anulación en un punto no crea por sí sola un segmento recto.

Un punto crítico: el segundo criterio de extremos. El signo de la segunda derivada sobre un intervalo informa de convexidad. Un dato distinto es conocer $f''(c)$ **solo en un punto** donde $f'(c) = 0$. ¿Puede ese dato clasificar el punto crítico? La clave es volver a la definición de derivada, esta vez aplicada a f', y obtener signos de la primera derivada a ambos lados de c.

Teorema 10.5. **Teorema (segundo criterio de la derivada para extremos locales).** Sean $A \subseteq \mathbb{R}$, $c \in A^\circ$ y $f\colon A \to \mathbb{R}$. Supongamos que f es diferenciable en algún intervalo abierto $(c - r, c + r) \subseteq A$, que $f'(c) = 0$ y que existe la segunda derivada finita $f''(c)$. Entonces:

- si $f''(c) > 0$, f tiene en c un mínimo local estricto;
- si $f''(c) < 0$, f tiene en c un máximo local estricto.

No se exige que f'' esté definida ni sea continua en los demás puntos del intervalo.

Demostración. Por hipótesis,

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

Si $f''(c) > 0$, la definición de límite asegura que existe $0 < \delta < r$ tal que, para todo x con $0 < |x - c| < \delta$,

$$\frac{f'(x) - f'(c)}{x - c} > 0.$$

Como $f'(c) = 0$, multiplicar por $x - c$ —atendiendo a su signo— da

$$f'(x) < 0 \quad (c-\delta < x < c), \qquad f'(x) > 0 \quad (c < x < c+\delta).$$

La diferenciabilidad de f en la vecindad garantiza su continuidad en c y en los puntos laterales. Se cumplen, pues, todas las hipótesis del primer criterio Teorema 10.2: el signo pasa de negativo a positivo y c es un mínimo local estricto. Si $f''(c) < 0$, el cociente anterior es negativo para todos los $x \neq c$ suficientemente próximos. Puesto que $f'(c) = 0$, tenemos entonces $f'(x) > 0$ a la izquierda y $f'(x) < 0$ a la derecha. El mismo primer criterio prueba que c es un máximo local estricto. □

Observemos que el argumento no aplica el teorema de la segunda derivada *a todo el intervalo*. Solo utiliza que el cociente incremental de f' tiene un límite estrictamente positivo o negativo **en** c, y transforma esa información en los signos de f' que requiere el primer criterio. Así evitamos presuponer continuidad de f'' o siquiera su existencia cerca del punto.

Antes de aplicar el segundo criterio
Comprueba que c sea interior, que f sea diferenciable en una vecindad, que $f'(c) = 0$ y que exista $f''(c)$ como número real. Si esta última derivada es estrictamente positiva o negativa, el criterio clasifica un extremo **local estricto**. No demuestra por sí solo que sea un extremo absoluto sobre todo el dominio; esa pregunta requiere las comparaciones de §10.5.

Ejemplo 10.19. **Una segunda derivada positiva y un mínimo que también es absoluto.** Definamos

$$p(x) = (x-1)^2 + (x-1)^4 \qquad (x \in \mathbb{R}).$$

Las derivadas son

$$p'(x) = 2(x-1) + 4(x-1)^3, \qquad p''(x) = 2 + 12(x-1)^2 > 0.$$

En particular, $p'(1) = 0$ y $p''(1) = 2 > 0$: el segundo criterio demuestra que 1 es un mínimo local estricto. Como $p'' > 0$ en todo $\mathbb{R}$, Teorema 10.4 también demuestra que p es estrictamente convexa. Sin embargo, la conclusión **absoluta** necesita su propia comprobación: para cualquier $x \neq 1$ se cumple

$$p(x) = (x-1)^2(1 + (x-1)^2) > 0 = p(1).$$

Así, p tiene un mínimo absoluto estricto en 1. Hemos separado deliberadamente los dos argumentos: el criterio puntual prueba lo local; la desigualdad global prueba lo absoluto.

Cuando $f''(c) = 0$, el criterio guarda silencio. Si $f'(c) = 0$ y $f''(c) = 0$, el cociente incremental de f' tiende a cero. De ahí **no se obtiene** un signo uniforme para f' a la izquierda o a la derecha. Los tres casos de Ejemplo 10.5 proporcionan una refutación completa de cualquier clasificación automática. En $c = 0$,

Función	Segunda derivada	Comportamiento en 0
x^4	$12x^2$	Mínimo local estricto.
$-x^4$	$-12x^2$	Máximo local estricto.
x^3	$6x$	Ningún extremo local.

En los tres casos, $f'(0) = f''(0) = 0$. La primera y la segunda derivadas tienen, por tanto, **exactamente los mismos valores puntuales**, pero los resultados sobre extremos son incompatibles. En tal situación debemos regresar al signo de f' en intervalos laterales o comparar directamente $f(x)$ con $f(c)$; no hemos obtenido una nueva regla de clasificación a partir de un cero.

Tampoco conviene invertir el segundo criterio: el mínimo estricto de x^4 en 0 no obliga a que $f''(0) > 0$. Más aún, el criterio no clasifica puntos donde $f'(c) \neq 0$, por grande que sea la segunda derivada. Para $g(x) = x + x^2$ tenemos $g''(0) = 2 > 0$, pero $g'(0) = 1$ y hay valores de g tanto menores como mayores que $g(0)$ arbitrariamente cerca de 0.

El carácter puntual del teorema no es una continuidad oculta.

Ejemplo 10.20. **Un mínimo detectable aunque f'' cambie de signo arbitrariamente cerca.** Usaremos el mismo polinomio por tramos B y los intervalos $I_n = [a_n, a_n + a_n^3]$ de Ejemplo 10.6. Definamos una perturbación par:

$$v(x) = \begin{cases} a_n^5 B\Big(\dfrac{|x| - a_n}{a_n^3}\Big), & |x| \in I_n, \\ 0, & |x| \notin \bigcup_{n\geq 2} I_n, \end{cases} \qquad u(x) = x^2 + v(x).$$

El valor $v(0)$ es cero. En un intervalo de soporte, $|x| \geq a_n$, y $0 \leq v(x) \leq a_n^5/64$; por tanto $|v(x)/x| \leq a_n^4/64 \to 0$. Fuera de los soportes el cociente es cero. Existe $v'(0) = 0$, y entonces $u'(0) = 0$. Como B' es un polinomio sobre $[0,1]$, hay una constante M tal que $|B'| \leq M$. En los soportes no nulos, la regla de la cadena da $|v'(x)| \leq M a_n^2$, y, en consecuencia,

$$\left|\frac{v'(x) - v'(0)}{x}\right| \leq M a_n \longrightarrow 0.$$

En los huecos $v' = 0$. En los extremos de los soportes, $B = B' = B'' = 0$ garantiza los empalmes hasta la segunda derivada. Así $v''(0) = 0$ y $u''(0) = 2$. El segundo criterio puntual garantiza un mínimo estricto; también lo vemos directamente porque $u(x) = x^2 + v(x) \geq x^2 > 0 = u(0)$ para $x \neq 0$. Sin embargo, en el interior de cada soporte positivo,

$$v''(x) = \frac{1}{a_n} B''\Big(\frac{x - a_n}{a_n^3}\Big).$$

El cálculo polinómico da $B''(1/2) = -3/8$ y $B''(1/10) = 297/1000$. Para $r_n = a_n + a_n^3/2$ y $s_n = a_n + a_n^3/10$ resulta

$$u''(r_n) = 2 - \frac{3}{8a_n} < 0 \quad \text{para } n \text{ grande}, \qquad u''(s_n) = 2 + \frac{297}{1000 a_n} > 0.$$

La paridad produce los mismos valores en $-r_n$ y $-s_n$. Hay ambos signos arbitrariamente cerca del origen en cada lado, y u'' no es continua en cero. El segundo criterio **puntual** sigue siendo válido: no exige continuidad de la segunda derivada.

Este último ejemplo distingue dos métodos que no deben mezclarse. Para garantizar convexidad en un intervalo usaríamos Teorema 10.4 y necesitaríamos un signo de f'' en **todo** su interior. Para clasificar un extremo en c usamos Teorema 10.5 y solo necesitamos un valor no nulo de $f''(c)$, además de las hipótesis de diferenciabilidad en torno al punto. Una función puede pasar esta segunda prueba sin satisfacer la primera en ninguna vecindad completa del punto.

Laboratorio: reparar el diagnóstico de la curvatura. Considera cuatro afirmaciones y localiza en cada una la hipótesis que falta o la conversa indebida: «$f''(c) > 0$, luego c es un mínimo», sin comprobar $f'(c) = 0$; «$f'' > 0$ en el interior de un intervalo cerrado, luego hay convexidad en el cerrado», sin comprobar continuidad en el borde; «f es estrictamente convexa, por tanto $f'' > 0$ en todo punto», a pesar de Ejemplo 10.18; y «$f'(c) = f''(c) = 0$, luego no hay extremo», frente a Ejemplo 10.5. Para reparar cada razonamiento, escribe el dominio, los cuantificadores y el resultado preciso que autoriza la conclusión.

Antes de seguir. ¿Por qué la prueba de Teorema 10.5 deduce signos de f' sin conocer los signos de f'' alrededor del punto? ¿Por qué los ceros aislados de f'' no impiden necesariamente la convexidad estricta? ¿Qué aporta el *cambio* de convexidad a ambos lados de un punto que no proporciona el dato $f''(c) = 0$? Esta última cuestión nos conducirá a §10.9, donde definiremos rigurosamente los puntos de inflexión y auditaremos sus criterios.

10.9. Puntos de inflexión y errores de diagnóstico

La segunda derivada nos permitió reconocer convexidad y concavidad sobre intervalos. Pero decir que una gráfica tiene un *punto de inflexión* exige algo más que calcular una derivada en una abscisa: debemos demostrar que cambia su comportamiento a uno y otro lado de un punto **que pertenece al dominio**. Precisaremos primero el significado del cambio y solo entonces deduciremos qué pueden aportar las derivadas.

Una definición que excluye las falsas inflexiones. Una función afín es simultáneamente convexa y cóncava por Proposición 10.6. Por tanto, la frase «convexa a la izquierda y cóncava a la derecha» sería ambigua si aceptáramos como inflexión cualquier punto de una recta: las dos desigualdades débiles se cumplen allí con igualdad. Adoptaremos una convención deliberadamente inequívoca, basada en las versiones estrictas ya definidas en Definición 10.2.

Definición 10.3. **Definición (punto de inflexión por cambio estricto de convexidad).** Sean $A \subseteq \mathbb{R}$, $f\colon A \to \mathbb{R}$ y $c \in A^\circ$. Diremos que c **es un punto de inflexión de** f si existe $r > 0$ tal que $[c - r, c + r] \subseteq A$, f es continua en c y se cumple una de las dos alternativas:

1. f es estrictamente cóncava en $[c - r, c]$ y estrictamente convexa en $[c, c + r]$;
2. f es estrictamente convexa en $[c - r, c]$ y estrictamente cóncava en $[c, c + r]$.

Llamaremos **punto de inflexión de la gráfica** al punto $(c, f(c))$ correspondiente, para distinguirlo de su abscisa c.

La elección de las desigualdades estrictas es una convención de este libro. Hay exposiciones que admiten otras formulaciones del cambio de concavidad; al comparar resultados debemos comprobar cuál se está utilizando. Nuestra definición exige curvatura estricta en cada intervalo lateral completo, pero **no** exige tangente ni derivada segunda en c. Tampoco confunde un intervalo rectilíneo con una transición de curvatura.

La continuidad central es una hipótesis expresa: que una función tenga una forma a cada lado no autoriza a unirlas mediante un valor aislado arbitrario. Además, c es interior del dominio y $f(c)$ debe estar definido. Si el dominio excluye c, no tenemos un punto de la gráfica que pueda recibir ese nombre según la definición.

El punto no es la pareja de signos de una tabla
Una prueba de inflexión debe nombrar un intervalo izquierdo y otro derecho, justificar la convexidad estricta y la concavidad estricta en ellos, y comprobar dominio y continuidad en c. El signo de f'' puede proporcionar estas desigualdades **si se conoce en intervalos enteros**; un dato aislado en c no reemplaza esa demostración.

Un criterio suficiente: dos intervalos, dos signos. El criterio de la segunda derivada Teorema 10.4 y la caracterización estricta Proposición 10.9 permiten construir una prueba de inflexión sin tomar el valor $f''(c)$ como punto de partida.

Teorema 10.6. **Teorema (criterio de inflexión mediante signos laterales de la segunda derivada).** Sea c un punto interior del dominio de f. Supongamos que existe $r > 0$ con $[c - r, c + r]$ contenido en el dominio, que f es continua en ese segmento y que f'' existe en $(c - r, c) \cup (c, c + r)$. Entonces cada una de las siguientes condiciones es suficiente para que c sea un punto de inflexión:

1. $f''(x) < 0$ para todo $x \in (c - r, c)$ y $f''(x) > 0$ para todo $x \in (c, c + r)$;
2. $f''(x) > 0$ para todo $x \in (c - r, c)$ y $f''(x) < 0$ para todo $x \in (c, c + r)$.

No es necesario que $f''(c)$ exista ni que $f'(c)$ sea cero.

Demostración. Bajo la primera condición, consideremos $I_- = [c-r, c]$. La función es continua allí y posee segunda derivada estrictamente negativa en su interior $(c - r, c)$. La versión cóncava de Teorema 10.4 demuestra que f es estrictamente cóncava en I_-. En $I_+ = [c, c + r]$, las mismas hipótesis y la positividad de f'' demuestran convexidad estricta. Por definición Definición 10.3, c es punto de inflexión. Bajo la segunda condición se intercambian las dos conclusiones y se obtiene la alternativa restante. En ningún paso se ha diferenciado f en c ni se ha evaluado $f''(c)$. □

El carácter estricto de los signos facilita su comprobación, pero **no es indispensable**. Bajo la misma continuidad y existencia de f'' en los intervalos laterales, basta suponer $f'' \leq 0$ a la izquierda y $f'' \geq 0$ a la derecha, siempre que en *ningún intervalo abierto no vacío de cada lado* sea f'' idénticamente cero. Por Proposición 10.9, esto da concavidad y convexidad estrictas, respectivamente. La variante con signos intercambiados es análoga. Esta precisión distingue «algunos ceros» de «un tramo completo de curvatura nula»; los primeros no invalidan automáticamente una inflexión.

Un ejemplo en el que la tangente no es horizontal.

Ejemplo 10.21. **El cubo y una inflexión con pendiente no nula.** Para $p(x) = x^3$ se tiene $p''(x) = 6x$. En cualquier segmento simétrico $[-r, r]$, con $r > 0$, la segunda derivada es negativa en $(-r, 0)$ y positiva en $(0, r)$. Por Teorema 10.6, 0 es punto de inflexión. La función p es estrictamente cóncava a la izquierda y estrictamente convexa a la derecha.

Ahora consideremos $q(x) = x + x^3$. Su segunda derivada sigue siendo $q''(x) = 6x$, así que el mismo argumento muestra una inflexión en 0. Pero

$$q'(0) = 1, \qquad q(0) = 0.$$

La tangente es $y = x$, no una recta horizontal. Una inflexión no tiene por qué ser un extremo ni un punto crítico: la función q tiene derivada $1 + 3x^2 > 0$ en toda la recta y es estrictamente creciente.

Inflexión no es una clasificación de extremos
El primer y segundo criterio de extremos estudian cómo se comparan $f(x)$ y $f(c)$ cerca de c; una inflexión compara la *convexidad de restricciones laterales*. Ni la palabra «inflexión» implica $f'(c) = 0$, ni un cero de f' garantiza un cambio de convexidad. En x^3 coinciden una inflexión y una derivada nula, mientras que en $x + x^3$ hay inflexión sin derivada nula.

Qué es necesario cuando existe la segunda derivada en el punto. Podemos justificar la regla habitual $f''(c) = 0$, pero su dirección lógica y sus hipótesis deben quedar visibles.

Proposición 10.10. **Proposición (condición necesaria en un punto de inflexión dos veces diferenciable).** Si c es un punto de inflexión según Definición 10.3 y existe la derivada segunda finita $f''(c)$, entonces

$$f''(c) = 0.$$

Demostración. La existencia de $f''(c)$ significa que f' está definida en alguna vecindad $(c - \delta, c + \delta)$ y es diferenciable, por tanto continua, en c. Supongamos que f es estrictamente cóncava a la izquierda y estrictamente convexa a la derecha, como en la primera alternativa de la definición. Elegimos $0 < \rho < \min\{r, \delta\}$. Las restricciones de f a $[c - \rho, c]$ y $[c, c + \rho]$ son, respectivamente, estrictamente cóncava y estrictamente convexa; además son continuas en esos intervalos cerrados, pues f es diferenciable en una vecindad de c. Ahora sí podemos aplicar Teorema 10.3 a ambas restricciones: f' es no creciente en $(c - \rho, c)$ y no decreciente en $(c, c + \rho)$. No se ha supuesto continuidad de f en los extremos exteriores originales $c - r$ y $c + r$.

Para cada $x \in (c - \rho, c)$ y cada y con $x < y < c$, tenemos $f'(x) \geq f'(y)$. Al hacer $y \to c^-$, la continuidad de f' en c da $f'(x) \geq f'(c)$. Por consiguiente, si $h < 0$ es suficientemente pequeño,

$$\frac{f'(c + h) - f'(c)}{h} \leq 0.$$

Análogamente, la monotonía no decreciente de f' a la derecha y la continuidad en c implican $f'(c + h) \geq f'(c)$ si $h > 0$ es pequeño; entonces

$$\frac{f'(c + h) - f'(c)}{h} \geq 0.$$

Los límites laterales de esos cocientes, si existe su límite bilateral $f''(c)$, son simultáneamente no positivo y no negativo. Han de coincidir, de modo que $f''(c) = 0$. La otra alternativa de la definición invierte todos los signos y produce la misma conclusión. □

Esta proposición es **necesaria bajo existencia de** $f''(c)$, no una caracterización. No afirma que debamos calcular previamente esa segunda derivada: cuando no existe, todavía podemos probar la inflexión por las cuerdas o por los signos laterales.

Ejemplo 10.22. **Un cero que no produce inflexión.** Sea $f(x) = x^4$ sobre $\mathbb{R}$. Tenemos

$$f''(x) = 12x^2, \qquad f''(0) = 0.$$

Sin embargo, f es estrictamente convexa en toda la recta por Ejemplo 10.18, por lo que también lo es en cada intervalo lateral alrededor de 0. No puede ser estrictamente cóncava en ninguno de ellos: una función estrictamente convexa y estrictamente cóncava en un intervalo con más de dos puntos exigiría desigualdades estrictas opuestas para la misma cuerda. Así, 0 **no** es punto de inflexión. La igualdad $f''(0) = 0$ señala únicamente un candidato que debemos investigar.

La segunda derivada puede no existir en una inflexión auténtica.

Ejemplo 10.23. **Una inflexión con primera derivada, pero sin segunda derivada en el punto.** Definamos $g\colon \mathbb{R} \to \mathbb{R}$ mediante

$$g(x) = x|x| = \begin{cases} -x^2, & x < 0, \\ x^2, & x \geq 0. \end{cases}$$

La función es continua en 0 y

$$g'(0) = \lim_{h\to 0} \frac{h|h|}{h} = \lim_{h\to 0} |h| = 0.$$

Fuera de 0 obtenemos $g''(x) = -2$ para $x < 0$ y $g''(x) = 2$ para $x > 0$. El criterio Teorema 10.6 demuestra que 0 es un punto de inflexión. No obstante, para $x \neq 0$ se tiene $g'(x) = 2|x|$, y por tanto $\frac{g'(h)-g'(0)}{h} = 2\frac{|h|}{h}$. Los límites laterales son -2 y 2: $g''(0)$ **no existe**. En consecuencia, la existencia de la segunda derivada en el candidato no es necesaria para una inflexión según nuestra definición.

Podríamos construir incluso una inflexión sin primera derivada en c, empalmando de manera continua dos ramas de convexidades estrictas opuestas con pendientes laterales diferentes. La definición Definición 10.3 no exige diferenciabilidad: lo que hay que probar son las dos desigualdades de convexidad sobre intervalos laterales y la continuidad central.

Cambios de signo observados no equivalen a intervalos de signo. Existe otra trampa, más sutil que el simple cero de f'': encontrar valores positivos y negativos de f'' arbitrariamente próximos a c **no** demuestra que toda una mitad del entorno sea convexa y la otra cóncava. Las dos mitades podrían contener oscilaciones de curvatura.

Ejemplo 10.24. **Oscilación de f'' sin inflexión.** Recuperemos la perturbación $v = u - x^2$ construida en Ejemplo 10.20. Allí probamos que $v'(0) = v''(0) = 0$ y que, para $r_n = a_n + a_n^3/2$ y $s_n = a_n + a_n^3/10$,

$$v''(r_n) = -\frac{3}{8a_n} < 0, \qquad v''(s_n) = \frac{297}{1000a_n} > 0.$$

Como v es par y dos veces diferenciable, v'' es par. Los puntos $-r_n$ y $-s_n$ dan los mismos signos a la izquierda. Las cuatro sucesiones convergen a cero. En cualquier intervalo lateral abierto que termine en cero, v'' adopta valores positivos y negativos. Si v fuese convexa en uno de esos intervalos, la condición necesaria de Teorema 10.4 impondría $v'' \geq 0$ en todo su interior; si fuese cóncava, impondría $v'' \leq 0$. Ambas posibilidades quedan refutadas. Por ello, **cero no es punto de inflexión** según Definición 10.3, pese a que $v''(0) = 0$ y existen signos de ambos tipos arbitrariamente cerca.

Este ejemplo enseña a distinguir un **cambio lateral uniforme** de una mera alternancia de signos en sucesiones. Para el criterio suficiente debemos establecer cuantificadores como «para todo $x \in (c - r, c)$» y «para todo $x \in (c, c + r)$». Calcular algunos valores o trazar una curva aproximada no demuestra esas condiciones.

Un punto excluido no puede ser inflexión de la gráfica.

Ejemplo 10.25. **Una gráfica con un hueco exactamente donde cambiaría la curvatura.** Sea $h\colon \mathbb{R}\setminus\{0\}\to\mathbb{R}$, $h(x)=x^3$. Sus restricciones a $(-\infty,0)$ y $(0,\infty)$ son estrictamente cóncava y estrictamente convexa, respectivamente. No obstante, $0\notin\operatorname{dom}(h)$ y el punto $(0,h(0))$ ni siquiera está definido. Nuestra definición no permite llamar a 0 punto de inflexión de h. Si extendemos h a 0 mediante $h(0)=0$, obtenemos otra función, ahora continua allí, y la prueba de Ejemplo 10.21 sí es aplicable.

Tampoco basta que 0 pertenezca formalmente al dominio: si definiéramos $k(x)=x^3$ para $x\neq 0$ pero $k(0)=1$, se rompería la continuidad central. Las desigualdades laterales en sus intervalos *abiertos* no sustituyen la comprobación del valor en el punto central.

Protocolo de diagnóstico: primero el cambio, después las derivadas. Para estudiar un candidato c, empieza por comprobar que sea interior del dominio y que f sea continua allí. Elige un mismo entorno y demuestra, mediante cuerdas o criterios ya justificados, concavidad estricta a un lado y convexidad estricta al otro. Si decides utilizar f'', verifica su existencia y su signo **en todos los puntos de cada intervalo lateral**; puede faltar en c. Solo después clasifica el punto como inflexión.

La igualdad $f''(c)=0$, si puede calcularse, es una consecuencia necesaria de una inflexión dos veces diferenciable, no la prueba del cambio. Los ejemplos de x^4, $x|x|$ y la función oscilatoria muestran, respectivamente, un cero sin inflexión, una inflexión sin segunda derivada y una alternancia de signos que no configura ninguna de las dos mitades exigidas.

Tres proposiciones que no son equivalentes
«$f''(c)=0$», «f'' toma signos diferentes cerca de c» y «f cambia de convexidad estricta a concavidad estricta a través de c» son afirmaciones distintas. Únicamente la tercera —con dominio y continuidad comprobados— corresponde a nuestra definición de inflexión. No debe confundirse esta clasificación geométrica con los criterios de extremos locales de §§10.4 y 10.8.

Laboratorio: auditar cinco diagnósticos. Considera estas frases y determina, para cada una, qué hipótesis falta o qué implicación se invirtió: «$f''(0)=0$, luego hay inflexión»; «f'' es negativa en ciertos puntos a la izquierda y positiva en ciertos puntos a la derecha, luego hay inflexión»; «una inflexión tiene tangente horizontal»; «no existe $f''(0)$, luego no puede haber inflexión»; y «las restricciones de x^3 son cóncava y convexa a ambos lados del hueco, luego 0 es una inflexión del dominio perforado». Repara las afirmaciones verdaderamente reparables utilizando Teorema 10.6 y los ejemplos Ejemplo 10.21, Ejemplo 10.22, Ejemplo 10.23, Ejemplo 10.24 y Ejemplo 10.25; para las demás, exhibe un contraejemplo y di cuál es su función lógica.

Un signo aislado, un signo válido en todo un intervalo y un cambio de convexidad proporcionan conclusiones diferentes. Al reunir dominio, límites, monotonía, extremos e inflexiones registraremos esas conclusiones con sus ámbitos antes de representar la gráfica.

10.10. Reconstrucción de gráficas y laboratorio de pruebas

Para estudiar una función concreta reuniremos afirmaciones justificadas sobre su dominio, límites, monotonía, extremos y convexidad. El bosquejo representará ese conjunto de resultados; las relaciones entre ellos deberán respetar los intervalos y las condiciones en que cada prueba es válida.

Un estudio comienza por el dominio, no por una tabla de signos. Sea $f\colon D \to \mathbb{R}$. Antes de calcular derivadas, determinamos D y distinguimos sus intervalos componentes, los puntos excluidos y los extremos que sí pertenecen a él. Si la fórmula involucra un denominador, sus ceros quedan excluidos; si contiene una raíz real que impone restricciones, estas también pertenecen a la determinación del dominio. Una función definida por tramos exige además comprobar qué valor se asignó a los puntos de empalme.

Después registramos la información que pueda obtenerse directamente: intersecciones con los ejes cuando existan, signo de f donde se pueda demostrar, y simetrías. Para afirmar que f es par o impar, el dominio debe ser simétrico respecto de cero y deben verificarse, **para todo** $x \in D$, respectivamente $f(-x) = f(x)$ o $f(-x) = -f(x)$. La simetría permite trasladar hechos ya demostrados de una parte del dominio a otra, pero no autoriza añadir puntos que falten en él.

Los límites aportan datos sobre los extremos del dominio y sobre el comportamiento cuando $x \to \pm\infty$. Si a es un punto excluido y alguno de los límites laterales de f es infinito, una recta vertical $x = a$ describe ese comportamiento; tal información no incorpora $(a, f(a))$ a la gráfica. Si $\lim_{x\to+\infty} f(x) = L$, la distancia vertical entre la gráfica y la recta $y = L$ tiende a cero por la derecha; puede llamarse a esta última asíntota horizontal en ese extremo. Las dos direcciones del infinito deben comprobarse por separado y no determinan, por sí solas, la forma de la curva en un intervalo finito.

Qué afirma un límite y qué no afirma
La igualdad $\lim_{x\to a} f(x) = L$ no asegura que $a \in D$ ni, si pertenece, que $f(a) = L$: esta última conclusión requiere continuidad en a. Del mismo modo, una asíntota es una afirmación de límite, no una barrera que la gráfica tenga prohibido cruzar. Al estudiar una gráfica distinguimos valores efectivamente alcanzados de valores únicamente aproximados.

De los cálculos a certificados sobre intervalos. Una vez fijado el dominio, los cálculos diferenciales pueden organizarse en dos registros independientes. Para estudiar monotonía, buscamos dónde existe f' y determinamos su signo en **cada intervalo completo**; para estudiar convexidad, podemos partir de las cuerdas o, si contamos con las hipótesis apropiadas, investigar los signos de f''. Los ceros de ambas derivadas desempeñan papeles diferentes. Un cero de f' es candidato a extremo interior, pero necesita clasificación; un cero de f'' puede ser candidato a cambio de convexidad, pero no prueba una inflexión.

Tampoco debemos perder los lugares donde las derivadas no existan. Un extremo interior puede situarse en una esquina (Ejemplo 10.4), mientras que una inflexión puede existir sin segunda derivada en el punto (Ejemplo 10.23). Cuando una derivada no está definida en un punto que sí pertenece al dominio, el criterio diferencial puede aplicarse a los intervalos abiertos laterales y debe completarse con continuidad y comparaciones en el punto central.

Dato demostrado	Resultado que autoriza	Lo que todavía no autoriza
$f' > 0$ en todo el interior de un intervalo y continuidad en sus bordes incluidos	Crecimiento estricto en ese intervalo por Teorema 10.1	Crecimiento entre componentes disjuntas.
$f'(c) = 0$	Inclusión de c entre los candidatos interiores	Máximo, mínimo o inflexión en c.
$f'' > 0$ en todo el interior y continuidad en el intervalo	Convexidad estricta por Teorema 10.4	Que f sea creciente o que exista un máximo absoluto.

Dato demostrado	Resultado que autoriza	Lo que todavía no autoriza
Signos opuestos adecuados de f'' en intervalos laterales, con dominio y continuidad centrales	Inflexión por Teorema 10.6	Que $f'(c) = 0$ o que haya un extremo local.
Continuidad de f en $[a, b]$	Existencia de máximo y mínimo absolutos por Teorema 5.5	Su valor o la unicidad de los puntos donde se alcanzan.

El registro debe consignar también dónde falla cada hipótesis. Una tabla de signos obtenida mediante una factorización exacta es una prueba abreviada si se han justificado los signos de todos los factores; una tabla obtenida a partir de cinco valores numéricos es solo una exploración. No confundiremos una conclusión que vale para cualquier punto de un intervalo con una sucesión de observaciones particulares.

Un bosquejo es una consecuencia, no una demostración
El orden lógico es: dominio y continuidad $\longrightarrow$ valores y límites $\longrightarrow$ derivadas e intervalos de signo $\longrightarrow$ monotonía, extremos y convexidad $\longrightarrow$ eventual inflexión $\longrightarrow$ bosquejo. Cada flecha designa una tarea de justificación; no implica que siempre debamos usar derivadas. Cuando una desigualdad algebraica proporciona una conclusión global directamente, puede ser más informativa que la inspección del dibujo.

Primer estudio completo: un polinomio con dos mínimos y dos inflexiones.

Ejemplo 10.26. **Reconstrucción razonada de $f(x) = x^4 - 2x^2$.** Consideremos $f\colon \mathbb{R} \to \mathbb{R}$. Su dominio es toda la recta y, al ser un polinomio, es continua y diferenciable dos veces en cada punto. Además, $f(-x) = (-x)^4 - 2(-x)^2 = f(x)$, por lo que es par. La identidad $f(x) = x^2(x^2 - 2)$ localiza los ceros $-\sqrt{2}$, 0 y $\sqrt{2}$. Si $0 < |x| < \sqrt{2}$, el primer factor es positivo y el segundo negativo: $f(x) < 0$. Si $|x| > \sqrt{2}$, ambos son positivos y $f(x) > 0$. En 0 hay un cero sin cambio de signo.

Primera derivada y monotonía. Obtenemos

$$f'(x) = 4x^3 - 4x = 4x(x - 1)(x + 1).$$

Los únicos ceros son -1, 0 y 1. Los signos, justificados por los tres factores, son

Intervalo	$(-\infty, -1)$	$(-1, 0)$	$(0, 1)$	$(1, \infty)$
Signo de f'	$-$	$+$	$-$	$+$

Como f es continua también en los puntos que separan esos intervalos, Teorema 10.1 demuestra que es estrictamente decreciente en $(-\infty, -1]$ y $[0, 1]$, y estrictamente creciente en $[-1, 0]$ y $[1, \infty)$. Por Teorema 10.2, hay mínimos locales estrictos en -1 y 1, y un máximo local estricto en 0. Los valores correspondientes son

$$f(-1) = f(1) = -1, \qquad f(0) = 0.$$

La conclusión acerca de los **extremos absolutos** requiere otro argumento. Completando el cuadrado en x^2 encontramos $f(x) = (x^2 - 1)^2 - 1 \geq -1$. La igualdad se cumple si, y solo si, $x^2 = 1$. Así, -1 es el mínimo absoluto y se alcanza exactamente en -1 y 1. No hay máximo absoluto:

para $|x| \geq 2$ se tiene $f(x) = x^2(x^2-2) \geq 2x^2$, y sus valores no están acotados superiormente cuando $|x| \to \infty$. El máximo en 0 es, por tanto, únicamente local.

Segunda derivada y convexidad. Calculamos $f''(x) = 12x^2 - 4 = 4(3x^2-1)$.

Sus ceros son $c_- = -1/\sqrt{3}$ y $c_+ = 1/\sqrt{3}$. De la desigualdad $3x^2 - 1 > 0$ cuando $|x| > 1/\sqrt{3}$ y $3x^2 - 1 < 0$ cuando $|x| < 1/\sqrt{3}$, y de Teorema 10.4, concluimos que f es estrictamente convexa en $(-\infty, c_-]$ y $[c_+, \infty)$, y estrictamente cóncava en $[c_-, c_+]$. La continuidad permite incluir los extremos de esos intervalos en las conclusiones. En cada uno de los dos puntos c_- y c_+ el signo de f'' cambia sobre intervalos laterales completos. Como la función es continua en ellos, Teorema 10.6 demuestra que son abscisas de inflexión. Sus ordenadas coinciden:

$$f(c_-) = f(c_+) = \frac{1}{9} - \frac{2}{3} = -\frac{5}{9}.$$

No son extremos locales: $f'(c_\pm) = 4c_\pm(c_\pm^2 - 1) = -8c_\pm/3 \neq 0$. De hecho, ambos se encuentran dentro de intervalos de monotonía estricta ya identificados.

Límites y forma certificada. Puesto que $f(x) = x^2(x^2-2)$, se obtiene $\lim_{x\to+\infty} f(x) = \lim_{x\to-\infty} f(x) = +\infty$. El conjunto de resultados permite un bosquejo simétrico respecto del eje vertical: dos mínimos de altura -1, un máximo local central de altura 0, dos inflexiones de altura $-5/9$, tres intersecciones con el eje horizontal y ramas que ascienden sin cota a ambos extremos. Cada elemento de esa descripción está respaldado por una identidad, una desigualdad, un límite o un criterio demostrado; ninguna propiedad se ha inferido únicamente de la apariencia del bosquejo.

Una distinción que el ejemplo obliga a conservar
En $x = 0$ tenemos $f'(0) = 0$ y $f''(0) = -4 < 0$: hay un máximo local, pero **no** una inflexión. En $c_\pm$ tenemos $f''(c_\pm) = 0$ y un cambio probado de convexidad, pero $f'(c_\pm) \neq 0$ y no hay extremo. Los dos tipos de punto responden a preguntas diferentes.

Segundo estudio: la misma derivada negativa, tres componentes separadas.

Ejemplo 10.27. **Una función racional con dos huecos en el dominio.** Sea

$$g\colon D \longrightarrow \mathbb{R}, \qquad g(x) = \frac{x}{x^2-1}, \qquad D = \mathbb{R} \setminus \{-1, 1\}.$$

El dominio es la unión de tres intervalos: $(-\infty, -1)$, $(-1, 1)$ y $(1, \infty)$. Es simétrico respecto de cero y $g(-x) = -g(x)$; por tanto, g es impar. Su único cero es $x = 0$, pues el denominador allí vale -1.

Límites sin inventar puntos de la gráfica. Factorizando el denominador, $x^2 - 1 = (x-1)(x+1)$, obtenemos

$$\lim_{x\to-1^-} g(x) = -\infty, \quad \lim_{x\to-1^+} g(x) = +\infty,$$
$$\lim_{x\to1^-} g(x) = -\infty, \quad \lim_{x\to1^+} g(x) = +\infty.$$

Por ejemplo, si $x \to -1^-$, el numerador es negativo y el denominador positivo y tiende a cero; las otras tres direcciones se verifican mediante los signos de los dos factores. Las rectas $x = -1$ y $x = 1$ describen estos límites verticales, pero $g(-1)$ y $g(1)$ no existen. Dividiendo numerador y denominador por x^2 resulta

$$g(x) = \frac{1/x}{1 - 1/x^2},$$

así que $\lim_{x\to\pm\infty} g(x) = 0$. La recta $y = 0$ es una asíntota horizontal en ambos sentidos, aunque el origen $(0, 0)$ sí pertenece a la gráfica y se encuentra sobre ella. Una asíntota horizontal no es una recta que la gráfica tenga prohibido alcanzar.

Monotonía y ausencia de extremos. En todos los puntos del dominio,

$$g'(x) = -\frac{x^2 + 1}{(x^2 - 1)^2} < 0.$$

La continuidad y Teorema 10.1 permiten concluir que g es estrictamente decreciente en **cada uno de los tres intervalos componentes**. La derivada nunca se anula, no hay puntos no diferenciables dentro de D y cada componente es abierta; por Proposición 10.3, no hay extremos locales en el dominio. Los límites infinitos muestran, además, que la función no está acotada superior ni inferiormente y carece de extremos absolutos. El mismo signo de g' no autoriza un decrecimiento en toda la unión D. En efecto, $-2 < 0$ pero

$$g(-2) = -\frac{2}{3} < 0 = g(0),$$

lo que contradice la desigualdad exigida para una función no creciente sobre todo D. La tentativa de aplicar Lagrange a $[-2, 0]$ es inválida porque contiene $-1 \notin D$.

Convexidad e inflexión. Otra derivación proporciona

$$g''(x) = \frac{2x(x^2 + 3)}{(x^2 - 1)^3}.$$

Como $x^2 + 3 > 0$, el signo depende del signo de x y del denominador. Obtenemos la siguiente lectura, válida en intervalos completos:

Intervalo	$(-\infty, -1)$	$(-1, 0)$	$(0, 1)$	$(1, \infty)$
Signo de g''	$-$	$+$	$-$	$+$
Geometría demostrada	Estrictamente cóncava	Estrictamente convexa	Estrictamente cóncava	Estrictamente convexa

Como g es continua en $[-r, r]$ para cualquier $0 < r < 1$, y g'' es positiva en $(-r, 0)$ y negativa en $(0, r)$, Teorema 10.6 demuestra que 0 es punto de inflexión. Su tangente no es horizontal: $g'(0) = -1$. En cambio, ni -1 ni 1 pueden ser puntos de inflexión según Definición 10.3, porque no pertenecen al dominio.

Con esta información podemos describir tres ramas separadas, cada una estrictamente decreciente, con los cuatro comportamientos verticales y los dos horizontales ya demostrados; dentro de la rama central se produce una inflexión en el origen. Un trazo que uniese las ramas a través de $x = \pm 1$ representaría **otra** cosa y contradiciría el dominio comprobado.

Qué no decide un conjunto finito de puntos del dibujo.

Ejemplo 10.28. **Los mismos tres datos, comportamientos distintos.** Supongamos que una tabla solo proporciona $f(-1) = f(0) = f(1) = 0$. La función constante $u(x) = 0$ y el polinomio $v(x) = x(x^2 - 1)$ satisfacen exactamente esos tres datos. Sin embargo, u es constante en $\mathbb{R}$, mientras que $v(2) = 6$ y $v(1/2) = -3/8$. Sus gráficos y sus propiedades globales son diferentes. Por ello, ni una interpolación visual ni la concordancia con una tabla finita prueban monotonía, convexidad o ausencia de extremos. Para decidir cualquiera de esas propiedades necesitamos los cuantificadores de sus definiciones o las hipótesis de un criterio demostrado.

Los dos estudios completos muestran otra precaución: disponer de más información no autoriza mezclarla sin revisar sus ámbitos. En Ejemplo 10.26, el dominio es un intervalo y el signo de cada derivada controla los tramos correspondientes, con continuidad en sus extremos incluidos. En Ejemplo 10.27, el dominio está perforado y las conclusiones diferenciales se restringen a cada componente. Los límites describen el comportamiento hacia los huecos sin completarlos; el único punto de inflexión está donde efectivamente hay función y continuidad.

Laboratorio de pruebas: localizar y reparar cinco errores. Una reconstrucción rigurosa puede utilizar el siguiente control. Cada diagnóstico incorrecto debe repararse **sin dibujar primero** la gráfica.

1. «$f' = 0$ en tres puntos; por tanto, hay tres extremos». Identifica la confusión entre candidatos y extremos, y utiliza x^3 de Ejemplo 10.5 para refutar la inferencia puntual.
2. «$g' < 0$ en todo su dominio; por tanto, g decrece en toda la unión». Señala el segmento que no está contenido en D y el par de Ejemplo 10.27 que refuta la conclusión.
3. «La gráfica toca una asíntota horizontal, luego el cálculo de límites es incorrecto». Escribe la definición mediante el límite y comprueba el valor $g(0)$ del mismo ejemplo.
4. «$f''(c) = 0$, así que debo dibujar una inflexión». Contrasta la cuártica x^4 de Ejemplo 10.22 con los dos puntos de Ejemplo 10.26, donde sí hay cambio demostrado en intervalos laterales.
5. «f tiene un máximo local en 0, luego el bosquejo no puede superar esa altura». En Ejemplo 10.26, demuestra con la desigualdad de crecimiento no acotado por qué el máximo local central no es absoluto.

La entrega de un estudio de gráfica es un conjunto de pruebas
Un estudio se considera justificado cuando especifica el dominio y las continuidades utilizadas, identifica exactamente los intervalos sobre los que se afirman signos y monotonía, clasifica los extremos con el alcance local o absoluto correcto, certifica convexidad e inflexiones con sus hipótesis y consigna límites probados. Si algún paso no puede demostrarse con las herramientas disponibles, debe permanecer como conjetura de bosquejo, no convertirse silenciosamente en afirmación matemática.

Para determinar un extremo absoluto tendremos que comparar todos los candidatos admisibles; para deducir monotonía mediante el signo de la derivada necesitaremos un intervalo de comparación; para afirmar inflexión habrá que comprobar el cambio de convexidad. Los ejercicios reúnen estas decisiones y sus contraejemplos. El estudio de Ejemplo 10.26 permite comparar los candidatos a extremo absoluto; Ejemplo 10.27 muestra por qué un hueco del dominio impide comparar mediante el teorema valores situados en componentes distintas.

10.11. Ejercicios y soluciones

El objetivo del banco es justificar afirmaciones globales a partir de hipótesis comprobadas, no adivinar la gráfica. Los cuarenta ejercicios originales se distribuyen en siete niveles: $6A + 6B + 6C + 6D + 6E + 6F + 4G = 40$. Cada enunciado tiene una solución desarrollada identificada por el mismo número. Se utilizan las herramientas construidas hasta §10.10; no se requieren Taylor, integración ni procedimientos numéricos posteriores.

Nivel A — Definiciones, cuantificadores y dominios. Ejercicio A1. Dos testigos distintos. Demuestra que $f(x) = x^2$ no es monótona en $[-2, 2]$. Proporciona un par que refute el carácter no decreciente y otro que refute el carácter no creciente. Determina los dos intervalos máximos de monotonía con sentido estricto.

Ejercicio A2. La inyectividad elimina las igualdades. Sea $D \subseteq \mathbb{R}$ y $f : D \to \mathbb{R}$ no decreciente. Demuestra que f es estrictamente creciente si y solo si es inyectiva. Formula y prueba la versión para una función no creciente.

Ejercicio A3. Dos ramas no componen una regla. Para $g(x) = (x + 1)/(x - 1)$, con dominio $D = \mathbb{R} \setminus \{1\}$, determina la monotonía en cada componente usando derivadas; decide si g es no creciente en todo D y justifica la respuesta con dos puntos concretos.

Ejercicio A4. Constancia y derivada. Sea I un intervalo y $f : I \to \mathbb{R}$ continua en I y diferenciable en I°. Demuestra que f es constante en I si y solo si $f'(t) = 0$ para todo $t \in I^\circ$. Explica por qué la misma conclusión no vale para cualquier dominio desconectado.

Ejercicio A5. Un extremo discontinuo. Sea $q : [0, 1] \to \mathbb{R}$ dada por $q(0) = 2$ y $q(x) = x$ para $0 < x \leq 1$. Calcula la derivada interior; determina si q es no decreciente en $[0, 1]$ y localiza la hipótesis faltante del criterio por derivadas.

Ejercicio A6. Tres observaciones no prueban monotonía. Construye dos polinomios distintos que satisfagan $f(-1) = f(0) = f(1) = 0$, uno no decreciente en $\mathbb{R}$ y otro que no sea monótono en $[-1, 1]$. Demuestra las propiedades afirmadas sin apoyarte en un dibujo.

Nivel B — Signos de la derivada y conversas. Ejercicio B1. Positividad uniforme. Demuestra que $f(x) = x^5 + x$ es estrictamente creciente en $\mathbb{R}$. Justifica el uso del teorema del valor medio y determina si podría tener dos ceros diferentes.

Ejercicio B2. Crecer con derivada cero. Prueba que $f(x) = x^5$ es estrictamente creciente en $\mathbb{R}$ pese a $f'(0) = 0$. Identifica qué conversa del criterio $f' > 0$ queda refutada.

Ejercicio B3. Un cuadro de signos demostrado. Para $p(x) = x^3 - 3x + 1$ en $[-2, 2]$, determina los puntos donde la derivada se anula y los intervalos cerrados correspondientes de crecimiento y decrecimiento estrictos. Explica por qué el dominio entero no es un intervalo de monotonía.

Ejercicio B4. Un criterio sin falsos estrictos. Sea f continua en $[a, b]$, diferenciable en (a, b) y con $f' \geq 0$. Demuestra que f es estrictamente creciente si y solo si no existe un subintervalo abierto no vacío de (a, b) en el cual f' sea idénticamente cero.

Ejercicio B5. Un dato puntual insuficiente. Para $f(x) = x^3 - 3x$, verifica que $f'(2) > 0$, pero refuta que f sea no decreciente en $\mathbb{R}$. Indica una hipótesis suficiente que repararía la inferencia.

Ejercicio B6. Igualdad de signos, diferente monotonía global. En $D = \mathbb{R} \setminus \{0\}$ compara $u(x) = -1/x$ y $v(x) = x$. Ambas tienen derivada positiva en cada punto del dominio. Demuestra que solo una es estrictamente creciente en D y explica por qué no se contradice el teorema del valor medio.

Nivel C — Extremos interiores, bordes y candidatos. Ejercicio C1. La frontera supera al máximo interior. Estudia $p(x) = x^3 - 6x^2 + 9x$ sobre $[0, 5]$. Determina todos los puntos críticos interiores, clasifica sus extremos locales y calcula los valores y lugares de los extremos absolutos.

Ejercicio C2. Dos esquinas y una meseta. Para $q(x) = |x - 1| + |x + 1|$ en $[-2, 2]$, encuentra todos los puntos no diferenciables interiores y describe exhaustivamente los puntos de mínimo y máximo absolutos. Explica por qué no basta buscar $q' = 0$.

Ejercicio C3. Extremos sin Weierstrass. Sea $h : [-1, 1] \to \mathbb{R}$ dada por $h(0) = 5$ y $h(x) = x$ cuando $x \neq 0$. Encuentra máximo y mínimo absolutos y todos sus puntos de realización. Explica por qué Weierstrass no es aplicable, aunque existan los extremos.

Ejercicio C4. Tres críticos y dos bordes. Determina todos los extremos locales interiores y los extremos absolutos de $p(x) = x^4 - 4x^2$ en $[-3, 3]$. Comprueba cada valor candidato y registra todos los empates.

Ejercicio C5. Un dominio no compacto con extremos. Demuestra que $f(x) = x/(1+x^2)$ alcanza un máximo y un mínimo absolutos sobre $\mathbb{R}$. Encuentra sus valores y puntos mediante derivadas y da una comprobación independiente por una desigualdad algebraica.

Ejercicio C6. El mismo cero, tres destinos. Para $a(x) = x^3$, $b(x) = x^2$ y $c(x) = -x^2$ sobre $\mathbb{R}$, comprueba que sus derivadas valen cero en el origen y demuestra directamente que allí se dan, respectivamente, ningún extremo, un mínimo absoluto estricto y un máximo absoluto estricto.

Nivel D — Convexidad por cuerdas y tangentes. Ejercicio D1. La separación de una cuerda. Para $f(x) = x^2 + 3x - 1$ en $\mathbb{R}$, calcula $(1 - t)f(x) + tf(y) - f((1 - t)x + ty)$ y deduce, sin derivar, que f es estrictamente convexa.

Ejercicio D2. Convexidad sin tangente. Prueba que $f(x) = |x - 2|$ es convexa en $\mathbb{R}$, pero no estrictamente convexa ni diferenciable en 2. Señala por qué la ausencia de derivada no refuta la definición por cuerdas.

Ejercicio D3. Igualdad en todas las cuerdas. Demuestra que $f(x) = mx + b$ es convexa y cóncava en cualquier intervalo. Si el intervalo contiene dos puntos distintos, ¿puede ser estrictamente convexa? Justifica usando la definición, no la segunda derivada.

Ejercicio D4. La convexidad depende del intervalo. Estudia $f(x) = x^3$: demuestra concavidad estricta en $(-\infty, 0]$, convexidad estricta en $[0, \infty)$ y refuta convexidad y concavidad globales con sendas cuerdas explícitas.

Ejercicio D5. Una tangente estrictamente inferior. Para $f(x) = x^4$, demuestra para cualquier $c \in \mathbb{R}$ que $f(x) - f(c) - f'(c)(x - c) = (x - c)^2(x^2 + 2cx + 3c^2) \geq 0$. Estudia cuándo hay igualdad e interpreta el resultado.

Ejercicio D6. Una función convexa que salta en los bordes. Define $h : [0, 1] \to \mathbb{R}$ por $h(0) = h(1) = 1$ y $h(x) = 0$ si $0 < x < 1$. Prueba directamente que es convexa y señala exactamente dónde es discontinua.

Nivel E — Segunda derivada e inflexión. Ejercicio E1. Una segunda derivada que se anula. Calcula f'' para $f(x) = x^6$ y demuestra que f es estrictamente convexa en $\mathbb{R}$, pese a que $f''(0) = 0$.

Ejercicio E2. La ecuación $f''(0) = 0$ no decide. Comprueba que x^4 y x^3 tienen segunda derivada nula en 0. Demuestra que solo una de las dos tiene una inflexión allí según la definición del capítulo.

Ejercicio E3. Inflexión sin segunda derivada. Sea $f(x) = x|x|$. Determina f' y las segundas derivadas laterales en puntos no nulos; comprueba que $f''(0)$ no existe y prueba, sin embargo, la inflexión en 0.

Ejercicio E4. Inflexión con tangente horizontal. Estudia $f(x) = x^5 + x^3$ sobre $\mathbb{R}$. Demuestra que 0 es una inflexión, determina su tangente y prueba que no es un extremo local.

Ejercicio E5. Criterio suficiente frente a necesario. Refuta la frase «si una función dos veces diferenciable es estrictamente convexa, su segunda derivada es positiva en todos los puntos». Da un ejemplo concreto y explica cuál es el criterio exacto de §10.8 bajo $f'' \geq 0$.

Ejercicio E6. Dos pruebas que requieren datos diferentes. Para $f(x) = x^4 + 2x^2$, localiza y clasifica todos los críticos. Determina qué prueba aporta $f''(0)$ y qué conclusión global adicional requiere una desigualdad válida para todo x.

Nivel F — Reconstrucción y reparación de pruebas. Ejercicio F1. Certificado de una cúbica. Estudia $f(x) = x^3 - 3x^2$ en $\mathbb{R}$: ceros, signos de f', extremos locales, convexidad, inflexión, valores al infinito y existencia de extremos absolutos. Justifica cada afirmación.

Ejercicio F2. Dos ramas separadas. Para $g(x) = x + 1/x$ sobre $\mathbb{R} \setminus \{0\}$, determina límites en el hueco y al infinito, signos de g' y g'', extremos locales e inflexiones. Comprueba si es monótona en todo su dominio y si posee extremos absolutos.

Ejercicio F3. Un mínimo alcanzado y un supremo ausente. Estudia $f(x) = x^2/(1 + x^2)$ sobre $\mathbb{R}$: paridad, monotonía, extremos, supremo, convexidad e inflexiones; comprueba si alcanza su supremo.

Ejercicio F4. Cuatro límites laterales. Para $g(x) = x/(x^2-1)$, determina dominio, límites laterales en ± 1, asíntota horizontal y si puede cortarla, monotonía por componentes, existencia de extremos e inflexiones. Explica por qué no debes unir las tres ramas.

Ejercicio F5. Curvatura sin segunda derivada central. Define $f(x) = x^2$ para $x \le 0$ y $f(x) = 2x^2$ para $x > 0$. Prueba que es diferenciable en toda la recta y estrictamente convexa, calcula si existe $f''(0)$ y decide si allí hay inflexión.

Ejercicio F6. Tres frases, tres reparaciones. Diagnostica y corrige: (i) «$f' > 0$ donde existe, luego f crece en su dominio»; (ii) «si una asíntota horizontal es $y = 0$, la gráfica nunca cruza el eje»; (iii) «$f''(c) = 0$ demuestra una inflexión». Aporta un contraejemplo explícito para cada frase y formula la conclusión válida.

Nivel G — Desafíos integradores. Ejercicio G1. Una familia y dos certificados de convexidad. Para $f(x) = x^2 + ax + b$, con $a, b \in \mathbb{R}$, demuestra convexidad estricta mediante cuerdas y tangentes; determina el único mínimo absoluto y muestra que no existe máximo absoluto.

Ejercicio G2. Un parámetro cambia la geometría. Estudia $f_a(x) = x^4 - 2ax^2$ sobre $\mathbb{R}$, distinguiendo $a < 0$, $a = 0$ y $a > 0$. Determina convexidad estricta global o sus intervalos de convexidad y concavidad, puntos de inflexión, extremos locales y absolutos y ausencia o existencia de máximo absoluto.

Ejercicio G3. El cero de la derivada ocupa medio dominio. Sea $f(x) = 0$ si $x \le 0$ y $f(x) = x^2$ si $x > 0$. Determina $f'(0)$, monotonía global y en $[0, \infty)$, convexidad global y estricta, y los extremos absolutos. Explica cómo interviene la presencia de un intervalo entero con derivada cero.

Ejercicio G4. De Fermat a las tangentes de apoyo. Sea $f : \mathbb{R} \to \mathbb{R}$ diferenciable y convexa. Demuestra que c es punto de mínimo absoluto si y solo si $f'(c) = 0$. Si f es estrictamente convexa, demuestra la unicidad del mínimo cuando existe. Explica con x^3 y con $f(x) = x$ por qué no se pueden suprimir, respectivamente, convexidad y la cláusula «cuando existe».

Soluciones desarrolladas.

Soluciones del nivel A. **Solución A1.** $-1 < 0$ pero $f(-1) = 1 > 0 = f(0)$, lo que refuta «no decreciente». Por otra parte, $0 < 1$ y $f(0) = 0 < 1 = f(1)$ refutan «no creciente». Para $x < y \le 0$, $f(y) - f(x) = (y - x)(x + y) < 0$; para $0 \le x < y$, ambos factores son positivos. Por tanto, decrece estrictamente en $[-2, 0]$ y crece estrictamente en $[0, 2]$. No pueden ampliarse estos intervalos, conservando su sentido, a través de 0.

Solución A2. Si f es no decreciente e inyectiva, $x < y$ implica $f(x) \le f(y)$, pero $f(x) \neq f(y)$; entonces $f(x) < f(y)$. A la inversa, si f es estrictamente creciente y $x \neq y$, uno de los dos argumentos es menor y sus imágenes son distintas: es inyectiva. Si f es no creciente, el mismo razonamiento reemplaza $\le$ por $\ge$ y produce la equivalencia entre inyectividad y decrecimiento estricto. No se necesitan continuidad ni hipótesis de intervalo.

Solución A3. La regla del cociente da $g'(x) = -2/(x-1)^2 < 0$ para $x \neq 1$. La función es continua y diferenciable en cada componente, así que Lagrange prueba decrecimiento estricto en $(-\infty, 1)$ y $(1, \infty)$. Pero $0 < 2$ mientras $g(0) = -1 < 3 = g(2)$; esto refuta que sea no creciente en D. El segmento $[0, 2]$ contiene $1 \notin D$, de modo que el criterio diferencial no lo compara.

Solución A4. Si f es constante, cada cociente incremental interior vale cero y su límite es $f'(t) = 0$. Recíprocamente, para cualesquiera $x < y$ de I, el intervalo contiene $[x, y]$; continuidad en el segmento y diferenciabilidad interior permiten escribir $f(y) - f(x) = f'(c)(y - x) = 0$ para algún $c \in (x, y)$. Todos los pares tienen igual imagen, luego f es constante. En $D = \mathbb{R} \setminus \{0\}$, la función que vale 0 para $x < 0$ y 1 para $x > 0$ tiene derivada cero en todo D, pero no es globalmente constante: falta el segmento que une las componentes.

Solución A5. En $(0, 1)$ la función coincide con la identidad y $q'(t) = 1 > 0$. No obstante, $0 < 1/2$ y $q(0) = 2 > 1/2 = q(1/2)$; falla incluso el crecimiento débil en el dominio completo. En 0, el límite relativo por la derecha es $0 \neq q(0)$: falta continuidad en el extremo incluido, precisamente la hipótesis requerida para aplicar Lagrange a segmentos que comiencen allí.

Solución A6. Tomemos $u(x) = 0$ y $v(x) = x^3 - x$. Ambos polinomios valen cero en $-1, 0, 1$. La función u es no decreciente porque todas sus imágenes son iguales. Para v, $-1/2 < 0$ pero $v(-1/2) = 3/8 > 0 = v(0)$, y $1/2 < 1$ pero $v(1/2) = -3/8 < 0 = v(1)$. El primer par refuta no decrecimiento y el segundo refuta no crecimiento. La concordancia en tres observaciones no certifica una propiedad cuantificada sobre todos los pares.

Soluciones del nivel B. **Solución B1.** La función es polinómica, por lo que es continua y diferenciable en todo intervalo cerrado. $f'(x) = 5x^4 + 1 \geq 1 > 0$ para todo real. Dados $x < y$, Lagrange proporciona $f(y) - f(x) = f'(c)(y - x) > 0$: crece estrictamente. Si $f(a) = f(b) = 0$ con $a < b$, contradiríamos $f(a) < f(b)$; por tanto, hay a lo sumo un cero. Como $f(0) = 0$, este es el único.

Solución B2. $f'(x) = 5x^4 \geq 0$ y se anula únicamente en $x = 0$, luego no es idénticamente nula sobre ningún intervalo abierto. La equivalencia de Proposición 10.2 demuestra crecimiento estricto. Alternativamente, para $x < y$ con ambos del mismo signo, la potencia quinta conserva la desigualdad; si $x < 0 < y$, entonces $x^5 < 0 < y^5$, y los casos con un extremo cero se tratan igual. Sin embargo, $f'(0) = 0$, de modo que es falsa la supuesta necesidad de $f' > 0$ *en cada punto* para el crecimiento estricto.

Solución B3. $p'(x) = 3x^2 - 3 = 3(x-1)(x+1)$ y sus únicos ceros son $-1, 1$. Para $x < -1$, los dos factores son negativos y su producto positivo; entre -1 y 1 tienen signos opuestos; después de 1, ambos son positivos. Continuidad en todo $[-2, 2]$ y el criterio por derivadas dan crecimiento estricto en $[-2, -1]$ y $[1, 2]$, decrecimiento estricto en $[-1, 1]$. Los pares $-2 < -1$ con $p(-2) = -1 < p(-1) = 3$ y $-1 < 0$ con $p(-1) = 3 > p(0) = 1$ refutan, respectivamente, no crecimiento y no decrecimiento globales.

Solución B4. Como $f' \geq 0$ en (a, b), Lagrange demuestra que f es no decreciente en $[a, b]$. Si no es estrictamente creciente, existen $x < y$ con $f(x) = f(y)$; para cada $z \in [x, y]$, $f(x) \leq f(z) \leq f(y)$, así que f es constante en $[x, y]$ y $f' = 0$ en el intervalo abierto no vacío (x, y). Recíprocamente, si f' fuese idénticamente cero en un intervalo abierto interior, Lagrange mostraría que f es constante en ese intervalo, contradiciendo el crecimiento estricto. El argumento incluye los extremos a, b gracias a la continuidad allí.

Solución B5. $f'(x) = 3x^2 - 3$ y $f'(2) = 9 > 0$. Sin embargo, $-1 < 0$ mientras $f(-1) = 2 > 0 = f(0)$: no es no decreciente en $\mathbb{R}$. Conocer $f'(2)$ no controla el punto intermedio desconocido de Lagrange para otro par. Una reparación suficiente consistiría en demostrar $f'(t) \geq 0$ para **todo** t interior de un intervalo donde f sea continua y diferenciable, y limitar la conclusión de monotonía a ese intervalo.

Solución B6. Para $x \neq 0$, $u'(x) = 1/x^2 > 0$ y $v'(x) = 1 > 0$. Si $-1 < 1$, $u(-1) = 1 > -1 = u(1)$, luego u no es no decreciente en D. Para v, la fórmula $v(x) = x$ implica directamente $x < y \Rightarrow v(x) < v(y)$ para cualquier par en D, incluso entre componentes. Lagrange se puede aplicar a cada componente, pero no a $[-1, 1]$, donde falta el punto 0; el resultado global sobre v viene de su definición algebraica, no de atravesar ese hueco con el teorema.

Soluciones del nivel C. **Solución C1.** El polinomio es continuo en el compacto, así que los extremos absolutos existen. $p'(x) = 3x^2 - 12x + 9 = 3(x-1)(x-3)$; los críticos interiores son 1 y 3. El signo es positivo en $(0, 1)$, negativo en $(1, 3)$ y positivo en $(3, 5)$: hay máximo local estricto en 1 y mínimo local estricto en 3. Debemos añadir los bordes. Calculamos $p(0) = 0$, $p(1) = 4$, $p(3) = 0$ y $p(5) = 20$. Por la proposición de candidatos, el máximo absoluto es 20, solo en 5; el mínimo es 0, en 0 y 3. El máximo local interior 4 no es global.

Solución C2. Si $x < -1$, $q(x) = -(x-1)-(x+1) = -2x$; si $-1 \leq x \leq 1$, $q(x) = 1-x+x+1 = 2$; si $x > 1$, $q(x) = 2x$. Las derivadas laterales en -1 son -2 y 0, y en 1 son 0 y 2: ambas esquinas son no diferenciables. La función disminuye hasta -1, permanece constante en $[-1, 1]$ y aumenta después de 1. El mínimo absoluto es 2 en **todos** los puntos de $[-1, 1]$; el máximo es 4 en ambos bordes $-2, 2$. Buscar solo $q' = 0$ descubriría el interior de la meseta, pero perdería sus extremos ± 1 y los candidatos de borde para el máximo.

Solución C3. Para $x \neq 0$ con $x \in [-1, 1]$, $-1 \leq x \leq 1$, mientras $h(0) = 5$. Por tanto, el máximo absoluto es 5, únicamente en 0; el mínimo absoluto es -1, únicamente en -1. La función no es continua en 0, pues $\lim_{x\to 0} h(x) = 0 \neq 5$. No se cumplen las hipótesis de Weierstrass, aunque la comparación directa prueba que los extremos existen; la inaplicabilidad de un teorema de existencia no implica la inexistencia de los objetos.

Solución C4. $p'(x) = 4x(x^2 - 2)$ se anula exactamente en $x = -\sqrt{2}, 0, \sqrt{2}$. Sus signos, de izquierda a derecha, son $-, +, -, +$. Así, $\pm\sqrt{2}$ son mínimos locales estrictos y 0 un máximo local estricto. Weierstrass y la proposición de candidatos obligan a comparar además los bordes: $p(\pm 3) = 81 - 36 = 45$, $p(\pm\sqrt{2}) = 4 - 8 = -4$ y $p(0) = 0$. Los mínimos absolutos valen -4 y se realizan en ambos $\pm\sqrt{2}$; los máximos absolutos valen 45 y se realizan en ambos ± 3. El máximo interior 0 es solo local.

Solución C5. $f'(x) = (1 - x^2)/(1 + x^2)^2$, de signo negativo para $|x| > 1$ y positivo para $|x| < 1$. Decrece en $(-\infty, -1]$, crece en $[-1, 1]$ y decrece en $[1, \infty)$. Como $\lim_{x\to\pm\infty} f(x) = 0$, y $f(-1) = -1/2$, $f(1) = 1/2$, la comparación de cada tramo demuestra mínimo absoluto $-1/2$ únicamente en -1 y máximo absoluto $1/2$ únicamente en 1. Independientemente, $(|x| - 1)^2 \geq 0$ implica $2|x| \leq 1 + x^2$, luego $|f(x)| \leq 1/2$. La igualdad exige $|x| = 1$, y los signos seleccionan el punto de cada extremo. No invocamos compacidad del dominio no acotado.

Solución C6. $a'(x) = 3x^2$, $b'(x) = 2x$ y $c'(x) = -2x$: las tres valen cero en 0. Para cada $r > 0$ tomamos $0 < t < r$: $a(-t) = -t^3 < 0 = a(0) < t^3 = a(t)$, así que ningún entorno del origen exhibe un extremo de a. En cambio, $b(x) = x^2 \geq 0 = b(0)$ con igualdad solo en 0: mínimo absoluto estricto. Por último, $c(x) = -x^2 \leq 0 = c(0)$ con igualdad solo en 0: máximo absoluto estricto. La ecuación $f'(0) = 0$ identifica candidatos, no los clasifica.

Soluciones del nivel D. **Solución D1.** Denotemos $z = (1-t)x + ty$. Al expandir y usar que $(1-t) + t = 1$, los términos afines $3z - 1$ se cancelan exactamente:

$$(1-t)f(x) + tf(y) - f(z) = (1-t)x^2 + ty^2 - z^2 = t(1-t)(x-y)^2.$$

Para $t \in [0,1]$ la diferencia es no negativa. Si $x \neq y$ y $0 < t < 1$, los tres factores son positivos, por lo que la desigualdad es estricta. Es la definición de convexidad estricta, sin necesidad de calcular derivadas. Los términos afines cambian la posición de la gráfica, no la separación entre una cuerda y sus valores interpolados.

Solución D2. Para $z = (1-t)x + ty$, la desigualdad triangular proporciona $|z-2| = |(1-t)(x-2) + t(y-2)| \leq (1-t)|x-2| + t|y-2|$. Es convexa. Al tomar $x = 3$, $y = 4$ y $0 < t < 1$, todos los argumentos de los valores absolutos son positivos y se obtiene igualdad, de modo que no es estrictamente convexa. Para $h \neq 0$, el cociente $[f(2+h) - f(2)]/h = |h|/h$ vale -1 a la izquierda y 1 a la derecha: no existe derivada en 2. La definición de convexidad solo compara alturas sobre cuerdas; no exige tangentes.

Solución D3. Para $z = (1-t)x + ty$ y cualquier $t \in [0,1]$, $f(z) = m((1-t)x + ty) + b = (1-t)(mx+b) + t(my+b) = (1-t)f(x) + tf(y)$. Hay igualdad en todas las cuerdas, así que las desigualdades débil convexa y débil cóncava se cumplen simultáneamente. Si existen $x \neq y$ en el intervalo, elegir $t = 1/2$ da igualdad en una situación donde la convexidad estricta exigiría desigualdad: ninguna función afín sobre tal intervalo es estrictamente convexa.

Solución D4. $f'(x) = 3x^2$ es estrictamente decreciente en $(-\infty, 0)$ y estrictamente creciente en $(0, \infty)$: si $x < y \leq 0$ entonces $x^2 > y^2$, mientras que $0 \leq x < y$ implica $x^2 < y^2$. Puesto que f es continua también en los extremos incluidos, el criterio de Teorema 10.3 da concavidad estricta en $(-\infty, 0]$ y convexidad estricta en $[0, \infty)$. Globalmente, la cuerda entre -2 y 0 tiene altura -4 en -1, pero $f(-1) = -1 > -4$: no es convexa. La cuerda entre 0 y 2 tiene altura 4 en 1, pero $f(1) = 1 < 4$: no es cóncava en toda $\mathbb{R}$.

Solución D5. $f'(c) = 4c^3$. Una expansión verifica

$$f(x) - f(c) - f'(c)(x-c) = x^4 - 4c^3x + 3c^4 = (x-c)^2(x^2 + 2cx + 3c^2).$$

El segundo factor es $(x+c)^2 + 2c^2$, no negativo. Si $x \neq c$ y $c \neq 0$, es positivo; si $c = 0$ y $x \neq c$, vale $x^2 > 0$. Por tanto, la diferencia se anula exactamente cuando $x = c$, y la tangente $y = c^4 + 4c^3(x-c)$ apoya la gráfica estrictamente por debajo en todos los demás puntos. Esto verifica el criterio de tangentes de apoyo y demuestra convexidad; la desigualdad estricta global frente a cada tangente es más fuerte que la mera condición débil de apoyo.

Solución D6. Para $z = (1-t)x + ty$, si $z \in (0,1)$, entonces $h(z) = 0$ y $(1-t)h(x) + th(y) \geq 0$ porque todos los valores de h son no negativos. Si $z = 0$, con $x, y \in [0,1]$ y coeficientes no negativos solo puede contribuir con peso positivo un extremo igual a 0; el miembro derecho vale $h(0) = 1$. Lo mismo ocurre si $z = 1$. Así se verifica la desigualdad de convexidad en todos los casos. Sin embargo, $\lim_{x\to 0^+} h(x) = 0 \neq h(0) = 1$ y $\lim_{x\to 1^-} h(x) = 0 \neq h(1) = 1$. Es continua en todo punto interior y discontinua en ambos bordes. La convexidad en un cerrado no asegura continuidad allí.

Soluciones del nivel E. **Solución E1.** $f'(x) = 6x^5$ y $f''(x) = 30x^4 \geq 0$, con igualdad únicamente en 0. La función es continua en todo $\mathbb{R}$ y dos veces diferenciable. Por Teorema 10.4 es convexa; como f'' no se anula idénticamente en ningún intervalo abierto, Proposición 10.9 asegura convexidad **estricta**. También puede comprobarse que $f'(x) = 6x^5$ es estrictamente creciente, lo que permite aplicar Teorema 10.3. Un cero aislado de f'' no produce un segmento afín ni impide convexidad estricta.

Solución E2. Para $u(x) = x^4$, $u''(x) = 12x^2$, luego $u''(0) = 0$ y $u''(x) > 0$ a ambos lados del origen. Es estrictamente convexa en cada intervalo lateral; no hay cambio de convexidad a concavidad y 0 no es inflexión. Para $v(x) = x^3$, $v''(x) = 6x$ y $v''(0) = 0$, pero $v'' < 0$ a la izquierda y $v'' > 0$ a la derecha. La continuidad en 0 y Teorema 10.6 dan una inflexión allí. La condición puntual compartida $f''(0) = 0$ es necesaria cuando existe segunda derivada en una inflexión, pero no suficiente.

Solución E3. $f(x) = -x^2$ para $x < 0$ y $f(x) = x^2$ para $x \geq 0$. En 0, $f(h)/h = |h| \to 0$, por lo que $f'(0) = 0$. Para $x < 0$, $f'(x) = -2x$ y $f''(x) = -2$; para $x > 0$, $f'(x) = 2x$ y $f''(x) = 2$. El cociente $[f'(h) - f'(0)]/h$ vale -2 para $h < 0$ y 2 para $h > 0$: $f''(0)$ no existe. No obstante, f es continua en el origen, estrictamente cóncava a su izquierda y estrictamente convexa a su derecha por los signos constantes de las segundas derivadas. Es inflexión según Definición 10.3: la existencia de $f''(0)$ no forma parte de esa definición.

Solución E4. $f'(x) = 5x^4 + 3x^2 = x^2(5x^2 + 3) \geq 0$, con cero únicamente en 0; Proposición 10.2 demuestra crecimiento estricto en toda la recta. Por tanto, 0 no es extremo local: valores inmediatamente a su izquierda son menores y a su derecha mayores. $f''(x) = 20x^3 + 6x = 2x(10x^2 + 3)$ es negativa para $x < 0$ y positiva para $x > 0$, de modo que 0 es inflexión. Como $f(0) = f'(0) = 0$, la recta tangente allí es $y = 0$, horizontal. Tener tangente horizontal no equivale a tener un extremo.

Solución E5. $f(x) = x^4$ es estrictamente convexa: $f'(x) = 4x^3$ es estrictamente creciente en $\mathbb{R}$ y se aplica Teorema 10.3. Pero $f''(0) = 12 \cdot 0^2 = 0$, lo que refuta la afirmación. Bajo continuidad en el intervalo y segunda diferenciabilidad interior, la condición $f'' \geq 0$ caracteriza convexidad; **añadir** que f'' no sea idénticamente cero en ningún intervalo abierto interior caracteriza convexidad estricta por Proposición 10.9. La condición más fuerte $f'' > 0$ es suficiente, no necesaria.

Solución E6. $f'(x) = 4x^3 + 4x = 4x(x^2 + 1)$, que se anula solo en 0. La segunda derivada $f''(x) = 12x^2 + 4$ satisface $f''(0) = 4 > 0$, y el segundo criterio prueba mínimo local estricto en 0. Para demostrar que es **absoluto** no basta ese dato puntual; observamos $f(x) = x^4 + 2x^2 = x^2(x^2 + 2) \geq 0 = f(0)$ para todo real, con igualdad únicamente en 0. La función es no acotada superiormente, pues $f(x) \geq x^4 \to \infty$ cuando $|x| \to \infty$; no tiene máximo absoluto. Quedan separados diagnóstico local, comparación global y existencia de máximo.

Soluciones del nivel F. **Solución F1.** Factorizamos $f(x) = x^2(x-3)$: sus ceros son 0 y 3, y 0 tiene multiplicidad dos. La derivada $f'(x) = 3x(x-2)$ es positiva en $(-\infty, 0)$, negativa en $(0, 2)$ y positiva en $(2, \infty)$. Aplicando el primer criterio, 0 es máximo local estricto de valor 0 y 2 es mínimo local estricto de valor $f(2) = -4$. Por continuidad se incluyen los puntos críticos en los intervalos de monotonía: crece en $(-\infty, 0]$ y $[2, \infty)$, decrece en $[0, 2]$.

La segunda derivada es $f''(x) = 6(x-1)$: es negativa antes de 1 y positiva después. Así, f es estrictamente cóncava en $(-\infty, 1]$, estrictamente convexa en $[1, \infty)$ y 1 es inflexión, con $f(1) = -2$ y $f'(1) = -3$, tangente $y = -2 - 3(x-1)$. Finalmente, $f(x) = x^3(1 - 3/x)$ tiene límite $-\infty$ al ir $x \to -\infty$ y $+\infty$ cuando $x \to +\infty$. La imagen no está acotada ni superior ni inferiormente: los dos extremos encontrados son solamente locales, y no existen extremos absolutos.

Solución F2. El dominio tiene componentes $(-\infty, 0)$ y $(0, \infty)$. Se cumple $\lim_{x\to 0^-} g(x) = -\infty$ y $\lim_{x\to 0^+} g(x) = +\infty$, mientras $\lim_{x\to\pm\infty} g(x) = \pm\infty$ en los sentidos respectivos, porque $g(x) - x = 1/x \to 0$. Además de la asíntota vertical $x = 0$, la recta $y = x$ describe la aproximación oblicua en ambos infinitos.

Para $x \neq 0$, $g'(x) = 1 - 1/x^2 = (x^2-1)/x^2$, positivo cuando $|x| > 1$ y negativo cuando $0 < |x| < 1$. Crece en $(-\infty, -1]$ y $[1, \infty)$, decrece en $[-1, 0)$ y $(0, 1]$. El cambio de + a − en -1 da máximo local estricto $g(-1) = -2$; el de − a + en 1 da mínimo local estricto $g(1) = 2$. Como $g''(x) = 2/x^3$, es estrictamente cóncava en la componente negativa y estrictamente convexa en la positiva; no existe inflexión, pues el posible lugar del cambio, 0, no pertenece al dominio. Los límites infinitos descartan extremos absolutos. Tampoco es monótona en toda la unión: $-2 < -1$ con $g(-2) = -5/2 < -2 = g(-1)$ refuta no crecimiento, mientras $-1 < -1/2$ con $g(-1) = -2 > -5/2 = g(-1/2)$ refuta no decrecimiento. Ninguna prueba diferencial cruza el hueco.

Solución F3. $f(-x) = f(x)$, por lo que es par. Es continua en toda la recta. Sus derivadas son

$$f'(x) = \frac{2x}{(1+x^2)^2}, \qquad f''(x) = \frac{2(1-3x^2)}{(1+x^2)^3}.$$

El signo de f' coincide con el de x: decrece estrictamente en $(-\infty, 0]$ y crece estrictamente en $[0, \infty)$. Como $f(x) \geq 0 = f(0)$, el único mínimo absoluto es 0 en $x = 0$. La identidad $f(x) = 1 - 1/(1+x^2)$ muestra que $f(x) < 1$ para todo x y $\lim_{x\to\pm\infty} f(x) = 1$; por tanto, su supremo es 1, pero no tiene máximo absoluto. La segunda derivada es positiva para $|x| < 1/\sqrt{3}$ y negativa para $|x| > 1/\sqrt{3}$. En $[-1/\sqrt{3}, 1/\sqrt{3}]$ es estrictamente convexa; en $(-\infty, -1/\sqrt{3}]$ y $[1/\sqrt{3}, \infty)$ es estrictamente cóncava. Los puntos $\pm 1/\sqrt{3}$ son inflexiones de altura $1/4$, demostradas por cambios de signo y continuidad. El dato de supremo requiere el límite, no una lista finita de críticos.

Solución F4. El denominador se anula exactamente en -1 y 1, de modo que $D = \mathbb{R} \setminus \{-1, 1\}$. Factorizar $(x-1)(x+1)$ y controlar signos da

$$\lim_{x\to -1^-} g(x) = -\infty, \quad \lim_{x\to -1^+} g(x) = +\infty, \quad \lim_{x\to 1^-} g(x) = -\infty, \quad \lim_{x\to 1^+} g(x) = +\infty.$$

Además $g(x) = (1/x)/(1 - 1/x^2) \to 0$ en ambos infinitos. La asíntota horizontal es $y = 0$ y sí la corta: $g(0) = 0$. Derivando se obtiene $g'(x) = -(x^2+1)/(x^2-1)^2 < 0$ para todo $x \in D$. Por continuidad y Lagrange, decrece estrictamente en **cada** una de las tres componentes, pero no en toda la unión: $-2 < 0$ y $g(-2) = -2/3 < 0 = g(0)$ refutan no crecimiento global. Como ningún

punto del dominio es crítico, no hay extremos locales interiores; los límites prueban ausencia de extremos absolutos.

La segunda derivada es $g''(x) = 2x(x^2+3)/(x^2-1)^3$, de signos $-, +, -, +$ en $(-\infty, -1)$, $(-1, 0)$, $(0, 1)$ y $(1, \infty)$ respectivamente. Los correspondientes tramos son, en ese orden, estrictamente cóncavo, convexo, cóncavo y convexo. En 0 hay inflexión por continuidad y cambio lateral; en ± 1 no hay función, luego tampoco inflexiones. Unir las ramas representaría puntos que no existen y ocultaría las asíntotas verticales.

Solución F5. La función es continua en 0, donde ambas fórmulas valen cero. Para $h < 0$, $[f(h) - f(0)]/h = h$; para $h > 0$, el cociente es $2h$. Ambos tienden a cero, así que $f'(0) = 0$. Fuera del origen, $f'(x) = 2x$ para $x < 0$ y $f'(x) = 4x$ para $x > 0$. Es estrictamente creciente dentro de cada semirrecta y, al cruzar el origen, los valores de f' pasan de negativos a cero y luego a positivos: también es estrictamente creciente en toda $\mathbb{R}$. Como f es diferenciable en el intervalo completo, Teorema 10.3 prueba convexidad estricta global.

Para $h < 0$, $[f'(h) - f'(0)]/h = 2$, y para $h > 0$ vale 4. Los límites laterales difieren: $f''(0)$ no existe. Sin embargo, a ambos lados hay convexidad estricta (segundas derivadas 2 y 4); **no** hay cambio a concavidad, luego 0 no es inflexión. Es un mínimo absoluto estricto porque $f(x) > 0 = f(0)$ para todo $x \neq 0$.

Solución F6. (i) Para $u(x) = -1/x$ en $\mathbb{R} \setminus \{0\}$, $u' > 0$ en todo su dominio, pero $-1 < 1$ y $u(-1) = 1 > -1 = u(1)$. La conclusión válida se restringe a cada intervalo conectado donde se cumplan continuidad y diferenciabilidad; las comparaciones entre componentes se prueban aparte. (ii) Para $v(x) = x/(1 + x^2)$, $\lim_{x\to\pm\infty} v(x) = 0$, por lo que $y = 0$ es una asíntota horizontal; sin embargo $v(0) = 0$. Una asíntota expresa comportamiento límite al infinito, no prohibición de intersección. (iii) $w(x) = x^4$ satisface $w''(0) = 0$ y $w''(x) = 12x^2 > 0$ a ambos lados; no hay inflexión. Una prueba suficiente correcta verifica convexidad estricta en un intervalo lateral y concavidad estricta en el otro, continuidad e interioridad del punto, por ejemplo mediante signos estrictos opuestos de f'' en intervalos laterales completos.

Soluciones del nivel G. **Solución G1.** Para $z = (1 - t)x + ty$, una expansión cancela la contribución afín y da

$$(1 - t)f(x) + tf(y) - f(z) = t(1 - t)(x - y)^2.$$

La diferencia es positiva si $x \neq y$ y $0 < t < 1$: convexidad estricta por cuerdas. Además, $f'(c) = 2c + a$ y

$$f(x) - f(c) - f'(c)(x - c) = (x - c)^2 \geq 0,$$

con igualdad solo en $x = c$, lo que proporciona un certificado independiente de tangentes de apoyo. Completando el cuadrado, $f(x) = \left(x + \frac{a}{2}\right)^2 + b - \frac{a^2}{4}$. Así, su único mínimo absoluto es $b - a^2/4$, alcanzado en $x = -a/2$. Dado cualquier M, escoger $|x|$ suficientemente grande hace que el cuadrado supere $M - b + a^2/4$: la función no está acotada superiormente y no tiene máximo absoluto.

Solución G2. Las derivadas son

$$f_a'(x) = 4x(x^2 - a), \qquad f_a''(x) = 12x^2 - 4a.$$

Si $a < 0$, $f_a''(x) = 12x^2 + 4|a| > 0$ para todo x, de modo que la función es estrictamente convexa y no tiene inflexiones. Además $f_a(x) = x^4 + 2|a|x^2 \geq 0$, con igualdad únicamente en 0: es su único mínimo absoluto y local; f_a' solo se anula allí. **Si** $a = 0$, $f_a(x) = x^4$, $f_a'' = 12x^2 \geq 0$ y no se anula en ningún intervalo abierto completo: sigue siendo estrictamente convexa, sin inflexiones y con mínimo absoluto único 0 en el origen.

Si $a > 0$, llamemos $s = \sqrt{a/3}$. $f_a'' < 0$ en $(-s, s)$ y $f_a'' > 0$ fuera de $[-s, s]$. Por los criterios demostrados, la función es estrictamente cóncava en $[-s, s]$ y estrictamente convexa en $(-\infty, -s]$ y $[s, \infty)$; los puntos $\pm s$ son inflexiones, con valor $f_a(\pm s) = a^2/9 - 2a^2/3 = -5a^2/9$. Los críticos de f_a' son $-\sqrt{a}, 0, \sqrt{a}$; sus signos son $-, +, -, +$ al recorrer la recta de izquierda a derecha. Por el primer criterio, 0 es máximo local estricto con valor 0, y $\pm\sqrt{a}$ son mínimos locales estrictos de valor $-a^2$. La identidad $f_a(x) = (x^2 - a)^2 - a^2 \geq -a^2$ prueba que los dos mínimos son absolutos. En los tres casos del parámetro, $f_a(x) \to +\infty$ cuando $|x| \to \infty$, por lo que **nunca existe máximo absoluto**. Para $a > 0$, la función no es globalmente convexa ni globalmente cóncava, pues sus intervalos laterales exhiben signos opuestos de segunda derivada.

Solución G3. La función es continua en el origen, y $[f(h) - f(0)]/h$ vale 0 para $h < 0$ y h para $h > 0$; ambos límites son 0, por lo que $f'(0) = 0$. Para $x < 0$, $f'(x) = 0$, y para $x > 0$, $f'(x) = 2x > 0$. El criterio asegura que es no decreciente en toda $\mathbb{R}$. No es estrictamente creciente allí: $-2 < -1$ pero $f(-2) = f(-1) = 0$. En $[0, \infty)$ es estrictamente creciente, pues $0 \leq x < y$ implica $x^2 < y^2$, aunque la derivada se anule en el borde 0.

La derivada es no decreciente en toda la recta: vale cero hasta 0 y aumenta después. Por Teorema 10.3, f es convexa globalmente; no es estrictamente convexa, porque en el semieje negativo coincide con una función afín constante y hay igualdad sobre cualquiera de sus cuerdas. Su mínimo absoluto vale 0 y se alcanza en todo $(-\infty, 0]$. Para $x \to +\infty$, $f(x) = x^2 \to +\infty$, por lo que no hay máximo absoluto. El intervalo de ceros de f' explica por qué el criterio exacto de crecimiento estricto **global** no se satisface.

Solución G4. Si $f'(c) = 0$, el criterio de tangentes de apoyo para una función diferenciable y convexa (Proposición 10.8) da, para todo $x \in \mathbb{R}$, $f(x) \geq f(c) + f'(c)(x - c) = f(c)$, así que c es mínimo absoluto. A la inversa, un mínimo absoluto es local y su argumento es interior de $\mathbb{R}$; como f es diferenciable, Fermat (Teorema 9.1) implica $f'(c) = 0$. Si hubiera dos minimizadores diferentes $u < v$ en una función estrictamente convexa, en su punto medio se obtendría $f\left(\frac{u+v}{2}\right) < \frac{f(u)+f(v)}{2} = m$, contradiciendo que m sea el valor mínimo; el minimizador, si existe, es único. La función x^3 tiene derivada cero en 0 y no tiene allí un mínimo, porque no es convexa en toda la recta: muestra por qué no se elimina la hipótesis de convexidad. La función afín $f(x) = x$ es convexa y diferenciable, pero no alcanza mínimo absoluto en $\mathbb{R}$; la equivalencia establece una caracterización de los **puntos que existen**, no garantiza su existencia previa.

Capítulo 11

Derivadas superiores y fórmula de Taylor con resto

La caracterización por linealización local (Teorema 6.2) expresa la variación de una función diferenciable como un término lineal más un residuo de orden menor. Si f es diferenciable en a, tenemos $f(a+h) = f(a) + f'(a)h + o(h)$. La recta tangente aparece así como un modelo de primer orden: después de retirar de la variación de f el término lineal $f'(a)h$, el residuo es pequeño comparado con h. Esta afirmación es local, cuantitativa y mucho más precisa que decir simplemente que la gráfica «se parece» a su tangente.

Pero la propia fórmula deja una pregunta abierta. Si f' también puede derivarse, y después podemos derivar de nuevo, ¿podemos utilizar esa información adicional para construir modelos locales mejores que una recta? ¿Existe un procedimiento que produzca un polinomio de grado 2, 3 o n a partir de las derivadas de f en un solo punto? Y, sobre todo, ¿cómo sabremos cuánto nos estamos equivocando al reemplazar la función por ese polinomio?

Las derivadas sucesivas permitirán elegir los coeficientes de un polinomio que comparta con la función sus datos diferenciales en el centro. El **resto** será la diferencia exacta entre ambos. Para controlar el error necesitaremos precisar qué hipótesis garantizan una estimación de ese resto. Hay una frontera que conviene fijar desde el comienzo. Para un entero n concreto estudiaremos una suma finita

$$P(x) = c_0 + c_1(x-a) + \cdots + c_n(x-a)^n$$

y una identidad de la forma $f(x) = P(x) + R_n(x)$. El grado permanecerá fijado en cada aproximación y el término de error formará parte de la afirmación. La igualdad con una serie infinita exigiría estudiar además convergencia y analiticidad; esa cuestión corresponde al Tomo II.

11.1. De la aproximación lineal a la aproximación polinómica

La recta tangente utiliza dos datos. La aproximación lineal en a puede escribirse como $L_a(x) = f(a) + f'(a)(x-a)$. Para construirla necesitamos exactamente dos números:

$$f(a), \qquad f'(a).$$

El primero obliga a la recta a pasar por $(a, f(a))$; el segundo fija su pendiente. Dicho de otra manera,

$$L_a(a) = f(a), \qquad L_a'(a) = f'(a).$$

La aproximación lineal reproduce en el punto central el valor de la función y su primera derivada. Esta observación sugiere una estrategia general: si disponemos también de $f''(a)$, ¿podremos construir un polinomio cuadrático que reproduzca **tres** datos? Si conocemos sucesivamente más derivadas, ¿podremos exigir

$$P(a) = f(a), \quad P'(a) = f'(a), \quad P''(a) = f''(a), \quad \ldots?$$

Antes de responder para una función arbitraria, conviene invertir el problema y estudiar un caso en el que ya sabemos cuál debe ser la respuesta.

Un polinomio revela el patrón de los coeficientes.

EJEMPLO 11.1. **Leer los coeficientes mediante derivadas.** Fijemos $a \in \mathbb{R}$ y consideremos

$$p(x) = 2 - 3(x-a) + 5(x-a)^2 - 2(x-a)^3.$$

Evaluar en el centro elimina todos los términos que contienen $x - a$: $p(a) = 2$.
Derivando una vez,

$$p'(x) = -3 + 10(x-a) - 6(x-a)^2,$$

de modo que $p'(a) = -3$.
Una segunda y una tercera derivación dan

$$p''(x) = 10 - 12(x-a), \qquad p'''(x) = -12,$$

y por tanto

$$p''(a) = 10, \qquad p'''(a) = -12.$$

Ahora recuperemos los coeficientes originales:

$$2 = p(a), \qquad -3 = p'(a), \qquad 5 = \frac{p''(a)}{2}, \qquad -2 = \frac{p'''(a)}{6}.$$

Como $2 = 2!$ y $6 = 3!$, el mismo polinomio puede escribirse como

$$p(x) = p(a) + p'(a)(x-a) + \frac{p''(a)}{2!}(x-a)^2 + \frac{p'''(a)}{3!}(x-a)^3.$$

La aparición de los factoriales no es decorativa. Cada vez que derivamos $(x-a)^k$ disminuye la potencia y aparece un factor entero; después de k derivaciones, al evaluar en a, sobrevive exactamente el factor $k! = k(k-1)\cdots 2 \cdot 1$.

Este cálculo permite anticipar una regla que más adelante demostraremos en general. Si

$$P(x) = c_0 + c_1(x-a) + \cdots + c_n(x-a)^n,$$

entonces las derivadas sucesivas evaluadas en a aíslan los coeficientes uno a uno:

$$P^{(k)}(a) = k!\, c_k, \qquad 0 \le k \le n,$$

si adoptamos la convención $P^{(0)} = P$ y $0! = 1$. Por consiguiente, si queremos que un polinomio reproduzca los datos $f(a), f'(a), \ldots, f^{(n)}(a)$, el candidato natural tendrá necesariamente coeficientes $c_k = \frac{f^{(k)}(a)}{k!}$. Los coeficientes surgen de exigir que el polinomio comparta un número prescrito de derivadas con f en el centro. Para justificar la construcción tendremos que definir las derivadas de orden arbitrario y demostrar la existencia y unicidad del polinomio correspondiente.

Coincidir en el centro no controla todavía el error. Aquí aparece una distinción esencial. Supongamos que encontramos un polinomio P que satisface

$$P^{(k)}(a) = f^{(k)}(a) \qquad (0 \le k \le n).$$

Eso describe una coincidencia de datos **en el punto** a. No demuestra por sí solo cuánto se separan $P(x)$ y $f(x)$ cuando $x \neq a$. Para no ocultar este problema detrás del símbolo $\approx$, definiremos siempre el error por una identidad exacta: $R_n(x) := f(x) - P(x)$. Entonces $f(x) = P(x) + R_n(x)$ es una igualdad, no una aproximación informal. La verdadera tarea analítica consiste en estudiar R_n. En el caso lineal ya conocemos una respuesta. Si

$$P(x) = L_a(x) = f(a) + f'(a)(x - a),$$

el teorema de linealización afirma que, escribiendo $x = a + h$, $R_1(a + h) = o(h)$. Así, el residuo es pequeño frente a la primera potencia del incremento. Es razonable esperar que, al hacer coincidir más derivadas, aparezcan potencias más altas de h y que podamos obtener un control más fuerte. Pero esa expectativa debe convertirse en un teorema: **igualar derivadas no sustituye una estimación del resto**.

El resto forma parte de la fórmula
Una expresión como

$$f(x) \approx f(a) + f'(a)(x - a) + \frac{f''(a)}{2!}(x - a)^2$$

solo tiene contenido matemático cuando se especifica qué significa el símbolo $\approx$ o se proporciona una cota para la diferencia. En este capítulo preferiremos la forma exacta $f(x) = P(x) + R(x)$ y preguntaremos después qué sabemos demostrar sobre $R(x)$.

Tres preguntas diferentes que no debemos fusionar. El estudio de Taylor separará tres problemas. El primero es **algebraico**: dados los números $f(a), f'(a), \ldots, f^{(n)}(a)$,

¿cuál es el polinomio de grado a lo sumo n que reproduce esos datos? Aquí aparecerán los factoriales y la unicidad. El segundo es **analítico**: bajo qué hipótesis sobre f podemos escribir el error en una forma que dependa de una derivada de orden superior en algún punto intermedio. Aquí reaparecerá el teorema de Rolle, aplicado de manera iterada. El tercero es **cuantitativo**: si conocemos una cota para esa derivada superior en todo un intervalo, ¿qué cota concreta obtenemos para $|R_n(x)|$? Esta será la pieza que permita hablar rigurosamente de precisión.

Confundir estas preguntas produce errores típicos. Calcular correctamente un polinomio no prueba que el resto sea pequeño; saber que el resto tiende a cero no proporciona necesariamente una cota numérica; y una cota válida para una derivada en el centro no controla automáticamente su valor en todos los puntos intermedios que pueda seleccionar un teorema de existencia.

Una fórmula finita no es una serie infinita. Supongamos, por ejemplo, que para cada entero fijo n logramos construir un polinomio P_n. De ahí no se sigue que podamos escribir

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

ni mucho menos que una suma infinita de los coeficientes obtenidos converja a $f(x)$. La variable n introduce un **nuevo proceso de límite**, diferente del límite $x \to a$ que define la aproximación local. En este tomo mantendremos n fijo. La fórmula de Taylor que demostraremos tendrá la arquitectura

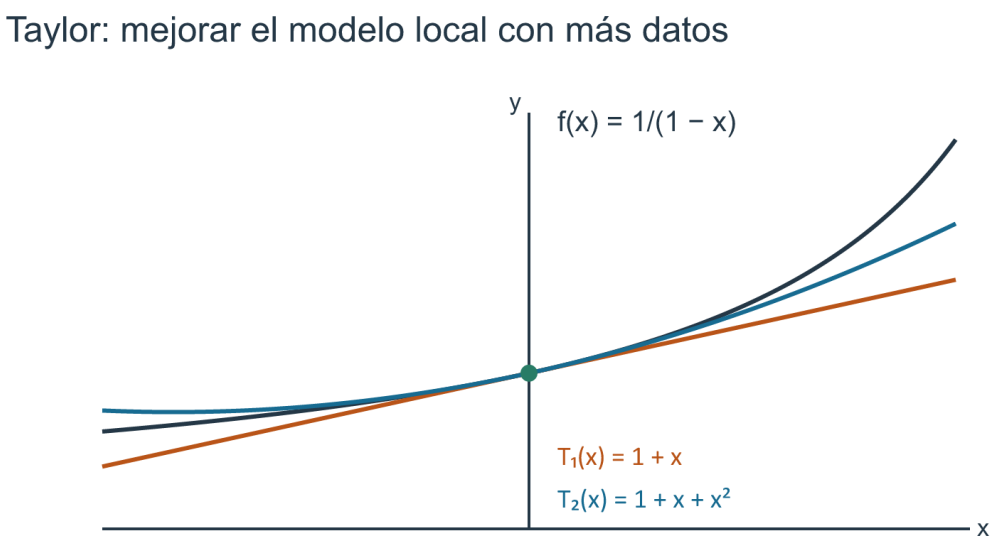


FIGURA 11.1. Para $f(x) = 1/(1-x)$ cerca de 0, los modelos $T_1(x) = 1 + x$ y $T_2(x) = 1 + x + x^2$ incorporan sucesivamente más datos locales. El control riguroso de la diferencia requiere una fórmula para el resto.

$$\boxed{\text{función} = \text{polinomio finito determinado por derivadas} + \text{resto controlado}}.$$

El paso hacia $n \to \infty$ requerirá estudiar sucesiones y series de funciones, convergencia y analiticidad. No es una consecuencia automática de poseer derivadas de todos los órdenes.

Continuidad con la aproximación lineal
El caso de grado 1 no es una teoría distinta. El polinomio $f(a)+f'(a)(x-a)$ que ya conocemos es precisamente el primer miembro no trivial de la jerarquía que vamos a construir. Taylor no reemplaza la linealización: la prolonga a órdenes finitos superiores y añade un mecanismo explícito para estudiar el error.

Hacia las derivadas iteradas. La expresión $f^{(n)}(a)$ que hemos utilizado de manera anticipada contiene una exigencia que todavía debemos hacer explícita. Para hablar de una tercera derivada no basta con que $f'(a)$ y $f''(a)$ aparezcan como números aislados: la segunda derivada debe existir como una función en puntos próximos para que pueda volver a derivarse. A cada iteración puede cambiar el dominio en el que la siguiente derivada está definida. Por eso el siguiente paso no será memorizar el polinomio de Taylor. En §11.2 definiremos cuidadosamente las derivadas de orden superior, sus dominios y la notación $f^{(n)}$. Solo entonces podremos demostrar qué operaciones sobreviven al repetir la derivación y, finalmente, qué polinomio queda determinado por esos datos.

Antes de seguir. En Ejemplo 11.1, ¿por qué la tercera derivada permite recuperar el coeficiente cúbico pero elimina todos los términos de grado menor? Si un polinomio P coincide con f en sus primeras n derivadas en a, ¿qué parte de esa información es puramente algebraica

y qué parte falta para justificar que $P(x)$ sea una buena aproximación cuando x se aleja de a? ¿Qué nuevo problema aparece si intentamos dejar que n tienda a infinito?

11.2. Derivadas de orden superior: iteración, dominios y notación

Para repetir la derivación, la derivada anterior debe estar definida como función en los puntos próximos que requiere el nuevo límite. Al introducir cada orden registraremos el dominio en el que puede efectuarse la siguiente derivación.

Derivar una función y derivar un número no son la misma operación. Saber que $f'(a)$ existe proporciona un número. Para preguntar por $f''(a)$ necesitamos algo más: f' debe estar definida en un conjunto que tenga a a como punto de acumulación. Solo entonces tiene sentido formar el cociente incremental de f'. Por ejemplo, si una función está definida y es diferenciable únicamente en el punto a dentro de cierto dominio muy disperso, la frase «derivemos otra vez» puede carecer de significado. La iteración siempre actúa sobre una **función derivada**, no sobre el valor aislado de una derivada.

Definición 11.1. **Derivadas sucesivas.** Sea $I \subseteq \mathbb{R}$ un intervalo abierto y sea $f : I \to \mathbb{R}$.

Definimos recursivamente

$$f^{(0)} := f.$$

Si $f^{(k)}$ está definida en un intervalo abierto $I_k \subseteq I$ y es diferenciable en los puntos de un intervalo abierto $I_{k+1} \subseteq I_k$, definimos

$$f^{(k+1)} := (f^{(k)})' \qquad \text{en } I_{k+1}.$$

Cuando las primeras n iteraciones están definidas en un mismo intervalo abierto $J \subseteq I$, decimos que f es n **veces diferenciable en** J. Escribimos también

$$f' = f^{(1)}, \qquad f'' = f^{(2)}, \qquad f''' = f^{(3)}.$$

Para $n \geq 4$ preferiremos la notación $f^{(n)}$, que evita una proliferación de primas.

La definición se ha formulado para intervalos abiertos porque las pruebas de Taylor trabajarán en un entorno de un segmento y aplicarán Rolle en puntos interiores. Cuando estudiemos restricciones a intervalos cerrados, las derivadas que intervengan serán las del intervalo abierto correspondiente o las de una extensión ya disponible; no ocultaremos derivadas laterales detrás de la notación ordinaria.

La cadena de dominios puede encogerse. Si definimos

$$D_0 := \operatorname{dom} f, \qquad D_1 := \operatorname{dom} f', \qquad D_2 := \operatorname{dom} f'', \quad \ldots,$$

entonces necesariamente $D_{k+1} \subseteq D_k$.

No hay razón para que la inclusión sea una igualdad. Una función puede ser diferenciable en un punto en el que su derivada deje de ser diferenciable.

Ejemplo 11.2. **Una primera derivada existe donde la segunda falla.** Consideremos $f(x) = x|x|$.

Como vimos en el capítulo anterior,

$$f(x) = \begin{cases} -x^2, & x < 0, \\ x^2, & x \geq 0, \end{cases}$$

y $f'(0) = 0$. Para $x \neq 0$, $f'(x) = 2|x|$.

Por tanto f' existe en toda la recta. Sin embargo, $\frac{f'(h)-f'(0)}{h} = \frac{2|h|}{h}$ vale -2 para $h < 0$ y 2 para $h > 0$. Luego $f''(0)$ no existe. Aquí

$$D_1 = \mathbb{R}, \qquad D_2 = \mathbb{R} \setminus \{0\}.$$

El ejemplo impide interpretar f'' como una operación meramente formal aplicada a una expresión: primero debe existir la función f' en los puntos relevantes y después debemos comprobar su diferenciabilidad.

Existencia y continuidad son niveles diferentes. Para Taylor necesitaremos distinguir «tener derivadas» de «tener derivadas continuas».

Definición 11.2. **Clase C^n.** Sea I un intervalo abierto y $n \in \mathbb{N}$. Diremos que

$$f \in C^n(I)$$

si existen $f^{(k)}$ en todo I para $0 \leq k \leq n$ y cada una de esas funciones es continua en I. Por convención, $C^0(I)$ significa simplemente continuidad.

La notación

$$f \in C^n$$

siempre debe entenderse respecto de un dominio indicado o claro por el contexto.

Ser n veces diferenciable no obliga, por definición, a que $f^{(n)}$ sea continua. En cambio, si existe $f^{(n+1)}$ en un intervalo, entonces $f^{(n)}$ es diferenciable allí y por Teorema 6.1 es automáticamente continua allí. Esta observación será útil para separar qué continuidad es consecuencia y cuál es una hipótesis adicional.

Polinomios: la iteración termina. Para un monomio $(x-a)^m$, las reglas de derivación ya demostradas dan, mientras $0 \leq k \leq m$,

$$\frac{d^k}{dx^k}(x-a)^m = \frac{m!}{(m-k)!}(x-a)^{m-k}.$$

En particular,

$$\left.\frac{d^k}{dx^k}(x-a)^m\right|_{x=a} = \begin{cases} 0, & k < m, \\ m!, & k = m. \end{cases}$$

y para $k > m$ la derivada es idénticamente cero. Así, un polinomio de grado m posee derivadas de todos los órdenes, pero toda derivada de orden mayor que m es la función cero.

Ejemplo 11.3. **Un grado se detecta por la última derivada no nula.** Sea $p(x) = 4x^5 - 3x^3 + 7x - 2$.

Las derivadas cuarta, quinta y sexta son

$$p^{(4)}(x) = 480x, \qquad p^{(5)}(x) = 480, \qquad p^{(6)}(x) = 0.$$

El hecho de que $p^{(6)}$ sea cero no significa que el proceso de derivación deje de estar definido: la función cero vuelve a derivarse y sigue dando cero. Así, $p \in C^n(\mathbb{R})$ para todo entero $n \geq 0$.

Qué información contiene $f^{(n)}(a)$. El valor $f^{(n)}(a)$ es local: describe la tasa de cambio de $f^{(n-1)}$ en a. Pero la fórmula de Taylor con resto seleccionará un punto intermedio c que no conocemos de antemano. Por eso, cuando queramos una cota numérica del error, un único valor $f^{(n+1)}(a)$ no bastará: necesitaremos controlar $f^{(n+1)}$ en todo el segmento relevante.

No confundir tres afirmaciones
Las proposiciones

1. «$f^{(n)}(a)$ existe»;
2. «$f^{(n)}$ existe en un intervalo alrededor de a»;
3. «$f \in C^n$ en ese intervalo»

no son sinónimas. La primera es puntual; la segunda permite seguir trabajando con $f^{(n)}$ como función; la tercera añade continuidad. Cada teorema del capítulo especificará cuál necesita.

Antes de seguir. Si $f''(a)$ existe, ¿qué función tuvo que ser derivada para producir ese número? ¿Por qué la existencia de $f^{(n+1)}$ en un intervalo fuerza continuidad de $f^{(n)}$ allí? ¿Qué parte de la cadena de dominios falló en Ejemplo 11.2?

11.3. Álgebra de derivadas superiores

Las reglas del capítulo 7 pueden iterarse. Para sumas, la repetición es inmediata: si las derivadas pertinentes existen, $(af + bg)^{(n)} = af^{(n)} + bg^{(n)}$. El producto es más interesante. Ya sabemos que $(fg)' = f'g + fg'$. Al derivar una vez más aparecen tres términos después de reunir los dos términos centrales: $(fg)'' = f''g + 2f'g' + fg''$. La siguiente derivación produce coeficientes $1, 3, 3, 1$. El patrón es el triángulo de Pascal, y la razón no es casual: en cada paso cada término puede recibir la nueva derivada en el factor izquierdo o en el derecho.

Proposición 11.1. **Fórmula de Leibniz para la derivada n-ésima de un producto.** Sea I un intervalo abierto, $n \in \mathbb{N}$ y sean $f, g : I \to \mathbb{R}$ n veces diferenciables. Entonces fg es n veces diferenciable y

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

Demostración. Procedemos por inducción sobre n.

Para $n = 0$ la fórmula dice simplemente

$$fg = \binom{0}{0} f^{(0)} g^{(0)}.$$

Supongamos válida la identidad para cierto n. Como f y g poseen una derivada adicional, podemos derivar término a término:

$$(fg)^{(n+1)} = \sum_{k=0}^{n} \binom{n}{k} \left(f^{(k+1)} g^{(n-k)} + f^{(k)} g^{(n-k+1)} \right)$$
$$= \sum_{j=1}^{n+1} \binom{n}{j-1} f^{(j)} g^{(n+1-j)} + \sum_{j=0}^{n} \binom{n}{j} f^{(j)} g^{(n+1-j)}.$$

Los términos extremos tienen coeficiente 1. Para $1 \le j \le n$, la identidad de Pascal

$$\binom{n}{j-1} + \binom{n}{j} = \binom{n+1}{j}$$

permite reunir las sumas. Obtenemos

$$(fg)^{(n+1)} = \sum_{j=0}^{n+1} \binom{n+1}{j} f^{(j)} g^{(n+1-j)},$$

que es la fórmula para $n+1$. La inducción queda cerrada. □

Dónde se usan las hipótesis. La inducción no necesita continuidad de las derivadas. Necesita que cada término que se deriva esté realmente definido: para llegar al orden n debemos disponer de $f^{(k)}$ y $g^{(k)}$ hasta ese orden. La identidad combinatoria de Pascal organiza coeficientes; las reglas de producto y suma justifican el paso analítico.

Ejemplo 11.4. **Derivar un producto sin expandirlo.** Sea $h(x) = x^3(x-1)^4$.

Para calcular $h^{(5)}$ con Leibniz solo pueden contribuir los índices para los que $k \le 3$ y $5-k \le 4$, es decir $k = 1, 2, 3$. Por tanto,

$$h^{(5)} = \binom{5}{1} (x^3)'((x-1)^4)^{(4)} + \binom{5}{2} (x^3)''((x-1)^4)^{(3)} + \binom{5}{3} (x^3)'''((x-1)^4)''.$$

Como

$$(x^3)' = 3x^2, \quad (x^3)'' = 6x, \quad (x^3)''' = 6,$$

y

$$((x-1)^4)^{(4)} = 24, \quad ((x-1)^4)^{(3)} = 24(x-1), \quad ((x-1)^4)'' = 12(x-1)^2,$$

resulta

$$h^{(5)}(x) = 360x^2 + 1440x(x-1) + 720(x-1)^2.$$

La expansión final puede servir como control, pero no era necesaria para organizar la derivación.

Una identidad útil para potencias centradas. Para $m, n \in \mathbb{N}$,

$$((x-a)^m)^{(n)} = \begin{cases} \dfrac{m!}{(m-n)!}(x-a)^{m-n}, & n \le m, \\ 0, & n > m. \end{cases}$$

Esta fórmula será la pieza algebraica central de la unicidad del polinomio de Taylor. Al evaluar en a, todas las potencias centradas desaparecen excepto aquella cuyo grado coincide exactamente con el número de derivaciones.

COROLARIO 11.1. **Derivadas de un polinomio centrado.** Si

$$P(x) = \sum_{j=0}^{m} c_j (x-a)^j,$$

entonces, para $0 \le k \le m$,

$$\boxed{P^{(k)}(a) = k!\, c_k.}$$

Además $P^{(k)} \equiv 0$ para todo $k > m$.

Demostración. Por linealidad, derivamos término a término. Al evaluar en a, los términos con $j > k$ conservan una potencia positiva de $x - a$ y valen cero; los términos con $j < k$ ya han desaparecido; el término $j = k$ vale $k!c_k$. □

Una advertencia sobre composiciones. No existe una versión de la regla de la cadena para derivadas superiores tan simple como $(f \circ g)^{(n)} = (f^{(n)} \circ g)(g')^n$.

Ya para $n = 2$ aparece

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

En este capítulo no necesitamos desarrollar la fórmula combinatoria general para composiciones. Cuando una derivada superior de una composición concreta sea necesaria, aplicaremos repetidamente las reglas ya probadas y controlaremos cada término. La fórmula de Taylor puede construirse sin introducir una maquinaria combinatoria adicional.

Antes de seguir. ¿Por qué los coeficientes binomiales aparecen en Leibniz? ¿Qué términos sobreviven al evaluar $P^{(k)}$ en el centro a? ¿En qué punto de la demostración se usa que existen derivadas hasta orden n y no solo sus valores en un único punto?

11.4. El polinomio de Taylor y su unicidad

La pregunta algebraica de §11.1 ya puede resolverse exactamente. Dados los números $f(a), f'(a), \dots, f^{(n)}(a)$, buscamos un polinomio de grado a lo sumo n que reproduzca todos esos datos en a.

DEFINICIÓN 11.3. **Polinomio de Taylor finito.** Sea f una función cuyas derivadas $f^{(k)}(a)$ existen para $0 \le k \le n$. Definimos el **polinomio de Taylor de grado n de f centrado en a** por

$$\boxed{T_{n,a}f(x) := \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k.}$$

Usamos las convenciones $f^{(0)} = f$ y $0! = 1$.

La palabra «de grado n» describe la truncación elegida; el grado algebraico efectivo puede ser menor si $f^{(n)}(a) = 0$ u otros coeficientes finales se anulan.

TEOREMA 11.1. **Caracterización y unicidad del polinomio de Taylor.** Bajo las hipótesis de Definición 11.3, $T_{n,a}f$ es el único polinomio P de grado a lo sumo n que satisface

$$P^{(k)}(a) = f^{(k)}(a) \qquad (0 \le k \le n).$$

Demostración. Escribamos

$$T_{n,a}f(x) = \sum_{j=0}^{n} c_j (x-a)^j, \qquad c_j := \frac{f^{(j)}(a)}{j!}.$$

Por Corolario 11.1,

$$(T_{n,a}f)^{(k)}(a) = k!c_k = f^{(k)}(a),$$

así que el polinomio construido posee la propiedad requerida.
Para la unicidad, sea

$$P(x) = \sum_{j=0}^{n} d_j (x-a)^j$$

un polinomio con las mismas derivadas en a. De nuevo por Corolario 11.1,

$$k!d_k = P^{(k)}(a) = f^{(k)}(a),$$

por lo que

$$d_k = \frac{f^{(k)}(a)}{k!} = c_k$$

para cada k. Todos los coeficientes coinciden y $P = T_{n,a}f$. □

El caso lineal reaparece exactamente. Para $n = 1$, $T_{1,a}f(x) = f(a) + f'(a)(x-a)$, que es la linealización del capítulo 6. Para $n = 2$,

$$T_{2,a}f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2.$$

Nada se ha dicho todavía sobre el error de usar esos polinomios fuera del centro. Teorema 11.1 es un teorema de **interpolación de derivadas en un punto**, no un teorema de aproximación cuantitativa.

Ejemplo 11.5. **Dos centros producen dos polinomios diferentes.** Para $f(x) = \frac{1}{x}$, el dominio excluye 0. En $a = 1$,

$$f(1) = 1, \qquad f'(1) = -1, \qquad f''(1) = 2,$$

y por tanto $T_{2,1}f(x) = 1 - (x-1) + (x-1)^2$.
En $a = 2$,

$$f(2) = \frac{1}{2}, \qquad f'(2) = -\frac{1}{4}, \qquad f''(2) = \frac{1}{4},$$

de modo que

$$T_{2,2}f(x) = \frac{1}{2} - \frac{1}{4}(x-2) + \frac{1}{8}(x-2)^2.$$

Un polinomio de Taylor siempre lleva incorporado un centro. Cambiar a cambia los datos que debe reproducir y, en general, cambia el polinomio.

Funciones distintas pueden compartir el mismo polinomio. Sea n fijo y consideremos, centrados en a,

$$f(x) = 0, \qquad g(x) = (x-a)^{n+1}.$$

Para todo $0 \le k \le n$, $f^{(k)}(a) = g^{(k)}(a) = 0$.

Por unicidad, $T_{n,a}f = T_{n,a}g = 0$.

Sin embargo f y g no son la misma función. El polinomio de orden n contiene exactamente un número finito de datos locales; no codifica la función completa.

El resto se define por una identidad. A partir de ahora escribiremos $R_{n,a}f(x) := f(x) - T_{n,a}f(x)$.

Entonces

$$\boxed{f(x) = T_{n,a}f(x) + R_{n,a}f(x)}$$

es exacta siempre que ambos miembros estén definidos. Todo contenido aproximativo vendrá de demostrar propiedades de $R_{n,a}f$.

Unicidad no equivale a precisión
El teorema de unicidad responde: **¿qué polinomio de grado $\le n$ comparte esas derivadas?** No responde: **¿cuán cerca está de la función en otro punto?** La segunda pregunta requiere el resto y, por tanto, información sobre la función entre el centro y el punto evaluado.

Antes de seguir. ¿Por qué $T_{n,a}f$ puede tener grado efectivo menor que n? Si dos funciones comparten sus primeras n derivadas en a, ¿qué podemos afirmar sobre sus polinomios de Taylor? ¿Qué no podemos afirmar todavía sobre sus valores fuera de a?

11.5. Un lema de Rolle iterado

Rolle produce un cero de la derivada entre dos ceros de una función. Si además una derivada se anula en un extremo, podemos aplicar de nuevo el mismo argumento. La iteración permitirá obtener la forma de Lagrange del resto.

Del primer cero a una cadena de ceros. Supongamos que $F(a) = F(x) = 0$. Rolle produce un punto c_1 entre a y x con $F'(c_1) = 0$. Si también sabemos que $F'(a) = 0$, podemos aplicar Rolle a F' entre a y c_1 y encontrar c_2 con $F''(c_2) = 0$. Si además $F''(a) = 0$, repetimos. La geometría de los intervalos se va anidando:

$$a < c_n < \cdots < c_2 < c_1 < x$$

cuando $a < x$. Si $x < a$, el mismo argumento se aplica con la orientación invertida. Lo relevante es la palabra **entre**, no el orden elegido al escribir los puntos.

LEMA 11.1. **Lema de Rolle iterado.** Sea J un intervalo abierto, sean $a, x \in J$ con $a \neq x$, y sea $F : J \to \mathbb{R}$ una función $n+1$ veces diferenciable en J. Supongamos que

$$F(a) = F'(a) = \cdots = F^{(n)}(a) = 0$$

y además

$$F(x) = 0.$$

Entonces existe un punto c estrictamente entre a y x tal que

$$\boxed{F^{(n+1)}(c) = 0.}$$

Demostración. Basta tratar $a < x$; si $x < a$, intercambiamos el orden de los extremos en cada aplicación de Rolle.

Como F es diferenciable en J, es continua en $[a, x]$. De $F(a) = F(x) = 0$, Teorema 9.2 proporciona $c_1 \in (a, x)$ con

$$F'(c_1) = 0.$$

Ahora $F'(a) = 0$ por hipótesis. La función F' es diferenciable, luego continua, en el intervalo que contiene $[a, c_1]$. Aplicando Rolle a F' entre a y c_1 obtenemos $c_2 \in (a, c_1)$ con

$$F''(c_2) = 0.$$

Continuamos. Tras r pasos hemos obtenido

$$a < c_r < \cdots < c_1 < x, \qquad F^{(r)}(c_r) = 0.$$

Como $F^{(r)}(a) = 0$ para $r \leq n$, Rolle aplicado a $F^{(r)}$ en $[a, c_r]$ produce c_{r+1} con

$$F^{(r+1)}(c_{r+1}) = 0.$$

Al completar $n + 1$ aplicaciones obtenemos un punto c_{n+1} entre a y x tal que

$$F^{(n+1)}(c_{n+1}) = 0.$$

Tomamos $c = c_{n+1}$. □

Auditoría de hipótesis. Cada aplicación de Rolle necesita continuidad en el intervalo cerrado usado y diferenciabilidad en su interior. No hemos supuesto por separado todas esas continuidades: la existencia de la derivada siguiente garantiza que la derivada anterior es diferenciable y, por Teorema 6.1, continua. La hipótesis $F^{(k)}(a) = 0$ para $k = 0, \ldots, n$ no es ornamental. Después del primer paso, cada nueva aplicación de Rolle compara el cero recién encontrado con el cero ya disponible en a para la derivada del mismo orden.

Ejemplo 11.6. **Un caso de orden tres.** Supongamos que F es tres veces diferenciable,

$$F(a) = F'(a) = F''(a) = 0, \qquad F(x) = 0, \qquad a < x.$$

Rolle sobre F produce $c_1 \in (a, x)$ con $F'(c_1) = 0$. Rolle sobre F' entre a y c_1 produce c_2 con $F''(c_2) = 0$. Rolle sobre F'' entre a y c_2 produce c_3 con $F'''(c_3) = 0$. Obsérvese que no hemos supuesto tres ceros distintos de F. Un solo cero en a, acompañado por ceros de sus derivadas, actúa como un cero de orden diferencial alto.

Qué no dice el lema. El lema no localiza c ni afirma que sea único. Tampoco afirma que todo punto intermedio satisfaga la ecuación. Es una herramienta de existencia. En Taylor, precisamente esa falta de localización obligará a usar una cota válida en **todo** el intervalo si queremos reemplazar $f^{(n+1)}(c)$ por un número conocido.

Antes de seguir. ¿Dónde se usa $F'(a) = 0$ después de la primera aplicación de Rolle? ¿Por qué no necesitamos postular continuidad separada de $F^{(n)}$? ¿Qué cambiaría en la notación, pero no en la lógica, si $x < a$?

11.6. Fórmula de Taylor con resto de Lagrange

Ya tenemos todas las piezas. Fijaremos un grado n, un centro a y un punto x. El polinomio $T_{n,a}f$ reproduce las derivadas hasta orden n en a. Para medir el error construiremos una función auxiliar que tenga un cero adicional en x sin destruir esos ceros de derivadas en a.

TEOREMA 11.2. **Teorema de Taylor con resto de Lagrange.** Sea J un intervalo abierto, sean $a, x \in J$ y sea $n \in \mathbb{N}_0$. Supongamos que $f : J \to \mathbb{R}$ posee derivadas hasta orden $n+1$ en todo J. Si $x \neq a$, existe un punto c estrictamente entre a y x tal que

$$\boxed{f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}.}$$

Equivalentemente,

$$\boxed{R_{n,a}f(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}.}$$

No se supone continuidad de $f^{(n+1)}$.

Demostración. Fijemos $x \neq a$ y escribamos

$$P(t) := T_{n,a}f(t), \qquad R := f(x) - P(x).$$

Definamos

$$\lambda := \frac{R}{(x-a)^{n+1}}$$

y la función auxiliar

$$F(t) := f(t) - P(t) - \lambda(t-a)^{n+1}.$$

Por la elección de λ,

$$F(x) = f(x) - P(x) - R = 0.$$

Por Teorema 11.1, $P^{(k)}(a) = f^{(k)}(a)$ para $0 \leq k \leq n$. Además, para esos órdenes,

$$\Big((t-a)^{n+1}\Big)^{(k)}\Big|_{t=a} = 0.$$

Por tanto

$$F^{(k)}(a) = 0 \qquad (0 \leq k \leq n).$$

La función F posee $n+1$ derivadas en J: f las posee por hipótesis y los otros términos son polinomios. Aplicamos Lema 11.1. Existe c estrictamente entre a y x tal que

$$F^{(n+1)}(c) = 0.$$

Ahora P tiene grado a lo sumo n, luego $P^{(n+1)} \equiv 0$, mientras

$$\Big((t-a)^{n+1}\Big)^{(n+1)} = (n+1)!.$$

Así,

$$0 = F^{(n+1)}(c) = f^{(n+1)}(c) - \lambda(n+1)!.$$

Despejamos

$$\lambda = \frac{f^{(n+1)}(c)}{(n+1)!}.$$

Recordando que $R = \lambda(x-a)^{n+1}$ obtenemos

$$R = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1},$$

que es la fórmula afirmada. □

Dónde se usa cada hipótesis. La existencia de las derivadas hasta orden n permite construir P y garantizar el contacto en a. La derivada de orden $n+1$ permite aplicar Rolle una vez más y evaluar $F^{(n+1)}$ en el punto intermedio. El hecho de trabajar dentro de un intervalo abierto asegura que el segmento entre a y x permanece en el dominio y que las derivadas requeridas están disponibles alrededor de sus puntos interiores. No necesitamos que $f^{(n+1)}$ sea continua. El teorema afirma que existe un c donde esa derivada toma el valor necesario; la prueba obtiene ese punto por Rolle, no por el teorema del valor intermedio aplicado a $f^{(n+1)}$.

Los primeros casos. Para $n = 0$, el polinomio es la constante $f(a)$ y Taylor da $f(x) = f(a) + f'(c)(x-a)$, que es exactamente el teorema del valor medio de Lagrange escrito de otra forma.

Para $n = 1$,

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(c)}{2}(x-a)^2.$$

La recta tangente aparece acompañada por un error cuadrático cuyo coeficiente usa la segunda derivada en un punto intermedio.

Para $n = 2$,

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + \frac{f'''(c)}{6}(x-a)^3.$$

Ejemplo 11.7. **Taylor exacto para un polinomio.** Sea $f(x) = x^4 - 2x^2 + 3$ y tomemos $a = 1$, $n = 4$. Como f ya es un polinomio de grado 4, Teorema 11.1 implica $T_{4,1}f = f$ como polinomios. Además $f^{(5)} \equiv 0$, de modo que el resto de Lagrange es cero para cualquier punto intermedio. En este caso la «aproximación» de grado 4 es una identidad global exacta.

Si usamos solamente $n = 2$, el resto ya no tiene por qué anularse: hemos decidido conservar menos información aun cuando la función disponga de más.

El punto c depende del problema. La notación

$$R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}$$

no define en general una función explícita $c = c(x)$. El teorema asegura para cada x la existencia de al menos un punto intermedio adecuado. Dos valores diferentes de x pueden requerir puntos distintos, y para un mismo x puede haber más de uno. Esta es la razón por la que una estimación útil del error no intentará «encontrar c». Buscará una cota que funcione para todos los posibles puntos intermedios.

Taylor finito es una igualdad exacta
La fórmula del teorema no dice simplemente que $f(x) \approx T_{n,a}f(x)$. Dice exactamente $f(x) = T_{n,a}f(x) + R_{n,a}f(x)$, y representa el resto mediante una derivada de orden superior en un punto intermedio. El símbolo de aproximación solo será legítimo después de estimar ese resto.

Antes de seguir. ¿Por qué elegimos λ dividiendo por $(x-a)^{n+1}$? ¿Qué propiedad del polinomio P hace desaparecer su derivada de orden $n+1$? ¿En qué paso sería incorrecto reemplazar $f^{(n+1)}(c)$ por $f^{(n+1)}(a)$?

11.7. Cotas del resto y control cuantitativo del error

La fórmula de Lagrange contiene un punto c desconocido. Para convertirla en una herramienta cuantitativa debemos dominar la derivada de orden superior en todo el trayecto que ese punto podría ocupar.

Corolario 11.2. **Cota uniforme del resto de Taylor.** Bajo las hipótesis de Teorema 11.2, supongamos además que existe $M \geq 0$ tal que

$$|f^{(n+1)}(t)| \leq M$$

para todo t situado entre a y x. Entonces

$$\boxed{|R_{n,a}f(x)| \leq \frac{M}{(n+1)!}|x-a|^{n+1}.}$$

Demostración. El teorema de Taylor proporciona un punto intermedio c con

$$R_{n,a}f(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}.$$

Tomando valores absolutos y usando $|f^{(n+1)}(c)| \leq M$ obtenemos la desigualdad. □

La cota debe cubrir el segmento completo. Si $a < x$, el punto c pertenece a (a, x); si $x < a$, pertenece a (x, a). Una estimación de $f^{(n+1)}$ válida solo en a no controla necesariamente ese punto. Tampoco basta una cota verificada en un subconjunto que el segmento pueda abandonar.

Ejemplo 11.8. **Aproximar una función racional con certificado de error.** Tomemos $f(x) = \frac{1}{1+x}$ en el intervalo $[-1/2, 1/2]$ y centremos en $a = 0$. Sus derivadas iniciales son

$$f(0) = 1, \qquad f'(0) = -1, \qquad f''(0) = 2,$$

de modo que $T_{2,0}f(x) = 1 - x + x^2$.

Además $f'''(t) = -\frac{6}{(1+t)^4}$.

Si $|x| \le 1/4$, todo punto entre 0 y x satisface $t \in [-1/4, 1/4]$, y por tanto $1 + t \ge \frac{3}{4}$.

Así, $|f'''(t)| \le \frac{6}{(3/4)^4} = \frac{512}{27}$.

Por Corolario 11.2,

$$\left| \frac{1}{1+x} - (1 - x + x^2) \right| \le \frac{1}{6}\frac{512}{27}|x|^3 = \frac{256}{81}|x|^3$$

para $|x| \le 1/4$.

La estimación no exige conocer el punto c. Exige controlar una región que necesariamente lo contiene.

Error absoluto y objetivo de precisión. Si queremos garantizar $|R_{n,a}f(x)| < \varepsilon$, una condición suficiente es $\frac{M}{(n+1)!}|x-a|^{n+1} < \varepsilon$. Podemos usar esa desigualdad de dos maneras: fijado el grado n, determinar cuán cerca de a debemos trabajar; o, en contextos donde sepamos controlar las derivadas de órdenes sucesivos, comparar grados finitos diferentes. En este capítulo no convertiremos la segunda posibilidad en un procedimiento infinito ni en una teoría de series.

Una cota puede ser correcta sin ser óptima. Si sabemos que $|f^{(n+1)}(t)| \le M_1$ y también una cota menor $M_2 < M_1$, ambas producen certificados válidos; la segunda es simplemente más precisa. El propósito lógico de una cota es ser **demostrada** y cubrir todos los puntos posibles. Encontrar la mejor constante es un problema adicional.

El máximo de la derivada no se evalúa solo en el centro
La sustitución

$$|R_n(x)| \le \frac{|f^{(n+1)}(a)|}{(n+1)!}|x-a|^{n+1}$$

no se deduce de Taylor salvo que hayamos demostrado previamente que $|f^{(n+1)}(t)| \le |f^{(n+1)}(a)|$ en todo el segmento relevante. El teorema selecciona c, no necesariamente a.

Escala del error para un grado fijo. Si existe una cota M de $|f^{(n+1)}|$ en un entorno fijo de a, entonces para x en ese entorno

$$|R_{n,a}f(x)| \le C|x-a|^{n+1}, \qquad C := \frac{M}{(n+1)!}.$$

Esto se abrevia diciendo que el resto es de orden a lo sumo $|x-a|^{n+1}$ localmente. En la sección siguiente volveremos a la notación de pequeño-o del capítulo 6 para expresar una forma más fina de contacto.

Antes de seguir. Si el punto c es desconocido, ¿por qué una cota uniforme resuelve el problema? ¿Qué diferencia hay entre obtener una cota correcta y obtener la mejor cota? ¿Cómo

se transforma una cota del resto en una condición explícita sobre $|x-a|$ para alcanzar una tolerancia dada?

11.8. Orden de contacto y residuo normalizado

La fórmula de Lagrange proporciona una estimación de orden $n+1$ cuando existe una derivada adicional. Pero la idea de «compartir derivadas hasta orden n» admite otra lectura: después de quitar el polinomio de Taylor, el residuo se vuelve despreciable frente a $(x-a)^n$ bajo una hipótesis de continuidad natural.

Recordatorio: pequeño-o. Escribimos

$$r(x) = o((x-a)^n) \qquad (x \to a)$$

si

$$\lim_{x\to a,\, x\neq a} \frac{r(x)}{(x-a)^n} = 0.$$

Cuando n es impar, el denominador cambia de signo; eso no altera la definición. Equivalentemente puede usarse $|x-a|^n$ en el denominador al hablar del valor absoluto del cociente.

Proposición 11.2. **Forma de Peano bajo regularidad C^n.** Sea I un intervalo abierto, $a \in I$, $n \geq 1$ y $f \in C^n(I)$. Entonces

$$\boxed{f(x) = T_{n,a}f(x) + o((x-a)^n) \qquad (x \to a).}$$

Demostración. Para $x \neq a$ suficientemente próximo a a, aplicamos Teorema 11.2 de orden $n-1$. Como f posee n derivadas, existe un punto c_x entre a y x tal que

$$f(x) = T_{n-1,a}f(x) + \frac{f^{(n)}(c_x)}{n!}(x-a)^n.$$

Restamos el término de orden n del polinomio completo:

$$T_{n,a}f(x) = T_{n-1,a}f(x) + \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

Por tanto

$$\frac{f(x) - T_{n,a}f(x)}{(x-a)^n} = \frac{f^{(n)}(c_x) - f^{(n)}(a)}{n!}.$$

Como c_x está entre a y x,

$$|c_x - a| \leq |x-a|,$$

de modo que $c_x \to a$ cuando $x \to a$. La continuidad de $f^{(n)}$ en a implica

$$f^{(n)}(c_x) - f^{(n)}(a) \to 0.$$

El cociente anterior tiende a cero, que es precisamente la afirmación. □

Por qué la continuidad aparece aquí. Taylor–Lagrange de orden $n-1$ nos entrega $f^{(n)}(c_x)$, mientras el polinomio de grado n contiene $f^{(n)}(a)$. Para demostrar que la diferencia de ambos coeficientes desaparece cuando $x \to a$, necesitamos que los valores de $f^{(n)}$ cerca de a se acerquen a $f^{(n)}(a)$. Esa es exactamente la continuidad utilizada en la prueba. Si disponemos de $f^{(n+1)}$ acotada en un entorno, Corolario 11.2 da incluso $R_{n,a}f(x) = O((x-a)^{n+1})$, que implica la conclusión pequeño-o respecto de $(x-a)^n$. Pero Proposición 11.2 muestra que una derivada adicional no es necesaria bajo la hipótesis C^n.

Contacto de dos funciones. Supongamos que $f, g \in C^n(I)$ y

$$f^{(k)}(a) = g^{(k)}(a) \qquad (0 \le k \le n).$$

Entonces $T_{n,a}f = T_{n,a}g$.

Aplicando Proposición 11.2 a ambas y restando, $f(x) - g(x) = o((x-a)^n)$. En este sentido, compartir derivadas hasta orden n significa poseer contacto de orden al menos n en el centro. La afirmación es local; no dice que las funciones coincidan en un intervalo.

Ejemplo 11.9. **Mismo contacto cuadrático, funciones distintas.** Tomemos en $a = 0$

$$f(x) = x^2, \qquad g(x) = x^2 + x^3.$$

Se cumple

$$f(0) = g(0) = 0, \qquad f'(0) = g'(0) = 0, \qquad f''(0) = g''(0) = 2.$$

Por tanto comparten $T_{2,0}(x) = x^2$. Su diferencia es $g(x) - f(x) = x^3 = o(x^2)$. Sin embargo, para todo $x \neq 0$ suficientemente pequeño siguen siendo valores distintos. El orden de contacto mide la rapidez con que se separan, no convierte dos funciones en una.

El primer término que no coincide. Si $f, g \in C^m$ y sus derivadas coinciden hasta orden $m-1$, pero $f^{(m)}(a) \neq g^{(m)}(a)$, entonces sus polinomios de Taylor empiezan a diferir precisamente en el término $\frac{f^{(m)}(a)-g^{(m)}(a)}{m!}(x-a)^m$. Bajo las hipótesis de Proposición 11.2,

$$f(x) - g(x) = \frac{f^{(m)}(a) - g^{(m)}(a)}{m!}(x-a)^m + o((x-a)^m).$$

Esta fórmula permite comparar escalas locales sin recurrir todavía a una teoría asintótica general.

Antes de seguir. ¿Dónde se usa la continuidad de $f^{(n)}$ en la prueba de Peano? ¿Por qué $c_x \to a$ aunque no tengamos una fórmula explícita para c_x? Si dos funciones comparten el polinomio de Taylor de grado n, ¿qué información local aporta y qué información global sigue ausente?

11.9. Hipótesis, fallos y frontera con las series infinitas

Taylor es una herramienta poderosa precisamente porque sus conclusiones son precisas. Esa precisión se pierde si borramos las hipótesis o si reemplazamos una afirmación finita por otra infinita que todavía no ha sido demostrada.

Tener un polinomio no garantiza el orden siguiente. La definición de $T_{n,a}f$ solo utiliza derivadas hasta orden n en a. El resto de Lagrange de orden n, en cambio, requiere una derivada de orden $n+1$ en el intervalo relevante.

Consideremos $f(x) = |x|^3$.

La función es dos veces diferenciable en 0 y

$$f(0) = f'(0) = f''(0) = 0,$$

por lo que $T_{2,0}f = 0$.

Sin embargo la tercera derivada no existe en 0: para $x > 0$, $f'''(x) = 6$, y para $x < 0$, $f'''(x) = -6$. No podemos invocar Teorema 11.2 con $n = 2$ en un intervalo abierto que contenga el origen. El polinomio existe; esa forma particular del resto, no.

Una coincidencia finita nunca determina toda la función. Para cualquier grado fijo n,

$$f(x) = P(x), \qquad g(x) = P(x) + (x-a)^{n+1}$$

comparten todas las derivadas hasta orden n en a, pero son funciones distintas. Por tanto ningún conjunto finito de datos $f(a), \dots, f^{(n)}(a)$ determina la función en un entorno.

«Para cada n» y «al dejar $n \to \infty$» son afirmaciones distintas. Supongamos que una función posee derivadas de todos los órdenes y que, para cada entero fijo n, podemos formar $T_{n,a}f$. Eso produce una sucesión de polinomios $T_{0,a}f, T_{1,a}f, T_{2,a}f, \dots$ pero no demuestra por sí solo que para un $x \neq a$ $T_{n,a}f(x) \to f(x)$. Para concluirlo necesitaríamos estudiar el comportamiento de $R_{n,a}f(x)$ cuando **el grado** n tiende a infinito. Ese es un límite nuevo: ya no es $x \to a$ con n fijo. La teoría correspondiente exige herramientas sobre sucesiones y series de funciones y pertenece al Tomo II.

Fórmula finita y serie de Taylor no son sinónimos
La identidad $f(x) = T_{n,a}f(x) + R_{n,a}f(x)$ para cada n fijo no autoriza a borrar el resto y escribir una suma infinita. El paso $n \to \infty$ debe justificarse por separado. En este tomo no lo daremos por supuesto.

Una cota local tampoco es una afirmación global. Si Corolario 11.2 controla el resto para $|x-a| \leq r$, no hemos probado nada para puntos fuera de ese intervalo. El centro, el grado y la región de control son parte de los datos del certificado.

Ejemplo 11.10. **Una aproximación puede mejorar localmente y empeorar lejos.** Sea $f(x) = \frac{1}{1+x}$ centrada en 0. El polinomio cuadrático es $P_2(x) = 1 - x + x^2$.

La identidad algebraica $\frac{1}{1+x} - (1 - x + x^2) = -\frac{x^3}{1+x}$ muestra el error exacto donde $x \neq -1$. Cerca de 0, el numerador cúbico hace pequeño el error. Pero al aproximarnos a -1, el denominador se hace arbitrariamente pequeño y el error crece sin límite. Un dato local no atraviesa una singularidad del dominio.

La inaplicabilidad de un teorema no prueba la negación de su conclusión. Si falta $f^{(n+1)}$, no podemos usar el resto de Lagrange de orden n. Eso no implica automáticamente que $T_{n,a}f$ sea una mala aproximación ni que ninguna otra estimación sea posible. Como en capítulos anteriores, distinguir «el teorema no aplica» de «la conclusión es falsa» evita muchas demostraciones inválidas.

Antes de seguir. ¿Qué hipótesis falta para aplicar Taylor de orden dos a $|x|^3$ alrededor de 0? ¿Por qué una fórmula válida para todo n fijo no responde al problema $n \to \infty$? ¿Qué enseña Ejemplo 11.10 sobre la región donde una estimación local puede utilizarse?

11.10. Laboratorio de demostraciones y aproximaciones finitas

Esta sección reúne las decisiones que más fácilmente se automatizan mal. No añadiremos un nuevo teorema general; reconstruiremos la cadena de razonamiento y auditaremos qué puede concluirse en cada etapa.

Estación 1 — Elegir primero el objeto que se quiere controlar. Antes de calcular derivadas, fijemos cuatro datos:

- el centro a;
- el punto o intervalo donde queremos evaluar;
- el grado n;
- la tolerancia o el tipo de conclusión buscada.

Sin esos datos, «calcular Taylor» es una instrucción incompleta. El mismo f genera polinomios distintos según el centro y el grado.

Estación 2 — Reconstruir un polinomio a partir de sus derivadas. Supongamos que un polinomio P de grado a lo sumo 3 satisface

$$P(2) = 1, \quad P'(2) = -4, \quad P''(2) = 6, \quad P'''(2) = 12.$$

Por Teorema 11.1, sin resolver un sistema lineal en la base $1, x, x^2, x^3$,

$$P(x) = 1 - 4(x-2) + 3(x-2)^2 + 2(x-2)^3.$$

La base centrada convierte los datos derivados en coeficientes directamente.

Estación 3 — Separar el cálculo del polinomio de la prueba del error. Sea

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

Podemos calcular $T_{2,1}f$ usando solo derivadas en 1. Pero para certificar su error en $x = 6/5$ necesitamos estudiar $f'''(t) = -\frac{6}{t^4}$ en **todo** $[1, 6/5]$. Como $t \geq 1$, $|f'''(t)| \leq 6$, y entonces

$$|R_{2,1}f(6/5)| \leq \frac{6}{3!}\left(\frac{1}{5}\right)^3 = \frac{1}{125}.$$

El cálculo del polinomio fue puntual; el certificado del error fue intervalar.

Estación 4 — Auditar una prueba circular. Considérese el argumento:

> «El resto de Taylor tiene un cero de orden $n+1$ en a, así que debe ser igual a $f^{(n+1)}(c)(x-a)^{n+1}/(n+1)!$.»

La primera parte solo describe derivadas nulas en el centro. La segunda parte ya es prácticamente la conclusión del teorema y no se sigue de manera algebraica. La reparación consiste en construir la función auxiliar de §11.6 y aplicar el lema de Rolle iterado. El punto intermedio c debe ser **producido por un teorema**, no introducido por intuición.

Estación 5 — Dos aproximaciones con la misma información central. Para

$$f(x) = \frac{1}{1+x}, \qquad a = 0,$$

los grados 1 y 2 son

$$T_{1,0}f(x) = 1 - x, \qquad T_{2,0}f(x) = 1 - x + x^2.$$

En $x = 1/10$, las identidades exactas $\frac{1}{1+x} - (1-x) = \frac{x^2}{1+x}$,

$$\frac{1}{1+x} - (1 - x + x^2) = -\frac{x^3}{1+x}$$

muestran que el segundo error tiene un factor adicional $|x|$. Esto concuerda con la teoría del orden, pero aquí puede comprobarse sin aproximación decimal.

Estación 6 — Reparar una cota incompleta. Supongamos que alguien escribe:

$$|f'''(a)| \le 10 \quad \Longrightarrow \quad |R_2(x)| \le \frac{10}{6}|x-a|^3.$$

La inferencia es inválida. Debe reemplazarse la hipótesis puntual por $|f'''(t)| \le 10$ para todo t entre a y x. Solo entonces el punto c seleccionado por Taylor queda cubierto.

Estación 7 — Comparar funciones por el primer término diferente. Sean

$$f(x) = x^2 + x^4, \qquad g(x) = x^2 - 3x^4.$$

En 0 coinciden valor, primera, segunda y tercera derivadas. La primera diferencia aparece en orden 4: $f(x) - g(x) = 4x^4$. No hace falta dibujar para saber que, para $x \neq 0$, $f(x) > g(x)$. El análisis del contacto identifica exactamente la escala en la que se separan.

Estación 8 — Mapa de decisión. Ante un problema de Taylor finito, una ruta segura es:

1. comprobar dominio y segmento entre a y x;
2. fijar el orden n;
3. verificar la existencia de las derivadas requeridas;
4. construir $T_{n,a}f$ con datos en a;
5. definir el resto como diferencia exacta;
6. si se requiere la forma de Lagrange, verificar la derivada de orden $n+1$ en el intervalo;
7. si se requiere una cota, dominar esa derivada en todo el segmento;
8. solo entonces traducir la desigualdad del resto a una afirmación de precisión.

Esta ruta impide tres atajos comunes: introducir una serie infinita, evaluar la derivada superior solo en el centro o tratar el símbolo $\approx$ como si fuera una conclusión matemática.

Frontera del capítulo
El laboratorio termina en aproximaciones polinómicas **finitas** y certificados del error. El uso sistemático de esas aproximaciones para construir algoritmos iterativos, incluido Newton, queda para `T1-C13`. La pregunta de convergencia cuando el grado tiende a infinito queda para el Tomo II.

Al utilizar Taylor conviene distinguir la construcción del polinomio, la identidad exacta que define el resto y la estimación que garantiza una precisión. Conocer los coeficientes resuelve la primera tarea; la cota del error requiere las hipótesis y el control adicional de derivadas que especifica el teorema.

11.11. Ejercicios y soluciones

El banco entrena por separado la iteración de derivadas, la combinatoria de Leibniz, la construcción del polinomio, la prueba de Taylor y el control del resto. Los cuarenta ejercicios son originales y se distribuyen en siete niveles:

$$6A + 6B + 6C + 6D + 6E + 6F + 4G = 40.$$

Cada enunciado tiene una solución desarrollada con el mismo número. Solo se utilizan herramientas disponibles hasta §11.10; no se requieren series infinitas, L'Hôpital, integración ni método de Newton.

Nivel A — Derivadas superiores, notación y dominios. Ejercicio A1. La derivación termina algebraicamente, no operacionalmente. Para $p(x) = x^7 - 3x^4 + 2x$, calcula $p^{(4)}$, $p^{(7)}$, $p^{(8)}$ y $p^{(10)}$. Explica por qué tiene sentido hablar de todas ellas aunque a partir de cierto orden sean idénticamente cero.

Ejercicio A2. Una segunda derivada pierde un punto. Para $f(x) = x|x|$, determina D_0, D_1 y D_2, donde D_k es el dominio de $f^{(k)}$. Justifica la existencia de $f'(0)$ y la inexistencia de $f''(0)$ mediante cocientes incrementales.

Ejercicio A3. Dos derivadas continuas y una tercera ausente. Estudia $f(x) = |x|^3$. Demuestra que $f \in C^2(\mathbb{R})$, pero que $f'''(0)$ no existe.

Ejercicio A4. Un empalme C^2 que no admite tercera derivada. Define

$$f(x) = \begin{cases} 2x^3, & x < 0, \\ x^3, & x \geq 0. \end{cases}$$

Demuestra que $f \in C^2(\mathbb{R})$ y que $f'''(0)$ no existe. Localiza exactamente qué derivada deja de poder iterarse en el origen.

Ejercicio A5. La derivada siguiente fuerza continuidad. Sea I un intervalo abierto. Demuestra que si $f^{(n+1)}$ existe en todo I, entonces $f^{(n)}$ es continua en I. Explica por qué la conversa es falsa usando el ejercicio A3 con $n = 2$.

Ejercicio A6. Aislar un coeficiente. Para enteros $m, k \geq 0$, calcula

$$\frac{d^k}{dx^k}(x-a)^m\Big|_{x=a}$$

en los tres casos $k < m$, $k = m$ y $k > m$. Deduce directamente Corolario 11.1 para un polinomio centrado.

Nivel B — Leibniz, inducción y cálculo iterado. Ejercicio B1. El triángulo de Pascal aparece antes de la fórmula general. Partiendo solo de la regla del producto, deriva explícitamente $(fg)'''$ y $(fg)^{(4)}$. Comprueba que los coeficientes son $1,3,3,1$ y $1,4,6,4,1$.

Ejercicio B2. Derivada n-ésima de una potencia. Demuestra por inducción que, para $0 \le n \le m$, $\frac{d^n}{dx^n} x^m = \frac{m!}{(m-n)!} x^{m-n}$, y demuestra que todas las derivadas posteriores son cero.

Ejercicio B3. Quinta derivada sin expansión inicial. Calcula mediante Proposición 11.1 la quinta derivada de $h(x) = x^4(x+2)^3$.

Simplifica el resultado y úsalo para comprobar que es un polinomio cuadrático.

Ejercicio B4. El grado controla las derivadas del producto. Sean P, Q polinomios de grados a lo sumo r y s. Usa Leibniz para demostrar que $(PQ)^{(n)} \equiv 0$ para todo $n > r + s$, sin expandir el producto.

Ejercicio B5. Dos órdenes de anulación se suman. Supón que f y g poseen derivadas suficientes alrededor de a y que

$$f^{(j)}(a) = 0 \ (0 \le j < r), \qquad f^{(r)}(a) \neq 0,$$

$$g^{(j)}(a) = 0 \ (0 \le j < s), \qquad g^{(s)}(a) \neq 0.$$

Demuestra con Leibniz que $(fg)^{(n)}(a) = 0$ para $n < r + s$ y que

$$(fg)^{(r+s)}(a) = \binom{r+s}{r} f^{(r)}(a) g^{(s)}(a) \neq 0.$$

Ejercicio B6. La cadena de segundo orden tiene dos términos. Deduce a partir de la regla de la cadena y la del producto que

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

Verifica la fórmula para $f(t) = t^3$ y $g(x) = x^2 + 1$ comparándola con una derivación directa de $(x^2+1)^3$.

Nivel C — Polinomio de Taylor y unicidad. Ejercicio C1. Centro distinto de cero. Calcula $T_{3,1} f$ para $f(x) = x^5 - 2x^3 + x + 4$.

Escríbelo en potencias de $x - 1$ y comprueba directamente que comparte con f las derivadas de órdenes $0, 1, 2, 3$ en 1.

Ejercicio C2. El recíproco alrededor de uno. Para $f(x) = 1/x$, calcula $T_{3,1} f$. Sin usar series infinitas, verifica algebraicamente la identidad exacta $\frac{1}{x} - T_{3,1} f(x) = \frac{(x-1)^4}{x}$ para $x \neq 0$.

Ejercicio C3. Una familia de coeficientes. Para $f(x) = 1/(1+x)$, demuestra por inducción que $f^{(k)}(x) = (-1)^k \frac{k!}{(1+x)^{k+1}}$.

Deduce $T_{4,0} f$ y comprueba por multiplicación que $\frac{1}{1+x} - T_{4,0} f(x) = -\frac{x^5}{1+x}$.

Ejercicio C4. Reconstrucción sin sistema lineal. Determina el único polinomio P de grado a lo sumo 4 que satisface

$$P(2) = 3, \quad P'(2) = -1, \quad P''(2) = 8, \quad P'''(2) = 0, \quad P^{(4)}(2) = -48.$$

Justifica la unicidad mediante Teorema 11.1.

Ejercicio C5. El mismo polinomio no determina la función. Fija $n \geq 1$, $a \in \mathbb{R}$ y un polinomio P de grado $\leq n$. Define

$$f(x) = P(x), \qquad g(x) = P(x) + (x-a)^{n+1}.$$

Demuestra que $T_{n,a}f = T_{n,a}g = P$, pero $f \neq g$.

Ejercicio C6. Un mismo polinomio visto desde dos centros. Para $q(x) = 2x^2 - 3x + 5$, calcula $T_{2,0}q$ y $T_{2,2}q$. Comprueba que las expresiones centradas son diferentes pero que, al expandirlas, ambas coinciden exactamente con q.

Nivel D — Rolle iterado y resto de Lagrange. Ejercicio D1. Tres aplicaciones escritas una por una. Sea F tres veces diferenciable en un intervalo abierto que contiene $[a, b]$, con $a < b$, y supón

$$F(a) = F'(a) = F''(a) = F(b) = 0.$$

Demuestra sin citar Lema 11.1 que existe $c \in (a, b)$ con $F'''(c) = 0$.

Ejercicio D2. Un cero múltiple y otro cero. Sea $F(x) = (x-a)^3(x-b)$ con $a < b$. Verifica las hipótesis del ejercicio D1 y concluye que F''' tiene un cero en (a, b). Calcúlalo explícitamente y comprueba la conclusión.

Ejercicio D3. El resto de una cúbica. Para $f(x) = x^3$, $a = 0$ y $n = 2$, aplica Teorema 11.2. Determina $T_{2,0}f$, el resto exacto y explica por qué cualquier punto c entre 0 y x satisface la fórmula de Lagrange.

Ejercicio D4. En este caso el punto intermedio puede localizarse. Sea $f(x) = 1/x$, $a = 1$, $n = 1$ y $x > 0$, $x \neq 1$. Compara el resto exacto con la forma de Lagrange y demuestra que puede tomarse $c = x^{1/3}$.

Prueba además que c está estrictamente entre 1 y x.

Ejercicio D5. Un polinomio de Taylor sin resto de Lagrange de ese orden. Para $f(x) = |x|^3$ demuestra que $T_{2,0}f = 0$. Explica exactamente por qué Teorema 11.2 con $n = 2$ no es aplicable en ningún intervalo abierto que contenga 0.

Ejercicio D6. Taylor de orden cero es Lagrange. Demuestra que Teorema 11.2 para $n = 0$ es equivalente al teorema del valor medio de Lagrange Teorema 9.3.

Nivel E — Cotas de error y orden de contacto. Ejercicio E1. Un certificado sobre un intervalo. Para $f(x) = 1/(1+x)$ y $T_{2,0}f = 1 - x + x^2$, demuestra que para $|x| \le 1/5$, $|R_{2,0}f(x)| \le \frac{125}{64}|x|^3$.

Compara esta cota con el error exacto $|x|^3/|1+x|$.

Ejercicio E2. Convertir una cota en una distancia admisible. Supón que en $|t-a| \le r$ se cumple $|f'''(t)| \le M$. Encuentra una condición suficiente sobre $|x-a|$ que garantice $|f(x) - T_{2,a}f(x)| < \varepsilon$.

Incluye la restricción necesaria para no salir del intervalo donde vale la cota.

Ejercicio E3. Un error exactamente nulo. Sea P un polinomio de grado $\le n$. Demuestra que para cualquier centro a, $T_{n,a}P = P$ y $R_{n,a}P \equiv 0$. Da dos pruebas: una por unicidad y otra usando que $P^{(n+1)} \equiv 0$ en Teorema 11.2.

Ejercicio E4. Una cota uniforme no necesita localizar c. Para $f(x) = 1/x$, centro $a = 1$ y grado 1, demuestra que si $|x-1| \le 1/4$, entonces $|R_{1,1}f(x)| \le \frac{64}{27}|x-1|^2$.

Indica de dónde proviene la constante.

Ejercicio E5. Contacto cúbico. Para

$$f(x) = x^3 + x^5, \qquad g(x) = x^3 - 2x^5,$$

calcula sus derivadas hasta orden 3 en 0, demuestra que comparten $T_{3,0}$ y comprueba directamente que $f(x) - g(x) = o(x^3)$.

Ejercicio E6. La primera derivada diferente fija la escala. Sean $f, g \in C^m(I)$, $a \in I$, y supón que

$$f^{(k)}(a) = g^{(k)}(a) \quad (0 \le k < m),$$

pero $f^{(m)}(a) \neq g^{(m)}(a)$. Deduce de Proposición 11.2 que

$$\lim_{x \to a} \frac{f(x) - g(x)}{(x-a)^m} = \frac{f^{(m)}(a) - g^{(m)}(a)}{m!}.$$

Nivel F — Hipótesis, contraejemplos y reparación de pruebas. Ejercicio F1. Existencia del polinomio no implica existencia del resto de Lagrange. Usa $f(x) = |x|^3$ en 0 para refutar la frase: «si existe $T_{n,a}f$, entonces Taylor con resto de Lagrange de orden n es aplicable».

Ejercicio F2. El punto intermedio no puede sustituirse por el centro. Para $f(x) = 1/x$, $a = 1$ y $n = 1$, muestra que en general es falsa la identidad $R_{1,1}f(x) = \frac{f''(1)}{2}(x-1)^2$.

Encuentra el resto exacto y señala qué inferencia inválida se habría cometido.

Ejercicio F3. Una estimación local no atraviesa una singularidad. Para $f(x) = 1/(1+x)$ y $T_{2,0}f = 1 - x + x^2$, demuestra que el error tiende a infinito en valor absoluto cuando $x \to -1$, aunque $R_{2,0}f(x) = O(x^3)$ cuando $x \to 0$. Explica por qué no hay contradicción.

Ejercicio F4. Mismos datos finitos, valores distintos. Construye dos polinomios distintos que tengan el mismo polinomio de Taylor de grado 4 en $a = 2$. Demuestra la afirmación mediante derivadas en el centro.

Ejercicio F5. Muchas derivadas nulas no fuerzan nulidad local. Para un entero fijo $n \geq 1$, da una función polinómica no nula tal que

$$f(a) = f'(a) = \cdots = f^{(n)}(a) = 0.$$

Explica por qué esto refuta cualquier inferencia que pretenda determinar una función a partir de un número finito de derivadas en un punto.

Ejercicio F6. Localiza el paso no demostrado. Un estudiante escribe: «para cada n, $f = T_{n,a}f + R_{n,a}f$; por tanto, haciendo $n \to \infty$, resulta $f = \sum_{k\geq 0} f^{(k)}(a)(x-a)^k/k!$». Identifica el paso que falta justificar y formula en términos de $R_{n,a}f(x)$ la condición que sería necesaria para ese x. No uses teoría de series que todavía no se ha desarrollado.

Nivel G — Desafíos integradores. Ejercicio G1. Polinomio y certificado en un solo problema. Sea $f(x) = \frac{1}{2-x}$.

Calcula $T_{3,0}f$ y demuestra que para $|x| \leq 1/2$, $|f(x) - T_{3,0}f(x)| \leq \frac{32}{243}|x|^4$.

Justifica cada hipótesis de Taylor y cada paso de la cota.

Ejercicio G2. Un parámetro decide el primer orden de separación. Sean

$$f_a(x) = x^4 + ax^5, \qquad g_b(x) = x^4 + bx^5.$$

Para $a, b \in \mathbb{R}$, determina el mayor orden $n \leq 5$ para el que necesariamente comparten $T_{n,0}$ y estudia qué cambia cuando $a = b$ o $a \neq b$. En el segundo caso calcula

$$\lim_{x\to 0} \frac{f_a(x) - g_b(x)}{x^5}.$$

Ejercicio G3. Derivadas nulas y divisibilidad. Sea P un polinomio y $a \in \mathbb{R}$. Demuestra que

$$P(a) = P'(a) = \cdots = P^{(n)}(a) = 0$$

si y solo si existe un polinomio Q tal que $P(x) = (x-a)^{n+1}Q(x)$. Usa la expansión centrada de P y Corolario 11.1.

Ejercicio G4. De un jet nulo a una desigualdad cuantitativa. Sea f $n+1$ veces diferenciable en un intervalo abierto que contiene el segmento entre a y x. Supón

$$f^{(k)}(a) = 0 \qquad (0 \leq k \leq n)$$

y que $|f^{(n+1)}(t)| \leq M$ para todo t entre a y x. Demuestra que $|f(x)| \leq \frac{M}{(n+1)!}|x-a|^{n+1}$. Explica por qué la conclusión combina una afirmación algebraica en a con una hipótesis analítica sobre todo el segmento.

Soluciones desarrolladas.

Soluciones del nivel A. **Solución A1.** Derivando cuatro veces, $p^{(4)}(x) = 840x^3 - 72$.

El término x^7 produce $7! = 5040$ en la séptima derivada y los términos de grados 4 y 1 ya han desaparecido, así que $p^{(7)}(x) = 5040$. La derivada de una constante es cero: $p^{(8)} \equiv 0$. Derivar la función cero vuelve a dar la función cero, luego también $p^{(10)} \equiv 0$. El proceso no deja de estar definido cuando alcanza cero; simplemente queda estacionado en la función cero. Por eso un polinomio pertenece a $C^n(\mathbb{R})$ para todo n.

Solución A2. La función está definida en toda la recta, luego $D_0 = \mathbb{R}$. Para $h \neq 0$, $\frac{f(h)-f(0)}{h} = \frac{h|h|}{h} = |h| \to 0$, así que $f'(0) = 0$. Fuera del origen,

$$f'(x) = \begin{cases} -2x, & x < 0, \\ 2x, & x > 0, \end{cases} = 2|x|,$$

y por tanto $D_1 = \mathbb{R}$. En cambio, $\frac{f'(h)-f'(0)}{h} = \frac{2|h|}{h}$ tiende a -2 por la izquierda y a 2 por la derecha. No existe $f''(0)$. Para $x \neq 0$ sí existe y vale -2 a la izquierda, 2 a la derecha, de modo que $D_2 = \mathbb{R} \setminus \{0\}$.

Solución A3. Para $x > 0$, $f(x) = x^3$; para $x < 0$, $f(x) = -x^3$. En el origen, $\frac{f(h)}{h} = \frac{|h|^3}{h} \to 0$, así que $f'(0) = 0$. Fuera de cero,

$$f'(x) = \begin{cases} -3x^2, & x < 0, \\ 3x^2, & x > 0. \end{cases}$$

El cociente $f'(h)/h$ tiende a 0 por ambos lados, de modo que $f''(0) = 0$. Para $x \neq 0$, $f''(x) = 6|x|$, y esta misma fórmula vale en 0; por tanto f'' es continua. Así $f \in C^2(\mathbb{R})$. Finalmente, $\frac{f''(h)-f''(0)}{h} = 6\frac{|h|}{h}$ tiene límites laterales -6 y 6. Luego $f'''(0)$ no existe.

Solución A4. Ambas ramas valen 0 en el origen. Los cocientes incrementales muestran $f'(0) = 0$. Para $x \neq 0$,

$$f'(x) = \begin{cases} 6x^2, & x < 0, \\ 3x^2, & x > 0, \end{cases}$$

y estas expresiones tienden a 0, luego f' es continua en 0. De nuevo por cocientes,

$$f''(0) = \lim_{h\to 0} \frac{f'(h)}{h} = 0.$$

Fuera del origen,

$$f''(x) = \begin{cases} 12x, & x < 0, \\ 6x, & x > 0, \end{cases}$$

y ambas ramas tienden a $0 = f''(0)$. Por tanto $f \in C^2(\mathbb{R})$. Para la tercera derivada en el origen,

$$\frac{f''(h) - f''(0)}{h} = \begin{cases} 12, & h < 0, \\ 6, & h > 0, \end{cases}$$

por lo que el límite no existe. La iteración falla al intentar derivar f'' en 0.

Solución A5. Si $f^{(n+1)}$ existe en todo I, entonces por definición $f^{(n)}$ es diferenciable en todo I. Toda función diferenciable es continua en el punto de diferenciabilidad, por Teorema 6.1. Por tanto $f^{(n)}$ es continua en I. La conversa «$f^{(n)}$ continua $\Rightarrow$ $f^{(n+1)}$ existe» es falsa: en A3, $f''(x) = 6|x|$ es continua en toda la recta, pero no es diferenciable en 0, de modo que $f'''(0)$ no existe.

Solución A6. Si $k < m$, después de k derivaciones queda un múltiplo de $(x-a)^{m-k}$, que al evaluar en a vale cero. Si $k = m$, obtenemos exactamente $m!$. Si $k > m$, la derivada ya es la función cero. Por tanto

$$\frac{d^k}{dx^k}(x-a)^m\Big|_{x=a} = \begin{cases} 0, & k < m, \\ m!, & k = m, \\ 0, & k > m. \end{cases}$$

Para $P(x) = \sum_{j=0}^{m} c_j (x-a)^j$, la linealidad permite evaluar término a término. En $P^{(k)}(a)$ solo sobrevive el sumando $j = k$, que aporta $k!c_k$. Así $P^{(k)}(a) = k!c_k$.

Soluciones del nivel B. **Solución B1.** A partir de $(fg)' = f'g + fg'$, $(fg)'' = f''g + 2f'g' + fg''$. Derivando una vez más,

$$(fg)''' = f'''g + 3f''g' + 3f'g'' + fg'''.$$

Otra derivación da

$$(fg)^{(4)} = f^{(4)}g + 4f'''g' + 6f''g'' + 4f'g''' + fg^{(4)}.$$

Los coeficientes interiores aparecen porque cada uno recibe contribuciones de dos términos adyacentes del orden anterior: $1 + 2 = 3$, $3 + 3 = 6$, etc. Es la identidad de Pascal que organiza Proposición 11.1.

Solución B2. Para $n = 0$ la fórmula es $x^m = x^m$. Supongamos $\frac{d^n}{dx^n}x^m = \frac{m!}{(m-n)!}x^{m-n}$ con $n < m$. Derivando,

$$\frac{d^{n+1}}{dx^{n+1}}x^m = \frac{m!}{(m-n)!}(m-n)x^{m-n-1} = \frac{m!}{(m-n-1)!}x^{m-n-1}.$$

Esto cierra la inducción hasta $n = m$, donde resulta $m!$. La derivada siguiente de esa constante es cero y todas las posteriores siguen siendo cero.

Solución B3. En Leibniz, $(x^4)^{(k)}$ se anula para $k > 4$ y $((x+2)^3)^{(5-k)}$ se anula si $5 - k > 3$, es decir, si $k < 2$. Solo contribuyen $k = 2, 3, 4$:

$$\begin{aligned} h^{(5)} &= \binom{5}{2}(12x^2)(6) + \binom{5}{3}(24x)(6(x+2)) + \binom{5}{4}(24)(3(x+2)^2) \\ &= 720x^2 + 1440x(x+2) + 360(x+2)^2 \\ &= 2520x^2 + 4320x + 1440. \end{aligned}$$

Como h tiene grado 7, su quinta derivada debe tener grado a lo sumo 2, coherente con el resultado.

Solución B4. Por Leibniz,

$$(PQ)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} P^{(k)} Q^{(n-k)}.$$

Si $n > r + s$, para cada k es imposible tener simultáneamente $k \le r$ y $n - k \le s$, pues eso implicaría $n \le r + s$. Por tanto, en cada término ocurre $k > r$ o $n - k > s$. En el primer caso $P^{(k)} \equiv 0$; en el segundo $Q^{(n-k)} \equiv 0$. Todos los sumandos son cero y $(PQ)^{(n)} \equiv 0$.

Solución B5. Para $n < r + s$, cada sumando $\binom{n}{k} f^{(k)}(a) g^{(n-k)}(a)$ tiene $k < r$ o $n - k < s$: si ambas desigualdades fallaran, tendríamos $n = k + (n - k) \ge r + s$. Por las hipótesis de anulación, cada sumando vale cero, luego $(fg)^{(n)}(a) = 0$.

Para $n = r + s$, un término puede ser no nulo solo si $k \ge r$ y $r + s - k \ge s$, lo que fuerza $k = r$. Así $(fg)^{(r+s)}(a) = \binom{r+s}{r} f^{(r)}(a) g^{(s)}(a)$. El coeficiente binomial es positivo y ambos factores finales son no nulos, por lo que el producto no es cero.

Solución B6. La cadena da $(f \circ g)' = (f' \circ g) g'$.

Derivamos como producto:

$$(f \circ g)'' = ((f' \circ g)') g' + (f' \circ g) g''.$$

Aplicando cadena al primer factor, $(f' \circ g)' = (f'' \circ g) g'$, y obtenemos la fórmula. Para $f(t) = t^3$, $f'(t) = 3t^2$, $f''(t) = 6t$, mientras $g' = 2x$, $g'' = 2$. Entonces

$$(f \circ g)'' = 6(x^2 + 1)(2x)^2 + 3(x^2 + 1)^2 \cdot 2 = 24x^2(x^2 + 1) + 6(x^2 + 1)^2.$$

Al expandir $(x^2 + 1)^3 = x^6 + 3x^4 + 3x^2 + 1$, su segunda derivada es $30x^4 + 36x^2 + 6$, y la expresión anterior se simplifica al mismo polinomio.

Soluciones del nivel C. **Solución C1.** Calculamos $f(1) = 4$,

$$f'(x) = 5x^4 - 6x^2 + 1, \qquad f'(1) = 0,$$

$$f''(x) = 20x^3 - 12x, \qquad f''(1) = 8,$$

$$f'''(x) = 60x^2 - 12, \qquad f'''(1) = 48.$$

Por tanto

$$T_{3,1} f(x) = 4 + 4(x - 1)^2 + 8(x - 1)^3.$$

Derivando esta expresión y evaluando en 1 obtenemos, respectivamente, $4, 0, 8, 48$, los mismos cuatro datos de f. Esa coincidencia es exactamente la caracterización de Teorema 11.1.

Solución C2. Para $f(x) = x^{-1}$, $f^{(k)}(1) = (-1)^k k!$, para $k = 0, 1, 2, 3$. Así

$$T_{3,1}f(x) = 1 - (x-1) + (x-1)^2 - (x-1)^3.$$

Poniendo $u = x - 1$, tenemos $x = 1 + u$ y $(1+u)(1-u+u^2-u^3) = 1 - u^4$.

Por tanto

$$\frac{1}{x} - T_{3,1}f(x) = \frac{1 - xT_{3,1}f(x)}{x} = \frac{(x-1)^4}{x}.$$

La verificación es finita y algebraica; no se ha usado ninguna serie geométrica infinita.

Solución C3. Para $k = 0$ la fórmula es la propia definición. Si $f^{(k)}(x) = (-1)^k k!(1+x)^{-k-1}$, entonces

$$f^{(k+1)}(x) = (-1)^k k!(-k-1)(1+x)^{-k-2} = (-1)^{k+1}(k+1)!(1+x)^{-k-2}.$$

La inducción queda probada. En 0, $f^{(k)}(0) = (-1)^k k!$, de modo que $T_{4,0}f(x) = 1 - x + x^2 - x^3 + x^4$.

Multiplicando,

$$(1+x)(1-x+x^2-x^3+x^4) = 1 + x^5.$$

Luego

$$\frac{1}{1+x} - T_{4,0}f(x) = \frac{1 - (1+x)T_{4,0}f(x)}{1+x} = -\frac{x^5}{1+x}.$$

Solución C4. El teorema de unicidad obliga a

$$P(x) = 3 - (x-2) + \frac{8}{2!}(x-2)^2 + \frac{0}{3!}(x-2)^3 + \frac{-48}{4!}(x-2)^4.$$

Por tanto

$$\boxed{P(x) = 3 - (x-2) + 4(x-2)^2 - 2(x-2)^4.}$$

Teorema 11.1 prueba simultáneamente que esta expresión posee los datos prescritos y que ningún otro polinomio de grado ≤ 4 puede poseerlos todos.

Solución C5. La diferencia $g(x) - f(x) = (x-a)^{n+1}$ tiene derivadas de órdenes $0, 1, \ldots, n$ nulas en a. Por tanto $g^{(k)}(a) = f^{(k)}(a)$ para esos órdenes. Como P tiene grado $\leq n$, $T_{n,a}f = P$; y por los mismos datos, la unicidad da $T_{n,a}g = P$. Sin embargo, si $x \neq a$, $(x-a)^{n+1} \neq 0$, de modo que $f(x) \neq g(x)$. Un número finito de derivadas no determina la función completa.

Solución C6. En $a = 0$,

$$q(0) = 5, \quad q'(0) = -3, \quad q''(0) = 4,$$

por lo que $T_{2,0}q(x) = 5 - 3x + 2x^2 = q(x)$.

En $a = 2$,

$$q(2) = 7, \quad q'(2) = 5, \quad q''(2) = 4,$$

luego

$$T_{2,2}q(x) = 7 + 5(x-2) + 2(x-2)^2.$$

Al expandir,

$$7 + 5x - 10 + 2(x^2 - 4x + 4) = 2x^2 - 3x + 5.$$

Las bases centradas cambian, pero un polinomio de grado 2 es igual a su Taylor de grado 2 alrededor de cualquier centro.

Soluciones del nivel D. **Solución D1.** Como $F(a) = F(b) = 0$, Rolle aplicado a F en $[a, b]$ produce $c_1 \in (a, b)$ con $F'(c_1) = 0$. Como también $F'(a) = 0$, Rolle aplicado a F' en $[a, c_1]$ produce $c_2 \in (a, c_1)$ con $F''(c_2) = 0$. Finalmente $F''(a) = 0$, de modo que Rolle aplicado a F'' en $[a, c_2]$ produce $c \in (a, c_2)$ con $F'''(c) = 0$. Las continuidades requeridas provienen de la diferenciabilidad de la función correspondiente.

Solución D2. El factor $(x-a)^3$ asegura

$$F(a) = F'(a) = F''(a) = 0,$$

y el factor $(x-b)$ da $F(b) = 0$. Por D1, F''' tiene un cero entre a y b. Para localizarlo, escribimos $u = x - a$ y $d = b - a > 0$; entonces $F = u^3(u-d) = u^4 - du^3$.

Por tanto

$$F'''(x) = 24u - 6d = 6(4(x-a) - (b-a)).$$

Su cero es $c = a + \frac{b-a}{4} = \frac{3a+b}{4}$, que satisface $a < c < b$ porque $0 < (b-a)/4 < b - a$.

Solución D3. En 0 se anulan $f(0)$, $f'(0)$ y $f''(0)$, así que $T_{2,0}f = 0$.

El resto exacto es $R_{2,0}f(x) = x^3$. Como $f''' \equiv 6$, Taylor–Lagrange da $R_{2,0}f(x) = \frac{6}{3!}x^3 = x^3$. La expresión no depende de c porque la tercera derivada es constante. En consecuencia cualquier punto estrictamente intermedio cumple la representación.

Solución D4. El polinomio lineal es $T_{1,1}f(x) = 1 - (x-1) = 2 - x$.

El resto exacto es $\frac{1}{x} - (2 - x) = \frac{(x-1)^2}{x}$.

Como $f''(t) = 2/t^3$, la forma de Lagrange exige $\frac{(x-1)^2}{x} = \frac{1}{c^3}(x-1)^2$. Para $x \neq 1$ podemos cancelar $(x-1)^2$ y queda $c^3 = x$, es decir $c = x^{1/3}$. Si $x > 1$, entonces $1 < x^{1/3} < x$; si $0 < x < 1$, al elevar desigualdades positivas al cubo vemos $x < x^{1/3} < 1$. En ambos casos c está estrictamente entre 1 y x.

Solución D5. Por A3,

$$f(0) = f'(0) = f''(0) = 0,$$

por lo que $T_{2,0}f = 0$. Pero también vimos que $f'''(0)$ no existe. Teorema 11.2 con $n = 2$ exige derivadas hasta orden 3 en todo un intervalo abierto que contenga el segmento entre el centro y el punto considerado; cualquier intervalo abierto que contenga 0 incluye el punto donde falta la tercera derivada. La hipótesis falla. La existencia del polinomio de orden 2 no repara esa ausencia.

Solución D6. Para $n = 0$, $T_{0,a}f = f(a)$ y Teorema 11.2 afirma que existe c entre a y x tal que $f(x) = f(a) + f'(c)(x - a)$.

Si $x \neq a$, despejar produce $f'(c) = \frac{f(x)-f(a)}{x-a}$, que es exactamente la conclusión de Teorema 9.3 sobre el intervalo cuyos extremos son a y x. Recíprocamente, la fórmula del valor medio multiplicada por $x - a$ da Taylor de orden cero.

Soluciones del nivel E. **Solución E1.** Tenemos $f'''(t) = -\frac{6}{(1+t)^4}$.

Si $|x| \leq 1/5$, cualquier t entre 0 y x satisface $t \geq -1/5$, luego $1+t \geq \frac{4}{5}$ y $|f'''(t)| \leq \frac{6}{(4/5)^4} = \frac{1875}{128}$. Por el resto de grado 2, $|R_{2,0}f(x)| \leq \frac{1}{3!}\frac{1875}{128}|x|^3 = \frac{625}{256}|x|^3$. Esta cota es válida pero no coincide con la constante propuesta en el enunciado. De hecho, usando el error exacto

$$|R_{2,0}f(x)| = \frac{|x|^3}{|1+x|}$$

y $|1 + x| \geq 4/5$, obtenemos la cota más fuerte $|R_{2,0}f(x)| \leq \frac{5}{4}|x|^3$. Como $5/4 < 125/64$, queda demostrada también la desigualdad pedida. El ejercicio muestra que una cota de Lagrange no tiene por qué ser óptima.

Solución E2. Taylor de grado 2 da $|f(x) - T_{2,a}f(x)| \leq \frac{M}{3!}|x - a|^3$.

Si $M > 0$, basta exigir

$$|x - a| < \left(\frac{6\varepsilon}{M}\right)^{1/3}.$$

Además necesitamos permanecer donde la cota de la tercera derivada es válida, así que una condición suficiente completa es

$$|x - a| < \min\left\{r, \left(\frac{6\varepsilon}{M}\right)^{1/3}\right\}.$$

Si $M = 0$, el resto es cero en esa región y cualquier $|x - a| \leq r$ sirve.

Solución E3. Primera prueba: P mismo es un polinomio de grado $\leq n$ cuyas derivadas en a coinciden obviamente con las de P. Por la unicidad de Teorema 11.1, el polinomio de Taylor de grado n debe ser P; así el resto es cero.

Segunda prueba: $P^{(n+1)} \equiv 0$. Taylor–Lagrange da para cada $x \neq a$

$$R_{n,a}P(x) = \frac{P^{(n+1)}(c)}{(n+1)!}(x - a)^{n+1} = 0.$$

En $x = a$ el resto también es cero por definición. Por tanto $R_{n,a}P \equiv 0$.

Solución E4. Para $f(x) = 1/x$, $f''(t) = \frac{2}{t^3}$.

Si $|x - 1| \leq 1/4$, cualquier punto t entre 1 y x satisface $t \geq \frac{3}{4}$.

Por tanto $|f''(t)| \leq \frac{2}{(3/4)^3} = \frac{128}{27}$.

El resto de grado 1 queda acotado por

$$|R_{1,1}f(x)| \leq \frac{1}{2!}\frac{128}{27}|x - 1|^2 = \frac{64}{27}|x - 1|^2.$$

La constante proviene de una cota uniforme de la segunda derivada en todo el segmento posible, no de evaluar f'' solamente en 1.

Solución E5. Para ambas funciones,

$$f(0) = g(0) = 0, \quad f'(0) = g'(0) = 0, \quad f''(0) = g''(0) = 0, \quad f'''(0) = g'''(0) = 6.$$

Por unicidad, $T_{3,0}f(x) = T_{3,0}g(x) = x^3$.
La diferencia es $f(x) - g(x) = 3x^5$.
Entonces, para $x \neq 0$, $\frac{f(x)-g(x)}{x^3} = 3x^2 \to 0$.
Así $f - g = o(x^3)$.

Solución E6. Por Proposición 11.2, $f(x) = T_{m,a}f(x) + o((x-a)^m)$,

$$g(x) = T_{m,a}g(x) + o((x-a)^m).$$

Las derivadas de órdenes menores que m coinciden, de modo que al restar los polinomios solo queda

$$T_{m,a}f(x) - T_{m,a}g(x) = \frac{f^{(m)}(a) - g^{(m)}(a)}{m!}(x-a)^m.$$

La diferencia de dos términos $o((x-a)^m)$ sigue siendo pequeño-o de esa potencia. Dividiendo por $(x-a)^m$ y tomando límite obtenemos $\frac{f^{(m)}(a)-g^{(m)}(a)}{m!}$.

Soluciones del nivel F. **Solución F1.** Para $f(x) = |x|^3$, A3 muestra que $f(0) = f'(0) = f''(0) = 0$, de modo que existe $T_{2,0}f = 0$.

Pero $f'''(0)$ no existe. La fórmula de Lagrange de orden 2 requiere precisamente la tercera derivada en un intervalo abierto que contenga el centro y el punto de evaluación. Por tanto el antecedente «existe $T_{2,0}f$» es verdadero y la supuesta aplicabilidad del teorema es falsa. El polinomio usa derivadas hasta orden 2; el resto de Lagrange exige una más.

Solución F2. Ya calculamos $T_{1,1}f(x) = 2 - x$ y $R_{1,1}f(x) = \frac{(x-1)^2}{x}$.

Como $f''(1) = 2$, sustituir arbitrariamente c por 1 produciría $\frac{f''(1)}{2}(x-1)^2 = (x-1)^2$. Esto solo coincide con el resto exacto cuando $x = 1$, punto donde ambos restos son cero. Para $x \neq 1$ y $x \neq 0$, en general $1/x \neq 1$. El error lógico consiste en reemplazar el punto existencial c suministrado por Taylor por el centro a sin demostrar que la derivada superior tenga el mismo valor en ambos.

Solución F3. La identidad exacta de §11.9 es $R_{2,0}f(x) = -\frac{x^3}{1+x}$.

Al tender $x \to -1$, el numerador tiende a 1 en valor absoluto mientras $|1 + x| \to 0$, luego $|R_{2,0}f(x)| \to \infty$.

En cambio, cuando $x \to 0$, si por ejemplo $|x| \leq 1/2$, entonces $|1+x| \geq 1/2$ y $|R_{2,0}f(x)| \leq 2|x|^3$. Por tanto el resto es $O(x^3)$ localmente en 0. No hay contradicción: una cota local posee una región de validez y no puede atravesar la singularidad $x = -1$.

Solución F4. Un ejemplo sencillo es

$$P(x) = 1 + (x-2)^2, \qquad Q(x) = 1 + (x-2)^2 + (x-2)^5.$$

Para $0 \le k \le 4$, el término $(x-2)^5$ tiene derivada k-ésima nula al evaluar en 2. Por tanto $P^{(k)}(2) = Q^{(k)}(2)$ para esos cinco órdenes. Ambos tienen el mismo Taylor de grado 4, $1+(x-2)^2$, pero $Q - P = (x-2)^5$ no es la función cero.

Solución F5. Tomemos $f(x) = (x-a)^{n+1}$.

Es un polinomio no nulo. Para cada $0 \le k \le n$, su derivada k-ésima conserva un factor $(x-a)^{n+1-k}$ de exponente positivo; al evaluar en a obtenemos cero. Así todas las derivadas prescritas se anulan aunque $f(x) \neq 0$ para todo $x \neq a$. Un conjunto finito de datos derivados no puede determinar los valores vecinos de una función.

Solución F6. La identidad finita para cada n no justifica intercambiarla con un límite en el grado. Para obtener

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

sería necesario demostrar

$$\boxed{R_{n,a} f(x) \to 0 \quad (n \to \infty).}$$

Solo después podría identificarse el límite de los polinomios parciales con la suma infinita correspondiente, siempre que esa suma estuviera definida. El paso faltante es, por tanto, un teorema de convergencia del resto respecto del índice n, no una nueva manipulación algebraica de la fórmula finita.

Soluciones del nivel G. **Solución G1.** Por inducción, $f^{(k)}(x) = \frac{k!}{(2-x)^{k+1}}$.

En 0,

$$\frac{f^{(k)}(0)}{k!} = \frac{1}{2^{k+1}},$$

por lo que $T_{3,0} f(x) = \frac{1}{2} + \frac{x}{4} + \frac{x^2}{8} + \frac{x^3}{16}$.

Si $|x| \le 1/2$, todo punto t entre 0 y x satisface $t \le 1/2$, así $2 - t \ge \frac{3}{2}$. La cuarta derivada es $f^{(4)}(t) = \frac{24}{(2-t)^5}$, de modo que $|f^{(4)}(t)| \le \frac{24}{(3/2)^5} = \frac{256}{81}$. Taylor de grado 3 es aplicable porque la función racional tiene todas esas derivadas en el intervalo $[-1/2, 1/2]$, que evita su singularidad 2. Entonces $|R_{3,0} f(x)| \le \frac{1}{4!}\frac{256}{81}|x|^4 = \frac{32}{243}|x|^4$.

Solución G2. Ambos polinomios tienen los mismos términos hasta grado 4: en 0 sus derivadas de órdenes $0, 1, 2, 3$ son cero y la cuarta vale $4! = 24$. Por tanto siempre comparten $T_{4,0}(x) = x^4$.

La quinta derivada en 0 vale $5!a$ para f_a y $5!b$ para g_b. Si $a = b$, las funciones son idénticas y comparten también el Taylor de grado 5. Si $a \neq b$, ya no comparten el grado 5, por lo que el mayor orden $n \le 5$ común es 4. En ese caso $\frac{f_a(x) - g_b(x)}{x^5} = a - b$ para $x \neq 0$, y el límite es $a - b$.

Solución G3. Escribamos P en potencias centradas:

$$P(x) = \sum_{j=0}^{m} c_j (x-a)^j.$$

Por Corolario 11.1, $P^{(j)}(a) = j!c_j$. Si las derivadas de órdenes $0, \ldots, n$ son cero, entonces $c_0 = \cdots = c_n = 0$, y

$$P(x) = \sum_{j=n+1}^{m} c_j (x-a)^j = (x-a)^{n+1} \sum_{j=n+1}^{m} c_j (x-a)^{j-n-1}.$$

La suma final es un polinomio Q. Recíprocamente, si $P = (x-a)^{n+1}Q$, cada derivada de orden $k \leq n$ conserva al evaluar en a al menos un factor que se anula; equivalentemente, la expansión centrada empieza en grado $n+1$. Por tanto todas esas derivadas son cero.

Solución G4. Las hipótesis en a hacen que todos los coeficientes de Taylor hasta grado n sean cero: $T_{n,a}f \equiv 0$.

Por Teorema 11.2 existe c entre a y x tal que

$$f(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}.$$

Tomando valores absolutos y usando la cota uniforme, $|f(x)| \leq \frac{M}{(n+1)!}|x-a|^{n+1}$. La parte algebraica es la anulación del polinomio a partir de datos puntuales en a. La parte analítica es la cota de $f^{(n+1)}$ en todo el segmento, necesaria porque el punto c no está localizado de antemano.

Capítulo 12

Aproximación, método de Newton y problemas de optimización

En los capítulos anteriores el cálculo diferencial fue adquiriendo dos capacidades distintas. Por una parte, aprendimos a describir con rigor el comportamiento de una función: continuidad, monotonía, extremos, convexidad y derivadas sucesivas. Por otra, la linealización y la fórmula de Taylor nos permitieron reemplazar localmente una función por objetos más sencillos sin perder de vista el error. La aproximación y la optimización exigen convertir esa información en decisiones verificables. Consideremos tres problemas:

- una ecuación $f(x) = 0$ tiene una raíz que no conocemos exactamente: ¿cómo construir aproximaciones y saber si son buenas?;
- un cálculo produce varios decimales aparentemente estables: ¿qué demuestra que alguno de ellos es correcto?;
- un problema geométrico o cuantitativo pide maximizar o minimizar una magnitud: ¿cómo pasar de la descripción verbal a una función y cómo probar que el candidato obtenido es realmente óptimo?

Para evaluar una aproximación necesitamos relacionar el número calculado con el valor buscado mediante una cota de error. Una calculadora o varias iteraciones pueden sugerir cercanía; la justificación debe separar estos tres objetos:

$$\boxed{\text{valor buscado}}, \qquad \boxed{\text{aproximación calculada}}, \qquad \boxed{\text{cota demostrada del error}}.$$

El método de Newton proporcionará un ejemplo especialmente rico. Su fórmula surge de sustituir una curva por su recta tangente; por eso es natural desde el punto de vista del cálculo diferencial. Pero la geometría que inspira el método no demuestra por sí sola que las iteraciones converjan. Para obtener una afirmación matemática necesitaremos volver a Taylor, controlar derivadas y vigilar el intervalo donde se realizan los pasos. La optimización plantea una cautela análoga. Resolver $f'(x) = 0$ puede producir candidatos interiores, pero no demuestra ni que exista un óptimo global ni que el candidato encontrado sea uno. El dominio, los bordes, los puntos no diferenciables y las restricciones del problema forman parte de la demostración. Este capítulo será, por tanto, un laboratorio sobre una idea transversal:

calcular propone; las hipótesis y las cotas certifican.

No desarrollaremos aquí análisis numérico general ni optimización multivariable. Trabajaremos con una variable real y utilizaremos únicamente las herramientas ya construidas en este tomo.

12.1. Aproximar no es adivinar: valor, error y certificado

Tres afirmaciones que suelen confundirse. Supongamos que queremos conocer un número real α y hemos calculado otro número x. Pueden ocurrir al menos tres situaciones conceptualmente distintas.

1. **Sabemos que x está cerca de α.** Esto exige alguna información sobre $|x - \alpha|$.

2. **Una ecuación asociada casi se satisface.** Si α es una raíz de f, esto significa que $|f(x)|$ es pequeño.
3. **Los últimos decimales de varias aproximaciones coinciden.** Esto es una observación sobre el procedimiento de cálculo, no todavía una afirmación sobre α.

Ninguna de las dos últimas implica automáticamente la primera. El objetivo del capítulo será aprender qué hipótesis permiten convertir datos computables en información sobre el error real.

Definición 12.1. **Aproximación, error absoluto y certificado.** Sea $\alpha \in \mathbb{R}$ el valor que se desea aproximar y sea $x \in \mathbb{R}$ un valor calculado.

El **error absoluto** de x respecto de α es

$$E(x;\alpha) := |x - \alpha|.$$

Dado $\varepsilon > 0$, diremos que x es una **aproximación con tolerancia** ε si

$$|x - \alpha| \leq \varepsilon.$$

Llamaremos **certificado de error** a una demostración que establece una desigualdad de este tipo a partir de datos conocidos, sin necesitar conocer de antemano el valor exacto de α.

La definición permite evaluar una aproximación por la información disponible sobre su distancia al valor buscado. La cantidad de decimales escritos solo será útil si podemos justificar su precisión.

Un intervalo ya es un certificado. La forma más sencilla de certificar una aproximación consiste en encerrar el valor desconocido entre dos números conocidos. Si $a \leq \alpha \leq b$, entonces cualquier $x \in [a, b]$ satisface $|x - \alpha| \leq b - a$. Si elegimos el punto medio $m = \frac{a+b}{2}$, obtenemos una cota dos veces mejor: $|m - \alpha| \leq \frac{b-a}{2}$. No hace falta conocer α para demostrarla. La geometría del intervalo basta.

Ejemplo 12.1. **Un decimal certificado para $\sqrt{2}$.** La raíz cuadrada positiva de 2 está definida canónicamente por la construcción de raíces de `T1-C09` (Proposición 8.1). Comprobemos los cuadrados de dos racionales: $1.414^2 = 1.999396 < 2$, mientras que $1.415^2 = 2.002225 > 2$.

Como $x \mapsto x^2$ es estrictamente creciente en $[0, \infty)$, se sigue que $1.414 < \sqrt{2} < 1.415$. El punto medio del intervalo es $m = 1.4145$.

Por la observación anterior,

$$\boxed{|m - \sqrt{2}| < 0.0005.}$$

El número 1.4145 no se convierte en una buena aproximación porque «parece» cercano a $\sqrt{2}$ ni porque una calculadora muestre dígitos compatibles. La desigualdad anterior es el certificado.

El intervalo anterior se construyó mediante comparación exacta. Aunque Newton permita obtener aproximaciones con mayor rapidez, tendremos que justificar igualmente la traducción del cálculo a una cota para $|x - \alpha|$.

El residuo de una ecuación no es el error en la variable. Si α satisface $f(\alpha) = 0$, es natural medir cuánto falla una aproximación x mediante el número $|f(x)|$. A esta cantidad la llamaremos **residuo** de la ecuación en x. El residuo es computable si podemos evaluar $f(x)$; el error $|x - \alpha|$, en cambio, contiene el número desconocido α. Precisamente por eso es tentador sustituir uno por otro. Pero no son la misma cantidad. Para verlo sin ninguna complicación, fijemos una raíz $\alpha = 1$ y, para cada entero $N \geq 1$, consideremos $f_N(t) = \frac{t-1}{N}$. En $x = 2$ el error

respecto de la raíz es siempre $|2-1|=1$, mientras que el residuo vale $|f_N(2)|=\frac{1}{N}$. Podemos hacer el residuo tan pequeño como queramos aumentando N, sin mover en absoluto la aproximación $x=2$. Por tanto,

$$\boxed{\text{residuo pequeño} \not\Rightarrow \text{error pequeño}}$$

si no añadimos información sobre cómo transforma f las distancias cerca de la raíz. El teorema del valor medio nos dará precisamente una forma de añadir esa información: una cota inferior para $|f'|$ permitirá convertir un residuo en una cota de error. Reservaremos esa construcción para §12.6, después de estudiar el paso de Newton.

Cifras estables tampoco constituyen una demostración. Supongamos que un procedimiento produce

$$x_4 = 1.41421357, \qquad x_5 = 1.41421356, \qquad x_6 = 1.41421356.$$

La repetición de los últimos ocho dígitos es evidencia útil para orientar un cálculo, pero por sí sola no demuestra que $|x_6-\alpha|<10^{-8}$. Podría ocurrir que el procedimiento se hubiera estabilizado en un ciclo, que estuviera convergiendo a otro valor o que los errores de redondeo ocultaran el movimiento real. En §12.7 veremos ejemplos concretos en los que una iteración de Newton no se comporta como la intuición más optimista sugiere.

La precisión debe demostrarse
En este capítulo no declararemos correctos ciertos decimales únicamente porque dos aproximaciones consecutivas coincidan en ellos. Para afirmar una tolerancia buscaremos una desigualdad demostrada: un intervalo que contenga el valor, una cota derivada del teorema del valor medio, una estimación de Taylor u otro certificado equivalente.

Dos problemas, la misma disciplina. La misma distinción reaparecerá en optimización. Si una derivada se anula en c, $f'(c)=0$, hemos localizado un **candidato** interior. Todavía falta demostrar qué significa ese punto para el problema original. Puede ser un mínimo, un máximo o ninguna de las dos cosas; incluso si es un extremo local, puede no ser global. Además, el óptimo puede encontrarse en un borde o en un punto donde la derivada no existe. Así, tanto en aproximación de raíces como en optimización, seguiremos un patrón común:

$$\boxed{\text{construir un candidato} \longrightarrow \text{verificar hipótesis} \longrightarrow \text{producir un certificado}}.$$

Para una raíz, el certificado podrá ser una cota de $|x-\alpha|$. Para un problema de optimización, será una demostración de que ningún punto factible produce un valor mejor.

El siguiente paso: aislar antes de iterar. Antes de introducir Newton conviene resolver un problema más básico. Si una ecuación tiene varias raíces, una aproximación numérica no tiene significado completo hasta que sepamos **qué raíz** estamos siguiendo. Y si ni siquiera hemos demostrado que exista una raíz en la región examinada, una sucesión de números no puede reemplazar esa falta. En §12.2 volveremos al teorema del valor intermedio y a la monotonía para construir intervalos que contengan exactamente una raíz. Ese aislamiento proporcionará simultáneamente existencia, unicidad local y un primer certificado de error. Solo después reemplazaremos el refinamiento elemental del intervalo por la recta tangente de Newton.

Antes de seguir. Si sabemos que $\alpha\in[3,3.02]$, ¿qué error máximo tiene el punto medio del intervalo? ¿Por qué un residuo $|f(x)|=10^{-12}$ no permite, sin información adicional sobre f,

concluir que x está a distancia menor que 10^{-12} de una raíz? ¿Qué diferencia lógica hay entre «$f'(c) = 0$» y «c minimiza f en todo su dominio»?

12.2. Aislar una raíz antes de aproximarla

En §12.1 vimos que un intervalo que contiene al valor buscado ya proporciona un certificado de error. Para una ecuación $f(x) = 0$, esto sugiere un objetivo previo a cualquier iteración rápida: encontrar un intervalo en el que sepamos **que existe una raíz y que no hay otra**. Diremos informalmente que un intervalo $[a, b]$ **aísla una raíz** si contiene exactamente una solución de $f(x) = 0$. El aislamiento no dice todavía que conozcamos muchos decimales de esa raíz. Dice algo lógicamente anterior y más importante: sabemos **qué raíz** estamos intentando aproximar.

Cambio de signo: existencia, pero todavía no unicidad. Si f es continua en $[a, b]$ y $f(a)f(b) < 0$, entonces 0 está estrictamente entre $f(a)$ y $f(b)$. El teorema del valor intermedio (Teorema 5.2) garantiza por tanto algún $\alpha \in (a, b)$ tal que $f(\alpha) = 0$.

La conclusión es de existencia. No afirma que α sea la única raíz del intervalo.

Por ejemplo, $q(x) = x(x^2 - 1)$ satisface

$$q(-2) = -6 < 0 < 6 = q(2),$$

pero en $[-2, 2]$ tiene tres raíces:

$$-1, \qquad 0, \qquad 1.$$

El cambio de signo certifica que hay al menos una raíz; no basta para aislarla.

La monotonía convierte existencia en aislamiento.

Proposición 12.1. **Proposición (aislamiento por cambio de signo y monotonía estricta).** Sea

$$f\colon [a, b] \to \mathbb{R}$$

continua, con $a < b$. Supongamos que

$$f(a)f(b) < 0$$

y que f es estrictamente monótona en $[a, b]$. Entonces existe una única raíz

$$\alpha \in (a, b)$$

de la ecuación $f(x) = 0$.

En particular, $[a, b]$ aísla a α. Si

$$m = \frac{a + b}{2},$$

entonces

$$|m - \alpha| \leq \frac{b - a}{2}.$$

Demostración. Como f es continua y sus valores en los extremos tienen signos opuestos, Teorema 5.2 proporciona al menos una raíz $\alpha \in (a, b)$.

Falta la unicidad. Supongamos que hubiera dos raíces distintas $u < v$ en $[a, b]$. Si f es estrictamente creciente, entonces $f(u) < f(v)$, mientras que $f(u) = f(v) = 0$, contradicción. Si f es estrictamente decreciente, la desigualdad se invierte y aparece la misma contradicción. Por tanto la raíz es única. Como $\alpha \in [a, b]$, la cota del punto medio es exactamente la ya demostrada en §12.1:

$$\left| \frac{a+b}{2} - \alpha \right| \leq \frac{b-a}{2}.$$

Así, el mismo intervalo que aísla la raíz proporciona un primer certificado cuantitativo. □

La proposición separa dos responsabilidades:

$$\boxed{\text{continuidad + cambio de signo} \Rightarrow \text{existencia}},$$

$$\boxed{\text{monotonía estricta} \Rightarrow \text{unicidad}}.$$

Esta separación será útil más adelante. Un método numérico puede producir muchos candidatos; si previamente hemos aislado una raíz, podremos juzgar si esos candidatos permanecen vinculados al problema que queríamos resolver.

Cómo certificar la monotonía. A veces la monotonía se conoce directamente por álgebra. Otras veces resulta más cómodo usar la derivada. Si f es continua en $[a, b]$, diferenciable en (a, b) y

$$f'(x) > 0 \qquad \text{para todo } x \in (a, b),$$

entonces Teorema 10.1 demuestra que f es estrictamente creciente en $[a, b]$. Si en cambio $f'(x) < 0$ en todo el interior, resulta estrictamente decreciente. En ambos casos, un cambio de signo en los extremos permite aplicar Proposición 12.1. Conviene notar lo que **no** basta. Saber solo que $f'(c) > 0$ en un punto, o comprobar numéricamente que la derivada parece positiva en una tabla finita, no demuestra monotonía en todo el intervalo. La hipótesis debe cubrir cada punto del interior, o debe disponerse de otro argumento global equivalente.

Ejemplo 12.2. **Aislar y después refinar una raíz de un cúbico.** Consideremos $p(x) = x^3 + x - 1$.

En $[0, 1]$ tenemos

$$p(0) = -1 < 0 < 1 = p(1).$$

Además, $p'(x) = 3x^2 + 1 > 0$ para todo $x \in \mathbb{R}$. Por Teorema 10.1, p es estrictamente creciente; por Proposición 12.1 existe una única raíz $\alpha \in (0, 1)$. El intervalo $[0, 1]$ ya la aísla, aunque todavía sea ancho. Ahora podemos **refinar** el aislamiento sin cambiar de raíz. Evaluamos el punto medio:

$$p\left(\frac{1}{2}\right) = \frac{1}{8} + \frac{1}{2} - 1 = -\frac{3}{8} < 0,$$

por lo que

$$\alpha \in \left(\frac{1}{2}, 1\right).$$

El siguiente punto medio útil es 3/4:

$$p\left(\frac{3}{4}\right) = \frac{27}{64} + \frac{3}{4} - 1 = \frac{11}{64} > 0,$$

así que

$$\alpha \in \left(\frac{1}{2}, \frac{3}{4}\right).$$

Una nueva evaluación da

$$p\left(\frac{5}{8}\right) = \frac{125}{512} + \frac{5}{8} - 1 = -\frac{67}{512} < 0,$$

de modo que

$$\alpha \in \left(\frac{5}{8}, \frac{3}{4}\right).$$

Finalmente,

$$p\left(\frac{11}{16}\right) = \frac{1331}{4096} + \frac{11}{16} - 1 = \frac{51}{4096} > 0,$$

y obtenemos el intervalo certificado

$$\boxed{\frac{5}{8} < \alpha < \frac{11}{16}}.$$

Su punto medio es $m = \frac{21}{32}$, y el ancho del intervalo es 1/16. Por tanto,

$$\boxed{\left|\frac{21}{32} - \alpha\right| < \frac{1}{32}}.$$

No hemos usado ninguna cifra decimal de α. El certificado procede enteramente de signos exactos, continuidad y unicidad.

Aislar no es lo mismo que refinar. El ejemplo anterior permite distinguir dos tareas.

Aislar significa demostrar que un intervalo contiene exactamente una raíz. En el ejemplo, $[0, 1]$ ya cumplía esa función.

Refinar significa reemplazar ese intervalo por otro más pequeño que contiene la misma raíz. El teorema de bisección ya probado en Teorema 5.3 proporciona un procedimiento sistemático para hacerlo mientras se preserve el cambio de signo. Después de n divisiones de un intervalo inicial $[a_0, b_0]$, el intervalo retenido tiene longitud $\frac{b_0 - a_0}{2^n}$, y su punto medio queda a distancia a lo sumo $\frac{b_0 - a_0}{2^{n+1}}$ del punto seleccionado por la construcción.

Hay aquí una sutileza importante. El teorema de bisección no necesita que la raíz inicial sea única: construye una cadena de intervalos que converge a una raíz. Por eso, cuando queremos decir que estamos refinando **la raíz previamente identificada**, la unicidad obtenida mediante Proposición 12.1 cumple un papel adicional. En un intervalo aislante, todo subintervalo con cambio de signo necesariamente contiene esa misma raíz y ninguna otra.

Existencia, aislamiento y precisión son tres preguntas distintas
Para una ecuación $f(x) = 0$ conviene preguntar, en este orden:

1. **Existencia:** ¿hay al menos una raíz en la región?
2. **Aislamiento:** ¿podemos demostrar que allí hay exactamente una?
3. **Precisión:** ¿qué tan estrecho es el intervalo certificado que la contiene?

El IVT responde la primera pregunta; la monotonía estricta puede resolver la segunda; el ancho del intervalo responde cuantitativamente la tercera.

Por qué esto prepara el método de Newton. La bisección ofrece una ventaja lógica notable: mientras conservemos el cambio de signo, el nuevo intervalo sigue certificando la presencia de una raíz. Su costo es que la reducción del error es relativamente lenta: cada paso divide el ancho por dos. Newton tomará otra idea. En lugar de conservar siempre dos extremos con signos opuestos, utilizará información diferencial en un solo punto y sustituirá la curva por su aproximación afín. Esto puede producir mejoras mucho más rápidas, pero también elimina la seguridad automática que proporciona el encierro por signos. Por eso el orden conceptual importa:

primero aislar y saber qué raíz buscamos; después acelerar la aproximación.

En §12.3 construiremos el paso de Newton directamente desde la recta tangente y determinaremos con precisión dónde está definido.

Antes de seguir. Si una función continua cambia de signo en $[a, b]$, ¿qué afirmación garantiza el IVT y cuál no? ¿Qué hipótesis adicional de Proposición 12.1 convierte ese intervalo en aislante? Si un intervalo aislante tiene longitud 10^{-3}, ¿qué cota de error obtiene su punto medio? ¿Por qué una cadena de bisección puede converger a una raíz aunque el intervalo inicial contenga varias?

12.3. El paso de Newton desde la recta tangente

Una vez aislada la raíz, podemos usar el valor $f(x)$ y la pendiente $f'(x)$ en un punto de trabajo para proponer una aproximación nueva. La recta tangente indica cómo hacerlo. El método de Newton no comenzará como una fórmula que deba memorizarse. Lo construiremos resolviendo exactamente un problema más sencillo: encontrar el cero de la recta tangente en el punto actual. La dificultad consiste en recordar, durante toda la construcción, que el cero de esa recta no tiene por qué ser un cero de la función.

Fijar el punto actual y distinguirlo de la variable. Trabajemos con una función $f\colon I \to \mathbb{R}$ diferenciable en un intervalo abierto no vacío I. Esta elección permite utilizar derivadas bilaterales en cada punto de trabajo. Un intervalo cerrado $[a, b] \subset I$ podrá servir para aislar una raíz sin confundirse con el dominio completo de f. Fijemos $x \in I$. En este momento x es el punto conocido desde el cual construiremos el paso; usaremos t para el argumento variable de la recta. La linealización ya probada en Teorema 6.2 da $f(x+h) = f(x)+f'(x)h+r_x(h)$, para incrementos admisibles $x + h \in I$, donde podemos definir exactamente $r_x(h) = f(x + h) - f(x) - f'(x)h$.

La información analítica es

$$\lim_{h\to 0} \frac{r_x(h)}{h} = 0,$$

con $h \neq 0$ y $x + h \in I$. El subíndice recuerda que el residuo depende del centro x. La función afín correspondiente es

$$L_x(t) := f(x) + f'(x)(t-x), \qquad t \in \mathbb{R}.$$

Su gráfica es la recta tangente en $(x, f(x))$. La recta está definida para todo $t \in \mathbb{R}$, aunque f solo esté definida en I; la afirmación de aproximación, en cambio, se refiere a valores de t del dominio próximos a x. Conviene distinguir las funciones de las dos letras. Al variar t recorremos una recta fija. Al cambiar x construimos, en general, otra recta tangente.

Resolver la ecuación afín, no sustituir una aproximación por una igualdad. La ecuación que queremos resolver es $f(t) = 0$. Newton propone un candidato resolviendo en su lugar $L_x(t) = 0$.

Esta segunda ecuación es $f(x) + f'(x)(t-x) = 0$.

Si $f'(x) \neq 0$, las reglas algebraicas del cuerpo permiten despejar: $f'(x)(t-x) = -f(x)$,

$$t - x = -\frac{f(x)}{f'(x)},$$

$$\boxed{t = x - \frac{f(x)}{f'(x)}.}$$

La no anulación de la derivada se utiliza exactamente al dividir. Además, la solución de la ecuación afín es única: una recta con pendiente no nula corta al eje horizontal en un solo punto. El desplazamiento propuesto y el nuevo punto son objetos distintos:

$$h_x = -\frac{f(x)}{f'(x)}, \qquad x_{\text{nuevo}} = x + h_x.$$

Por ejemplo, un desplazamiento negativo significa que el nuevo punto queda a la izquierda del actual. No significa que el nuevo punto sea necesariamente negativo.

Qué igualdad hemos demostrado
La construcción demuestra exactamente $L_x(x_{\text{nuevo}}) = 0$. No demuestra $f(x_{\text{nuevo}}) = 0$. Si el nuevo punto pertenece a I, la identidad de linealización permite escribir $f(x_{\text{nuevo}}) = r_x(h_x)$. La parte afín se anuló; el residuo no se ha anulado por ese motivo. Para controlar el resultado necesitaremos información sobre ese término.

Tampoco basta invocar $r_x(h) = o(h)$ para declarar pequeño el error del paso calculado. Ese límite se toma cuando $h \to 0$ con el centro x fijo; todavía no hemos probado que h_x sea pequeño ni que exista un control uniforme al cambiar los centros. Esta es la diferencia entre motivar el método mediante linealización y demostrar su eficacia mediante estimaciones.

El operador de Newton y su dominio.

Definición 12.2. **Operador de Newton.** Sea $I \subseteq \mathbb{R}$ un intervalo abierto no vacío y sea $f\colon I \to \mathbb{R}$ diferenciable. Definimos el conjunto de puntos admisibles

$$D_f := \{x \in I : f'(x) \neq 0\}.$$

El **operador de Newton** asociado a f es la función

$$N_f\colon D_f \to \mathbb{R}, \qquad N_f(x) := x - \frac{f(x)}{f'(x)}.$$

Dado $x \in D_f$, llamaremos **paso de Newton desde** x al cálculo del nuevo valor $N_f(x)$.

El codominio escrito es $\mathbb{R}$, no I ni D_f. La fórmula produce un número real, pero no garantiza que ese número vuelva a pertenecer al dominio donde podemos calcular otro paso. No suponemos que f' sea continua: para definir $N_f(x)$ basta que la derivada en ese punto exista como número real y sea distinta de cero.

Tenemos así tres regiones diferentes: el dominio I de la función, el subconjunto D_f donde la división está permitida y un eventual intervalo aislante $[a, b] \subset I$. Ni la pertenencia a D_f garantiza permanencia en $[a, b]$, ni haber aislado una raíz convierte automáticamente a N_f en una aplicación de $[a, b]$ en sí mismo.

De un paso a una iteración. Elegido un punto inicial $x_0 \in D_f$, podemos calcular $x_1 = N_f(x_0)$. Para calcular después $x_2 = N_f(x_1)$ debemos comprobar que $x_1 \in D_f$. En general escribiremos

$$\boxed{x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}}$$

mientras $x_n \in D_f$. La regla no asegura por sí sola que estén definidos términos para todos los índices naturales. Para demostrar que existe una sucesión infinita de iteraciones habrá que justificar que el procedimiento puede continuar en cada paso. Cada nueva iteración reconstruye la tangente en el punto recién obtenido. No seguimos utilizando indefinidamente L_{x_0}: esa recta ya tenía un único cero, x_1. El paso siguiente utiliza L_{x_1}, con los valores $f(x_1)$ y $f'(x_1)$. Hay una observación exacta útil. Para todo $x \in D_f$,

$$N_f(x) = x \quad \Longleftrightarrow \quad \frac{f(x)}{f'(x)} = 0 \quad \Longleftrightarrow \quad f(x) = 0.$$

Así, dentro del dominio de Newton, los puntos que la fórmula deja fijos son precisamente las raíces. Esto no demuestra que las iteraciones alcancen uno de ellos ni que converjan a él. Tampoco una igualdad entre números redondeados equivale a la igualdad exacta $N_f(x) = x$. Si durante el cálculo se demuestra que $f(x_n) = 0$, el problema de encontrar una raíz ya está resuelto y puede detenerse el procedimiento. Esa comprobación es válida incluso si $f'(x_n) = 0$: en tal caso no aplicamos la fórmula ni asignamos un valor a $0/0$.

Dos tangentes sucesivas para la raíz ya aislada.

Ejemplo 12.3. **Newton para $p(t) = t^3 + t - 1$.** Retomemos la función de Ejemplo 12.2, definida en toda $\mathbb{R}$. Sabemos que tiene una única raíz α en $[0, 1]$ y que $\frac{5}{8} < \alpha < \frac{11}{16}$.

Como $p'(x) = 3x^2+1 > 0$ para todo real x, el dominio de su operador es $D_p = \mathbb{R}$. Simplificando la fórmula obtenemos $N_p(x) = x - \frac{x^3+x-1}{3x^2+1} = \frac{2x^3+1}{3x^2+1}$. Para mantener sencillas las primeras fracciones, tomemos $x_0 = 1$. Es un extremo del intervalo aislante inicial, pero es un punto interior del dominio $\mathbb{R}$ de p y pertenece a D_p.

En ese punto,

$$p(1) = 1, \qquad p'(1) = 4.$$

La primera recta es $L_1(t) = 1 + 4(t - 1) = 4t - 3$, y su cero proporciona $x_1 = 1 - \frac{1}{4} = \frac{3}{4}$. Sin embargo, $p(3/4) = 11/64 \neq 0$: hemos encontrado exactamente el cero de la recta, no el cero del cúbico. Reconstruyamos la tangente en $x_1 = 3/4$. Ahora

$$p\left(\frac{3}{4}\right) = \frac{11}{64}, \qquad p'\left(\frac{3}{4}\right) = \frac{43}{16},$$

por lo que

$$L_{3/4}(t) = \frac{11}{64} + \frac{43}{16}\left(t - \frac{3}{4}\right).$$

Su cero da el segundo paso:

$$\begin{aligned} x_2 &= \frac{3}{4} - \frac{11/64}{43/16} \\ &= \frac{3}{4} - \frac{11}{172} \\ &= \frac{59}{86}. \end{aligned}$$

Podemos comprobar por multiplicación cruzada de denominadores positivos que $\frac{5}{8} < \frac{59}{86} < \frac{11}{16}$. Además, una evaluación exacta muestra

$$p\left(\frac{59}{86}\right) = \frac{5687}{636056} > 0.$$

Por la monotonía estricta de p y $p(\alpha) = 0$, este último dato permite concluir $\alpha < x_2$. Esa conclusión procede del signo y de la monotonía, no solamente de haber aplicado Newton. La conversión cuantitativa del residuo en una cota de error se estudiará en §12.6. Los dos pasos son legítimos porque los puntos desde los cuales dividimos pertenecen a D_p. Aunque en este ejemplo $D_p = \mathbb{R}$ permite continuar la recurrencia indefinidamente en aritmética exacta, todavía no hemos demostrado que las iteraciones converjan ni a qué velocidad lo harían.

El ejemplo permite leer geométricamente cada cálculo sin apoyarse en un dibujo como prueba: el punto de contacto es $(x_n, p(x_n))$, la pendiente es $p'(x_n)$ y la intersección de esa recta con el eje horizontal es $(x_{n+1}, 0)$. Para construir el paso siguiente se vuelve a la gráfica, al punto $(x_{n+1}, p(x_{n+1}))$; no se supone que este último coincida con $(x_{n+1}, 0)$.

El caso en que un solo paso resuelve el problema.

Ejemplo 12.4. **Una función afín no constante.** Sea $f\colon \mathbb{R} \to \mathbb{R}$,

$$f(t) = at + b, \qquad a \neq 0.$$

Su derivada es $f'(x) = a$ y $D_f = \mathbb{R}$. Entonces, para todo punto inicial x, $N_f(x) = x - \frac{ax+b}{a} = -\frac{b}{a}$. El resultado es la raíz exacta, independientemente del punto inicial. Aquí la tangente no es solamente un modelo local:

$$L_x(t) = ax + b + a(t - x) = at + b = f(t).$$

El residuo $r_x(h)$ es idénticamente cero. Por eso resolver la ecuación de la tangente resuelve también la ecuación original. Esta explicación no puede trasladarse sin prueba a una función no afín.

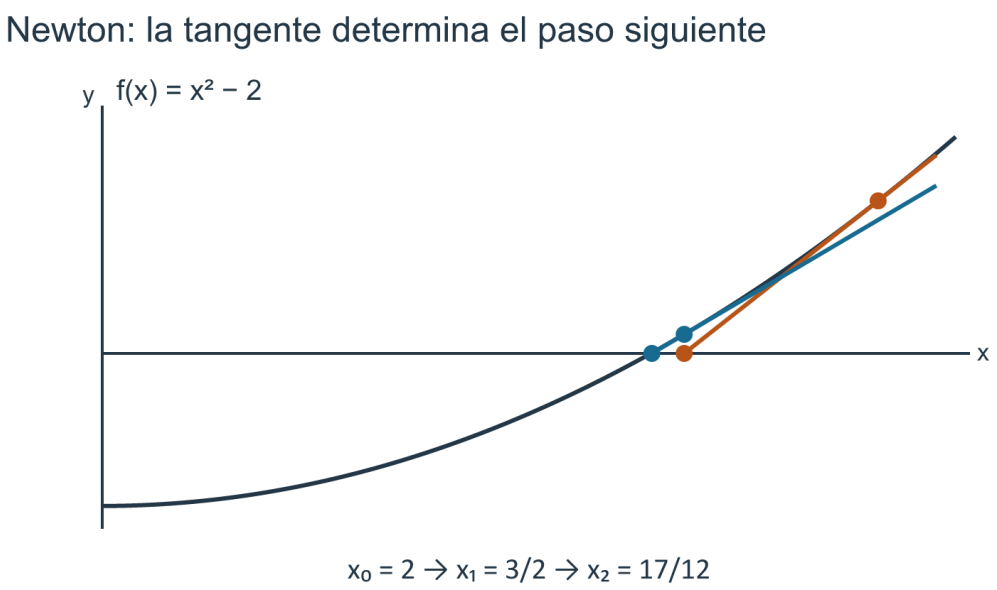


FIGURA 12.1. Para $f(x) = x^2 - 2$, las intersecciones de las tangentes con el eje horizontal dan $x_0 = 2$, $x_1 = 3/2$ y $x_2 = 17/12$. La construcción de los pasos precede a la justificación de su convergencia.

Qué impide construir o repetir el paso. Si $f'(x) = 0$, la ecuación afín se reduce a $f(x) = 0$, sin término variable. Cuando $f(x) \neq 0$, la recta horizontal no corta al eje horizontal. Cuando $f(x) = 0$, la recta coincide con ese eje y todos sus puntos son ceros: no hay una intersección única de la cual deducir la fórmula. La raíz ya encontrada permite detenerse, pero no legitima una división por cero.

Por otra parte, un paso puede estar bien definido y llevarnos fuera del dominio de la función.

EJEMPLO 12.5. **Un valor real que no admite el paso siguiente.** Consideremos

$$g\colon (0,\infty) \to \mathbb{R}, \qquad g(t) = \sqrt{t} - 1.$$

La función tiene la raíz 1. La derivada de la raíz cuadrada, ya establecida en Proposición 8.1, proporciona

$$g'(x) = \frac{1}{2\sqrt{x}} > 0, \qquad x > 0.$$

Así, $D_g = (0,\infty)$ y

$$N_g(x) = x - 2\sqrt{x}(\sqrt{x} - 1) = 2\sqrt{x} - x.$$

En particular, $N_g(9) = 6 - 9 = -3$.

La división en $x = 9$ es legítima. El problema aparece después: $-3 \notin (0,\infty)$, por lo que $g(-3)$ no está definido y no podemos calcular $N_g(-3)$. La fórmula define una aplicación hacia $\mathbb{R}$, pero no necesariamente hacia su propio dominio.

El ejemplo comprueba si la operación puede repetirse, pero deja abiertas la convergencia y la calidad de la elección inicial. Los ciclos y las raíces múltiples permitirán examinar esas dificultades por separado.

Laboratorio de reconstrucción. Reconstruir el desplazamiento. Alguien propone $x_{\text{nuevo}} = -f(x)/f'(x)$. Localiza el paso perdido: la ecuación afín determina primero $t - x$, no t. Para $f(t) = 2t - 6$ y $x = 5$, la corrección es -2, mientras que el nuevo punto es 3, la raíz. El valor -2 no es esa raíz.

Comprobar la identidad exacta. Sustituye $t = N_f(x)$ en $L_x(t)$. Los términos $f(x)$ y $-f(x)$ se cancelan, lo que demuestra $L_x(N_f(x)) = 0$. Para sustituir el mismo argumento en f, comprueba antes que $N_f(x) \in I$; aun entonces, la cancelación de la parte afín no elimina el residuo.

Separar dos afirmaciones sobre el dominio. «$f'(x_0) \neq 0$» autoriza un paso desde x_0. «Todos los términos permanecen en D_f» autoriza repetirlo indefinidamente. El ejemplo anterior refuta que la primera afirmación implique la segunda.

Precisar qué falta demostrar. Las igualdades $x_1 = N_f(x_0)$ y $x_2 = N_f(x_1)$ describen dos operaciones. No contienen una estimación de $|x_2 - \alpha|$, una relación entre errores sucesivos ni un argumento de convergencia. Esas afirmaciones requieren pruebas adicionales; no se obtienen contando los pasos realizados.

De la construcción al control del error. Ya conocemos el origen de la fórmula, su dominio y el significado de sus iteraciones. La cuestión siguiente será comparar el cero de la tangente con una raíz verdadera α. Taylor de grado uno, centrado en el punto actual x y evaluado en la raíz α, relacionará exactamente $N_f(x) - \alpha$ con el resto, porque $f(\alpha) = 0$. Para convertir esa identidad en una estimación de convergencia necesitaremos controlar las derivadas que aparecen en ella.

Antes de seguir. ¿En qué línea se utiliza $f'(x) \neq 0$? ¿Qué ecuación resuelve exactamente $N_f(x)$? ¿Por qué el codominio de N_f se ha escrito como $\mathbb{R}$? ¿Qué diferencia hay entre una raíz donde la derivada es cero y un punto donde la derivada es cero pero la función no se anula? ¿Por qué dos iteraciones calculadas correctamente no son todavía una prueba de convergencia?

12.4. La identidad del error de Newton

El paso de Newton resuelve exactamente una ecuación afín, pero nuestro objetivo sigue siendo resolver $f(t) = 0$. Necesitamos comparar el cero de la tangente con una raíz de la función, no solo comprobar que la fórmula de la iteración puede calcularse. Fijemos una raíz α y un punto actual x en el que $f'(x) \neq 0$. Escribiremos

$$x^+ := N_f(x), \qquad e := x - \alpha, \qquad e^+ := x^+ - \alpha.$$

Los números e y e^+ son errores con signo: indican también de qué lado de la raíz están los puntos. Sus valores absolutos son los errores absolutos definidos en Definición 12.1. Una fórmula para e^+ podrá contener información que se perdería al tomar inmediatamente valores absolutos. La pregunta rectora de esta sección es concreta: ¿qué queda del error después de que Newton cancela la parte afín? Taylor permitirá responderla mediante una identidad exacta.

Elegir correctamente el centro de Taylor. Hay dos puntos disponibles, x y α. La recta utilizada por Newton es la tangente en x, de modo que debemos desarrollar alrededor de x, no alrededor de la raíz desconocida. El punto en el que evaluaremos ese desarrollo será α. En la notación de Teorema 11.2 hacemos estas sustituciones:

Papel en Taylor	Elección para el paso de Newton
Centro del polinomio	El punto actual x
Punto de evaluación	La raíz α
Grado del polinomio	$n = 1$
Derivada que aparece en el resto	La segunda derivada
Incremento desde el centro	$\alpha - x$

No estamos buscando un polinomio cuadrático. Utilizamos el polinomio de grado uno, que es la tangente, y conservamos el resto de segundo orden. Si $x \neq \alpha$ y se cumplen las hipótesis de Taylor, obtenemos algún ξ estrictamente entre ambos puntos para el cual

$$f(\alpha) = f(x) + f'(x)(\alpha - x) + \frac{f''(\xi)}{2}(\alpha - x)^2.$$

La igualdad no es una aproximación. La curvatura que la recta no recoge permanece representada por el último término.

El enunciado de un solo paso.

TEOREMA 12.1. **Teorema (identidad del error de Newton).** Sea I un intervalo abierto y sea $f : I \to \mathbb{R}$ una función que posee primera y segunda derivadas en todo I. Supongamos que $\alpha \in I$ satisface $f(\alpha) = 0$.

Para cada $x \in I$ con $x \neq \alpha$ y $f'(x) \neq 0$, existe un punto

$$\xi \in (\text{mín}\{x, \alpha\}, \text{máx}\{x, \alpha\})$$

tal que

$$\boxed{N_f(x) - \alpha = \frac{f''(\xi)}{2f'(x)}(x - \alpha)^2.}$$

No se exige continuidad de f''. Tampoco se exige que $N_f(x)$ pertenezca a I: la conclusión compara dos números reales, sin evaluar f en el nuevo punto. Si $x = \alpha$ y $f'(\alpha) \neq 0$, se tiene directamente $N_f(\alpha) = \alpha$. Este caso se trata por separado, sin introducir un punto estrictamente intermedio entre dos puntos iguales.

Demostración. Supongamos primero que $x \neq \alpha$. Como I es un intervalo y contiene ambos puntos, contiene todo el segmento que los une. La existencia de las dos derivadas permite aplicar Teorema 11.2 con grado uno, centro x y evaluación en α. Existe, pues, el punto ξ indicado y

$$0 = f(x) + f'(x)(\alpha - x) + \frac{f''(\xi)}{2}(\alpha - x)^2,$$

porque $f(\alpha) = 0$.

Para auditar el signo, pasemos el término lineal al otro miembro. Puesto que $-(\alpha - x) = x - \alpha$ y $(\alpha - x)^2 = (x - \alpha)^2$,

$$f'(x)(x - \alpha) = f(x) + \frac{f''(\xi)}{2}(x - \alpha)^2.$$

La hipótesis $f'(x) \neq 0$ autoriza dividir esta igualdad:

$$x - \alpha = \frac{f(x)}{f'(x)} + \frac{f''(\xi)}{2f'(x)}(x-\alpha)^2.$$

Restando $f(x)/f'(x)$ y usando Definición 12.2,

$$\begin{aligned} N_f(x) - \alpha &= x - \alpha - \frac{f(x)}{f'(x)} \\ &= \frac{f''(\xi)}{2f'(x)}(x-\alpha)^2. \end{aligned}$$

Si $x = \alpha$ y $f'(\alpha) \neq 0$, no necesitamos Taylor:

$$N_f(\alpha) = \alpha - \frac{f(\alpha)}{f'(\alpha)} = \alpha.$$

Esto completa ambos casos. □

Qué hipótesis se usó y dónde. La demostración es breve, pero contiene autorizaciones diferentes que conviene no confundir.

Hipótesis o dato	Uso preciso
x, α pertenecen al mismo intervalo abierto I	El segmento entre ellos está en la región donde se aplica Taylor.
f posee dos derivadas en I	Autoriza el Taylor de grado uno con resto de Lagrange.
$f(\alpha) = 0$	Anula el valor de la función en el punto de evaluación, no el resto.
$f'(x) \neq 0$	Define el paso de Newton y permite dividir en el despeje.
$x \neq \alpha$	Hace no vacío el intervalo abierto en el que Taylor sitúa ξ.

La existencia de f'' significa que f' es diferenciable y, por tanto, continua, según Teorema 6.1. Eso no obliga a que f'' sea continua. El Taylor canónico utilizado aquí no añade esa última hipótesis y tampoco la necesita nuestra identidad. La existencia de una raíz se ha supuesto, no demostrado mediante Newton. Si la raíz se obtuvo antes por aislamiento, ese argumento proporciona el dato $f(\alpha) = 0$. La unicidad no interviene en el despeje: fijada una raíz, la identidad se refiere a ella.

Llamaremos *raíz simple* a una raíz α para la cual $f'(\alpha) \neq 0$. Es importante distinguir esta condición de $f'(x) \neq 0$. La identidad de un paso solo divide por la derivada en el punto actual x; no requiere que la derivada en la raíz sea no nula cuando $x \neq \alpha$. En cambio, la simplicidad será decisiva para obtener una cota inferior positiva de $|f'|$ cerca de α: como f' es continua, de $f'(\alpha) \neq 0$ se deduce que $|f'(x)| \geq |f'(\alpha)|/2$ en un entorno suficientemente pequeño. Esa observación prepara el control local de §12.5; no es todavía una demostración de convergencia. Si $f'(\alpha) = 0$, haber encontrado exactamente la raíz permite detener el procedimiento, como vimos en §12.3. No permite evaluar allí una expresión que divide por cero.

El punto intermedio no es un dato elegible. El teorema afirma que para cada x admisible existe algún ξ intermedio que hace verdadera la identidad. No afirma que ese punto sea único ni que se conozca. En general depende de x y de la raíz fijada. Por tanto no podemos sustituir libremente $f''(\xi)$ por $f''(x)$, por $f''(\alpha)$ o por el valor en el punto medio. Tampoco podemos suponer que el mismo ξ sirve para pasos distintos. El segundo ejemplo de esta sección permitirá ver explícitamente un punto intermedio que no es el punto medio. Para obtener una desigualdad útil no suele hacer falta localizar ξ. Basta dominar f'' en todos los lugares donde podría estar. Esta es la misma estrategia que utilizamos en Corolario 11.2 para acotar el resto de Taylor.

De la identidad a una cota cuadrática. Supongamos que en todo el segmento entre x y α tenemos una cota $|f''| \le M$, con $M \ge 0$. La identidad proporciona $|N_f(x) - \alpha| \le \frac{M}{2|f'(x)|}|x-\alpha|^2$. El denominador todavía depende del punto actual. Para controlar varios puntos con una misma constante podemos imponer una cota inferior uniforme.

Corolario 12.1. **Corolario (cota uniforme para un paso de Newton).** Bajo las hipótesis sobre I, f y α de Teorema 12.1, sea $J \subseteq I$ un intervalo que contiene a α. Supongamos que existen constantes $m > 0$ y $M \ge 0$ tales que, para todo $t \in J$,

$$|f'(t)| \ge m, \qquad |f''(t)| \le M.$$

Entonces $J \subseteq D_f$ y, para cada $x \in J$,

$$\boxed{|N_f(x) - \alpha| \le \frac{M}{2m}|x-\alpha|^2.}$$

La conclusión no afirma que $N_f(J) \subseteq J$.

Demostración. La desigualdad $|f'(t)| \ge m > 0$ muestra que todos los puntos de J pertenecen a D_f. Si $x \neq \alpha$, el segmento entre x y α está contenido en J porque J es un intervalo. En particular, el punto ξ de Teorema 12.1 pertenece a J. Por consiguiente,

$$\begin{aligned} |N_f(x) - \alpha| &= \frac{|f''(\xi)|}{2|f'(x)|}|x-\alpha|^2 \\ &\le \frac{M}{2m}|x-\alpha|^2. \end{aligned}$$

Para $x = \alpha$, la cota inferior implica $f'(\alpha) \neq 0$, y ambos miembros de la desigualdad son cero. □

El caso $M = 0$ también está incluido: si se cumplen las hipótesis con esa constante, el paso llega exactamente a α desde cualquier $x \in J$. Es coherente con el comportamiento de las funciones afines de Ejemplo 12.4.

En cambio, la existencia de f'' por sí sola no proporciona automáticamente una cota uniforme M en la región escogida. La cota debe demostrarse o figurar como hipótesis, al igual que la separación de f' respecto de cero.

Un ejemplo en el que el coeficiente se conoce exactamente.

Ejemplo 12.6. **El error de Newton para la raíz cuadrada de dos.** Sea

$$f(t) = t^2 - 2, \qquad I = (0, \infty), \qquad \alpha = \sqrt{2}.$$

En este intervalo $f'(x) = 2x > 0$ y $f''(t) = 2$. Como la segunda derivada es constante, la identidad de Teorema 12.1 se convierte en

$$\boxed{N_f(x) - \sqrt{2} = \frac{(x-\sqrt{2})^2}{2x}, \qquad x > 0.}$$

En $x = \sqrt{2}$ la igualdad sigue siendo válida por el caso separado del teorema. Para $x \neq \sqrt{2}$, cualquier punto estrictamente intermedio sirve para representar el valor constante de f''. Podemos comprobar la identidad sin Taylor. La fórmula del paso es $N_f(x) = x - \frac{x^2-2}{2x} = \frac{x^2+2}{2x}$.

Como $\alpha^2 = 2$, $\frac{x^2+2}{2x} - \alpha = \frac{x^2-2\alpha x+\alpha^2}{2x} = \frac{(x-\alpha)^2}{2x}$.

Esta segunda vía verifica de manera independiente tanto el signo como el factor $1/2$. Además, el nuevo punto siempre es mayor o igual que la raíz positiva: el numerador es un cuadrado y el denominador es positivo. Hay igualdad precisamente cuando $x = \alpha$. En $J = [1, 2]$, las cotas $|f'| \geq 2$ y $|f''| \leq 2$ permiten usar $m = 2$, $M = 2$ en Corolario 12.1. Así,

$$|N_f(x) - \alpha| \leq \frac{1}{2}|x - \alpha|^2, \qquad x \in [1, 2].$$

Veamos también un certificado de un paso concreto. Las comparaciones exactas

$$\left(\frac{7}{5}\right)^2 = \frac{49}{25} < 2 < \frac{9}{4} = \left(\frac{3}{2}\right)^2$$

muestran que $7/5 < \alpha < 3/2$. Desde $x = 3/2$ obtenemos

$$N_f\left(\frac{3}{2}\right) = \frac{17}{12}.$$

La identidad exacta, usando $2x = 3$ en este punto, da

$$0 < \frac{17}{12} - \alpha = \frac{(3/2-\alpha)^2}{3} < \frac{(1/10)^2}{3} = \frac{1}{300}.$$

La aproximación $17/12$ queda certificada sin consultar ningún desarrollo decimal de $\sqrt{2}$. Se ha probado una afirmación sobre este paso, no todavía un teorema sobre todas las iteraciones.

El cúbico anterior revela el punto intermedio.

Ejemplo 12.7. **Una identidad algebraica para el cúbico aislado.** Recuperemos $p(t) = t^3 + t - 1$ y su raíz $\alpha \in (0, 1)$, ya aislada en Ejemplo 12.2. Como $p'(x) = 3x^2 + 1$, el paso está definido para todo real y $N_p(x) = \frac{2x^3+1}{3x^2+1}$.

La relación $p(\alpha) = 0$ equivale a $1 = \alpha^3 + \alpha$. Sustituyéndola en el numerador de la diferencia,

$$\begin{aligned} N_p(x) - \alpha &= \frac{2x^3 + 1 - \alpha(3x^2+1)}{3x^2+1} \\ &= \frac{2x^3 - 3\alpha x^2 + \alpha^3}{3x^2+1} \\ &= \frac{(x-\alpha)^2(2x+\alpha)}{3x^2+1}. \end{aligned}$$

Por otro lado, $p''(t) = 6t$, de modo que la identidad de Taylor tiene coeficiente

$$\frac{p''(\xi)}{2p'(x)} = \frac{3\xi}{3x^2+1}.$$

Para $x \neq \alpha$, ambas expresiones concuerdan con $\xi = \frac{2x+\alpha}{3}$. Este punto es estrictamente intermedio: $\xi - x = (\alpha - x)/3$ y $\alpha - \xi = 2(\alpha - x)/3$, así que ambas diferencias tienen el mismo signo que $\alpha - x$. No es, en general, el punto medio. En efecto, $\frac{2x+\alpha}{3} = \frac{x+\alpha}{2}$ obligaría a $x = \alpha$, caso excluido en esta comparación. El ejemplo muestra dos maneras de usar el teorema. Podemos prescindir de la posición exacta de ξ y acotar p'' en un intervalo; o, cuando una identidad algebraica adicional lo permite, podemos identificarlo. La segunda posibilidad es una ventaja particular de este ejemplo, no un permiso para elegir el punto intermedio en cualquier función.

Por qué un cuadrado no basta para demostrar convergencia. La forma $e^+ = A(x,\alpha)e^2$ no garantiza por sí sola que el error disminuya. El coeficiente puede ser grande y el punto actual puede estar demasiado lejos de la raíz. La expresión tampoco garantiza que el punto nuevo pertenezca a la región donde pretendemos repetir las estimaciones. El primer ejemplo permite comprobarlo con un solo cálculo, sin desarrollar todavía una teoría de fallos. Para $f(t) = t^2 - 2$ y $x = 1/4$,

$$N_f\left(\frac{1}{4}\right) = \frac{33}{8}.$$

Como $1 < \alpha < 3/2$, tenemos

$$\left|\frac{1}{4} - \alpha\right| < \frac{5}{4}, \qquad \left|\frac{33}{8} - \alpha\right| > \frac{21}{8} > \frac{5}{4}.$$

El error aumentó, aunque la función posee dos derivadas, la raíz es simple y la identidad es correcta. No contradice la cota establecida en $[1,2]$: $1/4$ está fuera de ese intervalo. El ejemplo subraya que una región de validez es parte del enunciado, no una anotación opcional. Si contamos con una constante uniforme $C = M/(2m)$, la desigualdad $|e^+| \le C|e|^2$ permite afirmar mejora estricta cuando $e \neq 0$ y $C|e| < 1$. Pero para convertir mejoras de un paso en un teorema de convergencia todavía debemos garantizar que todos los puntos sucesivos permanezcan en la región controlada y justificar el límite de la sucesión. Esa tarea corresponde a §12.5.

Tampoco se deduce aquí un criterio de parada basado solo en el residuo $|f(x)|$. Las cotas de esta sección relacionan el error nuevo con el error anterior; cuando ya disponemos de un encierro inicial, como en el ejemplo de $17/12$, permiten transportarlo a un certificado nuevo. La conversión directa del residuo en una cota de error se desarrollará en §12.6.

Laboratorio de lectura y reparación de la prueba. Reconstruir sin memorizar. Escribe Taylor de grado uno con centro x y evaluación en α. Solo después introduce $f(\alpha) = 0$. Si aparece un signo menos delante del resto en la fórmula final para $N_f(x) - \alpha$, revisa el paso de $\alpha - x$ a $x - \alpha$: el término lineal cambia de signo, pero su cuadrado no.

Distinguir dos usos de la raíz. La condición $f(\alpha) = 0$ elimina el valor del miembro izquierdo de Taylor. La condición $f'(\alpha) \neq 0$, cuando está disponible, sirve para separar localmente la primera derivada de cero. No reemplaces la segunda por $f'(x) \neq 0$ ni la primera por «$f(x)$ es pequeño».

Reparar una elección ilegítima. «Tomemos $\xi = (x+\alpha)/2$» no está autorizado por Taylor. El cúbico anterior ofrece una refutación explícita para $x \neq \alpha$. Para una cota basta reemplazar esa elección por una estimación válida en todo el segmento.

Cerrar el caso omitido. Si $x = \alpha$, no existe un punto estrictamente entre ambos. La conclusión correcta es comprobar directamente $N_f(\alpha) = \alpha$ cuando la derivada allí es no nula; si es nula, se reconoce la raíz y se detiene el cálculo, sin definir artificialmente el cociente.

Comprobar una cota numérica. Si las hipótesis de Corolario 12.1 se cumplen con $m = 2$, $M = 6$ y el punto actual pertenece a J con $|x - \alpha| \le 1/10$, entonces

$$|N_f(x) - \alpha| \le \frac{6}{2 \cdot 2} \left(\frac{1}{10}\right)^2 = \frac{3}{200}.$$

La desigualdad está demostrada para el punto nuevo. Para aplicarla otra vez con las mismas constantes sigue faltando comprobar que ese punto pertenece a J.

Lo que queda preparado. La construcción de §12.3 y la identidad de esta sección encajan ahora sin saltos: Newton anula el polinomio afín y Taylor expresa exactamente el error que sobrevive. Cuando la pendiente está separada de cero y la segunda derivada está acotada, ese error queda dominado por una constante multiplicada por el cuadrado del error anterior. En §12.5 construiremos la región y la condición inicial que permiten repetir esta estimación. Allí se demostrará la convergencia local y se precisará el significado del régimen cuadrático; aquí hemos establecido la identidad y el control de un solo paso.

Antes de seguir. ¿Por qué Taylor se centra en x? ¿Dónde se utiliza que α es una raíz? ¿Por qué la identidad no necesita evaluar f en $N_f(x)$? ¿Qué dos cotas producen la constante $M/(2m)$? ¿Qué falta, además de esa desigualdad, para aplicarla indefinidamente?

12.5. Convergencia local y régimen cuadrático

La identidad de §12.4 explica qué error deja un paso de Newton. Ahora queremos demostrar algo que se refiere a una sucesión completa: que todos sus términos pueden calcularse, que permanecen en una región controlada y que convergen a la raíz elegida. Son tres afirmaciones diferentes. No debemos utilizar la tercera para justificar las dos primeras. La pregunta rectora será esta: ¿cómo convertir una cota válida para un punto en una garantía que se conserva al repetir el procedimiento?

La dificultad está en poder volver a usar la desigualdad. Fijemos una raíz α y recordemos Corolario 12.1. Cuando las cotas de las derivadas valen en un intervalo J que contiene a la raíz, tenemos

$$|N_f(x) - \alpha| \le K|x - \alpha|^2, \qquad K = \frac{M}{2m}, \qquad x \in J.$$

Para aplicar la misma fórmula al punto $N_f(x)$ necesitamos que ese punto pertenezca a J. Su pertenencia no está contenida en la cota anterior para un intervalo arbitrario. La idea es escoger un intervalo cuya geometría convierta directamente una cota de error en una pertenencia.

Tomemos

$$B_r = [\alpha - r, \alpha + r], \qquad r > 0.$$

Un número real y pertenece a B_r exactamente cuando $|y - \alpha| \le r$. Por tanto, si la cota de un paso vale en B_r y $|x - \alpha| \le r$, entonces

$$|N_f(x) - \alpha| \le K|x - \alpha|^2 \le Kr\,|x - \alpha| \le Kr^2.$$

Para que el último miembro sea menor que r basta exigir

$$\boxed{Kr < 1.}$$

Así aparece una condición cuantitativa sobre el tamaño de la región. No estamos diciendo vagamente que el punto inicial debe estar «bastante cerca»: hemos identificado una desigualdad suficiente que enlaza radio, pendiente y segunda derivada. Diremos que B_r es *invariante hacia adelante* para Newton si $B_r \subseteq D_f$ y $N_f(B_r) \subseteq B_r$. La primera inclusión garantiza que el paso está definido; la segunda permite repetirlo. Demostraremos ambas, antes de estudiar el límite.

Un criterio local con radio y constantes explícitos.

PROPOSICIÓN 12.2. **Proposición (región invariante y convergencia local de Newton).** Sea I un intervalo abierto y sea $f : I \to \mathbb{R}$ dos veces diferenciable. Sea $\alpha \in I$ una raíz de f. Supongamos que existen $r > 0$, $m > 0$ y $M \geq 0$ tales que

$$B_r = [\alpha - r, \alpha + r] \subseteq I,$$

$$|f'(t)| \geq m, \qquad |f''(t)| \leq M \qquad \text{para todo } t \in B_r.$$

Pongamos

$$K = \frac{M}{2m}, \qquad q = Kr,$$

y supongamos $q < 1$. Entonces se cumplen las afirmaciones siguientes.

1. El paso de Newton está definido en todo B_r y
$$N_f(B_r) \subseteq [\alpha - qr, \alpha + qr] \subseteq (\alpha - r, \alpha + r).$$
2. Para cualquier $x_0 \in B_r$, la recurrencia
$$x_{n+1} = N_f(x_n), \qquad n \geq 0,$$
define una sucesión con $x_n \in B_r$ para todo n, y $x_n \to \alpha$.
3. Escribiendo $E_n = |x_n - \alpha|$, se tiene
$$\boxed{E_{n+1} \leq K E_n^2} \qquad (n \geq 0),$$
y
$$\boxed{E_n \leq q^n E_0} \qquad (n \geq 1).$$
Si $K > 0$, además
$$\boxed{E_n \leq \frac{1}{K}(KE_0)^{2^n}} \qquad (n \geq 0).$$
Si $K = 0$, se tiene $x_1 = \alpha$ y todos los términos posteriores son iguales a α.

No se presupone continuidad de f'': sus valores están controlados por la cota explícita. Las conclusiones se refieren a la raíz y a la región indicadas, no a cualquier punto inicial de I.

Demostración. Organizaremos la prueba en el orden que evita la circularidad.

Primero: el paso está definido y conserva la región. La cota inferior implica $f'(t) \neq 0$ para todo $t \in B_r$. Por Definición 12.2, $B_r \subseteq D_f$. Podemos, pues, calcular $N_f(x)$ para cualquier $x \in B_r$.

Apliquemos Corolario 12.1 al intervalo B_r. Puesto que $|x - \alpha| \leq r$,

$$|N_f(x) - \alpha| \le K|x - \alpha|^2 \le Kr^2 = qr < r.$$

La desigualdad no estricta da la primera inclusión del enunciado; $q < 1$ da la segunda. En particular, $N_f(x) \in B_r$. Todavía no hemos hablado de convergencia: hemos probado una propiedad del paso para todas las entradas de la región.

Segundo: la recurrencia produce todos los términos. El término x_0 pertenece a B_r por hipótesis. Si $x_n \in B_r$, la primera parte garantiza que x_{n+1} está definido y que también pertenece a B_r. La inducción justifica la existencia de la sucesión completa y la permanencia de todos sus términos en ese intervalo. No inferimos su existencia de escribir formalmente una recurrencia.

Si alguna iteración llega exactamente a α, la cota $|f'(\alpha)| \ge m > 0$ permite seguir escribiendo Newton y obtener siempre $N_f(\alpha) = \alpha$. En un cálculo efectivo podemos detenernos allí; prolongar la sucesión de forma constante solo sirve para enunciar uniformemente el resultado.

Tercero: controlar los errores y demostrar el límite. Ya sabemos que $x_n \in B_r$ para cada n. Ahora sí podemos aplicar la estimación a todos los pasos:

$$E_{n+1} \le KE_n^2 \le KrE_n = qE_n.$$

Una nueva inducción da $E_n \le q^n E_0$ para $n \ge 1$. Si $q = 0$, la desigualdad obliga a $E_1 = 0$ y todos los errores posteriores son cero. Si $0 < q < 1$, el límite geométrico Corolario 3.2 da $q^n \to 0$. Como $0 \le E_n \le q^n E_0$, también $E_n \to 0$. Equivalentemente, para cada $\varepsilon > 0$ basta elegir un índice a partir del cual $q^n E_0 < \varepsilon$; entonces $|x_n - \alpha| < \varepsilon$. Esto demuestra $x_n \to \alpha$ sin suponer antes la existencia de un límite de (x_n).

Cuarto: conservar la información cuadrática. Si $K > 0$, definamos $u_n = KE_n$. Multiplicar la estimación cuadrática por K produce $u_{n+1} \le u_n^2$.

Todos estos números son no negativos. Por inducción, $u_n \le u_0^{2^n}$. El caso $n = 0$ es igualdad. Si la afirmación vale para n, entonces

$$u_{n+1} \le u_n^2 \le \left(u_0^{2^n}\right)^2 = u_0^{2^{n+1}}.$$

Dividir por $K > 0$ proporciona la última cota. Además, $u_0 = KE_0 \le Kr = q < 1$. Si $E_0 = 0$, la fórmula sigue siendo válida: todos los exponentes 2^n son positivos y ambos miembros son cero.

Finalmente, si $K = 0$, no efectuamos esa división: la cota original ya dice $E_1 \le 0$. Por tanto $x_1 = \alpha$. Quedan demostrados todos los casos. □

Dónde entró cada hipótesis. Las inducciones cumplen dos tareas diferentes: primero conservar el dominio y después transportar las cotas. Intercambiar ese orden escondería precisamente la laguna que queríamos cerrar.

Dato	Función en la demostración
$f(\alpha) = 0$	Proporciona la raíz respecto de la cual se aplica la identidad del error. No se demuestra existencia mediante la iteración.
Dos derivadas en el intervalo abierto I	Mantiene disponibles las hipótesis del Taylor utilizado en §12.4.
$B_r \subseteq I$	Sitúa cada segmento entre la raíz y una entrada controlada dentro del dominio.
$\lvert f'\rvert \ge m > 0$ en todo B_r	Define todos los pasos de la región, incluida la prolongación constante en la raíz.

Dato	Función en la demostración
$\lvert f''\rvert \le M$ en todo B_r	Da una misma constante cuadrática, aunque el punto intermedio cambie en cada paso.
$Kr < 1$	Lleva la región a su interior y proporciona un factor geométrico estrictamente menor que uno.
$x_0 \in B_r$	Inicia la inducción en una entrada para la que ya están probadas las garantías.

La raíz es única en B_r, aunque esa unicidad no fue necesaria como hipótesis. Si hubiera otra raíz $\beta \in B_r$, su derivada sería no nula y $N_f(\beta) = \beta$. La estimación probada daría $|\beta - \alpha| \le q|\beta - \alpha|$. Como $q < 1$, esto obliga a $\beta = \alpha$. No se deduce que f carezca de otras raíces fuera de la región. Tampoco hemos demostrado que N_f sea una contracción entre dos puntos arbitrarios de B_r. La desigualdad compara cada punto con la raíz fija. La prueba usa directamente inducción y límites de sucesiones, no un teorema abstracto de punto fijo.

Por qué existe una región adecuada cerca de una raíz simple. El criterio anterior es cuantitativo, pero plantea otra pregunta: ¿qué hipótesis cualitativas aseguran que podemos encontrar sus constantes?

Corolario 12.2. **Corolario (existencia de un radio de convergencia local).** Sea $f : I \to \mathbb{R}$ dos veces diferenciable en un intervalo abierto y sea $\alpha \in I$ una raíz simple. Supongamos que existen $R > 0$ y $M_0 \ge 0$ tales que

$$[\alpha - R, \alpha + R] \subseteq I, \qquad |f''(t)| \le M_0 \quad \text{en ese intervalo.}$$

Entonces existe $r \in (0, R]$ tal que, para todo $x_0 \in [\alpha - r, \alpha + r]$, Newton está definido en todas las iteraciones, permanece en ese intervalo y converge a α con las estimaciones de Proposición 12.2. En particular, esto ocurre si $f \in C^2(I)$ y $f(\alpha) = 0$, $f'(\alpha) \neq 0$.

Demostración. Como f'' existe en I, la función f' es diferenciable allí y, por Teorema 6.1, es continua en α. La simplicidad de la raíz permite fijar $m = \frac{|f'(\alpha)|}{2} > 0$.

Por continuidad existe $\delta > 0$ tal que, para $t \in I$ con $|t - \alpha| < \delta$, $|f'(t) - f'(\alpha)| < \frac{|f'(\alpha)|}{2}$. La desigualdad triangular inversa implica entonces $|f'(t)| > m$.

Si $M_0 > 0$, elijamos

$$r = \min\left\{\frac{R}{2}, \frac{\delta}{2}, \frac{m}{M_0}\right\} > 0.$$

El intervalo cerrado de radio r está dentro del intervalo inicial y satisface la condición estricta $|t - \alpha| < \delta$, también en sus extremos. En él valen ambas cotas y $\frac{M_0}{2m}\, r \le \frac{1}{2} < 1$. Podemos aplicar Proposición 12.2. Si $M_0 = 0$, tomamos $r = \min\{R/2, \delta/2\}$; el factor K es cero y la misma proposición da llegada exacta en un paso. Para la última afirmación, recordemos la definición de clase C^2 en Definición 11.2. La continuidad de f'' en α, aplicada con tolerancia 1, proporciona un entorno en el que $|f''(t)| \le |f''(\alpha)| + 1$. Al reducir su radio podemos incluir un intervalo cerrado en I. Así se obtiene la hipótesis de acotación local con $M_0 = |f''(\alpha)| + 1$, y se aplica la parte ya demostrada. □

La condición C^2 es una manera suficiente de conseguir la cota, no una hipótesis que hayamos añadido retrospectivamente a la identidad de §12.4. Si sabemos directamente que f'' está acotada

cerca de la raíz, no necesitamos exigir su continuidad para este criterio de convergencia. Este corolario tampoco transforma «hay una raíz simple» en un radio numérico conocido. Demuestra la existencia de un radio; para certificar una elección concreta debemos comprobar las cotas y la inclusión. En los ejemplos lo haremos usando encierros racionales, sin suponer que conocemos exactamente la raíz.

Qué significa el régimen cuadrático. La cota geométrica $E_n \le q^n E_0$ ya prueba convergencia. Sin embargo, elimina parte de la información al reemplazar el factor variable KE_n por la constante Kr. La desigualdad original conserva $E_{n+1} \le (KE_n)E_n$, y el factor KE_n tiende a cero. La reducción del error puede volverse mucho más intensa que la ofrecida por un mismo factor geométrico fijo.

Definición 12.3. **Convergencia al menos cuadrática y orden exactamente dos.** Sea (x_n) una sucesión que converge a α, y escribamos $E_n = |x_n - \alpha|$.

Diremos que tiene *convergencia al menos cuadrática* si existen $C > 0$ y un índice n_0 tales que

$$E_{n+1} \le CE_n^2 \qquad (n \ge n_0).$$

Si, además, los errores son positivos a partir de algún índice y existe un número L con $0 < L < \infty$ tal que

$$\lim_{n\to\infty} \frac{E_{n+1}}{E_n^2} = L,$$

diremos que la convergencia tiene *orden exactamente dos*. Una sucesión que alcanza la raíz y después permanece en ella satisface la primera desigualdad, pero los cocientes posteriores son $0/0$ y no están definidos. En ese caso informaremos *terminación finita*, no un orden exacto obtenido mediante esos cocientes.

La proposición proporciona convergencia al menos cuadrática con $C = K$ cuando $K > 0$. Si $K = 0$, hay terminación finita y cualquier constante $C > 0$ sirve en la desigualdad. Una cota superior no obliga a que el error nuevo sea igual a la cota ni a que el orden sea exactamente dos. La fórmula $KE_n \le (KE_0)^{2^n}$ explica el nombre de manera cuantitativa. Por ejemplo, si $KE_0 \le 1/10$, entonces

$$KE_1 \le 10^{-2}, \qquad KE_2 \le 10^{-4}, \qquad KE_3 \le 10^{-8}.$$

Son cotas para el error escalado, no una declaración automática de que se duplica cierta cantidad de decimales escritos. La constante K y la conversión entre error absoluto y cifras siguen formando parte de la certificación. Tampoco estamos analizando todavía errores de redondeo: la recurrencia del teorema utiliza operaciones exactas en $\mathbb{R}$.

La constante asintótica requiere una hipótesis adicional. Supongamos las condiciones de Proposición 12.2 y añadamos que f'' es continua en α. Consideremos una órbita que no llega exactamente a la raíz en un número finito de pasos. Sus errores E_n son entonces positivos: si alguno fuera cero, todos los posteriores también lo serían. Por Teorema 12.1, para cada n existe ξ_n estrictamente entre x_n y α con

$$\frac{x_{n+1} - \alpha}{(x_n - \alpha)^2} = \frac{f''(\xi_n)}{2f'(x_n)}.$$

Cualquiera que sea el punto proporcionado en cada paso, $|\xi_n - \alpha| \le |x_n - \alpha| \longrightarrow 0$. Así, $\xi_n \to \alpha$. La continuidad añadida da $f''(\xi_n) \to f''(\alpha)$. La continuidad de f' y su valor no nulo en α permiten pasar al límite en el denominador. Por tanto,

$$\lim_{n\to\infty} \frac{x_{n+1} - \alpha}{(x_n - \alpha)^2} = \frac{f''(\alpha)}{2f'(\alpha)},$$

y, tomando valores absolutos,

$$\boxed{\lim_{n\to\infty} \frac{E_{n+1}}{E_n^2} = \frac{|f''(\alpha)|}{2|f'(\alpha)|}.}$$

Si $f''(\alpha) \neq 0$, este límite es positivo y el orden es exactamente dos. Si $f''(\alpha) = 0$, el límite es cero: la conclusión justificada es que $E_{n+1}/E_n^2 \to 0$, no que el orden sea necesariamente tres. Una afirmación más precisa requeriría información adicional. Hemos localizado también dónde se usa la nueva continuidad: en el paso al límite de $f''(\xi_n)$, no en la cota uniforme ni en la inducción de invariancia.

De un encierro racional a una órbita certificada.

Ejemplo 12.8. **La raíz cuadrada de dos: certificar todas las iteraciones.** Recuperemos $f(t) = t^2 - 2$ sobre $I = (0, \infty)$, $\alpha = \sqrt{2}$ y $x_0 = 3/2$. En Ejemplo 12.6 probamos $\frac{7}{5} < \alpha < \frac{3}{2}$.

Elijamos $r = 1/4$. Sin conocer el valor exacto de α, ese encierro demuestra

$$\alpha - \frac{1}{4} > \frac{23}{20} > 1, \qquad \alpha + \frac{1}{4} < \frac{7}{4} < 2.$$

Por tanto $B_r \subseteq [1, 2]$. En el intervalo mayor tenemos $|f'(t)| = 2t \ge 2$ y $|f''(t)| = 2$, así que podemos tomar

$$m = 2, \qquad M = 2, \qquad K = \frac{1}{2}, \qquad q = Kr = \frac{1}{8} < 1.$$

Además, $0 < x_0 - \alpha < 1/10 < r$. Todas las hipótesis de Proposición 12.2 están verificadas. Ahora sí sabemos que cada paso existe, que todos los puntos permanecen en B_r y que $x_n \to \sqrt{2}$. Los primeros términos, calculados exactamente, son

$$x_0 = \frac{3}{2}, \qquad x_1 = \frac{17}{12}, \qquad x_2 = \frac{577}{408}, \qquad x_3 = \frac{665857}{470832}.$$

Como $E_0 < 1/10$, la cota cuadrática iterada da

$$E_n < 2\left(\frac{1}{20}\right)^{2^n}.$$

En particular,

$$E_1 < \frac{1}{200}, \qquad E_2 < \frac{1}{80000}, \qquad E_3 < \frac{1}{12800000000}.$$

La estimación particular $E_1 < 1/300$ obtenida en §12.4 es más ajustada que la primera cota uniforme de esta lista. No hay contradicción: aquella utilizó el denominador exacto del paso; estas cotas usan una misma constante válida para todos los puntos de la región. La identidad $N_f(x) - \alpha = (x - \alpha)^2/(2x)$ muestra que, partiendo de $x_0 \neq \alpha$, ningún paso llega exactamente a la raíz. Como $f'' \equiv 2$ es continua, la fórmula asintótica proporciona

$$\lim_{n\to\infty} \frac{E_{n+1}}{E_n^2} = \frac{1}{2\sqrt{2}} > 0.$$

Aquí el orden es exactamente dos. La constante asintótica $1/(2\sqrt{2})$ no debe confundirse con la constante superior $K = 1/2$ elegida para una región completa.

Ejemplo 12.9. **El cúbico: justificar la elección inicial y un prefijo de la órbita.** Sea $p(t) = t^3 + t - 1$ y sea α la raíz aislada en Ejemplo 12.2. Sabemos que $\frac{5}{8} < \alpha < \frac{11}{16}$.

Tomemos $r = 1/8$. Las desigualdades anteriores dan

$$\alpha - r > \frac{1}{2}, \qquad \alpha + r < \frac{13}{16} < 1,$$

por lo que $B_r \subseteq [1/2, 1]$. En ese intervalo mayor,

$$p'(t) = 3t^2 + 1 \geq \frac{7}{4}, \qquad |p''(t)| = 6t \leq 6.$$

Con $m = 7/4$ y $M = 6$ obtenemos

$$K = \frac{12}{7}, \qquad q = Kr = \frac{3}{14} < 1.$$

El punto $x_0 = 3/4$ está en B_r, pues $0 < \frac{3}{4} - \alpha < \frac{1}{8}$. Por Proposición 12.2, la órbita que comienza en $3/4$ está definida para siempre, permanece en B_r y converge a α. La cota de error es

$$E_n \leq \frac{7}{12}\left(\frac{12}{7}E_0\right)^{2^n} < \frac{7}{12}\left(\frac{3}{14}\right)^{2^n}.$$

En Ejemplo 12.3 habíamos calculado $1 \mapsto 3/4 \mapsto 59/86$. El punto 1 no pertenece a la región escogida, de modo que no podemos invocar el criterio directamente desde él. Sí podemos justificar separadamente que su primer paso está definido y es $3/4$. Desde ese momento la órbita entra en la región certificada y la proposición controla toda la cola. Añadir aquel primer término no cambia el límite de la sucesión, como estudiamos en el capítulo de sucesiones. La raíz no se ha obtenido suponiendo convergencia ni se ha seleccionado mirando decimales. Su existencia y su aislamiento precedieron a Newton; las cotas demuestran ahora que el procedimiento converge precisamente a ella.

Leer las condiciones suficientes sin convertirlas en equivalencias. El radio construido no tiene por qué ser el mayor radio de convergencia. Las cotas m y M pueden ser conservadoras, y otro argumento puede justificar puntos iniciales adicionales. Por eso, si con ciertas cotas resulta $Kr \geq 1$, la conclusión correcta es «este criterio no certifica esa región», no «Newton diverge allí».

La frontera $Kr = 1$ también aclara la lógica. La estimación geométrica se reduciría a $E_{n+1} \leq E_n$, que no obliga a $E_n \to 0$. Incluso la desigualdad abstracta $E_{n+1} \leq KE_n^2$, con $K > 0$, admite la sucesión constante $E_n = 1/K$. Esta observación refuta una deducción basada únicamente en la desigualdad; no pretende presentar esa sucesión como una órbita de Newton. Dentro de la región que sí hemos certificado, los errores absolutos disminuyen estrictamente mientras no sean cero, porque $E_{n+1} \leq qE_n < E_n$. Eso no demuestra que los puntos x_n sean siempre crecientes o siempre decrecientes: E_n mide distancia y no conserva el lado de la raíz en que está x_n.

Laboratorio: reconstruir el argumento completo. Localizar una circularidad. Alguien escribe: «Para todo n, aplicamos la cota de §12.4 a x_n; por tanto, los puntos no salen del intervalo». La primera frase ya requiere la pertenencia que intenta demostrar. Repara el orden: prueba el envío de toda la región a sí misma; deduce por inducción que cada término existe y pertenece a ella; solo entonces aplica la cota para todo n.

Elegir el radio sin dividir por cero. Supón que $|f'| \geq 3$ y $|f''| \leq 12$ en $[\alpha - R, \alpha + R]$, con $R > 0$. Aquí $K = 2$. Una elección segura es $r = \min\{R/2, 1/4\}$, que produce $Kr \leq 1/2$. Si la segunda derivada fuera idénticamente cero en la región, no escribirías un radio con $1/M$: usarías directamente el caso $K = 0$.

Separar límite de velocidad. La cota $E_n \leq q^n E_0$ demuestra el límite, pero no demuestra por sí sola convergencia cuadrática. Para esto necesitas recuperar $E_{n+1} \leq K E_n^2$. Para afirmar orden exactamente dos necesitas, además, un límite positivo del cociente de errores, o un argumento equivalente; no basta señalar el exponente 2 en una cota superior.

No trasladar una garantía al pasado. Si una iteración inicial entra en un intervalo certificado, el criterio controla la cola a partir de esa entrada. Hay que verificar separadamente que todos los pasos del prefijo estén definidos. El ejemplo del cúbico muestra exactamente esta división del trabajo.

Conocer la convergencia y su régimen cuadrático permite transportar una cota inicial mediante la recurrencia demostrada. Para convertir el residuo de una aproximación ya calculada en una tolerancia necesitaremos además una estimación a posteriori que relacione ese residuo con la distancia a la raíz.

Lo que ahora podemos afirmar. El paso de Newton ya no es solamente una fórmula sugerida por una tangente. Disponemos de una cadena completa:

raíz y cotas en una región
$\Downarrow$
radio con $Kr < 1$
$\Downarrow$
invariancia y existencia de todas las iteraciones
$\Downarrow$
convergencia y control cuadrático del error.

Cada flecha tiene una justificación diferente. Haber probado convergencia local no elimina la necesidad de comprobar esas condiciones cuando cambiamos la función, el intervalo o el punto inicial.

Antes de seguir. ¿Qué garantiza que x_{n+1} puede servir como entrada del siguiente paso? ¿Por qué el radio se compara con $1/K$ cuando $K > 0$? ¿Qué ocurre si $K = 0$ o si se alcanza exactamente la raíz? ¿Dónde fue necesaria la continuidad de f'' y dónde no? ¿Cómo se comprueba una región centrada en una raíz desconocida usando un encierro racional? ¿Qué diferencia hay entre una cota superior cuadrática y un orden exactamente dos?

12.6. Certificados a posteriori y criterios de parada

En §12.5 demostramos que Newton converge bajo condiciones precisas y transportamos una cota inicial a todos los errores posteriores. Ahora la pregunta cambia: después de calcular un punto concreto, ¿qué información accesible permite decidir si ya satisface la tolerancia pedida?

La distinción importa incluso cuando la convergencia está demostrada. Afirmar $x_n \to \alpha$ garantiza que los errores terminan siendo pequeños; para detener un cálculo en un índice particular necesitamos una cota que podamos verificar en ese índice. El número de iteraciones, el residuo de la ecuación y el tamaño del último paso no son certificados por sí solos.

Del control previo al examen del punto calculado. Una estimación *a priori* utiliza datos fijados antes de obtener las nuevas aproximaciones. Por ejemplo, las cotas de §12.5 combinan un encierro inicial y constantes de derivadas para dominar $E_n = |x_n - \alpha|$. Una estimación *a posteriori* utiliza además información obtenida al calcular la aproximación que examinamos: su residuo, un intervalo que ahora la acompaña o un paso efectivamente realizado.

La diferencia no es entre una prueba rigurosa y otra informal. Ambas deben ser demostraciones. Lo que cambia es la información utilizada y el momento en que está disponible. Tampoco llamaremos «error» a cualquier número que podamos calcular: primero habrá que probar que ese número domina la distancia a la raíz. Recuperemos el problema de §12.1. Podemos evaluar $f(x)$, pero no calcular directamente $x - \alpha$ porque desconocemos α. El teorema del valor medio proporciona el enlace que faltaba: para $x \neq \alpha$, bajo sus hipótesis en el segmento que une ambos puntos, $f(x) - f(\alpha) = f'(\xi)(x - \alpha)$ para algún ξ estrictamente intermedio. Si $f(\alpha) = 0$, entonces $|f(x)| = |f'(\xi)|\,|x - \alpha|$.

Así, una pendiente que no pueda hacerse demasiado pequeña impedirá que una distancia grande produzca un residuo engañosamente pequeño. Necesitamos una cota inferior válida dondequiera que pueda encontrarse ξ, no solamente la derivada en x.

La desigualdad que convierte residuo en error.

Proposición 12.3. **Proposición (certificado residual por el teorema del valor medio).** Sea $J \subseteq \mathbb{R}$ un intervalo que contiene al menos dos puntos. Sea $f : J \to \mathbb{R}$ continua en J, con la continuidad relativa en los extremos que pertenezcan a J, y diferenciable en su interior J°. Supongamos que existe $m > 0$ tal que

$$|f'(t)| \geq m \qquad (t \in J^\circ).$$

Entonces, para cualesquiera $x, y \in J$,

$$\boxed{m|x - y| \leq |f(x) - f(y)|.}$$

En particular, f es inyectiva en J. Si existe $\alpha \in J$ con $f(\alpha) = 0$, esa raíz es única en J y, para todo $x \in J$,

$$\boxed{|x - \alpha| \leq \frac{|f(x)|}{m}.}$$

Por consiguiente, para cualquier tolerancia $\varepsilon > 0$,

$$\boxed{|f(x)| \leq m\varepsilon \quad \Longrightarrow \quad |x - \alpha| \leq \varepsilon.}$$

La proposición no afirma por sí sola que exista una raíz en J.

Demostración. Si $x = y$, la primera desigualdad tiene ambos miembros nulos. Supongamos que $x \neq y$. Al ser J un intervalo, contiene todo el segmento cerrado entre ambos puntos. La restricción de f a ese segmento es continua, y es diferenciable en su interior: cualquier punto estrictamente entre dos puntos de J pertenece a J°.

Por Teorema 9.3 existe ξ estrictamente entre x e y tal que $f(x) - f(y) = f'(\xi)(x - y)$. Esta forma es válida en cualquiera de los dos órdenes de los extremos. Tomando valores absolutos y utilizando la cota inferior en ξ obtenemos

$$|f(x) - f(y)| = |f'(\xi)|\,|x - y| \geq m|x - y|.$$

Si $f(x) = f(y)$, resulta $m|x - y| \leq 0$. Como $m > 0$ y el valor absoluto es no negativo, $x = y$. Esto demuestra inyectividad y, por tanto, que no puede haber dos raíces distintas en J. Supongamos ahora que $f(\alpha) = 0$ para algún $\alpha \in J$. Aplicando la primera desigualdad al par x, α, $m|x - \alpha| \leq |f(x) - f(\alpha)| = |f(x)|$. Dividimos por $m > 0$ y obtenemos el certificado. Si además $|f(x)| \leq m\varepsilon$, entonces $|x - \alpha| \leq \varepsilon$. El caso $x = \alpha$ ya está cubierto por la primera parte, sin introducir un punto intermedio entre puntos iguales. □

La prueba no requiere segunda derivada, continuidad de la primera derivada ni un historial de iteraciones. Tampoco exige derivar en los extremos de J. Podemos certificar un punto producido por Newton, por bisección, por una búsqueda racional o por cualquier otra construcción, siempre que comprobemos su pertenencia al intervalo y las hipótesis pertinentes.

Qué debe comprobarse antes de usar el certificado. La primera tarea es localizar la raíz. El aislamiento de §12.2, por ejemplo, puede proporcionar $\alpha \in J$. Sin esa información la expresión «error respecto de la raíz de J» podría no tener referente: la función $f(t) = t + 1$ en $[0, 1]$ cumple $|f'| \geq 1$, pero no tiene raíz en ese intervalo. La cota de pendiente garantiza a lo sumo unicidad; no crea existencia.

La segunda tarea es comprobar que el punto calculado está en la misma región. Para $f(t) = t^2 - 2$, una cota demostrada en $[1, 2]$ no permite certificar una aproximación negativa de la raíz positiva. Un residuo pequeño puede corresponder a otra raíz de la misma ecuación. La elección de J identifica tanto las hipótesis como la raíz que seguimos.

La tercera tarea es justificar el número m. En la demostración aparece $f'(\xi)$ y no necesariamente $f'(x)$. Conocer $|f'(x)|$ en un solo punto no lo convierte en una cota inferior en todo J°. Tampoco basta comprobar varias entradas de una tabla. La condición debe valer para todos los puntos interiores del intervalo.

Dato	Papel en la prueba
J es un intervalo y $x, \alpha \in J$	El segmento entre ambos no abandona la región controlada.
Continuidad en J y diferenciabilidad en J°	Permiten aplicar el valor medio a ese segmento, incluidos sus bordes.
$f(\alpha) = 0$	Convierte la diferencia de valores en el residuo $\lvert f(x)\rvert$.
$\lvert f'\rvert \geq m > 0$ en todo J°	Domina la pendiente desconocida y autoriza dividir por m.

Si además disponemos de una cota superior $|f'(t)| \leq L$ en J°, el mismo valor medio da $m|x - \alpha| \leq |f(x)| \leq L|x - \alpha|$. En ese caso residuo y error se controlan en ambos sentidos, con constantes conocidas. La cota superior sola permitiría deducir residuo pequeño a partir de error pequeño; para la dirección que necesitamos al certificar, la información decisiva es la cota inferior.

Un residuo racional para la raíz cuadrada.

Ejemplo 12.10. **Certificar el punto calculado sin conocer el error inicial.** Consideremos $f(t) = t^2 - 2$ en $J = [1, 2]$ y su raíz positiva $\alpha = \sqrt{2}$. Sabemos que $\alpha \in J$ y que $f'(t) = 2t \geq 2$ en su interior. Podemos usar $m = 2$.

En Ejemplo 12.8 obtuvimos el punto $x = \frac{577}{408}$. Verificamos directamente $1 < x < 2$. Su residuo se calcula con enteros:

$$f(x) = \frac{577^2 - 2 \cdot 408^2}{408^2} = \frac{332929 - 332928}{166464} = \frac{1}{166464}.$$

Por Proposición 12.3,

$$\boxed{\left|\frac{577}{408} - \sqrt{2}\right| \leq \frac{1}{332928} < \frac{1}{300000}.}$$

La última desigualdad solo compara denominadores positivos. El punto satisface, por tanto, la tolerancia $1/300000$ sin consultar los decimales de $\sqrt{2}$ ni volver a estimar su distancia al punto inicial de Newton. Podemos comparar esta cota con la cota a priori $1/80000$ del ejemplo anterior. Ambas son correctas; aquí el residuo efectivamente calculado proporciona una cota más ajustada. No existe una regla que obligue a que todo certificado a posteriori sea siempre mejor: hay que comparar los valores obtenidos y la información usada en cada caso.

Un decimal es también un racional que puede examinarse.

Ejemplo 12.11. **Certificado para el cúbico, independiente de la procedencia de la aproximación.** Recuperemos $p(t) = t^3 + t - 1$ y su raíz $\alpha \in (5/8, 11/16)$ de Ejemplo 12.2. Tomemos ahora

$$J = [1/2, 1], \qquad x = 0.68233 = \frac{68233}{100000}.$$

Ambos, la raíz y el punto calculado, pertenecen a J. Para cada $t \in J^\circ$, $p'(t) = 3t^2 + 1 \geq \frac{7}{4}$, así que $m = 7/4$ es una cota inferior válida. El decimal finito escrito arriba se entiende como un racional exacto. Por cálculo entero,

$$p(x) = \frac{68233^3 + 68233 \cdot 100000^2 - 100000^3}{100000^3} = \frac{5263605337}{10^{15}} > 0.$$

El certificado resulta

$$|x - \alpha| \leq \frac{4}{7}\frac{5263605337}{10^{15}} = \frac{5263605337}{1750000000000000} < \frac{1}{300000}.$$

Para la última comparación basta comprobar

$$5263605337 \cdot 300000 = 1579081601100000 < 1750000000000000.$$

Así queda certificada la tolerancia $1/300000$. No hemos supuesto que este decimal pertenezca exactamente a una órbita de Newton. Su procedencia no interviene en la prueba: intervienen el intervalo, la cota de pendiente y su residuo. Como p es estrictamente creciente en J y $p(x) > 0 = p(\alpha)$, también sabemos $x > \alpha$. Por tanto el certificado puede leerse como un encierro de un solo lado, $x - |p(x)|/m \leq \alpha < x$. El signo aporta orientación; la cota aporta distancia.

El residuo depende de la escala de la ecuación. El ejemplo de §12.1 ya mostró que $f_N(t) = (t-1)/N$ puede tener un residuo muy pequeño en $x = 2$ sin que el error, igual a 1, disminuya. Ahora podemos identificar la compensación exacta: su cota de pendiente es $m_N = 1/N$, y

$$\frac{|f_N(2)|}{m_N} = \frac{1/N}{1/N} = 1.$$

El certificado no confunde una ecuación de escala pequeña con una aproximación precisa. Más generalmente, si reemplazamos f por $g = \lambda f$, con $\lambda \neq 0$, las raíces no cambian. Una cota $|f'| \geq m$ se transforma en $|g'| \geq |\lambda| m$, de modo que

$$\frac{|g(x)|}{|\lambda| m} = \frac{|f(x)|}{m}.$$

También el paso de Newton permanece igual donde está definido:

$$N_g(x) = x - \frac{\lambda f(x)}{\lambda f'(x)} = N_f(x).$$

El residuo bruto cambia con la escala; el cociente correctamente normalizado no. Por eso el umbral certificado es $m\varepsilon$, no una tolerancia para $|f(x)|$ elegida sin atender a la pendiente.

Un paso pequeño todavía puede dejar un error grande. Para un paso exacto de Newton, escribamos

$$x^+ = N_f(x), \qquad S = |x^+ - x|.$$

Siempre que $f'(x) \neq 0$,

$$S = \frac{|f(x)|}{|f'(x)|}.$$

Es tentador leer este cociente como una cota de $|x - \alpha|$. Pero Proposición 12.3 divide por una cota inferior uniforme m, no por una derivada escogida en el punto actual. Además, el tamaño S mide la distancia entre dos aproximaciones, no su distancia a la raíz.

Ejemplo 12.12. **Un paso arbitrariamente pequeño lejos de una raíz simple.** Para $0 < \delta < 1$, sea

$$g_\delta(t) = \frac{t-1}{t+\delta}, \qquad I = (-\delta, \infty).$$

Su única raíz es $\alpha = 1$, y

$$g'_\delta(t) = \frac{1+\delta}{(t+\delta)^2} > 0 \qquad (t \in I).$$

Desde $x = 0 \in I$, el paso está definido y da

$$x^+ = 0 - \frac{-1/\delta}{(1+\delta)/\delta^2} = \frac{\delta}{1+\delta}.$$

Tanto x como x^+ pertenecen a $[0,1] \subset I$. Sin embargo,

$$S = \frac{\delta}{1+\delta} < \delta, \qquad |x-\alpha| = 1, \qquad |x^+ - \alpha| = \frac{1}{1+\delta} > \frac{1}{2}.$$

Podemos hacer el paso menor que cualquier tolerancia positiva escogiendo δ suficientemente pequeño, mientras ambos errores permanecen grandes. La función es racional, la raíz es simple y ninguno de los dos puntos sale del dominio. Por tanto, la inferencia «paso pequeño implica aproximación precisa» necesita algo más que esas condiciones. El certificado residual no falla. En $J = [0,1]$ podemos usar $m = 1/(1+\delta)$, y el residuo inicial $1/\delta$ produce una cota $(1+\delta)/\delta$, que no es pequeña. El ejemplo no estudia la convergencia completa de esta familia: solo refuta el criterio de parada sin hipótesis adicionales.

Podemos reparar parcialmente la inferencia usando el certificado residual. Si $x \in J^\circ$ satisface las hipótesis de Proposición 12.3 y calculamos su paso exacto, entonces $|x-\alpha| \le \frac{|f(x)|}{m} = \frac{|f'(x)|}{m} S$. La desigualdad triangular proporciona además

$$|x^+ - \alpha| \le \left(1 + \frac{|f'(x)|}{m}\right) S.$$

El factor que faltaba puede ser grande. Esta última comparación no evalúa f en x^+ y no afirma que el punto nuevo pertenezca a J; para volver a aplicar allí el certificado residual habría que verificar esa pertenencia.

Un criterio basado en el paso dentro de la región certificada. La información de §12.5 permite una cota diferente. Allí ya demostramos $E_{n+1} \le qE_n$ con un mismo $q < 1$. Esa reducción conocida, y no la mera apariencia pequeña de un paso, es lo que ahora utilizaremos.

Corolario 12.3. **Corolario (certificados del paso en la región de convergencia).** Bajo todas las hipótesis y notaciones de Proposición 12.2, sea (x_n) la órbita exacta de Newton con $x_0 \in B_r$. Para $n \ge 0$, pongamos

$$S_n = |x_{n+1} - x_n|.$$

Entonces

$$\boxed{E_n \le \frac{S_n}{1-q}}, \qquad \boxed{E_{n+1} \le \frac{q}{1-q} S_n}.$$

En particular, para devolver el punto nuevo x_{n+1} con tolerancia $\varepsilon > 0$, basta comprobar

$$\boxed{qS_n \le (1-q)\varepsilon.}$$

Demostración. La proposición citada ya garantiza la existencia de los términos, su permanencia en B_r y la desigualdad $E_{n+1} \le qE_n$. Por la desigualdad triangular,

$$E_n = |x_n - \alpha| \le |x_n - x_{n+1}| + |x_{n+1} - \alpha| = S_n + E_{n+1} \le S_n + qE_n.$$

Restando qE_n y dividiendo por $1-q > 0$, $E_n \le \frac{S_n}{1-q}$.

Entonces $E_{n+1} \le qE_n \le qS_n/(1-q)$, y la condición de parada enunciada implica $E_{n+1} \le \varepsilon$. □

La prueba no utiliza una serie de pasos ni el teorema de Banach. Solo utiliza la reducción de error ya demostrada y la desigualdad triangular. No hemos supuesto una contracción de N_f entre puntos arbitrarios. Si $q = 0$, el criterio sigue siendo válido: el punto nuevo es la raíz. La condición de parada se escribió sin dividir por q precisamente para incluir ese caso. Si un

paso exacto tiene $S_n = 0$, la primera cota obliga a $E_n = 0$; esto no permite identificar un paso exactamente nulo con dos valores impresos que coinciden tras redondeo. Hay que leer también los índices. $S_n/(1-q)$ certifica x_n; $qS_n/(1-q)$ certifica x_{n+1}. Cambiar el punto devuelto sin cambiar la cota puede invalidar el informe del cálculo.

Por ejemplo, en la región de Ejemplo 12.8 tenemos $q = 1/8$. El paso $17/12 \mapsto 577/408$ mide

$$S_1 = \left|\frac{577}{408} - \frac{17}{12}\right| = \frac{1}{408}.$$

El corolario certifica para el punto nuevo un error a lo sumo

$$\frac{1/8}{1-1/8}\frac{1}{408} = \frac{1}{2856}.$$

La cota es válida, aunque resulta menos ajustada que $1/332928$, obtenida mediante el residuo del mismo punto en Ejemplo 12.10. Para elegir entre certificados podemos comparar las hipótesis que exigen, su costo de comprobación y la precisión concreta que proporcionan.

El valor evaluado también puede tener incertidumbre. Hasta aquí los residuos de los ejemplos se calcularon exactamente con racionales. Si solo conocemos un valor aproximado $\widetilde{r}$ de $f(x)$, necesitamos además una cota demostrada

$$|\widetilde{r} - f(x)| \le \eta, \qquad \eta \ge 0.$$

La desigualdad triangular da $|f(x)| \le |\widetilde{r}| + \eta$. Bajo las mismas hipótesis del certificado residual,

$$\boxed{|x - \alpha| \le \frac{|\widetilde{r}| + \eta}{m}.}$$

Por tanto una condición suficiente de parada es $|\widetilde{r}| + \eta \le m\varepsilon$. Si una evaluación entrega $\widetilde{r} = 0$, no hemos demostrado $f(x) = 0$ a menos que también sepamos $\eta = 0$. Con $\eta > 0$, la conclusión disponible es una cota η/m. Esta observación no presupone un método para obtener η. La cota de evaluación es un dato que debe justificarse; aquí mostramos cómo incorporarla cuando está disponible. No convertimos el modelo de aritmética exacta de Newton en un teorema sobre cualquier implementación redondeada. El número que entregamos al lector puede tener un segundo error. Si B certifica $|x - \alpha| \le B$ y reemplazamos x por un valor impreso y con $|y - x| \le \rho$, entonces

$$\boxed{|y - \alpha| \le \rho + B.}$$

Esta cota se obtiene otra vez por la desigualdad triangular. No requiere aplicar el teorema del valor medio en y, ni presupone que y sea un término exacto de la órbita. Para el punto $x = 577/408$, el valor $y = 1.414216$ satisface exactamente $y - x = \frac{1}{3187500} < \frac{1}{2000000}$. Al combinar esta cota de presentación con Ejemplo 12.10 obtenemos

$$|y - \sqrt{2}| \le \frac{1}{332928} + \frac{1}{2000000} = \frac{9113}{2601000000} < \frac{1}{250000} = 4 \cdot 10^{-6}.$$

Mostrar seis decimales no convirtió esa tolerancia en 10^{-6}. El objeto certificado es el número efectivamente entregado, con los errores que lo separan del punto exacto analizado. Tampoco una tolerancia absoluta, sin comprobar la posición respecto de los umbrales de redondeo, identifica automáticamente el redondeo correcto de la raíz a un número de decimales.

Elegir el certificado y justificar la parada. Podemos reunir las alternativas sin confundir qué punto certifica cada una:

Información comprobada	Punto y cota certificada
$\alpha \in [a,b]$ y $x \in [a,b]$	$\lvert x-\alpha\rvert \le \max\{x-a, b-x\}$; en el punto medio, $(b-a)/2$.
Hipótesis de Proposición 12.3 y residuo exacto	$\lvert x-\alpha\rvert \le \lvert f(x)\rvert/m$.
Mismas hipótesis y evaluación con error η	$\lvert x-\alpha\rvert \le (\lvert\widetilde{r}\rvert+\eta)/m$.
Hipótesis de Corolario 12.3 y paso exacto S_n	$\lvert x_{n+1}-\alpha\rvert \le qS_n/(1-q)$.

Si varios certificados se refieren al mismo punto y a la misma raíz, podemos tomar el mínimo de sus cotas. No podemos tomar el mínimo de números que certifican puntos o raíces diferentes como si fueran intercambiables. Una cota de presentación ρ se añade después, si cambiamos el valor que será entregado.

Un procedimiento de parada riguroso fija primero la tolerancia y la raíz aislada, comprueba las hipótesis en la región pertinente, calcula una cota para el punto que va a devolver y se detiene cuando esa cota no supera la tolerancia. Al informar el resultado conserva el punto, la cota y el argumento que la justifica. «Se realizaron diez iteraciones» describe un procedimiento; «el error es a lo sumo ε por tal desigualdad» describe una garantía.

La condición suficiente también debe distinguirse de una condición necesaria. Si $|f(x)| > m\varepsilon$, el certificado residual escogido no demuestra la tolerancia, pero tampoco demuestra que el error sea mayor que ε. Una cota conservadora puede ser demasiado grande aunque la aproximación sea buena. Podemos mejorar la cota de pendiente, estrechar la región o utilizar otro certificado válido.

En aritmética exacta, bajo las hipótesis de Proposición 12.2, el criterio residual termina satisfaciéndose para cualquier tolerancia positiva: ya sabemos $x_n \to \alpha$, y la continuidad da $f(x_n) \to f(\alpha) = 0$. Con el $m > 0$ fijo de esa región, eventualmente $|f(x_n)| < m\varepsilon$. Esta es una demostración de terminación del criterio dentro de aquel régimen; no una garantía universal para cualquier punto inicial o evaluación aproximada. Una incertidumbre de evaluación que no se reduzca puede impedir certificar tolerancias arbitrariamente pequeñas.

Laboratorio de certificación y reparación. Reconstruir el puente. Parte de $f(x) - f(\alpha) = f'(\xi)(x-\alpha)$. Explica por qué ξ está en la región donde conoces la cota de pendiente y por qué necesitas un límite inferior, no superior. Si $x = \alpha$, reconoce directamente la desigualdad sin invocar un intervalo no degenerado.

Auditar un umbral. Con $m = 3/2$ y tolerancia $\varepsilon = 1/10000$, basta exigir $|f(x)| \le 3/20000$. Si el residuo exacto es $1/5000$, este criterio no se cumple. No concluyas que la tolerancia es falsa: solo que esta cota no la certifica.

Incluir la evaluación. Si $m = 2$, $|\widetilde{r}| \le 3 \cdot 10^{-6}$ y $\eta \le 10^{-6}$, el error queda acotado por $2 \cdot 10^{-6}$. Si se imprime otro número con distancia a x a lo sumo 10^{-6}, la tolerancia certificada para ese número pasa a ser $3 \cdot 10^{-6}$.

Identificar el punto que se devuelve. En una región ya certificada con $q = 1/4$, un paso exacto $S_n \le 3 \cdot 10^{-5}$ da $E_{n+1} \le 10^{-5}$, pero solo da $E_n \le 4 \cdot 10^{-5}$. El mismo dato no autoriza el mismo umbral para los dos índices.

Reparar una falsa necesidad. Para $f(t) = 100(t-1)$ en $\mathbb{R}$, la cota $m = 1$ es válida aunque poco ajustada. Si $x = 1 + \varepsilon/2$, entonces $|x-1| = \varepsilon/2 \le \varepsilon$, pero $|f(x)| = 50\varepsilon > m\varepsilon$. El fallo del

umbral suficiente no refuta la precisión. Usar la cota mejor $m = 100$ recupera un certificado ajustado.

Lo que queda establecido. Disponemos ya de dos garantías distintas: §12.5 demuestra que una órbita converge dentro de una región controlada; esta sección permite certificar una aproximación concreta y decidir cuándo detenerse. El teorema del valor medio convierte un residuo en error cuando la pendiente está separada de cero. El tamaño de un paso solo se convierte en certificado después de añadir las hipótesis que permiten relacionarlo con la distancia a la raíz. En §12.7 examinaremos de manera sistemática qué puede ocurrir cuando esas condiciones faltan: derivada nula, salida del dominio, ciclos y raíces múltiples. Los certificados no eliminan esos problemas; nos enseñan a distinguir una aproximación demostrada de un cálculo que todavía no posee garantía.

Antes de seguir. ¿Por qué el certificado residual no prueba por sí solo existencia de una raíz? ¿Dónde necesitas controlar la pendiente? ¿Por qué multiplicar una ecuación por una constante pequeña no mejora la aproximación? ¿Qué punto certifica cada una de las dos cotas del paso? ¿Qué información falta cuando una evaluación aproximada devuelve cero? ¿Cómo cambia la tolerancia si redondeas el valor que entregas?

12.7. Cuándo Newton falla o engaña

En §12.5 demostramos convergencia bajo condiciones precisas y en §12.6 aprendimos a certificar un punto calculado. Ahora someteremos esas condiciones a una prueba distinta: construiremos situaciones en las que alguna falta y examinaremos qué conclusión deja de estar justificada. No todos los resultados desfavorables significan lo mismo. Puede faltar el paso siguiente; puede existir una órbita infinita que no converge; puede alcanzarse una raíz distinta de la buscada; o puede haber convergencia, pero no con la velocidad cuadrática esperada. También puede fallar únicamente el certificado elegido, sin que la aproximación sea mala. Confundir estos diagnósticos conduce a reparaciones igualmente confusas. La pregunta no será «¿es bueno o malo Newton?», sino esta:

¿qué afirmación queríamos hacer y qué hipótesis la autorizaba?

El dominio se comprueba en cada paso. Para una función diferenciable $f : I \to \mathbb{R}$ en un intervalo abierto, Definición 12.2 establece

$$D_f = \{x \in I : f'(x) \neq 0\}, \qquad N_f(x) = x - \frac{f(x)}{f'(x)} \quad (x \in D_f).$$

Hay dos comprobaciones diferentes: que x pertenezca a I, para poder evaluar la función y su derivada, y que $f'(x) \neq 0$, para poder formar el cociente. Una igualdad algebraica no elimina ninguna de ellas. Si se reconoce exactamente que $f(x) = 0$, el problema de encontrar una raíz ya está resuelto en ese punto. Se puede detener el cálculo aunque $f'(x) = 0$; no se necesita ni se autoriza evaluar $0/0$. En cambio, si $f(x) \neq 0$ y $f'(x) = 0$, la tangente es horizontal y no corta al eje horizontal: el modelo afín no proporciona el punto siguiente.

Ejemplo 12.13. **Una raíz existente no proporciona cualquier punto inicial.** Sea

$$f(t) = t^2 - 3, \qquad I = \mathbb{R}.$$

Sus raíces son $\pm\sqrt{3}$, construidas con las raíces algebraicas disponibles en el tomo. Sin embargo,

$$f(0) = -3, \qquad f'(0) = 0.$$

Newton no está definido desde $x_0 = 0$. No hemos refutado la existencia de raíces; hemos encontrado un punto inicial inadmisible para este procedimiento. La cercanía a ese punto también merece atención. Para cada entero $k \geq 1$, el punto $x = 1/k$ es admisible y

$$N_f\left(\frac{1}{k}\right) = \frac{(1/k)^2 + 3}{2/k} = \frac{3k}{2} + \frac{1}{2k}.$$

Los puntos iniciales tienden a cero, pero sus primeras imágenes no están acotadas superiormente. Todos esos pasos son operaciones reales legítimas. Lo que falta es una región en la que el cociente y el error queden controlados uniformemente.

No hay que transformar esta observación en otra regla falsa: una derivada pequeña, considerada aisladamente, no obliga a que Newton funcione mal. Para $f_\lambda(t) = \lambda(t-1)$, con $\lambda \neq 0$, se tiene $N_{f_\lambda}(x) = 1$ desde cualquier real, por pequeña que sea $|\lambda|$. Como vimos en §12.6, multiplicar toda la ecuación por una constante no nula no cambia el paso. Lo relevante es el cociente y las hipótesis que lo controlan, no un umbral arbitrario aplicado solo al valor de la derivada. Otra obstrucción ya quedó probada en Ejemplo 12.5. Para $g(t) = \sqrt{t} - 1$ sobre $(0, \infty)$, $N_g(9) = -3$.

El primer paso existe, pero la función no está definida en el punto nuevo. No corresponde reemplazarlo por 3, por su valor absoluto o por un extremo del dominio sin explicación: cualquiera de esas decisiones introduce un procedimiento distinto, que requiere su propia justificación. La condición $N_f(B_r) \subseteq B_r$ de Proposición 12.2 evitaba precisamente este problema.

Un ciclo exacto, incluso con una raíz aislada. Podríamos intentar salvar una promesa demasiado amplia de esta manera: «Si la raíz está aislada, las derivadas existen y todos los pasos se pueden calcular, entonces Newton debe converger». El siguiente ejemplo refuta esa afirmación sin errores de redondeo.

Ejemplo 12.14. **Dos tangentes que devuelven alternativamente los mismos puntos.** Consideremos

$$h(t) = t^3 - 5t = t(t^2 - 5), \qquad I = \mathbb{R}.$$

En $J = [-1, 1]$ la única raíz es $\alpha = 0$, pues $t^2 - 5 < 0$ en todo J. Además,

$$h(-1) = 4, \qquad h(1) = -4,$$

$$h'(t) = 3t^2 - 5 \leq -2 \quad (t \in J).$$

Tenemos, por tanto, una raíz aislada, cambio de signo y una derivada separada de cero en todo el intervalo. La función es un polinomio y admite todas las derivadas necesarias. Su operador de Newton es

$$N_h(x) = x - \frac{x^3 - 5x}{3x^2 - 5} = \frac{2x^3}{3x^2 - 5}, \qquad 3x^2 - 5 \neq 0.$$

Las evaluaciones exactas dan

$$\boxed{N_h(1) = -1, \qquad N_h(-1) = 1.}$$

Desde $x_0 = 1$, una inducción inmediata proporciona

$$x_{2n} = 1, \qquad x_{2n+1} = -1 \qquad (n \geq 0).$$

Esta órbita no converge a ningún real. Si convergiera a ℓ, para índices suficientemente grandes dos términos consecutivos estarían ambos a distancia menor que $1/2$ de ℓ, y entonces

$$2 = |x_{n+1} - x_n| \leq |x_{n+1} - \ell| + |x_n - \ell| < 1,$$

una contradicción. Se trata de un *ciclo de período dos*: el punto vuelve exactamente al mismo valor después de dos pasos, pero no después de uno. Más aún, J sí es invariante. Para $|x| \leq 1$,

$$|N_h(x)| = \frac{2|x|^3}{5 - 3x^2} \leq |x|^3 \leq 1.$$

La invariancia evita que el cálculo salga del intervalo, pero no obliga a que sus términos se acerquen a la raíz. En la órbita indicada,

$$|x_n - \alpha| = 1, \qquad |h(x_n)| = 4 \qquad (n \geq 0).$$

También podemos reconstruir el ejemplo en lugar de memorizarlo. Para la familia $h_A(t) = t^3 - At$, la condición $N_{h_A}(1) = -1$ se convierte en $\frac{2}{3-A} = -1$, de donde $A = 5$. La simetría impar da entonces el paso de vuelta. El contraejemplo se diseñó imponiendo las dos imágenes que queríamos obtener.

¿Por qué esto no contradice §12.5? En $[-1, 1]$ podemos tomar $m = 2$ y $M = 6$, pero con radio $r = 1$ resulta

$$K = \frac{M}{2m} = \frac{3}{2}, \qquad Kr = \frac{3}{2} > 1.$$

No se satisface la condición que convertía la cota cuadrática en reducción de error. En este caso el ciclo demuestra falta efectiva de convergencia; el solo incumplimiento de una condición suficiente, en otro problema, no bastaría para demostrarla. La reparación local puede hacerse explícita para la misma función y la misma raíz. En $[-1/2, 1/2]$,

$$|h'| \geq \frac{17}{4}, \qquad |h''| = |6t| \leq 3.$$

Con $r = 1/2$ obtenemos

$$K = \frac{6}{17}, \qquad q = Kr = \frac{3}{17} < 1.$$

Ahora sí se aplica Proposición 12.2 a todo punto inicial de ese intervalo. Por ejemplo, $1/2 \mapsto -1/17$ inicia una órbita cuya convergencia a cero está demostrada. No hemos «arreglado» el ciclo continuándolo durante más tiempo: hemos cambiado el punto inicial y justificado una región distinta.

Un residuo que tiende a cero mientras los puntos se alejan. El ciclo anterior conserva un residuo grande. ¿Bastaría, entonces, observar que los residuos tienden a cero? La hipótesis de existencia y la cota inferior uniforme de §12.6 siguen siendo indispensables para su certificado.

Ejemplo 12.15. **Iterar hacia una región sin raíz.** Sea

$$u(t) = \frac{1}{t}, \qquad I = (0, \infty).$$

La función es positiva en todo su dominio, de modo que no tiene ninguna raíz allí. Su derivada nunca se anula: $u'(t) = -\frac{1}{t^2}$. No obstante, Newton está definido en todo I, conserva ese intervalo y satisface

$$N_u(x) = x - \frac{1/x}{-1/x^2} = 2x.$$

Desde $x_0 = 1$, la órbita exacta es $x_n = 2^n$. Por Bernoulli, $2^n \geq n + 1$, de modo que la sucesión no está acotada y no posee límite real finito. En cambio, $|u(x_n)| = 2^{-n} \longrightarrow 0$ por el límite geométrico Corolario 3.2. El umbral $|u(x_n)| \leq \varepsilon$ acaba cumpliéndose para cualquier $\varepsilon > 0$, pero no certifica cercanía a una raíz: no existe tal raíz en I. Además, $|u'(x_n)| = 4^{-n} \to 0$, de modo que no hay un $m > 0$ válido como cota inferior de la pendiente en toda la región recorrida.

No sería legítimo escribir $u(\lim x_n) = 0$: ese límite real no existe. Tampoco podemos añadir un «punto en el infinito» al dominio como si fuera una raíz real. El ejemplo refuta el criterio basado únicamente en el residuo, no Proposición 12.3, que exige una raíz existente en el intervalo y una cota positiva uniforme.

Converger a una raíz distinta no resuelve la pregunta original. Aunque la convergencia esté demostrada, todavía debemos identificar su destino. Una ecuación puede tener varias raíces, y el método no recibe por separado la instrucción de escoger una de ellas.

Ejemplo 12.16. **La rama negativa de una misma ecuación.** Supongamos que buscamos la raíz positiva de $f(t) = t^2 - 2$, pero elegimos $x_0 = -3/2$. El operador, sobre $D_f = \mathbb{R} \setminus \{0\}$, es $N_f(x) = \frac{x^2+2}{2x}$.

Es impar: $N_f(-x) = -N_f(x)$. Sea (y_n) la órbita positiva desde $y_0 = 3/2$, cuya convergencia a $\sqrt{2}$ quedó demostrada en Ejemplo 12.8. Por inducción,

$$x_n = -y_n \qquad (n \geq 0).$$

Todos los pasos negativos están definidos y

$$x_n \longrightarrow -\sqrt{2}, \qquad |f(x_n)| = |f(y_n)| \longrightarrow 0.$$

Sin embargo, respecto de la raíz positiva solicitada,

$$|x_n - \sqrt{2}| = y_n + \sqrt{2} \longrightarrow 2\sqrt{2} \neq 0.$$

La órbita resuelve la ecuación, pero no aproxima la raíz que habíamos fijado. La información de aislamiento no puede abandonarse después de escribir la fórmula iterativa.

Conviene precisar el alcance de esta advertencia. Bajo regularidad suficiente, una órbita exacta que converge a un punto interior no puede converger a un valor donde la función sea no nula. En efecto, si $f \in C^1(I)$, todos los pasos están definidos y $x_n \to \ell \in I$, entonces $x_{n+1} \to \ell$ y

$$f(x_n) = f'(x_n)(x_n - x_{n+1}) \longrightarrow f'(\ell) \cdot 0 = 0.$$

Por continuidad de f, también $f(x_n) \to f(\ell)$, luego $f(\ell) = 0$. Aquí no necesitamos que $f'(\ell) \neq 0$; sí hemos usado la continuidad de f' en el límite y que este pertenezca al dominio. El

resultado no demuestra de antemano que la órbita converja ni identifica cuál de las raíces será su límite.

Una raíz múltiple puede reducir la velocidad de convergencia. Los resultados de §12.5 exigían una raíz simple o una cota que, en particular, implicaba $f'(\alpha) \neq 0$. No disponer de esa hipótesis no demuestra divergencia. El siguiente ejemplo permite ver exactamente qué cambia.

EJEMPLO 12.17. **Una raíz doble con convergencia lineal.** Sea

$$v(t) = (t-2)^2, \qquad I = \mathbb{R}, \qquad \alpha = 2.$$

Para $x \neq 2$, $v'(x) = 2(x-2) \neq 0$,

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

Si $x_0 \neq 2$, la inducción da

$$\boxed{x_n = 2 + 2^{-n}(x_0 - 2)}, \qquad \boxed{E_n = 2^{-n}|x_0 - 2| > 0}.$$

Ningún término llega exactamente a 2, así que todos los pasos están definidos. El límite geométrico prueba $x_n \to 2$, y $\frac{E_{n+1}}{E_n} = \frac{1}{2}$. El error se reduce siempre por el mismo factor; no se eleva efectivamente al cuadrado. En efecto, $\frac{E_{n+1}}{E_n^2} = \frac{1}{2E_n} \longrightarrow +\infty$. Por tanto no existe una constante $C > 0$ que haga válida eventualmente $E_{n+1} \leq CE_n^2$. La convergencia no es al menos cuadrática en el sentido de Definición 12.3.

Además, $|v(x_n)| = E_n^2$.

Un residuo a lo sumo 10^{-12} garantiza aquí un error a lo sumo 10^{-6}, pero no permite concluir que sea a lo sumo 10^{-12}. Para una tolerancia $\varepsilon > 0$, el certificado exacto de este ejemplo es $|v(x)| \leq \varepsilon^2$. No procede aplicar el certificado residual con una cota inferior positiva de $|v'|$ alrededor de 2, porque $v'(2) = 0$.

Diremos que una sucesión convergente tiene *convergencia lineal con factor asintótico* λ cuando sus errores son eventualmente positivos y

$$\frac{E_{n+1}}{E_n} \longrightarrow \lambda, \qquad 0 < \lambda < 1.$$

La palabra *lineal* se refiere a esta relación entre errores, no a que la función original sea afín. En el ejemplo, la función es cuadrática y la convergencia es lineal; para una función afín no constante habíamos obtenido terminación en un paso.

Lo que explica una factorización múltiple. No necesitamos desarrollar aquí una teoría general de multiplicidades. Examinaremos funciones para las que se ha comprobado una factorización local

$$f(t) = (t-\alpha)^p a(t), \qquad p \geq 2, \qquad a(\alpha) \neq 0.$$

El entero p cuenta las copias del factor $t - \alpha$ en esta factorización; no lo deducimos únicamente de observar $f'(\alpha) = 0$. La regularidad requerida del factor restante aparecerá expresamente en el resultado.

Proposición 12.4. **Proposición (efecto de un factor múltiple).** Sea I un intervalo abierto, $\alpha \in I$, $p \geq 2$ un entero y $a \in C^1(I)$ con $a(\alpha) \neq 0$. Definamos

$$f(t) = (t-\alpha)^p a(t).$$

Existe $r > 0$, con $[\alpha - r, \alpha + r] \subseteq I$, tal que desde cualquier x_0 con $0 < |x_0 - \alpha| \leq r$ la órbita de Newton está definida para todo n, permanece en ese entorno sin llegar exactamente a α y converge a α. Además,

$$\boxed{\lim_{n\to\infty} \frac{|x_{n+1}-\alpha|}{|x_n-\alpha|} = \frac{p-1}{p}.}$$

En particular, esta convergencia es lineal y no es al menos cuadrática. Si $x_0 = \alpha$, ya se dispone de la raíz, pero $N_f(\alpha)$ no está definido por el cociente de Newton.

Demostración. La regla del producto da

$$f'(x) = (x-\alpha)^{p-1}\big(pa(x) + (x-\alpha)a'(x)\big).$$

Escribamos $B(x) = pa(x) + (x-\alpha)a'(x)$.

La función B es continua y $B(\alpha) = pa(\alpha) \neq 0$. Por continuidad, en un entorno de α no se anulan ni a ni B. Para $x \neq \alpha$ en ese entorno, $f'(x) \neq 0$ y la cancelación de potencias no nulas proporciona

$$\begin{aligned} N_f(x) - \alpha &= (x-\alpha) - \frac{(x-\alpha)a(x)}{B(x)} \\ &= (x-\alpha)\frac{(p-1)a(x) + (x-\alpha)a'(x)}{pa(x) + (x-\alpha)a'(x)}. \end{aligned}$$

Denotemos por $A(x)$ el último cociente. Está definido y es continuo también en α, con $A(\alpha) = \lambda := \frac{p-1}{p} \in (0,1)$. Esto prolonga continuamente el *factor del error*, no redefine el operador de Newton en la raíz. Fijemos $q = (1+\lambda)/2 < 1$. La continuidad de A permite elegir un radio suficientemente pequeño para que $\frac{\lambda}{2} < A(x) < q$ en un intervalo cerrado $[\alpha - r, \alpha + r] \subseteq I$ en el que sigan sin anularse a y B. Se obtiene el cerrado tomando un radio menor que los radios abiertos proporcionados por continuidad. Si $0 < |x-\alpha| \leq r$, entonces el paso existe y

$$0 < |N_f(x) - \alpha| = A(x)|x-\alpha| < q|x-\alpha| \leq qr < r.$$

El nuevo punto no es la raíz, permanece dentro del intervalo y permite repetir el argumento. La inducción construye toda la órbita y da $0 < E_n \leq q^n E_0$. Por Corolario 3.2, $E_n \to 0$. Como A es continua,

$$\frac{E_{n+1}}{E_n} = A(x_n) \longrightarrow A(\alpha) = \lambda.$$

La cota inferior $A(x_n) > \lambda/2$ implica $\frac{E_{n+1}}{E_n^2} > \frac{\lambda}{2E_n} \longrightarrow +\infty$, lo que excluye cualquier cota cuadrática eventual con constante finita. Finalmente, la fórmula de f' da $f'(\alpha) = 0$ porque $p \geq 2$; ese punto se trata reconociendo la raíz, no formando el cociente. □

La prueba muestra el mecanismo sin invocar una expansión infinita: el error nuevo conserva una parte proporcional al error anterior. La continuidad de a' se utilizó para controlar los cocientes cerca de α; el factor no nulo $a(\alpha)$ mantuvo separado de cero el denominador después de cancelar.

No hemos probado que toda raíz con derivada nula admita esta factorización, ni trasladado el resultado a funciones que no cumplan sus hipótesis. En el caso particular $a \equiv 1$, la fórmula es exacta:

$$N_f(x) - \alpha = \left(1 - \frac{1}{p}\right)(x - \alpha).$$

Cuando la factorización se conoce, resolver una ecuación equivalente más sencilla puede ser preferible. Por ejemplo, $(t-2)^2 = 0$ equivale a $t - 2 = 0$; Newton aplicado a la segunda llega a 2 en un paso. Son ecuaciones con la misma raíz pero con operadores iterativos distintos. No se ha extendido por cancelación el dominio del primer operador en 2.

Diagnosticar antes de reparar. Los ejemplos permiten organizar una revisión concreta. En cada fila importa distinguir un hecho demostrado de una simple falta de garantía.

Situación comprobada	Diagnóstico y siguiente comprobación
$f(x) \neq 0$ y $f'(x) = 0$	El paso no existe. Debe elegirse otro punto admisible; no se informa una raíz.
El punto nuevo sale del dominio	El paso calculado no admite continuación para esa función. Hay que recuperar una región válida.
Dos valores distintos se repiten exactamente	Se ha probado un ciclo, no estabilización en una raíz. Más repeticiones del mismo ciclo no lo reparan.
El residuo disminuye sin raíz existente y sin cota de pendiente	Falta el puente del residuo al error. Se vuelve a existencia, aislamiento y cotas.
La órbita converge a otra raíz	Debe revisarse qué raíz se aisló y si la elección inicial conserva esa información.
Hay un factor múltiple verificado	Puede haber convergencia lineal. Se revisan el modelo del error y la tolerancia, no se exige sin prueba orden dos.
No se cumple una condición suficiente	Solo falta esa garantía. No se concluye divergencia ni error excesivo sin otro argumento.

Un encierro por signos que ya se haya demostrado puede conservarse como control independiente del candidato de Newton. Si este abandona la región permitida, el refinamiento por bisección de Teorema 5.3 sigue ofreciendo intervalos certificados bajo sus propias hipótesis. No afirmamos aquí un teorema nuevo para cualquier combinación de procedimientos: cada sustitución debe conservar el encierro o demostrar la nueva condición que utilizará.

Los errores de redondeo exigen otra distinción. Un ciclo exacto se prueba mediante igualdades, como en el cúbico; dos valores impresos iguales no prueban un ciclo exacto ni una raíz. Un paso pequeño sin la región de §12.5 tampoco autoriza Corolario 12.3, como mostró Ejemplo 12.12. Y una evaluación aproximada nula conserva la incertidumbre que se haya demostrado para ella. La comparación correcta sigue siendo con una tolerancia y un certificado del número efectivamente devuelto.

Laboratorio de contraejemplos y reparación. Una obstrucción no es inexistencia. Ante $f(t) = t^2 - 3$ y $x_0 = 0$, explica qué ha fallado. La respuesta es que no existe el cociente del paso, no que falten raíces reales. Elegir un punto distinto de cero permite calcular un paso, pero todavía no demuestra convergencia desde esa elección.

Refutar una promesa universal. «Una función polinómica con una raíz simple aislada y derivada no nula en el intervalo siempre admite convergencia de Newton desde cualquier punto del intervalo». Reconstruye el ciclo $1 \mapsto -1 \mapsto 1$ para $h(t) = t^3 - 5t$, comprueba $|h'| \geq 2$ y explica por qué el radio $1/2$, pero no el criterio usado con radio 1, queda certificado por §12.5.

No fabricar el límite que falta. Para $u(t) = 1/t$, sustituye $x_n = 2^n$ en la función y su derivada. El residuo tiende a cero, la pendiente también y los puntos no están acotados. Ninguno de esos cálculos produce una raíz positiva ni permite escribir $u(\lim x_n)$.

Separar destino y velocidad. La órbita negativa de $t^2 - 2$ converge a una raíz distinta de la positiva solicitada; la órbita de $(t-2)^2$ converge a la raíz correcta, pero linealmente. Son problemas diferentes: cambiar la tolerancia no convierte $-\sqrt{2}$ en $\sqrt{2}$, y observar una razón $1/2$ no refuta la convergencia a 2.

Leer correctamente la multiplicidad. Para $f(t) = (t-2)^4$, desde $x_0 \neq 2$ se tiene $E_n = (3/4)^n E_0$ y residuo E_n^4. Una tolerancia ε queda certificada exactamente por $|f(x)| \leq \varepsilon^4$. En cambio, el solo dato $f'(\alpha) = 0$ no autoriza asignar $p = 4$ ni aplicar Proposición 12.4 sin la factorización exigida.

Del procedimiento de raíces al problema de optimización. El estudio de Newton deja una disciplina que mantendremos al cambiar de problema: un candidato y un cálculo correcto no sustituyen las condiciones de la conclusión buscada. Hemos aprendido a separar existencia del paso, permanencia en una región, convergencia, identificación de la raíz, velocidad y certificado de un resultado concreto.

En §12.8 comenzaremos a construir funciones objetivo a partir de problemas verbales. Allí reaparecerán las mismas preguntas sobre el dominio y la existencia: antes de derivar habrá que identificar las variables, las restricciones y el conjunto de valores admisibles. Encontrar un punto crítico tampoco bastará, por sí solo, para haber resuelto un problema de optimización.

Antes de seguir. ¿Qué distingue una raíz con derivada nula de un punto no raíz con derivada nula? ¿Por qué invariancia no equivale a convergencia? ¿Qué refuta exactamente el ejemplo del recíproco? ¿Por qué un residuo pequeño no identifica cuál raíz buscamos? ¿Qué razón de errores caracteriza aquí la convergencia lineal? ¿Qué hipótesis permite pasar de un ejemplo de raíz doble al resultado para un factor múltiple?

12.8. De un problema verbal a una función objetivo

El cálculo puede ser correcto y el problema estar mal traducido. Hasta ahora buscábamos una raíz y exigíamos que el procedimiento no perdiera de vista su dominio ni su certificado. Al pasar a optimización aparece una dificultad anterior al cálculo: el problema puede no entregarnos todavía una función. Habla de dimensiones, material disponible, distancias o capacidades, y nos pide escoger la mejor configuración. Tenemos que decidir qué significan matemáticamente esas palabras antes de resolver una ecuación de derivadas.

Una respuesta puede contener una derivada impecable y, sin embargo, resolver otro problema. Ocurre, por ejemplo, si se admite una longitud negativa, si se confunde «a lo sumo» con «exactamente», si se incorpora una configuración degenerada que el enunciado excluye o si se trata como continua una cantidad que solo puede tomar valores enteros. La primera pregunta no será, pues, «¿qué derivo?», sino:

¿Qué elecciones son admisibles y qué cantidad estamos comparando?

El capítulo 2 ya enseñó que una fórmula no determina por sí sola una función. Aquí utilizaremos esa observación como una herramienta de modelización. Una misma expresión, restringida a dos conjuntos diferentes, puede tener diferentes soluciones óptimas o no tenerlas en uno de ellos. La teoría de extremos de los capítulos 5, 9 y 10 sigue vigente; no la sustituiremos por una receta verbal.

Variables, datos y unidades: tres papeles diferentes. Una *variable de decisión* representa algo que podemos elegir. Un *dato* o *parámetro* queda fijado mientras hacemos esa elección. Si un problema prescribe una longitud total $L > 0$, podemos estudiar varios valores de L para comparar problemas, pero dentro de cada uno de ellos L no cambia al variar la decisión x.

También debemos fijar las unidades. Si x e y son los valores numéricos de dos longitudes medidas en metros, $2x + y$ representa un número de metros y xy un número de metros cuadrados. La igualdad entre $2x + y$ y un presupuesto de longitud tiene sentido; sumarle directamente un área no lo tiene. Esta comprobación detecta errores de traducción antes de cualquier derivada.

Trabajaremos con los valores numéricos en las unidades declaradas. Así, una constante como 34 en una restricción de longitudes no es un número sin interpretación: significa 34 metros cuando las variables se han medido en metros. Cambiar las unidades exige cambiar coherentemente los datos y la interpretación de la función objetivo.

La pregunta «maximizar» o «minimizar» también requiere un complemento. Un diseño puede favorecer el área y perjudicar el perímetro; minimizar un costo total no equivale en general a minimizar un costo por unidad. Si se quieren comparar varias magnitudes sin especificar cómo resolver sus conflictos, todavía no se ha fijado un único problema de optimización del tipo que estudiaremos aquí.

El objeto matemático que debemos obtener.

Definición 12.4. **Conjunto factible y función objetivo en una variable.** Un modelo de optimización en una variable real especifica un conjunto $D \subseteq \mathbb{R}$, una función

$$F : D \longrightarrow \mathbb{R}$$

y el propósito de maximizar o minimizar sus valores. El conjunto D es el *conjunto factible*: recoge todas las restricciones impuestas a la variable. Sus elementos son las decisiones admisibles. La función F es la *función objetivo*: asigna a cada decisión la cantidad que se compara. Cuando $D \neq \varnothing$, una decisión óptima es un punto en el que F alcanza el extremo absoluto solicitado, en el sentido de Definición 9.1. El punto óptimo x_* y el valor óptimo $F(x_*)$ son objetos distintos. Puede haber varios puntos óptimos con el mismo valor. Si $D = \varnothing$, las restricciones son incompatibles y no hay decisión factible. No se asigna por ello un valor óptimo igual a cero.

En un máximo buscamos un punto $x_* \in D$ que satisfaga

$$F(x) \leq F(x_*) \qquad \text{para todo } x \in D.$$

No buscamos simplemente una solución de $F'(x) = 0$. Esa ecuación solo puede participar después de que sepamos qué función y qué dominio tenemos, y solo produce una condición necesaria en los puntos interiores diferenciables pertinentes. La comparación con todos los puntos del dominio continúa siendo parte de la conclusión. El dominio algebraico de una expresión y el conjunto factible cumplen funciones diferentes. Un polinomio puede evaluarse en cualquier real;

eso no convierte a cualquier real en una longitud admisible. A la inversa, un borde donde no existe una derivada bilateral puede ser una decisión perfectamente permitida. Las restricciones se leen en el problema, no en la comodidad de la fórmula.

Reducir a una variable sin perder configuraciones. Un enunciado puede mencionar varias magnitudes sin exigir cálculo diferencial de varias variables. Si una restricción permite expresar todas las demás en función de una sola decisión, podemos reducirlo a una función real de una variable. Pero el despeje debe conservar exactamente lo que se pretende comparar. Para examinar esa conservación usaremos únicamente el lenguaje de conjuntos, imágenes y composición del capítulo 2. Sea S un conjunto de configuraciones admisibles. Sus elementos pueden ser, por ejemplo, pares de longitudes; registrar un par no introduce todavía ninguna derivada en dos variables. Una función $G : S \to \mathbb{R}$ registra el valor que se compara en cada configuración. Maximizar G en S significa encontrar $s_* \in S$ con $G(s) \leq G(s_*)$ para toda $s \in S$; minimizar invierte esta comparación.

Proposición 12.5. **Proposición (reducción fiel mediante una parametrización).** Sean $S \neq \varnothing$ un conjunto, $G : S \to \mathbb{R}$ una función, $D \subseteq \mathbb{R}$ no vacío y $\varphi : D \to S$ una aplicación sobreyectiva. Definamos $F = G \circ \varphi$. Entonces

$$F(D) = G(S).$$

Además, para cualquier $x_* \in D$, el punto x_* maximiza F en D si y solo si la configuración $\varphi(x_*)$ maximiza G en S. La afirmación análoga vale para mínimos. En particular, un problema tiene un óptimo del tipo solicitado si y solo si lo tiene el otro, y sus valores óptimos coinciden.

Demostración. Para cada $x \in D$, el valor $F(x) = G(\varphi(x))$ pertenece a $G(S)$, porque $\varphi(x) \in S$. Por tanto, $F(D) \subseteq G(S)$. Recíprocamente, dada cualquier configuración $s \in S$, la sobreyectividad proporciona $x \in D$ con $\varphi(x) = s$. Así, $G(s) = F(x) \in F(D)$, lo que prueba la otra inclusión.

Si x_* maximiza F y $s \in S$, escogemos una preimagen x de s. Entonces

$$G(s) = F(x) \leq F(x_*) = G(\varphi(x_*)).$$

Como s era arbitraria, $\varphi(x_*)$ maximiza G. A la inversa, si esa configuración maximiza G, para todo $x \in D$ se tiene

$$F(x) = G(\varphi(x)) \leq G(\varphi(x_*)) = F(x_*).$$

Las mismas dos comparaciones, con las desigualdades invertidas, prueban la afirmación para mínimos. La igualdad de las imágenes y la sobreyectividad permiten trasladar también la existencia de un punto óptimo de un problema al otro. □

Cada hipótesis tiene un trabajo concreto. Que φ tome valores en S impide introducir configuraciones inadmisibles; su sobreyectividad impide omitir configuraciones permitidas; la identidad $F = G \circ \varphi$ garantiza que se está midiendo la misma cantidad. No se requiere inyectividad para conservar valores y extremos globales. Si una misma configuración tiene varios parámetros, puede tener varias representaciones óptimas. Para trasladar unicidad de la configuración a unicidad del parámetro hace falta, además, una representación inyectiva.

El resultado no afirma nada sobre extremos locales ni sobre derivadas de φ. Es una equivalencia de comparaciones globales. Si el despeje solo describe una parte de S, debe justificarse por otra razón que esa parte basta para el objetivo buscado. El ejemplo siguiente mostrará una reducción de ese segundo tipo.

Un modelo completo: el material disponible no es una igualdad automática.

EJEMPLO 12.18. **Un recinto rectangular con tres lados por cercar.** Consideremos el siguiente modelo idealizado. Una pared recta sirve de uno de los lados de un recinto rectangular y no limita la longitud de ese lado. Disponemos de a lo sumo 34 metros de cerco para los otros tres lados. La profundidad del recinto debe ser al menos 3 metros y su frente al menos 8 metros. No hay otras restricciones ni pérdidas de material. Se pide maximizar el área.

Llamemos x a la profundidad e y al frente, ambos medidos en metros. Los dos lados perpendiculares a la pared tienen longitud x y el lado que la enfrenta tiene longitud y. Por tanto, las configuraciones factibles son

$$S = \{(x, y) : x \geq 3,\ y \geq 8,\ 2x + y \leq 34\},$$

y el número de metros cuadrados de área es $G(x, y) = xy$. El conjunto no es vacío: por ejemplo, $(3, 8)$ es factible. Estas verificaciones preceden a cualquier intento de optimizar. La frase «a lo sumo» ha producido una desigualdad. No tenemos derecho a reemplazarla sin explicación por $2x + y = 34$. Podemos, sin embargo, demostrar que cualquier configuración que deje material sobrante no maximiza el área. En efecto, fijada una configuración factible, pongamos $y^* = 34 - 2x$. De $2x + y \leq 34$ se sigue que $y^* \geq y \geq 8$. La configuración (x, y^*) también es factible y

$$G(x, y^*) - G(x, y) = x(34 - 2x - y) \geq 0.$$

La desigualdad es estricta si queda material sobrante, porque $x \geq 3 > 0$. La pared no impone una cota adicional al frente, de modo que aumentar y no viola ninguna condición del modelo. Si el problema hubiera limitado también el frente, este argumento tendría que revisarse. Hemos probado dos cosas: cada configuración es superada o igualada por una que usa todo el material, y ningún máximo puede dejar material sobrante. Solo ahora reducimos la búsqueda del máximo a

$$S_= = \{(x, y) : x \geq 3,\ y \geq 8,\ 2x + y = 34\}.$$

En esta parte del conjunto, $y = 34 - 2x$. Las restricciones se traducen exactamente en

$$x \geq 3, \qquad 34 - 2x \geq 8,$$

es decir, $D = [3, 13]$.

Todo $x \in D$ produce una configuración permitida $(x, 34 - 2x)$ de $S_=$, y toda configuración de $S_=$ se obtiene así. La parametrización es incluso biyectiva. La función reducida es

$$A : [3, 13] \longrightarrow \mathbb{R}, \qquad A(x) = x(34 - 2x) = 34x - 2x^2.$$

Por Proposición 12.5, maximizar A equivale a maximizar el área dentro de $S_=$. El argumento de mejora anterior extiende esa equivalencia al máximo en todo S. La parametrización no recorre todo S: lo que justifica descartar el resto es la desigualdad de mejora, no una supuesta sobreyectividad sobre el conjunto original.

Existencia antes de localizar. El intervalo $[3, 13]$ es no vacío, cerrado y acotado. La función A es polinómica, luego continua. Por Teorema 5.5 alcanza un máximo en ese intervalo. La reducción ya demostrada garantiza, por tanto, una configuración de área máxima en el problema original. Hasta aquí no conocemos su posición y no hemos derivado.

En este caso podemos localizarla sin derivadas, completando un cuadrado:

$$A(x) = \frac{289}{2} - 2\left(x - \frac{17}{2}\right)^2 \leq \frac{289}{2}.$$

Hay igualdad exactamente en $x_* = 17/2$, que pertenece a $[3, 13]$. La otra dimensión es $y_* = 34 - 2x_* = 17$. La profundidad es 8.5 metros, el frente es 17 metros y el área máxima es 144.5 metros cuadrados. La comprobación de factibilidad incluye

$$\frac{17}{2} \geq 3, \qquad 17 \geq 8, \qquad 2 \cdot \frac{17}{2} + 17 = 34.$$

La comparación global y su caso de igualdad certifican el resultado; la continuidad había garantizado previamente su existencia. Ninguno de los dos argumentos se reduce a observar el aspecto de una gráfica.

La misma reducción no sirve para minimizar. En todo S se tiene $xy \geq 3 \cdot 8 = 24$, y la igualdad se alcanza en $(3, 8)$. Si impusiéramos gastar todo el material también para buscar el mínimo, eliminaríamos esta configuración. De hecho, para $x \in [3, 13]$, $A(x) - 84 = 2(x-3)(14-x) \geq 0$, con igualdad solo en $x = 3$. El mínimo en $S_=$ sería 84, no 24. La reducción que preserva el máximo no ha preservado el mínimo. Antes de eliminar decisiones hay que saber en qué sentido se está optimizando.

El ejemplo permite volver de los símbolos al relato. El término $2x$ cuenta dos lados; la condición $x \leq 13$ proviene de reservar al menos 8 metros para el frente; y la saturación del material utiliza que aumentar el frente mejora estrictamente el área. Son tres justificaciones distintas. Una equivocación en cualquiera de ellas cambiaría el problema aun si el desarrollo algebraico posterior fuese correcto. Las unidades también pueden auditarse al final. Si usamos centímetros, $X = 100x$ e $Y = 100y$, la restricción de material es $2X + Y \leq 3400$. La función reducida es $A_{\mathrm{cm}}(X) = X(3400 - 2X)$ y $A_{\mathrm{cm}}(100x) = 10000A(x)$. El factor positivo conserva qué recinto maximiza el área, pero cambia su valor numérico y sus unidades. Devolver 144.5 sin indicar metros cuadrados omitiría parte de la respuesta.

Un dominio abierto no autoriza a añadir sus bordes.

Ejemplo 12.19. **Una caja sin tapa y dos configuraciones que no son cajas.** De una lámina rectangular de 18 por 12 centímetros se cortan, en las cuatro esquinas, cuadrados del mismo lado x, y se pliegan las franjas restantes para formar una caja sin tapa. Supondremos espesor despreciable y ausencia de pestañas o pérdidas adicionales. Buscamos el mayor volumen entre las cajas con altura y ambos lados de la base estrictamente positivos.

El corte determina las dimensiones

$$\text{altura} = x, \qquad \text{largo} = 18 - 2x, \qquad \text{ancho} = 12 - 2x.$$

Cada dimensión de la base pierde dos segmentos de longitud x, uno en cada extremo. Las tres condiciones de positividad son

$$x > 0, \qquad x < 9, \qquad x < 6.$$

Por tanto, el conjunto factible exacto y la función objetivo son

$$D = (0, 6), \qquad V : D \longrightarrow \mathbb{R}, \qquad V(x) = x(18 - 2x)(12 - 2x).$$

La unidad de $V(x)$ es el centímetro cúbico. El dominio no es $\mathbb{R}$, aunque la expresión sea polinómica; tampoco es $[0,6]$. En $x=0$ falta altura y en $x=6$ uno de los lados de la base es cero. Ambos valores incumplen el modelo.

¿Existe un máximo en el dominio correcto? No podemos aplicar directamente Teorema 5.5 a $(0,6)$. Introduzcamos, como herramienta auxiliar y no como cambio de las decisiones permitidas, la extensión polinómica

$$\overline{V} : [0,6] \longrightarrow \mathbb{R}, \qquad \overline{V}(x) = x(18-2x)(12-2x).$$

Es continua en un intervalo cerrado y acotado, así que alcanza un máximo. Además,

$$\overline{V}(0) = \overline{V}(6) = 0, \qquad \overline{V}(3) = 3 \cdot 12 \cdot 6 = 216 > 0.$$

El máximo de la extensión es al menos 216, luego no puede alcanzarse en ninguno de los dos bordes. Algún punto de máximo pertenece a $(0,6)$. Allí V y $\overline{V}$ coinciden: ese punto es una caja admisible y maximiza su volumen también entre todas las decisiones originales. La prueba ha utilizado los bordes para certificar que el máximo está fuera de ellos. No los ha convertido en cajas permitidas. Hemos probado existencia, no unicidad ni las dimensiones del óptimo. La localización mediante los candidatos del capítulo 10 pertenece al análisis posterior.

Para el mínimo la conclusión es diferente. Toda caja admisible tiene $V(x) > 0$. Sin embargo, si $x_n = 1/n$ para $n \geq 1$, entonces $x_n \in (0,6)$ y

$$0 < V(x_n) = \frac{1}{n}\left(18 - \frac{2}{n}\right)\left(12 - \frac{2}{n}\right) \leq \frac{216}{n} \longrightarrow 0.$$

El ínfimo de los volúmenes es cero y no se alcanza. Si existiera una caja de volumen mínimo $v_* > 0$, los términos suficientemente tardíos de esta sucesión tendrían volumen menor que v_*, contradicción. En cambio, la extensión sí tiene mínimo cero en sus bordes. Confundir el dominio abierto con su cierre fabricaría una solución que el problema original prohíbe.

Este argumento no dice que cualquier problema sobre un dominio abierto deba cerrarse ni que cualquier extensión preserve sus óptimos. En la caja funcionó para el máximo porque la extensión existe, es continua y un valor interior supera a todos los valores añadidos. La sección siguiente estudiará cómo organizar estas decisiones de existencia y comparación en diferentes dominios.

Una restricción discreta cambia el problema.

Ejemplo 12.20. **El recinto cuando sus dimensiones deben ser números enteros de metros.** Modifiquemos el modelo de Ejemplo 12.18: ahora tanto la profundidad como el frente deben ser enteros. El material y las longitudes mínimas no cambian.

El argumento de saturación sigue siendo válido para el máximo: si x es entero, $y^* = 34 - 2x$ también lo es, y cualquier configuración con material sobrante sigue siendo mejorada por otra factible. La función objetivo conserva su expresión, pero el conjunto factible reducido pasa a ser $D_{\text{ent}} = \{3, 4, \ldots, 13\}$. La profundidad $17/2$ del problema continuo no pertenece a ese conjunto. No basta con declarar óptimo ese valor ni con redondearlo y trasladar sin prueba la conclusión anterior. La identidad del cuadrado proporciona una comparación exacta. Si x es entero, $2x - 17$ es un entero impar no nulo, de modo que

$$\left| x - \frac{17}{2} \right| \geq \frac{1}{2}.$$

Por tanto,

$$A(x) = \frac{289}{2} - 2\left(x - \frac{17}{2}\right)^2 \leq \frac{289}{2} - \frac{1}{2} = 144.$$

La igualdad ocurre exactamente en $x = 8$ y $x = 9$, ambos factibles. Sus frentes son 18 y 16 metros, respectivamente. Hay dos configuraciones óptimas y un solo valor máximo, 144 metros cuadrados. Aquí comparar los enteros próximos está justificado por la identidad global del objetivo. No hemos demostrado que redondear una solución continua sea un método general para restricciones enteras. Tampoco hace falta una derivada relativa en los puntos aislados de D_{ent}: la comparación finita y la desigualdad bastan.

Transformar el objetivo exige conservar el orden. A veces la reducción conserva las decisiones pero modifica la cantidad que calculamos. Es legítimo si se controla el efecto sobre las comparaciones. Si ψ es estrictamente creciente en un conjunto que contiene todos los valores de F, entonces, para decisiones x, y,

$$F(x) \leq F(y) \quad \Longleftrightarrow \quad \psi(F(x)) \leq \psi(F(y)).$$

La dirección directa viene de la monotonía; para la recíproca, si $F(x) > F(y)$, el crecimiento estricto produciría la desigualdad contraria entre sus imágenes. Por ello se conservan los puntos de máximo y mínimo. Sus valores numéricos cambian y deben reinterpretarse al responder. Si ψ es estrictamente decreciente, las comparaciones se invierten y máximos y mínimos intercambian sus papeles.

Un caso útil es sustituir una distancia $d(x) \geq 0$ por su cuadrado. Para $0 \leq u < v$, $v^2 - u^2 = (v - u)(v + u) > 0$, así que elevar al cuadrado conserva el orden de las distancias. La condición de no negatividad no es decorativa. La función $F(x) = x$ en $[-2, 1]$ se minimiza en -2, mientras que $F(x)^2 = x^2$ se minimiza en 0. Sin controlar el signo, cuadrar puede cambiar el óptimo. Análogamente, eliminar un factor positivo constante del objetivo conserva sus comparaciones; eliminar un factor negativo las invierte. El término «constante» importa: dividir por una cantidad que depende de la decisión puede cambiar el orden. No se reemplaza sin justificación un objetivo total por un cociente «por unidad».

Protocolo de traducción y control. Podemos condensar la preparación de un problema en preguntas que deben recibir respuestas verificables:

Pregunta	Objeto o comprobación que debe quedar escrito
¿Qué se puede elegir y qué está fijado?	Variables de decisión, parámetros y unidades.
¿Qué elecciones permite exactamente el enunciado?	Restricciones, conjunto factible y al menos una configuración admisible, o prueba de incompatibilidad.
¿Qué cantidad se compara y en qué sentido?	Función objetivo y elección entre máximo y mínimo.
¿Qué justifica pasar a una variable?	Recuperación de las configuraciones permitidas o prueba de que las descartadas no mejoran el objetivo solicitado.
¿Qué dominio queda después del despeje?	Intervalos y sus bordes, exclusiones, signos, denominadores y posibles restricciones discretas.

Pregunta	Objeto o comprobación que debe quedar escrito
¿Está garantizado que se alcance el valor buscado?	Teorema de existencia con hipótesis verificadas o argumento específico; no solo una lista de candidatos.
¿Cómo se devuelve la solución al relato?	Dimensiones u otras decisiones, valor óptimo, unidades y comprobación de todas las restricciones originales.

Este protocolo no obliga a utilizar derivadas cuando una desigualdad ya resuelve el problema. Su función es fijar aquello a lo que se aplicarían las herramientas de cálculo y evitar que la manipulación de una fórmula cambie silenciosamente el conjunto que estamos comparando. La existencia se examina por separado de la localización. En un intervalo cerrado y acotado, continuidad permite invocar Teorema 5.5. Si alguna hipótesis falta, no se concluye automáticamente inexistencia: habrá que probarla o refutarla por otro argumento, como hicimos con el dominio abierto de la caja. Solo después se organiza la búsqueda; Proposición 10.4 recuerda que los bordes y los puntos interiores no diferenciables no desaparecen por el hecho de que una derivada sea fácil de calcular.

Laboratorio de traducción y reparación. No imponer una igualdad por costumbre. En el recinto, identifica el paso que autoriza gastar todo el cerco. No es la palabra «disponible»: es la comparación $x(34 - 2x - y) \geq 0$ y su estrictez cuando queda material. Explica por qué el mismo paso no preserva el mínimo y verifica la configuración $(3, 8)$ que lo demuestra.

Detectar restricciones incompatibles. Si se exigiera profundidad al menos 14 metros y frente al menos 8 con el mismo presupuesto de 34, toda configuración requeriría $2x + y \geq 36 > 34$. El conjunto factible sería vacío. Ningún cálculo de derivadas puede convertirlo en no vacío, y un área máxima igual a cero no sería la respuesta correcta.

Revisar el despeje y su regreso al problema. En el recinto saturado se puede elegir el frente y como variable y escribir $x = (34 - y)/2$. Las restricciones dan $8 \leq y \leq 28$ y el objetivo $B(y) = y(34 - y)/2$. El valor $y = 17$ recupera $x = 17/2$. Cambia la parametrización, no el recinto óptimo ni su área. Comprueba ambas traducciones de las restricciones, no solo la igualdad de las fórmulas al sustituir.

No atribuir factibilidad a una extensión. En la caja, $x = 0$ y $x = 6$ pertenecen al dominio de $\overline{V}$, no al de V. Para transportar un máximo desde la extensión hubo que excluir esos bordes con el valor interior 216. Para el mínimo no podemos hacerlo: el valor cero de la extensión no corresponde a ninguna caja admisible.

Separar valor y decisión óptimos. En la variante entera, las profundidades 8 y 9 producen el mismo máximo 144. Una respuesta que solo entregue el área omite las configuraciones; una que solo entregue una profundidad omite la otra posibilidad. La unicidad del valor máximo no implica unicidad de la decisión que lo alcanza.

Lo que está listo para el análisis global. La modelización termina cuando hemos construido una función con dominio preciso, justificado la reducción y aclarado la cuestión de existencia que corresponda. El cálculo posterior ya no trabaja con «una fórmula que parece representar el problema», sino con un objeto cuyos puntos tienen una interpretación y cuyas restricciones siguen siendo visibles.

La solución matemática del modelo tampoco demuestra que sus idealizaciones describan perfectamente una situación material. Por ejemplo, el espesor de la lámina o una longitud

limitada de pared cambiarían las restricciones. En los ejemplos se indicaron esos supuestos: son parte del alcance de la conclusión, no efectos que las derivadas hayan eliminado.

En §12.9 sistematizaremos la búsqueda de óptimos globales en los dominios obtenidos: existencia en intervalos cerrados, candidatos interiores, bordes, puntos no diferenciables y comportamiento en dominios abiertos o no acotados. La teoría del capítulo 10 suministrará los criterios; la presente sección ha establecido cómo construir correctamente el problema al que se aplican.

Antes de seguir. ¿Por qué la fórmula y su dominio algebraico no bastan para especificar un problema de optimización? ¿Qué impide omitir configuraciones al parametrizar? ¿Cuándo se puede convertir «a lo sumo» en una igualdad? ¿Qué argumento transportó el máximo de la extensión de la caja al dominio original? ¿Por qué no transportó su mínimo? ¿Qué diferencia hay entre una configuración óptima, su parámetro y el valor que alcanza?

12.9. Optimización global en intervalos y dominios no compactos

Existencia, candidatos y comparación son preguntas distintas. La sección anterior terminó con un modelo ya traducido: conjunto factible preciso, función objetivo, unidades y una justificación de la reducción. Ahora aparece una segunda capa de trabajo. Aunque la función esté bien definida y podamos calcular su derivada, todavía debemos separar tres preguntas:

1. **¿existe** un máximo o un mínimo absoluto en el dominio considerado?;
2. **¿dónde podría estar** si existe?;
3. **¿cómo demostramos que un candidato es global**, es decir, mejor o peor que *todos* los demás puntos admisibles?

Estas preguntas se responden con herramientas distintas. El teorema de valores extremos Teorema 5.5 ofrece una garantía de existencia cuando una función continua está definida en un intervalo cerrado y acotado. La proposición de candidatos Proposición 10.4 indica dónde puede alcanzarse un extremo absoluto en ese contexto: bordes, puntos interiores con derivada nula y puntos interiores donde la derivada no existe como número real finito. Ninguno de los dos resultados sustituye al otro. Podemos resumir la lógica así:

$$\boxed{\text{existencia} \longrightarrow \text{candidatos} \longrightarrow \text{comparación global.}}$$

En un problema favorable, las tres etapas son cortas. En dominios abiertos o no acotados, en cambio, la primera puede requerir un argumento específico; además, los bordes que faltan y el comportamiento cuando $|x|$ crece pueden decidir si un supremo o un ínfimo se alcanza. Una condición suficiente que no podemos aplicar no se transforma por ello en una demostración de inexistencia. Un dominio no compacto puede tener extremos absolutos; un intervalo abierto puede tener uno de ellos; una función discontinua puede alcanzar máximo y mínimo. Lo que desaparece es la garantía automática de Weierstrass, no toda posibilidad de optimización.

Intervalos cerrados: el repertorio ya está completo. Sea $f : [a, b] \to \mathbb{R}$ continua. Por Teorema 5.5 existen mínimo y máximo absolutos. Si además podemos estudiar la derivada en (a, b), Proposición 10.4 reduce la localización a los candidatos pertinentes. El procedimiento no consiste en «resolver $f' = 0$». Hay que incluir:

- el extremo izquierdo a;
- el extremo derecho b;
- los puntos interiores donde $f'(x) = 0$;
- los puntos interiores donde la derivada no existe como número real finito.

Después se comparan los valores de la función en todos los candidatos que realmente pertenezcan al dominio. Si la lista es finita, la comparación es finita. Si hay infinitos candidatos —por ejemplo, una meseta— puede ser preferible usar monotonía o una desigualdad global antes que enumerarlos. La aparición de un punto no diferenciable no lo elimina. Para

$$F(x) = |x - 1|, \qquad x \in [-2, 3],$$

el mínimo absoluto se alcanza en $x = 1$, precisamente donde la derivada bilateral no existe. Los bordes dan $F(-2) = 3$ y $F(3) = 2$, mientras que $F(1) = 0$. La derivada sirve para localizar candidatos donde existe; no define por sí sola el conjunto de decisiones admisibles.

Localizar el máximo de la caja de §12.8.

EJEMPLO 12.21. **La caja de 18×12: existencia ya probada, localización ahora.** En Ejemplo 12.19 obtuvimos

$$V(x) = x(18 - 2x)(12 - 2x), \qquad 0 < x < 6,$$

y demostramos por extensión continua que el problema original posee al menos un máximo, aunque el dominio sea abierto. Quedó pendiente localizarlo. Expandimos solo para facilitar la derivación: $V(x) = 4x^3 - 60x^2 + 216x$.

Por tanto,

$$V'(x) = 12x^2 - 120x + 216 = 12(x^2 - 10x + 18).$$

Las raíces de la cuadrática son

$$5 - \sqrt{7}, \qquad 5 + \sqrt{7}.$$

Como $2 < \sqrt{7} < 3$,

$$2 < 5 - \sqrt{7} < 3 < 6 < 5 + \sqrt{7}.$$

Solo el primer cero pertenece al dominio. Además,

$$V'(x) = 12\big(x - (5 - \sqrt{7})\big)\big(x - (5 + \sqrt{7})\big).$$

De aquí se obtiene

$$V'(x) > 0 \quad \text{si } 0 < x < 5 - \sqrt{7},$$

y

$$V'(x) < 0 \quad \text{si } 5 - \sqrt{7} < x < 6.$$

Por el criterio de monotonía Teorema 10.1, V crece hasta $5 - \sqrt{7}$ y decrece después. Por consiguiente, el máximo cuya existencia ya estaba demostrada es único y se alcanza en

$$\boxed{x_* = 5 - \sqrt{7}.}$$

Las dimensiones de la caja son

$$\begin{aligned}\text{altura} &= 5 - \sqrt{7},\\ \text{largo} &= 18 - 2x_* = 8 + 2\sqrt{7},\\ \text{ancho} &= 12 - 2x_* = 2 + 2\sqrt{7}.\end{aligned}$$

El volumen máximo es

$$\boxed{V(x_*) = 80 + 56\sqrt{7}\ \text{cm}^3.}$$

El dominio sigue siendo $(0, 6)$. No hemos añadido 0 ni 6 a las decisiones permitidas; la extensión de §12.8 se utilizó para demostrar existencia y el signo de V' localizó después el punto óptimo dentro del dominio real del modelo.

Este ejemplo muestra por qué conviene conservar separadas las etapas. Si hubiéramos resuelto primero $V'(x) = 0$, habríamos obtenido un candidato interior correcto, pero todavía faltaría explicar por qué el problema abierto posee un máximo global. Esa existencia ya fue demostrada antes de derivar.

Un borde ausente puede impedir que el supremo sea un máximo. Los bordes que no pertenecen al dominio no son candidatos evaluables. Sin embargo, los valores de la función al aproximarse a ellos pueden controlar la existencia de un extremo.

Ejemplo 12.22. **Un mínimo alcanzado y un supremo no alcanzado en una semirrecta.** Consideremos

$$h : [0, \infty) \longrightarrow \mathbb{R}, \qquad h(x) = \frac{x}{1 + x}.$$

El denominador es positivo en todo el dominio y $h'(x) = \frac{1}{(1+x)^2} > 0$. Por Teorema 10.1, h es estrictamente creciente. El borde izquierdo sí pertenece al dominio, así que $h(0) = 0$ es el mínimo absoluto.

En cambio,

$$h(x) = 1 - \frac{1}{1 + x} < 1 \qquad (x \geq 0),$$

de modo que 1 no se alcanza. Además, dado cualquier $\varepsilon > 0$, basta elegir $x > 1/\varepsilon - 1$ para obtener

$$0 < 1 - h(x) = \frac{1}{1 + x} < \varepsilon.$$

Por tanto, $\sup h([0, \infty)) = 1$, pero no existe máximo absoluto. La ecuación $h'(x) = 0$ no tiene soluciones y, sin embargo, el problema sí posee un mínimo absoluto en el borde incluido. La ausencia de puntos críticos interiores tampoco demuestra que exista un máximo: el extremo derecho del dominio está «en el infinito» y solo produce un valor límite no alcanzado.

El ejemplo enseña a leer una semirrecta como un dominio con dos comportamientos distintos: un borde finito que puede pertenecer o no al conjunto y una cola no acotada. El análisis global debe controlar ambos.

No compacto no significa «sin extremos». Un dominio no acotado también puede contener los puntos que realizan máximo y mínimo absolutos. La compactitud proporciona una garantía universal; no es una condición necesaria para cada función particular.

Ejemplo 12.23. **Una función racional con máximo y mínimo absolutos en toda la recta.** Sea

$$g : \mathbb{R} \longrightarrow \mathbb{R}, \qquad g(x) = \frac{x}{1+x^4}.$$

El denominador es siempre positivo. La derivada es

$$g'(x) = \frac{1-3x^4}{(1+x^4)^2}.$$

Pongamos $a = 3^{-1/4} > 0$.

Entonces $3a^4 = 1$. Como el denominador de g' es positivo, el signo depende de $1 - 3x^4$:

x	$(-\infty, -a)$	$-a$	$(-a, a)$	a	(a, ∞)
$g'(x)$	$-$	0	$+$	0	$-$

Así, g decrece en $(-\infty, -a]$, crece en $[-a, a]$ y decrece en $[a, \infty)$. Para $x > 0$, $g(x) > 0$; para $x < 0$, $g(x) < 0$. En consecuencia, los puntos de la cola derecha no pueden competir con el valor negativo de $g(-a)$ para el mínimo, y los de la cola izquierda no pueden competir con el valor positivo de $g(a)$ para el máximo. La monotonía en los tres intervalos completa la comparación global.

Como $a^4 = 1/3$, $1 + a^4 = \frac{4}{3}$, y por tanto

$$g(a) = \frac{3a}{4} = \frac{3^{3/4}}{4}, \qquad g(-a) = -\frac{3^{3/4}}{4}.$$

Luego

$$\boxed{\max_{\mathbb{R}} g = \frac{3^{3/4}}{4} \text{ en } x = 3^{-1/4},}$$

mientras que

$$\boxed{\min_{\mathbb{R}} g = -\frac{3^{3/4}}{4} \text{ en } x = -3^{-1/4}.}$$

El dominio es no acotado, pero ambos extremos existen y se alcanzan. Lo que los demuestra aquí es el análisis global de signos y valores, no el teorema de valores extremos.

Qué hacer con bordes ausentes y con el infinito. Los tres tipos de dominio más frecuentes en los problemas de una variable pueden organizarse sin inventar un teorema nuevo.

Dominio	Garantía inicial	Información global que todavía puede ser necesaria
$[a, b]$	Continuidad + Teorema 5.5 garantizan mínimo y máximo.	Comparar bordes, críticos y no diferenciables mediante Proposición 10.4 o una desigualdad global.

Dominio	Garantía inicial	Información global que todavía puede ser necesaria
(a,b), $[a,b)$, $(a,b]$	No hay garantía automática por Weierstrass.	Estudiar candidatos interiores y el comportamiento al acercarse a cada borde excluido; comprobar qué valores se alcanzan.
Semirrecta o intervalo no acotado	No hay garantía automática por compacidad.	Estudiar candidatos, bordes finitos incluidos/excluidos y comportamiento cuando $x \to \pm\infty$ mediante límites, desigualdades o monotonía.

La palabra *borde* en esta tabla no autoriza a evaluar $f(a)$ cuando a no pertenece al dominio. Si el límite lateral existe, es información sobre valores próximos; puede coincidir con un supremo o ínfimo que no se alcance. Del mismo modo, «el valor en infinito» no es una evaluación de la función en un punto nuevo: resume el comportamiento de la cola. También debemos distinguir un límite finito de una conclusión de extremo. Si

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

puede ocurrir que L sea un supremo no alcanzado, como en Ejemplo 12.22; que algún punto interior produzca valores mayores que L y exista un máximo; o que la función oscile alrededor de L antes de estabilizarse. El límite es un dato de comparación, no una clasificación automática.

Un protocolo global que no confunde condiciones suficientes con respuestas. Después de la modelización de §12.8, una búsqueda rigurosa puede seguir este orden:

1. **Escribir el dominio exacto.** Marcar qué bordes pertenecen y cuáles están excluidos, y si el conjunto es no acotado.
2. **Preguntar por la existencia.** Si se cumplen las hipótesis de Teorema 5.5, usarlo. Si no, buscar un argumento particular: extensión controlada, desigualdad, monotonicidad, comportamiento en los bordes o en el infinito.
3. **Localizar candidatos admisibles.** Usar Proposición 10.4 cuando corresponda; en otros intervalos conservar la misma lógica local para puntos interiores, sin olvidar los bordes que sí pertenezcan al dominio.
4. **Estudiar los bordes ausentes y las colas.** Un límite o una desigualdad puede mostrar que un valor solo se aproxima y no se alcanza, o que ningún punto lejano puede superar a un candidato interior.
5. **Comparar globalmente.** La conclusión debe referirse a todos los puntos del conjunto factible, no solo a vecinos del candidato.
6. **Volver al problema original.** Traducir el parámetro a la configuración, las unidades y las restricciones iniciales.

Este orden no obliga a calcular una derivada. Si una desigualdad global identifica de inmediato el mejor valor, la derivada sería un procedimiento más largo, no una exigencia lógica. Tampoco obliga a decidir primero «máximo» o «mínimo» mediante un criterio de segunda derivada: la clasificación local es solo una parte posible de la prueba global.

Errores frecuentes al pasar de lo local a lo global. «Encontré $f'(c) = 0$, luego c es el máximo». La igualdad solo convierte a c en candidato interior bajo las hipótesis pertinentes. Puede ser mínimo, máximo local, punto estacionario sin extremo o un candidato superado por un borde.

«El dominio es abierto, luego no hay máximo ni mínimo». Falso. La caja de Ejemplo 12.21 tiene máximo absoluto en $(0, 6)$. El problema es que Weierstrass no lo garantiza automáticamente.

«El dominio es no acotado, luego no hay extremos». Falso. Ejemplo 12.23 posee máximo y mínimo absolutos en $\mathbb{R}$.

«El límite lateral vale L, así que L es un máximo». Un límite no crea un punto del dominio. En Ejemplo 12.22 el valor 1 es el supremo y no se alcanza.

«Si una función es continua, alcanza extremos». Falta especificar el conjunto. La continuidad sobre un compacto real proporciona la garantía de Teorema 5.5; en un dominio distinto debemos analizar qué reemplaza esa hipótesis.

Laboratorio de existencia, candidatos y bordes. Reparar una falsa aplicación de Weierstrass. Para $f(x) = x^2$ en $(-1, 1)$ no podemos citar Teorema 5.5 sobre el dominio dado. Sin embargo, $f(x) \geq 0 = f(0)$ prueba un mínimo absoluto en 0. No existe máximo: para cada $x \in (-1, 1)$, el número $y = (|x| + 1)/2$ pertenece a $(0, 1)$ y satisface $y > |x|$, luego $f(y) > f(x)$.

Localizar sin volver a demostrar existencia. En la caja de Ejemplo 12.21, identifica qué parte de la prueba pertenece a §12.8 y cuál a esta sección. La extensión a $[0, 6]$ con un testigo interior positivo demostró existencia. El signo de V' demuestra unicidad y localiza el máximo. Mezclar ambos pasos ocultaría qué hipótesis hace cada trabajo.

Separar supremo de máximo. Para Ejemplo 12.22, verifica directamente que $h(x) < 1$ para todo $x \geq 0$ y que para cualquier $c < 1$ existe un x con $h(x) > c$. El segundo paso puede obtenerse eligiendo $\varepsilon = 1 - c$ en la estimación ya demostrada. Concluye supremo 1, no máximo 1.

No olvidar un punto no diferenciable. En $F(x) = |x - 1|$ sobre $[-2, 3]$, la ecuación $F'(x) = 0$ no aporta el minimizador. La comparación de $F(-2)$, $F(1)$ y $F(3)$ muestra por qué Proposición 10.4 incorpora los puntos interiores donde la derivada no existe.

Auditar un dominio no acotado. En Ejemplo 12.23, explica por qué no basta con hallar los dos ceros de g'. La prueba global usa además el signo de g en cada semirrecta y la monotonía de los tres tramos para descartar valores mejores en las colas.

Lo que queda establecido. La optimización global en una variable no se reduce a una ecuación. En un intervalo cerrado y acotado, continuidad proporciona existencia y la teoría de C11 organiza los candidatos. En dominios abiertos o no acotados, los mismos candidatos interiores siguen importando, pero deben combinarse con un análisis de los bordes ausentes y de las colas. Un límite puede describir un valor frontera sin garantizar que se alcance; una función no compacta puede, pese a ello, poseer extremos absolutos. El principio del capítulo reaparece con otra forma:

un candidato local propone; la comparación con todo el dominio certifica.

En §12.10 volveremos al método de Newton, ahora aplicado a la ecuación $f'(x) = 0$. Obtener numéricamente un punto estacionario será solo una etapa: la clasificación mediante los criterios de C11 y la comparación global seguirán siendo necesarias para convertirlo en una conclusión de optimización.

Antes de seguir. ¿Qué resultado garantiza existencia en un intervalo cerrado? ¿Por qué esa garantía no se aplica directamente a $(0, 6)$? ¿Qué diferencia hay entre supremo y máximo en Ejemplo 12.22? ¿Por qué los dos ceros de la derivada de Ejemplo 12.23 sí terminan siendo extremos globales? ¿Dónde entran los puntos no diferenciables? ¿Qué información adicional debe estudiarse cuando el dominio no es acotado?

12.10. Newton sobre la derivada y laboratorio integrador

Buscar un punto estacionario es volver a resolver una ecuación. En las primeras siete secciones usamos Newton para buscar raíces de una función. Un problema de optimización introduce una ecuación distinta: si una función objetivo F es diferenciable y un extremo local se alcanza en un punto interior donde se cumplen las hipótesis de Fermat (Teorema 9.1), entonces ese punto debe satisfacer $F'(c) = 0$. Por tanto, cuando decidimos buscar **puntos estacionarios** mediante Newton, no estamos introduciendo un algoritmo nuevo. Aplicamos el mismo método a la función $G := F'$. La distinción importa. Newton puede ayudarnos a localizar una raíz de F'. La teoría de extremos de C11 decide después qué significa esa raíz para F, y §12.9 decide si el candidato participa en una conclusión global.

Supongamos, para fijar la fórmula, que F posee primera y segunda derivadas en un intervalo abierto I. En un punto $x \in I$ donde $F''(x) \neq 0$, el paso de Newton aplicado a $G = F'$ es

$$\boxed{x^+ = x - \frac{F'(x)}{F''(x)}.}$$

La fórmula se obtiene literalmente de Definición 12.2 sustituyendo f por F'. El numerador es el valor de la ecuación estacionaria y el denominador es su derivada. Hay que conservar dos casos separados.

- Si $F'(x) \neq 0$ y $F''(x) = 0$, el paso no está definido: la tangente a F' en ese punto es horizontal y no proporciona mediante esta fórmula un nuevo cero candidato.
- Si $F'(x) = 0$, ya hemos localizado un punto estacionario. Aunque también ocurra $F''(x) = 0$, no necesitamos ni debemos formar el cociente $0/0$. La tarea pasa a ser **clasificar** ese punto.

Así, el dominio del procedimiento de búsqueda y el conjunto de puntos estacionarios no son el mismo objeto.

Qué hereda el método y qué exige una derivada adicional. Todo lo que aprendimos sobre Newton puede reutilizarse, pero debe traducirse correctamente a la función $G = F'$. Para **calcular un solo paso** necesitamos evaluar F' y F'' y comprobar $F''(x) \neq 0$. Para aplicar la identidad de error de Teorema 12.1 a G, la función G debe poseer dos derivadas en la región pertinente. En términos de F, aparecen

$$G' = F'', \qquad G'' = F'''.$$

Por eso, nuestro criterio de convergencia local Proposición 12.2, aplicado sin modificaciones a la ecuación $F'(x) = 0$, requiere controlar una derivada más de la función objetivo: una cota inferior de $|F''|$ y una cota superior de $|F'''|$ en una región alrededor del punto estacionario buscado. Esta exigencia no debe confundirse con la clasificación posterior. El segundo criterio de la derivada Teorema 10.5 clasifica un punto estacionario cuando existe $F''(c)$ y su valor es positivo o negativo; no necesita por sí solo una tercera derivada. La tercera derivada entra aquí

únicamente porque estamos usando **nuestro teorema concreto de convergencia de Newton** sobre la ecuación $F' = 0$. Podemos resumir el traslado de hipótesis así:

Problema de raíces	Búsqueda de estacionarios de F
función cuya raíz buscamos: f	función cuya raíz buscamos: $G = F'$
paso $x - f/f'$	paso $x - F'/F''$
raíz simple: $f'(c) \neq 0$	raíz simple de F': $F''(c) \neq 0$
cota inferior de $\lvert f' \rvert$	cota inferior de $\lvert F'' \rvert$
cota superior de $\lvert f'' \rvert$	cota superior de $\lvert F''' \rvert$
convergencia a una raíz	convergencia a un punto estacionario

La última fila es deliberadamente asimétrica. Una raíz de F' solo dice $F'(c) = 0$. Todavía no dice «mínimo», «máximo» ni «óptimo global».

Un mínimo global cuya posición se aproxima y certifica.

Ejemplo 12.24. **Newton localiza un punto estacionario; la monotonía lo convierte en mínimo global.** Consideremos

$$F(x) = x^4 + x^2 + x, \qquad x \in \mathbb{R}.$$

La ecuación estacionaria es $g(x) := F'(x) = 4x^3 + 2x + 1 = 0$.

Su derivada es

$$g'(x) = F''(x) = 12x^2 + 2 > 0$$

para todo real. Por Teorema 10.1, g es estrictamente creciente. Además,

$$g\left(-\frac{2}{5}\right) = -\frac{7}{125} < 0, \qquad g\left(-\frac{3}{8}\right) = \frac{5}{128} > 0.$$

El IVT y la monotonía aíslan una única raíz

$$\alpha \in \left(-\frac{2}{5}, -\frac{3}{8}\right).$$

Esa raíz es el único punto estacionario de F. Apliquemos ahora Newton a g. Como $g'(x) = 12x^2 + 2$, el paso es $N_g(x) = x - \frac{4x^3+2x+1}{12x^2+2}$.

Desde $x_0 = 0$ obtenemos exactamente

$$x_1 = -\frac{1}{2}, \qquad x_2 = -\frac{2}{5}, \qquad x_3 = -\frac{27}{70}.$$

Los dos primeros pasos constituyen un prefijo que podemos comprobar directamente. Desde $x_2 = -2/5$ entraremos en una región certificada. Tomemos $r = 1/5$. El encierro anterior implica

$$\alpha - r > -\frac{3}{5}, \qquad \alpha + r < 0,$$

de modo que

$$B_r = [\alpha - r, \alpha + r] \subseteq \left[-\frac{3}{5}, 0\right].$$

En ese intervalo, $|g'(t)| = 12t^2 + 2 \geq 2$, y, como $g''(t) = F'''(t) = 24t$, tenemos $|g''(t)| \leq \frac{72}{5}$. Podemos escoger

$$m = 2, \qquad M = \frac{72}{5}, \qquad K = \frac{M}{2m} = \frac{18}{5}.$$

Entonces

$$q = Kr = \frac{18}{25} < 1.$$

Además,

$$\left| -\frac{2}{5} - \alpha \right| < \frac{1}{40} < r.$$

Todas las hipótesis de Proposición 12.2 se cumplen para Newton aplicado a g: desde x_2 la órbita está definida, permanece en B_r y converge a α al menos cuadráticamente. El tercer punto calculado puede certificarse sin conocer α. Como

$$g\left(-\frac{27}{70}\right) = -\frac{83}{85750}$$

y, en el intervalo aislante $[-2/5, -3/8]$, $g'(t) = 12t^2+2 \geq \frac{59}{16}$, la proposición residual Proposición 12.3 proporciona

$$\boxed{\left| -\frac{27}{70} - \alpha \right| \leq \frac{664}{2529625} < \frac{1}{3000}.}$$

Hasta aquí hemos localizado y certificado un punto estacionario. Falta clasificarlo. Como $F''(x) = 12x^2 + 2 > 0$ en toda la recta, Teorema 10.5 ya da un mínimo local estricto en α. Podemos decir más. La función $F' = g$ es estrictamente creciente y tiene una única raíz; por tanto

$$F'(x) < 0 \quad (x < \alpha), \qquad F'(x) > 0 \quad (x > \alpha).$$

El criterio de monotonía muestra que F decrece hasta α y crece después. Así,

$$\boxed{\alpha \text{ es el único punto de mínimo absoluto de } F \text{ en } \mathbb{R}.}$$

La iteración encontró una raíz de F'; la clasificación y la globalidad procedieron de argumentos adicionales.

El ejemplo combina casi todo el capítulo sin mezclar sus papeles: aislamiento para saber qué cero seguimos, Newton para aproximarlo, una región para demostrar convergencia, un residuo para certificar una aproximación y la teoría de C11 para interpretar el punto dentro de la función objetivo.

Convergencia a un estacionario que no es un extremo.

Ejemplo 12.25. **Newton converge, pero el punto límite no maximiza ni minimiza.** Sea

$$H(x) = x^3, \qquad x \in \mathbb{R}.$$

Entonces

$$H'(x) = 3x^2, \qquad H''(x) = 6x.$$

El único punto estacionario es $c = 0$. Para cada $x \neq 0$, Newton aplicado a H' da $N_{H'}(x) = x - \frac{3x^2}{6x} = \frac{x}{2}$. Si $x_0 \neq 0$, la órbita exacta es

$$\boxed{x_n = 2^{-n} x_0,}$$

y por el límite geométrico $x_n \longrightarrow 0$.

Todos los pasos anteriores al límite están definidos. En 0 la fórmula escrita como cociente sería $0/0$, pero no hace falta evaluarla: ya sabemos que $H'(0) = 0$ y hemos localizado el estacionario. Sin embargo, 0 no es un extremo local. Para todo $t > 0$,

$$H(-t) = -t^3 < 0 = H(0) < t^3 = H(t).$$

Hay valores mayores y menores que $H(0)$ arbitrariamente cerca del origen. El segundo criterio tampoco clasifica el punto porque $H''(0) = 0$.

La conclusión correcta es:

$$\boxed{\text{Newton ha convergido a un punto estacionario, no a un óptimo.}}$$

El ejemplo también recuerda §12.7: aquí la raíz de la ecuación $H' = 0$ es múltiple, y la convergencia es lineal con factor $1/2$, no cuadrática.

Este contraejemplo refuta una inferencia frecuente: «si Newton aplicado a la derivada converge, el límite es un mínimo». El método de raíces no contiene una etiqueta de clasificación. Solo busca un cero de la función a la que se aplica.

Máximo, mínimo o ninguna de ambas: la clasificación sigue perteneciendo a C11. Cuando Newton produce o aproxima un punto c con $F'(c) = 0$, recuperamos exactamente las herramientas ya construidas.

- Si $F''(c) > 0$, Teorema 10.5 da un mínimo local estricto.
- Si $F''(c) < 0$, el mismo teorema da un máximo local estricto.
- Si $F''(c) = 0$, el segundo criterio es inconcluso. Hay que estudiar el signo de F' a ambos lados, utilizar Teorema 10.2, una desigualdad u otra información disponible.

Los tres casos pueden verse con funciones elementales cuya ecuación estacionaria se resuelve inmediatamente:

$$F_1(x) = x^2, \qquad F_2(x) = -x^2, \qquad F_3(x) = x^3.$$

En las dos primeras, Newton aplicado a la derivada llega a 0 en un paso desde cualquier punto, porque F_1' y F_2' son afines no constantes. Sin embargo,

$$F_1''(0) = 2 > 0, \qquad F_2''(0) = -2 < 0,$$

y las clasificaciones son opuestas. En F_3, como acabamos de ver, el mismo destino estacionario no es extremo. El valor de la raíz de la ecuación $F' = 0$ no contiene la clasificación: hace falta información sobre F. Tampoco la clasificación local resuelve automáticamente el problema global. Una función puede tener un máximo local y ser no acotada superiormente, o un mínimo local superado por un borde permitido. La cadena completa para optimización sigue siendo

modelo
→ existencia
→ candidatos
→ aproximación si hace falta
→ clasificación local
→ comparación global.

Newton no ve todos los candidatos de optimización. La fórmula $x - F'(x)/F''(x)$ solo busca ceros de F' en puntos donde el cociente está definido. Por eso no puede reemplazar la proposición de candidatos de C11.

Bordes. Para $F(x) = x$ en $[0, 1]$, no existe ningún punto estacionario interior, pero 0 es el mínimo absoluto y 1 el máximo absoluto. Newton aplicado a F' ni siquiera tiene un paso: $F' = 1$ y $F'' = 0$. Los dos óptimos se encuentran en los bordes.

Puntos no diferenciables. Para $F(x) = |x|$ en $[-1, 1]$, el mínimo absoluto se alcanza en 0, donde F' no existe. La ecuación $F' = 0$ no puede localizar ese candidato.

Estacionarios con segunda derivada nula. Ejemplo 12.25 muestra que un estacionario puede existir aunque el operador aplicado a F' no esté definido en el propio punto. Llegar a ese punto por límite no autoriza dividir allí.

Estas tres situaciones explican por qué Definición 10.1 y Proposición 10.4 conservan bordes y no diferenciables junto a los ceros de la derivada. Newton puede ser una herramienta de **localización dentro de una clase de candidatos**; no genera la lista completa de candidatos de un problema global.

Volver a la caja: el cálculo numérico no sustituye la prueba global.

Ejemplo 12.26. **Dos pasos hacia el máximo de la caja.** Retomemos Ejemplo 12.21. Allí

$$V'(x) = 12(x^2 - 10x + 18), \qquad V''(x) = 24(x - 5).$$

Newton aplicado a V' tiene la forma

$$N_{V'}(x) = x - \frac{x^2 - 10x + 18}{2(x - 5)}, \qquad x \neq 5.$$

Desde $x_0 = 3$, $x_1 = 3 - \frac{-3}{-4} = \frac{9}{4}$.

Como

$$x_1^2 - 10x_1 + 18 = \frac{9}{16}, \qquad 2(x_1 - 5) = -\frac{11}{2},$$

el segundo paso es

$$x_2 = \frac{9}{4} - \frac{9/16}{-11/2} = \frac{207}{88}.$$

Los dos puntos quedan a la izquierda del crítico y el segundo está más cerca. En efecto, $233^2 - 7 \cdot 88^2 = 81 > 0$, por lo que $\sqrt{7} < 233/88 < 242/88 = 11/4$. Así, para $c = 5 - \sqrt{7}$,

$$0 < c - x_2 = \frac{233}{88} - \sqrt{7} < c - x_1 = \frac{11}{4} - \sqrt{7}.$$

Este acercamiento exacto sigue sin ser, por sí solo, una prueba de convergencia o de optimalidad. La mera apariencia decimal no fue la prueba de nada en §12.9. Allí ya habíamos demostrado que V' cambia de positivo a negativo en ese punto y que el máximo global del dominio abierto existe. También $V''(5-\sqrt{7}) = -24\sqrt{7} < 0$, por lo que Teorema 10.5 lo clasifica como máximo local estricto. Las tres capas son diferentes:

1. Newton produce candidatos numéricos para la ecuación $V' = 0$;
2. el criterio de la segunda derivada clasifica localmente el crítico exacto;
3. la prueba de §12.9 demuestra que ese punto es el máximo absoluto del problema de la caja.

Ninguna de las tres afirmaciones sustituye automáticamente a las otras dos.

Laboratorio integrador: auditar una solución de optimización con Newton. Para reunir las técnicas anteriores identificaremos qué resultado justifica cada paso de una solución:

1. Construir el problema antes del algoritmo. Especifica conjunto factible, función objetivo, unidades y sentido de la optimización. Si el problema verbal fue reducido a una variable, conserva la justificación de esa reducción.

2. Decidir qué candidatos faltan. Los bordes y puntos no diferenciables se registran antes de ejecutar Newton. La ecuación $F' = 0$ solo aborda los estacionarios interiores diferenciables.

3. Aislar el cero de F' que se desea seguir. Si hay varias raíces de la derivada, un punto inicial sin aislamiento puede llevar a otro estacionario. El IVT, monotonía u otro argumento puede separar candidatos antes de iterar.

4. Auditar cada paso. Para

$$x_{n+1} = x_n - \frac{F'(x_n)}{F''(x_n)}$$

comprueba que $F''(x_n) \neq 0$. Si ya sabes exactamente que $F'(x_n) = 0$, detén la búsqueda en vez de formar un cociente innecesario.

5. Si afirmas convergencia, identifica el teorema y sus derivadas. Usar Proposición 12.2 sobre $G = F'$ exige controlar F'' y F''' en la región correspondiente. Varias iteraciones favorables no sustituyen esas hipótesis.

6. Certificar el punto calculado. Una cota residual para la ecuación $F' = 0$ adopta la forma $|x - c| \leq \frac{|F'(x)|}{m}$, si se ha demostrado $|F''| \geq m > 0$ en un intervalo que contiene al estacionario c y al punto x. La cota certifica la **posición del estacionario**, no el error del valor óptimo $F(c)$; esa es una pregunta distinta.

7. Clasificar el punto. Usa Teorema 10.2, Teorema 10.5 u otro argumento válido. Si el segundo criterio es inconcluso, no conviertas $F''(c) = 0$ en una etiqueta.

8. Cerrar la comparación global. Vuelve a bordes, colas, no diferenciables y restantes candidatos. Solo entonces utiliza las palabras *máximo absoluto* o *mínimo absoluto*.

9. Regresar al modelo. Traduce el parámetro a la decisión original, con unidades y verificación de restricciones. Un número para x puede ser solo una coordenada intermedia y no la respuesta solicitada.

Cinco reparaciones finales. «Newton dio un punto donde F' casi vale cero; por tanto, ya tengo el mínimo». Repara la frase separando tres afirmaciones: primero certificar

cercanía a una raíz de F'; después clasificar esa raíz; finalmente compararla globalmente si el problema pide un extremo absoluto.

«Como $F''(c) = 0$, Newton no puede converger a c». Ejemplo 12.25 refuta la afirmación. Lo que falla en c es el paso definido por el cociente; una órbita puede aproximarse indefinidamente al punto sin evaluarlo allí.

«Para demostrar convergencia cuadrática de $x - F'/F''$ basta con que F sea dos veces diferenciable». No con el teorema desarrollado en este capítulo. Al aplicar Proposición 12.2 a $G = F'$, la cota de segunda derivada de G es una cota de F'''.

«Si no hay soluciones de $F' = 0$, no hay óptimos». Los bordes de $F(x) = x$ en $[0, 1]$ refutan la afirmación. Un punto no diferenciable también puede ser óptimo.

«Una vez clasificado un máximo local, ya resolví el problema». Solo si el problema pedía precisamente un máximo local. Para un máximo absoluto faltan las comparaciones del dominio completo. Ejemplo 12.26 separa explícitamente esas dos etapas.

Cierre del recorrido conceptual. El método de Newton y la optimización quedan ahora conectados sin identificarse. Newton resuelve —cuando sus hipótesis lo permiten— una ecuación de raíces. Aplicado a F', puede localizar puntos estacionarios. La teoría de C11 los clasifica y la modelización de §§12.8–12.9 determina qué significa un óptimo para el problema original. La cadena completa del capítulo puede leerse así:

aislar una raíz o un estacionario
$\Downarrow$
construir y controlar Newton
$\Downarrow$
certificar la aproximación
$\Downarrow$
clasificar el candidato
$\Downarrow$
comparar en todo el dominio y volver al modelo.

Los cuarenta ejercicios con soluciones razonadas combinan la construcción del procedimiento, la comprobación de sus hipótesis y la justificación del error o del carácter óptimo del resultado. Al aplicar Newton a la búsqueda de estacionarios de F, la función que recibe el método es F'. La no anulación de F'' permite formar el paso, y el teorema de convergencia exige además el control correspondiente de F'''. Localizar y clasificar un estacionario deja todavía pendiente la comparación global y la revisión de los candidatos que ese cociente no alcanza.

12.11. Ejercicios y soluciones

El banco reúne las dos líneas del capítulo —aproximación certificada y optimización— y obliga a distinguir cálculo, hipótesis y conclusión. Los cuarenta ejercicios son originales y se distribuyen según el diseño fino:

$$6A + 6B + 6C + 6D + 6E + 6F + 4G = 40.$$

Cada enunciado tiene una solución desarrollada con el mismo número. Solo se utilizan herramientas disponibles hasta §12.10; no se requiere integración, L'Hôpital, series infinitas, Newton multivariable ni optimización de varias variables.

Nivel A — Error, tolerancia, aislamiento y certificados. Ejercicio A1. Del encierro al certificado. Se sabe que $\alpha \in [7/5, 29/20]$.

1. Calcula el punto medio m y certifica una cota para $|m - \alpha|$.
2. Si se usa en cambio el extremo derecho $29/20$ como aproximación, ¿qué cota puede afirmarse usando solo el encierro dado?
3. Explica por qué ninguna de las dos cotas necesita conocer α exactamente.

Ejercicio A2. Aislar una raíz con datos exactos. Sea $q(x) = x^3 + 2x - 2$.

Demuestra que q tiene una única raíz α en $[1/2, 1]$. Después verifica

$$q\left(\frac{3}{4}\right) < 0 < q\left(\frac{4}{5}\right)$$

y deduce un intervalo aislante más estrecho y un certificado para su punto medio.

Ejercicio A3. Un residuo mejora el certificado del punto medio. Conserva la función y la raíz del ejercicio A2 y toma $x = \frac{31}{40}$.

1. Comprueba que $x \in [3/4, 4/5]$.
2. Calcula exactamente $q(x)$.
3. Usa una cota inferior uniforme para q' en $[1/2, 1]$ y Proposición 12.3 para certificar $|x - \alpha|$.
4. Compara la cota obtenida con la cota $1/40$ que proporciona solamente el ancho del intervalo $[3/4, 4/5]$.

Ejercicio A4. Cambiar la escala no cambia la precisión certificada. Supón que en un intervalo J una función f tiene una raíz α y satisface $|f'(t)| \geq m > 0$. Sea $\lambda \neq 0$ y define $g = \lambda f$.

Demuestra que las raíces de f y g en J coinciden, que una cota inferior válida para $|g'|$ es $|\lambda|m$ y que

$$\frac{|g(x)|}{|\lambda|m} = \frac{|f(x)|}{m}.$$

Concluye que multiplicar una ecuación por una constante muy pequeña puede hacer pequeño el residuo bruto sin mejorar el certificado del error.

Ejercicio A5. Un residuo calculado con incertidumbre. En un intervalo J hay una única raíz α y se sabe que $|f'| \geq 3$. Para cierto $x \in J$ una evaluación numérica entrega $\widetilde{r}$ con

$$|\widetilde{r}| \leq 2 \cdot 10^{-5}, \qquad |\widetilde{r} - f(x)| \leq 10^{-5}.$$

Obtén una cota certificada para $|x - \alpha|$. ¿Alcanza para garantizar tolerancia 10^{-5}?

Ejercicio A6. Certificar el número que realmente se imprime. Se ha demostrado $|x - \alpha| \leq 4 \cdot 10^{-6}$.

El valor que se publica es un número y que satisface $|y - x| \leq 5 \cdot 10^{-7}$.

1. Certifica $|y - \alpha|$.
2. Explica por qué mostrar seis decimales no permite sustituir esta demostración por la afirmación automática $|y - \alpha| < 10^{-6}$.

Nivel B — Construcción y primeras iteraciones de Newton. Ejercicio B1. Newton para $\sqrt{5}$. Para $f(t) = t^2 - 5$ en $(0, \infty)$:

1. deriva el operador de Newton;
2. partiendo de $x_0 = 2$, calcula exactamente x_1 y x_2;
3. comprueba que cada división utilizada está permitida.

Ejercicio B2. Una función afín se resuelve en un paso. Sea $f(t) = 7t - 5$. Demuestra directamente desde la ecuación de la tangente que para todo $x \in \mathbb{R}$, $N_f(x) = \frac{5}{7}$.

Identifica el desplazamiento $N_f(x) - x$ y explica por qué no debe confundirse con la raíz.

Ejercicio B3. Dos pasos hacia la raíz cúbica de 2. Para $f(t) = t^3 - 2$, calcula el operador de Newton y, desde $x_0 = 1$, obtiene x_1 y x_2 como fracciones irreducibles. No afirmes convergencia: indica qué información adicional haría falta para justificar una órbita completa.

Ejercicio B4. El primer paso existe y el segundo no. Considera

$$f(t) = \frac{1}{t} - 2, \qquad I = \mathbb{R} \setminus \{0\}.$$

1. Determina la raíz y el dominio D_f del operador de Newton.
2. Simplifica $N_f(x)$.
3. Calcula el paso desde $x_0 = 1$ y explica por qué la iteración no puede continuar desde el valor obtenido.

Ejercicio B5. Puntos fijos dentro del dominio de Newton. Sea $f : I \to \mathbb{R}$ diferenciable en un intervalo abierto y sea $D_f = \{x \in I : f'(x) \neq 0\}$. Demuestra que, para $x \in D_f$,

$$N_f(x) = x \quad \Longleftrightarrow \quad f(x) = 0.$$

Explica por qué esta equivalencia no autoriza a definir N_f en una raíz donde $f'(x) = 0$.

Ejercicio B6. Reconstruir la tangente antes de usar la fórmula. Para $f(t) = t^2 - 3$ y punto actual $x = 2$:

1. escribe la recta tangente $L_2(t)$;
2. resuelve $L_2(t) = 0$;
3. identifica por separado el desplazamiento y el punto nuevo;
4. verifica que el punto nuevo no es una raíz exacta de f.

Nivel C — Identidad del error y convergencia local. Ejercicio C1. Identidad exacta para $\sqrt{3}$. Sea $f(t) = t^2 - 3$ en $(0, \infty)$ y $\alpha = \sqrt{3}$. Demuestra algebraicamente, sin invocar Taylor, que

$$N_f(x) - \alpha = \frac{(x - \alpha)^2}{2x}, \qquad x > 0.$$

Explica cómo coincide esta identidad con Teorema 12.1.

Ejercicio C2. Construir una región certificada para $\sqrt{3}$. Demuestra primero $\frac{17}{10} < \sqrt{3} < \frac{7}{4}$.

Sea $r = 1/5$. Prueba que $B_r = [\sqrt{3} - r, \sqrt{3} + r] \subseteq [3/2, 2]$.

Para $f(t) = t^2 - 3$, encuentra constantes m, M, K, q que permitan aplicar Proposición 12.2 en B_r. Comprueba además que $x_0 = 7/4$ pertenece a B_r.

Ejercicio C3. Transportar una cota inicial. Bajo las condiciones del ejercicio C2, demuestra que si $E_0 < 1/20$, entonces

$$E_1 < \frac{1}{1200}, \qquad E_2 < \frac{1}{4320000}.$$

Indica en qué paso de la demostración se usa que las iteraciones permanecen en la región certificada.

Ejercicio C4. Un criterio con constantes dadas. Supón que una raíz α y un radio $r = 1/4$ satisfacen las hipótesis de Proposición 12.2 con

$$|f'(t)| \geq 4, \qquad |f''(t)| \leq 8$$

en B_r. Calcula K y q. Si $E_0 \leq 1/10$, obtiene cotas sucesivas para E_1, E_2, E_3 usando la desigualdad cuadrática, no solo la geométrica.

Ejercicio C5. Una desigualdad cuadrática aislada no basta. Refuta la inferencia

> «Si $E_{n+1} \leq 2E_n^2$ para todo n, entonces necesariamente $E_n \to 0$.»

Construye una sucesión no negativa que satisfaga la desigualdad y no tienda a cero. Explica qué condición adicional de §12.5 evita este problema para Newton.

Ejercicio C6. Constante asintótica para una raíz cuadrada. Sea $a > 0$ y $f(t) = t^2 - a$ en $(0, \infty)$. Supón que una órbita de Newton positiva converge a $\alpha = \sqrt{a}$ sin llegar exactamente a ella. Demuestra que

$$\lim_{n\to\infty} \frac{E_{n+1}}{E_n^2} = \frac{1}{2\sqrt{a}}.$$

Justifica tanto por la identidad exacta del error como por la fórmula general de §12.5.

Nivel D — Fallos, raíces múltiples y reparación de argumentos. Ejercicio D1. Un ciclo de período dos distinto del ejemplo del texto. Sea $f(t) = t^3 - 2t + 2$.

Demuestra que Newton está definido en 0 y en 1 y que $0 \longmapsto 1 \longmapsto 0$. Concluye que la órbita iniciada en 0 no converge. Explica por qué dos pasos perfectamente definidos no constituyen una prueba de convergencia.

Ejercicio D2. Una raíz existente no salva un punto inicial inadmisible. Para $f(t) = t^3 - 1$:

1. verifica que 1 es una raíz;
2. muestra que Newton no está definido desde $x_0 = 0$;
3. identifica exactamente cuál hipótesis de Definición 12.2 falla.

Ejercicio D3. Potencia pura y convergencia lineal. Sean $m \geq 2$ entero y $f(t) = (t-a)^m$. Para $x \neq a$, demuestra que $N_f(x) = a + \frac{m-1}{m}(x-a)$.

Deduce una fórmula explícita para $x_n - a$ y la razón E_{n+1}/E_n. Explica por qué no hay convergencia cuadrática.

Ejercicio D4. Residuos que tienden a cero sin ninguna raíz. Para $u(t) = 1/t$ en $(0, \infty)$:

1. demuestra que $N_u(x) = 2x$;
2. desde $x_0 = 1$, prueba $x_n = 2^n$;
3. demuestra $|u(x_n)| \to 0$;
4. explica por qué esto no certifica proximidad a una raíz.

Ejercicio D5. Un paso pequeño no es una tolerancia. Para $0 < \delta < 1$ considera

$$g_\delta(t) = \frac{t-1}{t+\delta}, \qquad I = (-\delta, \infty).$$

Desde $x = 0$, calcula el paso de Newton y demuestra simultáneamente

$$|N_{g_\delta}(0) - 0| < \delta, \qquad |N_{g_\delta}(0) - 1| > \frac{1}{2}.$$

Identifica qué información adicional convierte un tamaño de paso en certificado en Corolario 12.3.

Ejercicio D6. No estar inicialmente en la región certificada no implica divergencia. Para $f(t) = t^2 - 2$ parte de $x_0 = 3$.

1. Calcula x_1 y x_2 exactamente.
2. Usando $7/5 < \sqrt{2} < 3/2$, demuestra que x_2 pertenece a $[\sqrt{2} - 1/4, \sqrt{2} + 1/4]$.
3. Explica por qué desde x_2 sí puede aplicarse la región certificada de Ejemplo 12.8, aunque no pudiera aplicarse desde x_0.

Nivel E — Modelización y optimización en intervalos. Ejercicio E1. Rectángulo con perímetro fijado. Entre todos los rectángulos de lados positivos y perímetro 20 metros, determina el de área máxima. Debes:

1. elegir una variable y obtener el conjunto factible exacto;
2. justificar existencia del máximo aun cuando el dominio natural sea abierto;
3. demostrar la localización y devolver dimensiones, área y unidades.

Ejercicio E2. «A lo sumo» no empieza como igualdad. Un recinto rectangular usa una pared como cuarto lado. Para los otros tres lados hay a lo sumo 30 metros de cerco. La profundidad x debe ser al menos 2 metros y el frente y al menos 5 metros.

Modela el problema de **maximizar** el área, demuestra que un máximo debe usar todo el cerco y determina la configuración óptima. Explica por qué la saturación debe demostrarse antes de imponer $2x + y = 30$.

Ejercicio E3. Una caja desde una lámina de 20×14**.** Se cortan cuadrados de lado x en las cuatro esquinas y se forma una caja sin tapa.

1. Obtén el dominio factible y el volumen $V(x)$.
2. Demuestra que existe un máximo admisible usando una extensión al cierre.
3. Localiza el único máximo y demuestra que

$x_* = \frac{17-\sqrt{79}}{3}$.

No necesitas simplificar el volumen máximo.

Ejercicio E4. El mismo perímetro con lados enteros. Supón ahora que los lados de un rectángulo de perímetro 20 metros deben ser enteros positivos. Determina todas las configuraciones de área máxima. Justifica el resultado con una identidad algebraica, no por «redondear» la solución continua.

Ejercicio E5. Cuadrar una distancia y cuadrar una función no son la misma operación lógica.

1. Demuestra que si $d(x) \geq 0$ en el conjunto factible, entonces minimizar d equivale a minimizar d^2.
2. Da un contraejemplo mostrando que, sin la no negatividad, minimizar F y minimizar F^2 puede producir decisiones distintas. Usa $F(x) = x$ en $[-2, 1]$.

Ejercicio E6. Bordes y esquina cuentan. Determina máximo y mínimo absolutos de $F(x) = |x - 1|$ en $[-2, 3]$. Organiza la solución como existencia $\to$ candidatos $\to$ comparación y explica por qué resolver solamente $F'(x) = 0$ perdería el mínimo.

Nivel F — Dominios no compactos y Newton para puntos críticos. Ejercicio F1. Una semirrecta con mínimo pero sin máximo. Estudia

$$h(x) = \frac{x}{x+2}, \qquad x \in [0, \infty).$$

Determina sus extremos absolutos, su supremo y su ínfimo. Justifica cada afirmación sin tratar ∞ como un punto del dominio.

Ejercicio F2. Una recta completa sí puede tener ambos extremos. Para

$$g(x) = \frac{x}{1+x^2}, \qquad x \in \mathbb{R},$$

demuestra que tiene máximo absoluto $1/2$ en $x = 1$ y mínimo absoluto $-1/2$ en $x = -1$. Incluye el comportamiento cuando $|x| \to \infty$ en la comparación global.

Ejercicio F3. Un intervalo abierto puede tener máximo y carecer de mínimo. Sea

$$F(x) = x(1 - x), \qquad 0 < x < 1.$$

Demuestra que F tiene máximo absoluto $1/4$, que su ínfimo es 0 y que no tiene mínimo absoluto.

Ejercicio F4. Newton sobre la derivada y clasificación posterior. Sea $F(x) = x^4 + 2x^2 - 3$.

1. Calcula Newton aplicado a $G = F'$.
2. Desde $x_0 = 1$, calcula x_1 y x_2.
3. Demuestra por separado que el límite buscado 0 es el único mínimo absoluto de F.

Ejercicio F5. Un crítico simple para F' se alcanza en un paso. Sea $F(x) = (x-1)^2 + 2$.

Demuestra que Newton aplicado a F' lleva cualquier punto inicial a 1 en un solo paso. Después clasifica 1 y demuestra que es mínimo absoluto. Separa explícitamente «el algoritmo encontró un estacionario» de «el estacionario es mínimo».

Ejercicio F6. Newton sobre F' no ve los bordes. Considera $F(x) = x$ en $[0, 1]$.

1. Determina máximo y mínimo absolutos.
2. Comprueba que $F'(x) = 1$ en el interior y que no existe ningún punto estacionario.
3. Explica por qué el cociente F'/F'' tampoco produce candidatos y por qué esto no contradice la solución global.

Nivel G — Desafíos integradores. Ejercicio G1. Del aislamiento al paso certificado. Para $q(x) = x^3 + 2x - 2$ y su raíz α aislada en $[3/4, 4/5]$:

1. calcula un paso de Newton desde $x_0 = 4/5$ y demuestra que $x_1 = 27/35$;
2. calcula exactamente $q(x_1)$;
3. usando una cota inferior de q' en $[3/4, 4/5]$, demuestra

$|x_1 - \alpha| < \frac{1}{1900}$; 4. explica qué partes de la solución prueban existencia, identificación de la raíz, construcción de la aproximación y certificación de su error.

Ejercicio G2. Una caja completa: modelo, existencia y óptimo global. De una lámina de 30×20 centímetros se cortan cuadrados de lado x y se forma una caja sin tapa.

Desarrolla el problema completo: dominio factible, función volumen, existencia del máximo, candidatos, signo de la derivada y dimensiones óptimas. Demuestra que $x_* = \frac{25-5\sqrt{7}}{3}$ y que $V_{\text{máx}} = \frac{1000}{27}(10 + 7\sqrt{7})$ cm^3.

Ejercicio G3. Importar correctamente el teorema de convergencia a una ecuación estacionaria. Sea F tres veces diferenciable en un intervalo abierto y sea c un punto con $F'(c) = 0$. Supón que para

$$B_r = [c - r, c + r], \qquad 0 < r \leq \frac{1}{4},$$

se cumple

$$|F''(t)| \geq 4, \qquad |F'''(t)| \leq 8$$

para todo $t \in B_r$.

1. Aplica Proposición 12.2 a $G = F'$ y demuestra que Newton sobre F' converge a c desde cualquier $x_0 \in B_r$.
2. Obtén $E_{n+1} \leq E_n^2$ y una cota geométrica con factor a lo sumo $1/4$.

3. Si además $F''(c) > 0$, clasifica c mediante C11. Explica por qué esta última conclusión no procede del teorema de convergencia.

Ejercicio G4. Convergencia a un estacionario que no resuelve el problema global. En $[-1, 1]$ considera $F(x) = x^5 + x^3$.

1. Determina F', F'' y los puntos estacionarios.
2. Demuestra que Newton aplicado a F' satisface, para $x \neq 0$,

$N_{F'}(x) = x\,\frac{15x^2+3}{20x^2+6}$.

3. Prueba que $|N_{F'}(x)| \leq (3/4)|x|$ y concluye que cualquier órbita iniciada en $x_0 \neq 0$ y continuada por esa fórmula converge a 0.
4. Demuestra que F es estrictamente creciente en $[-1, 1]$, que 0 no es extremo y que los extremos absolutos están en los bordes.
5. Explica con precisión qué afirmación sobre optimización queda refutada.

Soluciones desarrolladas.

Soluciones del nivel A. **Solución A1.** El punto medio es

$$m = \frac{\frac{7}{5} + \frac{29}{20}}{2} = \frac{\frac{28}{20} + \frac{29}{20}}{2} = \frac{57}{40}.$$

El ancho del intervalo es $\frac{29}{20} - \frac{7}{5} = \frac{1}{20}$, de modo que la cota del punto medio es

$$\boxed{|m - \alpha| \leq \frac{1}{40}}.$$

Si usamos 29/20, solo sabemos que α está a su izquierda dentro del mismo intervalo; por tanto

$$\boxed{\left|\frac{29}{20} - \alpha\right| \leq \frac{1}{20}}.$$

Las dos conclusiones usan únicamente pertenencia a un intervalo. El valor exacto de α no interviene: la geometría del encierro domina todas sus posiciones posibles.

Solución A2. La función q es polinómica, luego continua. Además,

$$q\left(\frac{1}{2}\right) = \frac{1}{8} + 1 - 2 = -\frac{7}{8} < 0, \qquad q(1) = 1 > 0.$$

Por IVT existe al menos una raíz en $(1/2, 1)$. Como $q'(x) = 3x^2 + 2 > 0$ para todo real, Teorema 10.1 demuestra que q es estrictamente creciente; la raíz es única.

Para refinar,

$$q\left(\frac{3}{4}\right) = \frac{27}{64} + \frac{3}{2} - 2 = -\frac{5}{64} < 0,$$

mientras que

$$q\left(\frac{4}{5}\right) = \frac{64}{125} + \frac{8}{5} - 2 = \frac{14}{125} > 0.$$

Por unicidad,

$$\boxed{\frac{3}{4} < \alpha < \frac{4}{5}}.$$

El ancho es $1/20$, de modo que su punto medio

$$m = \frac{\frac{3}{4} + \frac{4}{5}}{2} = \frac{31}{40}$$

satisface

$$\boxed{|m - \alpha| < \frac{1}{40}}.$$

Solución A3. Como

$$\frac{3}{4} = \frac{30}{40} < \frac{31}{40} < \frac{32}{40} = \frac{4}{5},$$

el punto pertenece al intervalo aislante. Su residuo es

$$q\left(\frac{31}{40}\right) = \frac{31^3}{40^3} + 2\frac{31}{40} - 2 = \frac{29791 + 99200 - 128000}{64000} = \frac{991}{64000} > 0.$$

En $[1/2, 1]$,

$$q'(t) = 3t^2 + 2 \geq 3\left(\frac{1}{2}\right)^2 + 2 = \frac{11}{4}.$$

Por Proposición 12.3,

$$|x - \alpha| \leq \frac{|q(x)|}{11/4} = \frac{991}{64000}\frac{4}{11} = \boxed{\frac{991}{176000}}.$$

Como $\frac{991}{176000} < \frac{1}{40}$ porque $991 \cdot 40 = 39640 < 176000$, el certificado residual es más ajustado que el obtenido únicamente del ancho. Ambos son correctos; usan información distinta.

Solución A4. Si $g = \lambda f$ con $\lambda \neq 0$, entonces

$$g(x) = 0 \iff \lambda f(x) = 0 \iff f(x) = 0,$$

de modo que las raíces coinciden. Además, $g'(t) = \lambda f'(t)$, y por tanto $|g'(t)| = |\lambda|\,|f'(t)| \geq |\lambda| m$. Aplicando el certificado residual a g se obtiene

$$|x - \alpha| \leq \frac{|g(x)|}{|\lambda| m} = \frac{|\lambda||f(x)|}{|\lambda| m} = \boxed{\frac{|f(x)|}{m}}.$$

Así, hacer $|\lambda|$ pequeño reduce el residuo $|g(x)|$, pero reduce en la misma proporción la cota inferior disponible para la pendiente. La distancia certificada no mejora.

Solución A5. La desigualdad triangular da

$$|f(x)| \le |f(x) - \widetilde{r}| + |\widetilde{r}| \le 10^{-5} + 2 \cdot 10^{-5} = 3 \cdot 10^{-5}.$$

Como $|f'| \ge 3$, Proposición 12.3 produce

$$|x - \alpha| \le \frac{3 \cdot 10^{-5}}{3} = \boxed{10^{-5}}.$$

Sí alcanza para certificar tolerancia 10^{-5}, entendida con desigualdad no estricta. El hecho de que la evaluación aproximada pudiera haber devuelto un residuo menor no permitiría ignorar la incertidumbre 10^{-5}.

Solución A6. Por la desigualdad triangular,

$$|y - \alpha| \le |y - x| + |x - \alpha| \le 5 \cdot 10^{-7} + 4 \cdot 10^{-6} = \boxed{4.5 \cdot 10^{-6}}.$$

La cantidad de cifras impresas es una propiedad de la representación, no una cota matemática. Seis decimales pueden obtenerse por redondeo de un valor cuyo error sea mayor que 10^{-6}; para certificar una tolerancia hay que controlar tanto el error del punto analizado como el cambio introducido al presentarlo.

Soluciones del nivel B. **Solución B1.** Para $f(t) = t^2 - 5$, $f'(x) = 2x$, así que en $(0, \infty)$ nunca se anula. El operador es $N_f(x) = x - \frac{x^2-5}{2x} = \frac{x^2+5}{2x}$.

Desde $x_0 = 2$, $x_1 = \frac{4+5}{4} = \frac{9}{4}$.

Después,

$$x_2 = \frac{(9/4)^2 + 5}{2(9/4)} = \frac{81/16 + 80/16}{9/2} = \frac{161}{16}\frac{2}{9} = \boxed{\frac{161}{72}}.$$

Las dos divisiones son legítimas porque $x_0 = 2 > 0$ y $x_1 = 9/4 > 0$, de modo que $2x_0$ y $2x_1$ son no nulos. Esto verifica los pasos calculados; por sí solo todavía no es una prueba de convergencia.

Solución B2. La tangente en cualquier x es

$$L_x(t) = f(x) + f'(x)(t - x) = (7x - 5) + 7(t - x) = 7t - 5.$$

Como la función ya es afín, la tangente coincide con ella. Su cero es

$$7t - 5 = 0 \quad \Longrightarrow \quad t = \frac{5}{7}.$$

Por tanto

$$\boxed{N_f(x) = \frac{5}{7}}$$

para todo real x. El desplazamiento es $N_f(x) - x = \frac{5}{7} - x$, mientras que la raíz es el punto nuevo $5/7$. Solo cuando $x = 0$ ambos números coinciden accidentalmente.

Solución B3. Para $f(t) = t^3 - 2$, $f'(x) = 3x^2$, y Newton está definido para $x \neq 0$. Simplificando, $N_f(x) = x - \frac{x^3-2}{3x^2} = \frac{2x^3+2}{3x^2}$.

Desde $x_0 = 1$, $x_1 = \frac{4}{3}$.

Luego

$$x_2 = \frac{2(4/3)^3 + 2}{3(4/3)^2} = \frac{128/27 + 54/27}{16/3} = \frac{182}{27}\frac{3}{16} = \boxed{\frac{91}{72}}.$$

Para justificar una órbita completa haría falta, por ejemplo, una región donde el denominador esté separado de cero, las derivadas necesarias estén controladas y Newton preserve la región, como en Proposición 12.2. Dos cálculos exactos no contienen esa información.

Solución B4. La ecuación $\frac{1}{t} - 2 = 0$ tiene la única raíz $t = 1/2$. La función está definida para $t \neq 0$ y $f'(t) = -\frac{1}{t^2} \neq 0$ en todo su dominio. Por tanto $D_f = \mathbb{R} \setminus \{0\}$.

El operador es

$$N_f(x) = x - \frac{1/x - 2}{-1/x^2} = x + x - 2x^2 = \boxed{2x(1-x)}.$$

Desde $x_0 = 1$, $x_1 = 2(1)(0) = 0$.

El primer paso es legítimo porque $1 \in D_f$. Pero 0 no pertenece siquiera al dominio de f, así que $f(0)$ y $N_f(0)$ no están definidos. La iteración se detiene sin haber encontrado la raíz $1/2$.

Solución B5. Para $x \in D_f$ tenemos $f'(x) \neq 0$, luego

$$N_f(x) = x \iff x - \frac{f(x)}{f'(x)} = x \iff \frac{f(x)}{f'(x)} = 0 \iff f(x) = 0.$$

La última equivalencia usa que el denominador es no nulo. Si $f(x) = f'(x) = 0$, el punto es una raíz, pero $x \notin D_f$ y el cociente de Newton sería $0/0$. La raíz puede reconocerse y el cálculo detenerse; no se extiende la definición del operador por una manipulación formal.

Solución B6. En $x = 2$,

$$f(2) = 1, \qquad f'(2) = 4.$$

La tangente es $L_2(t) = 1 + 4(t-2) = 4t - 7$.

Su cero satisface $4t - 7 = 0$, de modo que

$$\boxed{x^+ = \frac{7}{4}}.$$

El desplazamiento es $x^+ - 2 = -\frac{1}{4}$.

Finalmente,

$$f\left(\frac{7}{4}\right) = \frac{49}{16} - 3 = \frac{1}{16} \neq 0.$$

Newton resolvió exactamente la ecuación de la tangente, no todavía la ecuación cuadrática.

Soluciones del nivel C. **Solución C1.** El operador es $N_f(x) = \frac{x^2+3}{2x}$.
Como $\alpha^2 = 3$,

$$\begin{aligned} N_f(x) - \alpha &= \frac{x^2 + \alpha^2 - 2\alpha x}{2x} \\ &= \boxed{\frac{(x-\alpha)^2}{2x}}. \end{aligned}$$

Por otro lado, $f'' \equiv 2$ y $f'(x) = 2x$. La identidad de Teorema 12.1 da

$$N_f(x) - \alpha = \frac{2}{2(2x)}(x-\alpha)^2 = \frac{(x-\alpha)^2}{2x},$$

exactamente la misma fórmula. En este caso la segunda derivada constante elimina toda dependencia del punto intermedio de Taylor.

Solución C2. Tenemos

$$\left(\frac{17}{10}\right)^2 = \frac{289}{100} < 3, \qquad \left(\frac{7}{4}\right)^2 = \frac{49}{16} > 3,$$

y el cuadrado es estrictamente creciente en los no negativos. Por tanto $\frac{17}{10} < \sqrt{3} < \frac{7}{4}$.
Con $r = 1/5$, $\sqrt{3} - r > \frac{17}{10} - \frac{1}{5} = \frac{3}{2}$, y

$$\sqrt{3} + r < \frac{7}{4} + \frac{1}{5} = \frac{39}{20} < 2.$$

Así $B_r \subseteq [3/2, 2]$. En ese intervalo,

$$|f'(t)| = 2t \geq 3, \qquad |f''(t)| = 2.$$

Podemos tomar

$$m = 3, \qquad M = 2, \qquad K = \frac{M}{2m} = \frac{1}{3}, \qquad q = Kr = \boxed{\frac{1}{15}} < 1.$$

Además,

$$0 < \frac{7}{4} - \sqrt{3} < \frac{7}{4} - \frac{17}{10} = \frac{1}{20} < \frac{1}{5},$$

de modo que $x_0 = 7/4 \in B_r$.

Solución C3. Del ejercicio anterior, $K = 1/3$. Por Proposición 12.2, todos los términos permanecen en la región donde vale $E_{n+1} \leq \frac{1}{3}E_n^2$.
Si $E_0 < 1/20$, entonces

$$E_1 < \frac{1}{3}\left(\frac{1}{20}\right)^2 = \boxed{\frac{1}{1200}}.$$

Una segunda aplicación da

$$E_2 < \frac{1}{3}\left(\frac{1}{1200}\right)^2 = \boxed{\frac{1}{4320000}}.$$

La permanencia se usa antes de la segunda aplicación: sin saber que x_1 sigue dentro de la región certificada, no tendríamos derecho a reutilizar la misma constante K.

Solución C4. Las constantes dan

$$K = \frac{8}{2 \cdot 4} = 1, \qquad q = Kr = \frac{1}{4} < 1.$$

Así la región es invariante y la desigualdad cuadrática puede iterarse. Si $E_0 \le 1/10$, $E_1 \le E_0^2 \le 10^{-2}$,

$$E_2 \le E_1^2 \le 10^{-4},$$

$$E_3 \le E_2^2 \le \boxed{10^{-8}}.$$

Estas cotas usan la estructura cuadrática. La cota geométrica $E_n \le (1/4)^n E_0$ también demostraría convergencia, pero sería mucho menos informativa para estos primeros errores.

Solución C5. Toma la sucesión constante $E_n = \frac{1}{2}$ para todo n. Entonces

$$E_{n+1} = \frac{1}{2} = 2\left(\frac{1}{2}\right)^2 = 2E_n^2,$$

así que satisface la desigualdad con igualdad, pero claramente no tiende a cero. En §12.5 no usamos únicamente $E_{n+1} \le KE_n^2$. Se exige además una región con radio r tal que $Kr = q < 1$ y $E_n \le r$. Entonces

$$E_{n+1} \le KE_n^2 \le KrE_n = qE_n,$$

lo que produce una reducción geométrica estricta. La condición local y la invariancia son parte esencial de la prueba.

Solución C6. La identidad exacta del ejercicio C1, con 3 reemplazado por a, es $N_f(x) - \alpha = \frac{(x-\alpha)^2}{2x}$.

Para una órbita positiva que no alcanza la raíz, $\frac{E_{n+1}}{E_n^2} = \frac{1}{2x_n}$.

Como $x_n \to \alpha = \sqrt{a}$,

$$\boxed{\lim_{n\to\infty} \frac{E_{n+1}}{E_n^2} = \frac{1}{2\sqrt{a}}}.$$

La fórmula general de §12.5 da el mismo resultado porque

$$f'(\alpha) = 2\sqrt{a}, \qquad f''(\alpha) = 2,$$

y por tanto

$$\frac{|f''(\alpha)|}{2|f'(\alpha)|} = \frac{2}{4\sqrt{a}} = \frac{1}{2\sqrt{a}}.$$

Soluciones del nivel D. **Solución D1.** Aquí $f'(x) = 3x^2 - 2$.

En 0 y 1 vale respectivamente -2 y 1, así que ambos puntos pertenecen al dominio de Newton. Ahora $N_f(0) = 0 - \frac{2}{-2} = 1$, y

$$N_f(1) = 1 - \frac{1 - 2 + 2}{3 - 2} = 1 - 1 = 0.$$

Por inducción,

$$x_{2n} = 0, \qquad x_{2n+1} = 1.$$

La sucesión no converge: dos subsucesiones constantes tienen límites distintos. Todos los pasos están bien definidos, pero la mera existencia de la recurrencia no proporciona una disminución del error ni una región con factor $q < 1$.

Solución D2. La función satisface $f(1) = 1 - 1 = 0$, así que posee la raíz real 1. Pero

$$f'(x) = 3x^2, \qquad f'(0) = 0,$$

y además $f(0) = -1 \neq 0$. El cociente

$$\frac{f(0)}{f'(0)}$$

no está definido. La hipótesis que falla es exactamente $f'(x_0) \neq 0$ en Definición 12.2. La existencia de una raíz en otro punto no convierte 0 en una entrada admisible.

Solución D3. Para $x \neq a$, $f'(x) = m(x-a)^{m-1} \neq 0$.

Entonces

$$\begin{aligned} N_f(x) &= x - \frac{(x-a)^m}{m(x-a)^{m-1}} \\ &= x - \frac{x-a}{m} \\ &= a + \frac{m-1}{m}(x-a). \end{aligned}$$

Por tanto $x_{n+1} - a = \frac{m-1}{m}(x_n - a)$, y una inducción produce

$$\boxed{x_n - a = \left(\frac{m-1}{m}\right)^n (x_0 - a)}.$$

Así

$$\boxed{\frac{E_{n+1}}{E_n} = \frac{m-1}{m}}.$$

La razón de errores tiende a una constante positiva menor que uno: la convergencia es lineal. Además, $\frac{E_{n+1}}{E_n^2} = \frac{m-1}{m}\frac{1}{E_n} \to \infty$, por lo que no puede existir una cota cuadrática eventual con constante finita.

Solución D4. Para $u(t) = 1/t$, $u'(t) = -\frac{1}{t^2}$, y

$$N_u(x) = x - \frac{1/x}{-1/x^2} = 2x.$$

Desde $x_0 = 1$, $x_n = 2^n$ por inducción. El residuo es $|u(x_n)| = \frac{1}{2^n} \to 0$. Sin embargo, $u(t) > 0$ para todo $t > 0$, así que no existe ninguna raíz en el dominio. El certificado de Proposición 12.3 exige previamente una raíz y una cota inferior positiva de $|u'|$ en una región pertinente. Aquí $|u'(2^n)| = 4^{-n} \to 0$. Un residuo pequeño aislado de esas hipótesis no mide distancia a una raíz inexistente.

Solución D5. En $x = 0$,

$$g_\delta(0) = -\frac{1}{\delta}, \qquad g'_\delta(0) = \frac{1+\delta}{\delta^2}.$$

Así $N_{g_\delta}(0) = \frac{\delta}{1+\delta}$.

El tamaño del paso es $|N_{g_\delta}(0) - 0| = \frac{\delta}{1+\delta} < \delta$. Pero la raíz es 1, de modo que $|N_{g_\delta}(0) - 1| = \frac{1}{1+\delta} > \frac{1}{2}$, porque $0 < \delta < 1$. En Corolario 12.3 el paso se convierte en certificado solo después de demostrar una región de convergencia con un factor uniforme $q < 1$ y la reducción $E_{n+1} \le qE_n$. Sin esa información, S_n solo mide la distancia entre dos iteraciones.

Solución D6. Para $f(t) = t^2 - 2$, $N_f(x) = \frac{x^2+2}{2x}$.

Desde $x_0 = 3$, $x_1 = \frac{9+2}{6} = \frac{11}{6}$.

Luego

$$x_2 = \frac{(11/6)^2 + 2}{11/3} = \frac{121/36 + 72/36}{11/3} = \frac{193}{36}\frac{3}{11} = \boxed{\frac{193}{132}}.$$

Además,

$$\left(\frac{193}{132}\right)^2 = \frac{37249}{17424} > 2,$$

pues $37249 > 34848$, así que $x_2 > \sqrt{2}$. Usando $\sqrt{2} > 7/5$,

$$0 < x_2 - \sqrt{2} < x_2 - \frac{7}{5} = \frac{193}{132} - \frac{7}{5} = \frac{41}{660} < \frac{1}{4}.$$

Por tanto $x_2 \in [\sqrt{2} - 1/4, \sqrt{2} + 1/4]$, la región certificada en Ejemplo 12.8. Desde ese índice la proposición de convergencia controla toda la cola. El hecho de que $x_0 = 3$ estuviera fuera de aquella región significaba solo que ese criterio no era aplicable desde el comienzo, no que Newton divergiera.

Soluciones del nivel E. **Solución E1.** Si los lados son $x > 0$ e $y > 0$, la condición de perímetro es $2x + 2y = 20$, por lo que $y = 10 - x$ y $0 < x < 10$.

El área es $A(x) = x(10 - x)$.

Aunque el dominio natural es abierto, extendemos la misma expresión a $[0, 10]$. Allí es continua y alcanza un máximo. Como

$$A(0) = A(10) = 0, \qquad A(5) = 25 > 0,$$

un máximo de la extensión no puede estar en los bordes; por tanto corresponde a un rectángulo admisible. Completando el cuadrado, $A(x) = 25 - (x-5)^2 \le 25$, con igualdad únicamente en $x = 5$. Entonces $y = 5$. El rectángulo óptimo es un cuadrado de 5 m por 5 m y su área máxima es

$$\boxed{25 \text{ m}^2}.$$

Solución E2. Las configuraciones factibles satisfacen

$$x \ge 2, \qquad y \ge 5, \qquad 2x + y \le 30,$$

y el área es xy. Si queda material sobrante, fijamos x y aumentamos el frente hasta $y^* = 30 - 2x \ge y$. Como $x > 0$, $xy^* - xy = x(30 - 2x - y) \ge 0$, con desigualdad estricta cuando sobra cerco. Por tanto todo máximo usa $2x + y = 30$. Ahora $y = 30 - 2x$ y $y \ge 5$ implica $x \le 25/2$. El dominio reducido es

$$D = \left[2, \frac{25}{2}\right].$$

La función objetivo es

$$A(x) = x(30 - 2x) = 30x - 2x^2 = \frac{225}{2} - 2\left(x - \frac{15}{2}\right)^2.$$

El máximo único se alcanza en

$$x = \frac{15}{2}, \qquad y = 15,$$

y vale

$$\boxed{\frac{225}{2} \text{ m}^2}.$$

La igualdad $2x + y = 30$ no procedía de la frase «a lo sumo», sino de la prueba de que toda configuración con sobrante puede mejorarse para el objetivo de maximizar.

Solución E3. Las dimensiones de la caja son

$$x, \qquad 20 - 2x, \qquad 14 - 2x.$$

Las tres deben ser positivas, por lo que $D = (0, 7)$.

El volumen es

$$V(x) = x(20 - 2x)(14 - 2x) = 4x^3 - 68x^2 + 280x.$$

La extensión polinómica a $[0, 7]$ es continua y alcanza máximo. Sus bordes valen cero, mientras $V(1) > 0$, así que cualquier máximo global positivo de la extensión se alcanza en el interior y corresponde a una caja admisible.

Derivamos:

$$V'(x) = 12x^2 - 136x + 280 = 4(3x^2 - 34x + 70).$$

Las raíces son $x = \frac{34 \pm \sqrt{316}}{6} = \frac{17 \pm \sqrt{79}}{3}$.

Como $8 < \sqrt{79} < 9$, la raíz grande es mayor que $25/3 > 7$, mientras que

$$0 < \frac{17 - \sqrt{79}}{3} < 3 < 7.$$

El coeficiente principal de V' es positivo; por tanto $V' > 0$ antes de la raíz pequeña y $V' < 0$ entre ella y 7. Así el máximo admisible es único y se alcanza en

$$\boxed{x_* = \frac{17 - \sqrt{79}}{3}}.$$

Solución E4. Si los lados enteros positivos son x, y, entonces $x + y = 10$.
Por tanto $x \in \{1, 2, \ldots, 9\}$ y $y = 10 - x$. El área satisface

$$A(x) = x(10 - x) = 25 - (x - 5)^2 \leq 25.$$

Como $x = 5$ es entero factible, se alcanza la igualdad. La única configuración, salvo intercambiar el nombre de los lados, es

$$\boxed{(5, 5)},$$

con área 25 m^2. No hemos redondeado una solución continua: la identidad demuestra directamente la desigualdad para todos los enteros factibles.

Solución E5. Para $u, v \geq 0$,

$$u < v \iff v^2 - u^2 = (v - u)(v + u) > 0 \iff u^2 < v^2.$$

Así la función $s \mapsto s^2$ es estrictamente creciente en $[0, \infty)$. Si $d(x) \geq 0$, las comparaciones entre $d(x)$ se conservan exactamente al cuadrar; por ello los minimizadores de d y d^2 coinciden. Sin la restricción de signo, toma $F(x) = x$ en $[-2, 1]$. El mínimo de F se alcanza en $x = -2$, con valor -2. En cambio $F(x)^2 = x^2$ tiene mínimo 0 en $x = 0$. Cuadrar cambió el orden porque la función cuadrado no es creciente en toda $\mathbb{R}$.

Solución E6. La función es continua en el intervalo cerrado $[-2, 3]$, de modo que Teorema 5.5 garantiza ambos extremos. En el interior es diferenciable salvo en $x = 1$; fuera de ese punto,

$$F'(x) = \begin{cases} -1, & x < 1, \\ 1, & x > 1. \end{cases}$$

No hay ceros de la derivada. Los candidatos relevantes son, por tanto,

$$-2, \qquad 1, \qquad 3.$$

Sus valores son

$$F(-2) = 3, \qquad F(1) = 0, \qquad F(3) = 2.$$

Así

$$\boxed{\text{mín}\, F = 0 \text{ en } x = 1}, \qquad \boxed{\text{máx}\, F = 3 \text{ en } x = -2}.$$

Resolver solo $F' = 0$ produciría una lista vacía y perdería precisamente el mínimo, alcanzado en un punto interior no diferenciable.

Soluciones del nivel F. **Solución F1.** El denominador es positivo y $h'(x) = \frac{2}{(x+2)^2} > 0$, así que h es estrictamente creciente. Como el borde izquierdo pertenece al dominio, $h(0) = 0$ es el mínimo absoluto y también el ínfimo.

Además, $h(x) = 1 - \frac{2}{x+2} < 1$ para todo $x \geq 0$. Dado $\varepsilon > 0$, si $x > 2/\varepsilon - 2$, entonces

$$0 < 1 - h(x) = \frac{2}{x+2} < \varepsilon.$$

Por tanto

$$\boxed{\sup h = 1},$$

pero 1 no se alcanza y no existe máximo absoluto. En resumen,

$$\boxed{\text{ínf } h = \text{mín } h = 0}, \qquad \boxed{\sup h = 1 \text{ sin máximo}}.$$

Solución F2. Derivando,

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

El denominador es siempre positivo. Por tanto $g' < 0$ en $(-\infty, -1)$, $g' > 0$ en $(-1, 1)$ y $g' < 0$ en $(1, \infty)$. Así -1 es mínimo local y 1 máximo local. Los valores son

$$g(-1) = -\frac{1}{2}, \qquad g(1) = \frac{1}{2}.$$

Además,

$$\lim_{x \to \pm\infty} \frac{x}{1 + x^2} = 0,$$

por ejemplo dividiendo numerador y denominador por x^2. En las dos colas los valores se acercan a 0, que queda estrictamente entre $-1/2$ y $1/2$. La monotonía en los tres intervalos cubre toda la recta, de modo que

$$\boxed{\text{mín } g = -\frac{1}{2} \text{ en } -1}, \qquad \boxed{\text{máx } g = \frac{1}{2} \text{ en } 1}.$$

Solución F3. Completando el cuadrado,

$$F(x) = x - x^2 = \frac{1}{4} - \left(x - \frac{1}{2}\right)^2 \leq \frac{1}{4}.$$

Como $1/2 \in (0, 1)$, el máximo se alcanza y

$$\boxed{\text{máx } F = \frac{1}{4}}.$$

Para todo $0 < x < 1$, ambos factores x y $1 - x$ son positivos, de modo que $F(x) > 0$. Sin embargo, para $x_n = 1/n$ con $n \geq 2$,

$$F(x_n) = \frac{1}{n}\left(1 - \frac{1}{n}\right) \to 0.$$

Así 0 es el ínfimo, pero no se alcanza en el dominio abierto. Por tanto

$$\boxed{\text{ínf } F = 0 \text{ y no existe mínimo absoluto}}.$$

Solución F4. Tenemos $F'(x) = 4x^3 + 4x = 4x(x^2 + 1)$,

$$F''(x) = 12x^2 + 4 = 4(3x^2 + 1) > 0.$$

Aplicando Newton a $G = F'$,

$$N_G(x) = x - \frac{4x^3 + 4x}{12x^2 + 4} = \boxed{\frac{2x^3}{3x^2 + 1}}.$$

Desde $x_0 = 1$, $x_1 = \frac{2}{4} = \frac{1}{2}$, y

$$x_2 = \frac{2(1/2)^3}{3(1/2)^2 + 1} = \frac{1/4}{7/4} = \boxed{\frac{1}{7}}.$$

La clasificación es independiente. Como $x^2 + 1 > 0$, el signo de F' coincide con el de x: $F' < 0$ para $x < 0$ y $F' > 0$ para $x > 0$. Por Teorema 10.1, F decrece en $(-\infty, 0]$ y crece en $[0, \infty)$. Por ello

$$\boxed{x = 0 \text{ es el único mínimo absoluto}},$$

con $F(0) = -3$.

Solución F5. Aquí

$$F'(x) = 2(x - 1), \qquad F''(x) = 2.$$

Newton aplicado a $G = F'$ da

$$N_G(x) = x - \frac{2(x - 1)}{2} = \boxed{1}.$$

Así cualquier punto inicial llega al estacionario 1 en un paso. Para clasificarlo usamos otra información: $F(x) = (x - 1)^2 + 2 \geq 2$, con igualdad exactamente en $x = 1$. Por tanto 1 es el único mínimo absoluto y el valor mínimo es 2. El primer argumento solo localiza una raíz de F'. El segundo compara valores de F y es el que demuestra el carácter mínimo y global.

Solución F6. En $[0, 1]$, $F(x) = x$ es estrictamente creciente, de modo que

$$\boxed{\text{mín } F = 0 \text{ en } x = 0}, \qquad \boxed{\text{máx } F = 1 \text{ en } x = 1}.$$

En el interior, $F'(x) = 1$, así que no existe ningún punto estacionario. Además $F''(x) = 0$, por lo que el cociente $F'(x)/F''(x)$ sería $1/0$ y Newton aplicado a F' no está definido. No hay contradicción: los extremos absolutos están en bordes del conjunto factible. Newton sobre la ecuación $F' = 0$ solo puede buscar raíces de la derivada donde su propio denominador sea no nulo; no reemplaza la lista global de candidatos.

Soluciones del nivel G. **Solución G1.** El operador para

$$q(x) = x^3 + 2x - 2, \qquad q'(x) = 3x^2 + 2$$

es $N_q(x) = x - \frac{x^3+2x-2}{3x^2+2} = \frac{2x^3+2}{3x^2+2}$.
Desde $x_0 = 4/5$,

$$x_1 = \frac{2(64/125) + 2}{3(16/25) + 2} = \frac{378/125}{98/25} = \boxed{\frac{27}{35}}.$$

El residuo es

$$q\left(\frac{27}{35}\right) = \frac{19683}{42875} + \frac{54}{35} - 2 = \frac{19683 + 66150 - 85750}{42875} = \boxed{\frac{83}{42875}} > 0.$$

En $[3/4, 4/5]$,

$$q'(t) = 3t^2 + 2 \geq 3\left(\frac{3}{4}\right)^2 + 2 = \frac{59}{16}.$$

Además $27/35$ pertenece al intervalo: $3/4 < 27/35 < 4/5$ por multiplicación cruzada. El certificado residual da

$$|x_1 - \alpha| \leq \frac{83/42875}{59/16} = \frac{1328}{2529625}.$$

Y $\frac{1328}{2529625} < \frac{1}{1900}$ porque $1328 \cdot 1900 = 2523200 < 2529625$. Por tanto

$$\boxed{|x_1 - \alpha| < \frac{1}{1900}}.$$

La existencia procede del cambio de signo y continuidad; la identificación de una única raíz, de la monotonía estricta; la aproximación $27/35$, de resolver la tangente de Newton; y la cota final, del residuo combinado con una pendiente uniformemente separada de cero.

Solución G2. Las dimensiones después de cortar cuadrados de lado x son

$$x, \qquad 30 - 2x, \qquad 20 - 2x.$$

La positividad exige $0 < x < 10$.
El volumen es

$$V(x) = x(30 - 2x)(20 - 2x) = 4x^3 - 100x^2 + 600x.$$

La extensión a $[0, 10]$ es continua, vale cero en los bordes y toma valores positivos en el interior, de modo que su máximo se alcanza en un punto admisible.
La derivada es

$$V'(x) = 12x^2 - 200x + 600 = 4(3x^2 - 50x + 150).$$

Los ceros son $x = \frac{50 \pm \sqrt{700}}{6} = \frac{25 \pm 5\sqrt{7}}{3}$.

Como $2 < \sqrt{7} < 3$, la raíz grande es mayor que $35/3 > 10$ y la pequeña pertenece a $(0, 5)$. Denotemos $x_* = \frac{25-5\sqrt{7}}{3}$. El coeficiente principal de V' es positivo; por tanto $V' > 0$ en $(0, x_*)$ y $V' < 0$ en $(x_*, 10)$. El máximo global es único.

Las otras dimensiones son $30 - 2x_* = \frac{10(4+\sqrt{7})}{3}$,

$$20 - 2x_* = \frac{10(1+\sqrt{7})}{3}.$$

Así

$$\begin{aligned} V_{\text{máx}} &= \frac{5(5-\sqrt{7})}{3}\,\frac{10(4+\sqrt{7})}{3}\,\frac{10(1+\sqrt{7})}{3} \\ &= \frac{500}{27}(5-\sqrt{7})(11+5\sqrt{7}) \\ &= \frac{500}{27}(20+14\sqrt{7}) \\ &= \boxed{\frac{1000}{27}(10+7\sqrt{7})\ \text{cm}^3}. \end{aligned}$$

Solución G3. Definamos $G = F'$.
Entonces

$$G' = F'', \qquad G'' = F'''.$$

La hipótesis $F'(c) = 0$ dice exactamente que c es una raíz de G. En B_r tenemos

$$|G'(t)| \geq 4, \qquad |G''(t)| \leq 8.$$

Por Proposición 12.2 podemos tomar

$$m = 4, \qquad M = 8, \qquad K = \frac{8}{2 \cdot 4} = 1, \qquad q = Kr = r \leq \frac{1}{4} < 1.$$

Por tanto Newton aplicado a G, es decir,

$$x_{n+1} = x_n - \frac{F'(x_n)}{F''(x_n)},$$

está definido para toda la órbita iniciada en B_r, permanece allí y converge a c. Además

$$\boxed{E_{n+1} \leq E_n^2}$$

y

$$\boxed{E_n \leq q^n E_0 \leq \left(\frac{1}{4}\right)^n E_0}.$$

Si $F''(c) > 0$, Teorema 10.5 clasifica c como mínimo local estricto. Esa conclusión usa el segundo criterio de la derivada para la función objetivo F; el teorema de convergencia solo dijo que la iteración se aproxima a una raíz de F'. Localizar y clasificar son inferencias distintas.

Solución G4. Derivando, $F'(x) = 5x^4 + 3x^2 = x^2(5x^2 + 3)$,

$$F''(x) = 20x^3 + 6x = 2x(10x^2 + 3).$$

Como $5x^2 + 3 > 0$, el único estacionario es $x = 0$. Para $x \neq 0$, Newton aplicado a $G = F'$ da

$$\begin{aligned} N_G(x) &= x - \frac{x^2(5x^2+3)}{2x(10x^2+3)} \\ &= x\left(1 - \frac{5x^2+3}{2(10x^2+3)}\right) \\ &= \boxed{x\frac{15x^2+3}{20x^2+6}}. \end{aligned}$$

El factor es positivo. Además,

$$4(15x^2 + 3) = 60x^2 + 12 \leq 60x^2 + 18 = 3(20x^2 + 6),$$

de modo que $|N_G(x)| \leq \frac{3}{4}|x|$.

Si $x_0 \neq 0$, ningún paso cambia de signo ni llega a cero por la fórmula, y por inducción

$$|x_n| \leq \left(\frac{3}{4}\right)^n |x_0| \to 0.$$

Así Newton sobre F' converge al único estacionario. Pero $F'(x) = x^2(5x^2 + 3) \geq 0$ en todo $[-1, 1]$, y no es idénticamente cero en ningún intervalo abierto. Por Proposición 10.2, F es estrictamente creciente. En particular, 0 no es máximo ni mínimo local: para $h > 0$ pequeño, $F(-h) < F(0) < F(h)$.

Los extremos absolutos son

$$F(-1) = -2, \qquad F(1) = 2,$$

por lo que

$$\boxed{\text{mín}\, F = -2 \text{ en } -1}, \qquad \boxed{\text{máx}\, F = 2 \text{ en } 1}.$$

El ejemplo refuta la afirmación «si Newton aplicado a F' converge a un punto estacionario, ese punto resuelve el problema global de optimización». La iteración localiza una raíz de la derivada; clasificación y comparación global requieren argumentos adicionales.

Cierre del banco
Los cuarenta pares anteriores entrenan una misma disciplina en contextos distintos: **candidato, hipótesis, certificado y alcance de la conclusión**. El banco no introduce teoría nueva; obliga a reutilizar y combinar resultados ya demostrados en §§12.1–12.10 y en los capítulos previos.

Capítulo 13

Del área y las sumas a la integral

La Parte III comienza con una pregunta más antigua que el propio cálculo: **¿cómo asignar un número a una región de borde curvo de manera que ese número pueda calcularse y justificarse?** Para rectángulos y triángulos disponemos de fórmulas geométricas exactas. Pero una región limitada por la gráfica de una función ya no se descompone, en general, en un número finito de figuras elementales cuya suma conozcamos de antemano. Empezaremos con rectángulos y sumas finitas, cuyas áreas podemos calcular exactamente. Las aproximaciones inferiores y superiores encerrarán la cantidad buscada; al refinar las descomposiciones podremos estudiar cuándo la incertidumbre se reduce hasta determinar un único número. Este orden separa dos niveles que conviene no confundir:

1. la **motivación geométrica**, que explica por qué aparecen rectángulos y áreas;
2. la **construcción analítica**, que debe formularse con números, desigualdades y límites sin depender de un dibujo.

Las particiones, las funciones escalonadas y sus sumas proporcionarán los objetos con los que podremos formular después la integral de Riemann de una función general.

13.1. Del área geométrica al problema analítico

Exhaución: encerrar antes de identificar. Una estrategia muy anterior a Newton y Leibniz consiste en no intentar medir inmediatamente la región curva. Se la **encierra** entre figuras sencillas cuyas áreas sí conocemos. Si una figura interior tiene área L y una exterior tiene área U, la monotonicidad geométrica sugiere $L \le A \le U$, donde A representa el área buscada. El dato importante no es solamente que existan dos aproximaciones. Necesitamos poder hacerlas cada vez más precisas. Si conseguimos una familia de encierros para la cual $U - L$ puede hacerse tan pequeño como queramos, entonces queda abierta la posibilidad de que el área esté determinada por esas cotas.

La **exhaución** consiste en reducir la diferencia entre las aproximaciones que encierran la cantidad buscada. Como en el aislamiento de raíces, la aproximación debe venir acompañada de una justificación: aquí la proporcionan dos sumas y una cota para su separación.

Los rectángulos convierten geometría en sumas. Supongamos por el momento que $f : [a, b] \to \mathbb{R}$ es no negativa. La región intuitiva situada bajo su gráfica es

$$S_f = \{(x, y) : a \le x \le b,\ 0 \le y \le f(x)\}.$$

No tomaremos esta fórmula como una definición de área. Solo identifica el conjunto geométrico que queremos estudiar.

Si dividimos el intervalo $[a, b]$ en franjas verticales, cada franja puede compararse con rectángulos. Un rectángulo de base Δx y altura h tiene área exacta $h\,\Delta x$. Una unión finita de rectángulos no superpuestos se traduce, por tanto, en una suma finita de términos de esta forma. Aquí aparece el núcleo aritmético de la integración:

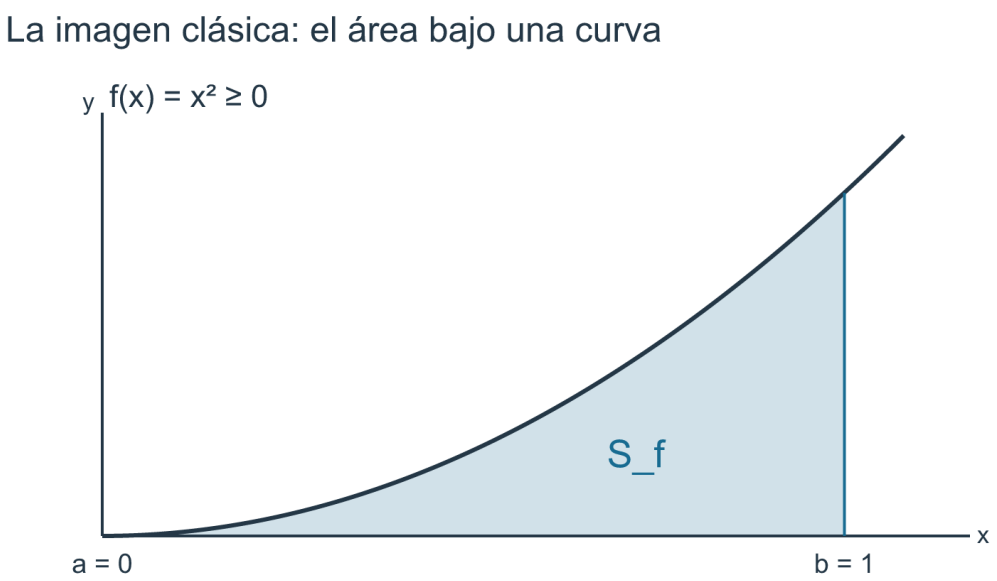


FIGURA 13.1. La imagen clásica del área bajo la curva: para $f(x) = x^2$ en $[0, 1]$, la región S_f queda entre la gráfica y el eje horizontal. El sombreado identifica el problema geométrico; todavía no define una integral.

$$\boxed{\text{altura} \times \text{anchura} \quad \longrightarrow \quad \sum \text{altura} \times \text{anchura}.}$$

La curva todavía no ha sido integrada. Lo único que hemos hecho es reemplazar un problema geométrico por el cálculo exacto de ciertas figuras escalonadas.

EJEMPLO 13.1. **El triángulo bajo $f(x) = x$ visto mediante rectángulos.** Consideremos en $[0, 1]$ la función $f(x) = x$.

La región bajo su gráfica es un triángulo rectángulo de base y altura iguales a 1, de modo que la geometría elemental proporciona como referencia $A = \frac{1}{2}$. Dividamos ahora $[0, 1]$ en n intervalos iguales de anchura $1/n$. En el intervalo

$$\left[\frac{k-1}{n}, \frac{k}{n}\right], \qquad 1 \le k \le n,$$

la función es creciente. El rectángulo interior más alto que utiliza todo ese subintervalo como base tiene altura $(k-1)/n$, mientras que un rectángulo exterior de altura k/n cubre la franja correspondiente de la región. La suma de las áreas interiores es

$$\begin{aligned} L_n &= \sum_{k=1}^{n} \frac{k-1}{n}\frac{1}{n} \\ &= \frac{1}{n^2}\sum_{j=0}^{n-1} j \\ &= \frac{1}{n^2}\frac{n(n-1)}{2} \\ &= \frac{n-1}{2n}. \end{aligned}$$

Análogamente, la suma de las áreas exteriores es

$$\begin{aligned} U_n &= \sum_{k=1}^{n} \frac{k}{n}\frac{1}{n} \\ &= \frac{1}{n^2}\frac{n(n+1)}{2} \\ &= \frac{n+1}{2n}. \end{aligned}$$

Por inclusión de regiones, $L_n \leq A \leq U_n$.

En este ejemplo podemos comprobar directamente que $\frac{n-1}{2n} \leq \frac{1}{2} \leq \frac{n+1}{2n}$. Lo decisivo es la anchura del encierro:

$$\boxed{U_n - L_n = \frac{1}{n}.}$$

Cuando n crece, esta diferencia tiende a cero. Las dos familias de rectángulos obligan al valor intermedio a quedar encerrado en intervalos cada vez más estrechos. No hemos usado una integral para obtener 1/2: el valor ya era conocido por geometría. El ejemplo sirve para aislar el mecanismo que necesitaremos cuando el área no sea conocida de antemano.

El cálculo anterior contiene ya varios ingredientes de la teoría futura: una división finita del intervalo, alturas elegidas en cada pieza, una suma y una comparación entre sumas. Sin embargo, utilizamos una división particularmente cómoda y una función creciente. Una teoría general no puede depender de esas ventajas.

El dibujo no decide qué suma es correcta. En una función arbitraria, incluso acotada, varias decisiones aparecen inmediatamente.

- ¿Cómo se permite dividir $[a, b]$?
- ¿Qué significa que una división sea más fina que otra?
- ¿Debemos tomar el valor de la función en el extremo izquierdo, en el derecho o en otro punto del subintervalo?
- Si la función no alcanza un mínimo o un máximo en una pieza, ¿qué alturas sirven para construir cotas rigurosas?
- ¿Dos divisiones distintas pueden producir cotas incompatibles?
- ¿Qué significa cuantitativamente que las piezas «se hagan pequeñas»?

Una figura puede sugerir respuestas, pero no resolver estas preguntas. Necesitamos definiciones que no dependan de cómo se haya dibujado la gráfica. Esta observación será especialmente importante cuando abandonemos las funciones continuas y no negativas. Una función puede tomar

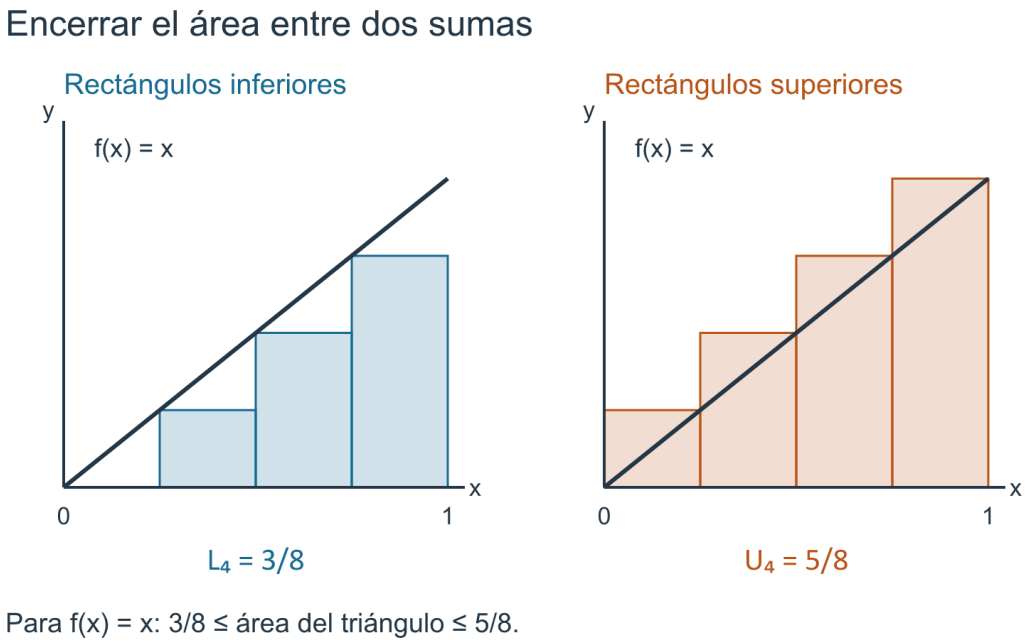


FIGURA 13.2. Cuatro franjas encierran el área del triángulo entre 3/8 y 5/8.

valores negativos, tener discontinuidades o no alcanzar sus extremos en ciertos subconjuntos. El objeto analítico que construiremos debe seguir teniendo sentido aun cuando la imagen geométrica de «área ordinaria» deje de ser suficiente.

Todavía no hemos definido la integral
Las cantidades L_n y U_n del ejemplo son **sumas finitas de áreas de rectángulos**. Que ambas se acerquen a $1/2$ explica qué esperamos de una teoría integral, pero no constituye todavía la definición de la integral de una función general. Antes debemos demostrar que las construcciones son coherentes frente a cambios de partición y de representación. Esa tarea ocupará este capítulo; la definición general de integrabilidad de Riemann quedará para el siguiente.

De la aproximación especial al lenguaje general. El ejemplo utilizó n subintervalos iguales. Esa uniformidad no es esencial. Para adaptarnos a una función que cambia rápidamente en una zona y lentamente en otra necesitaremos intervalos de longitudes distintas. El objeto adecuado será una **partición**: una colección finita y ordenada de puntos que contiene los dos extremos del intervalo. Una vez fijada la partición podremos construir dos clases de objetos complementarios. Por una parte, las **funciones escalonadas** codificarán exactamente las uniones finitas de rectángulos. Su suma altura por anchura podrá definirse sin recurrir a un proceso límite. Por otra, para una función acotada cualquiera podremos asociar a cada subintervalo cotas inferiores y superiores. Sus sumas proporcionarán encierros numéricos análogos a L_n y U_n. El problema decisivo será entonces algebraico y de orden:

¿qué relaciones entre esas sumas sobreviven cuando cambiamos o refinamos la partición?

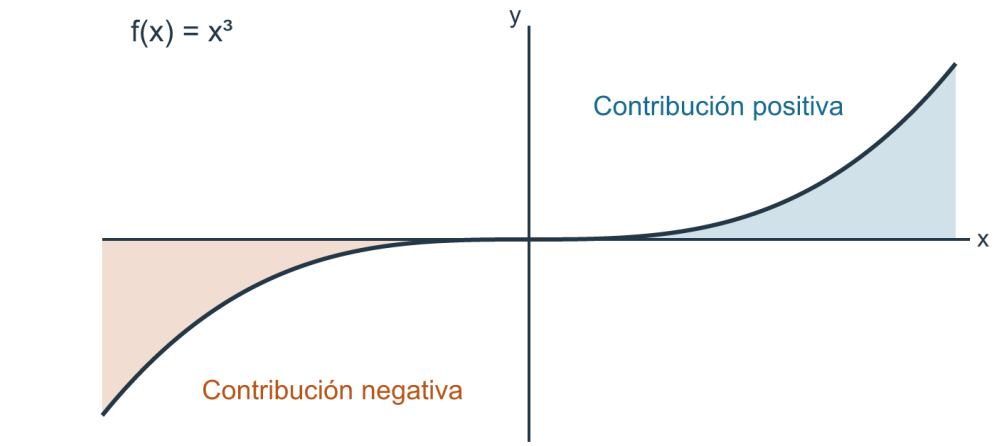


FIGURA 13.3. La imagen de área positiva tiene un límite: en $f(x) = x^3$ sobre $[-1, 1]$, las dos regiones tienen área geométrica positiva, pero sus contribuciones analíticas llevan signos opuestos. La teoría posterior debe distinguir ambas lecturas.

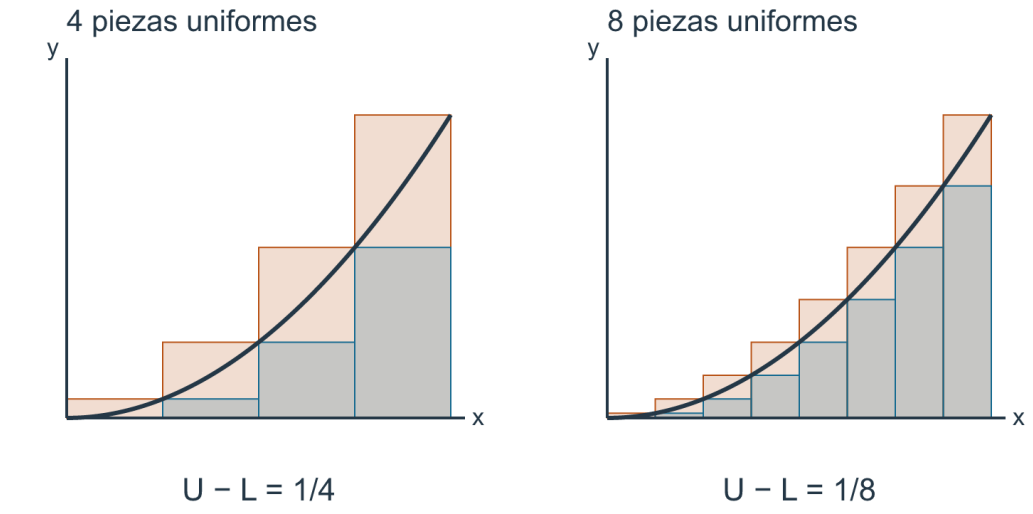


FIGURA 13.4. Un borde curvo también puede encerrarse entre escalones. Para $f(x) = x^2$ en $[0, 1]$, cuatro y ocho subintervalos uniformes producen diferencias $U_n - L_n = 1/4$ y $1/8$, respectivamente.

Responder esa pregunta es anterior a definir la integral. Si las cotas dependieran de la partición de manera incompatible, no habría un número único que pudiera merecer ese nombre.

Lo que cambia al entrar en la Parte III. En el cálculo diferencial partíamos de una función y examinábamos su comportamiento alrededor de un punto. La integración invierte parcialmente la escala de observación: dividiremos un intervalo completo en piezas, obtendremos información local en cada una y **sumaremos** esas contribuciones. El control se refiere aquí a cada subintervalo de una descomposición finita. Al considerar familias de descomposiciones más finas podremos estudiar un límite, conservando las estimaciones que relacionan sus piezas. El principio rector de este primer capítulo de integración será, por tanto:

la integral no se adivina a partir de una figura; se construye controlando sumas finitas.

En §13.2 formalizaremos el primer objeto de esa construcción: particiones, longitudes de subintervalos, refinamientos y malla. Solo después introduciremos funciones escalonadas y las sumas que pueden calcularse exactamente.

Antes de seguir. En el ejemplo de $f(x) = x$, ¿qué propiedad permite afirmar $L_n \le A \le U_n$? ¿Por qué la igualdad $U_n - L_n = 1/n$ aporta más información que saber solamente que ambas sumas parecen cercanas en una gráfica? ¿Qué parte del argumento dejaría de funcionar si la función no fuera creciente? ¿Por qué una teoría general necesita permitir particiones no uniformes?

13.2. Particiones, subintervalos y refinamientos

La descomposición uniforme de $[0, 1]$ facilitó el primer cálculo, pero una función puede variar de maneras muy distintas en zonas diferentes. Para admitir anchos desiguales necesitamos describir una descomposición finita de $[a, b]$ sin fijar todavía qué valores de la función usaremos dentro de cada pieza. El objeto básico será una **partición**. La palabra no designará una figura ni una suma: designará únicamente los puntos de corte del intervalo. Mantener separados estos papeles será importante. Primero decidimos dónde cortar; después construiremos funciones escalonadas, cotas o sumas sobre los subintervalos producidos por esos cortes.

Cortar un intervalo sin perderlo. Sea $a < b$. Si elegimos puntos interiores

$$a < x_1 < \cdots < x_{n-1} < b,$$

obtenemos los subintervalos consecutivos $[a, x_1],\ [x_1, x_2], \ldots, [x_{n-1}, b]$.

La unión de todos ellos es $[a, b]$. Dos subintervalos consecutivos comparten solamente su extremo común. Ese solapamiento en puntos de corte no crea una longitud adicional: la información cuantitativa que utilizaremos será la anchura de cada pieza. Para evitar que la notación oculte estas condiciones, incorporaremos los extremos, el orden estricto, las longitudes y la malla en una sola definición.

Definición 13.1. **Partición, subintervalos, malla y refinamiento.** Sean $a, b \in \mathbb{R}$ con $a < b$.

Una **partición** de $[a, b]$ es un conjunto finito

$$P = \{x_0, x_1, \ldots, x_n\},$$

cuyos elementos, escritos en orden creciente, satisfacen

$$a = x_0 < x_1 < \cdots < x_n = b, \qquad n \ge 1.$$

Los intervalos

$$I_k = [x_{k-1}, x_k], \qquad 1 \le k \le n,$$

se llaman **subintervalos de la partición**, y sus longitudes se denotan por

$$\Delta x_k := x_k - x_{k-1} > 0.$$

La **malla** de P es el número

$$\boxed{\|P\| := \max_{1\le k\le n} \Delta x_k.}$$

Si P y Q son particiones del mismo intervalo $[a, b]$, diremos que Q es un **refinamiento** de P si

$$P \subseteq Q.$$

En ese caso escribiremos $P \preceq Q$. Si además $P \neq Q$, diremos que el refinamiento es estricto.

La finitud cumple dos funciones. Primero, solo aparecen finitísimas piezas y las sumas que construiremos en este capítulo serán sumas finitas. Segundo, la lista positiva $\Delta x_1, \ldots, \Delta x_n$ tiene un máximo; por eso la malla queda definida sin recurrir a un supremo de una familia infinita. El orden estricto también importa. No admitimos un punto repetido, pues produciría una pieza de longitud cero y rompería la correspondencia simple entre puntos consecutivos y subintervalos. Los extremos a y b son obligatorios: una colección de cortes interiores que omita uno de ellos no es, por sí sola, una partición de todo $[a, b]$.

Por ahora suponemos $a < b$. El intervalo degenerado $[a, a]$ requerirá una convención aparte cuando estudiemos la aditividad y la orientación de las integrales; no necesitamos anticiparla para construir la teoría de particiones no degeneradas.

Dos identidades que conviene ver antes de sumar funciones. La primera información de una partición no depende de ninguna función. Las longitudes de sus piezas recuperan exactamente la longitud total:

$$\begin{aligned}\sum_{k=1}^{n} \Delta x_k &= \sum_{k=1}^{n} (x_k - x_{k-1}) \\ &= (x_1 - x_0) + (x_2 - x_1) + \cdots + (x_n - x_{n-1}) \\ &= x_n - x_0 \\ &= b - a.\end{aligned}$$

La cancelación es telescópica. Más adelante, cuando una función constante de valor c aparezca sobre cada pieza, esta identidad convertirá inmediatamente la suma de rectángulos en $c(b - a)$. La segunda observación es una cota elemental. Cada anchura satisface $0 < \Delta x_k \le \|P\|$. La malla controla, pues, el **peor tamaño** de las piezas. No afirma que todas tengan la misma longitud ni informa cuántas alcanzan el máximo.

Particiones uniformes: un caso especial, no la definición. La partición utilizada en Ejemplo 13.1 es un caso particularmente simétrico. Para un entero $n \ge 1$, definamos

$$P_n = \left\{ a + k\frac{b-a}{n} : 0 \le k \le n \right\}.$$

Sus n subintervalos tienen la misma longitud: $\Delta x_k = \frac{b-a}{n}$, y por tanto

$$\boxed{\|P_n\| = \frac{b-a}{n}.}$$

Cuando n crece, esta malla tiende a cero. Esa afirmación pertenece a la sucesión de particiones (P_n), no a una partición aislada. Una partición individual tiene una malla fija. Las particiones uniformes serán útiles para ejemplos y cálculos explícitos, pero no constituyen la noción general. Si una función presenta variación intensa cerca de un punto y es casi constante en otro tramo, puede ser más eficiente colocar muchos cortes en la primera zona y pocos en la segunda. La definición permite precisamente esa flexibilidad.

Refinar significa conservar todos los cortes anteriores. Si $P \preceq Q$, cada punto de corte de P sigue presente en Q. El refinamiento puede insertar nuevos puntos, pero no desplaza ni elimina los antiguos. En consecuencia, cada subintervalo de Q está contenido en algún subintervalo de P. Esta última frase es el hecho geométrico que más tarde permitirá comparar sumas. Un rectángulo construido sobre una pieza de P puede ser reemplazado por varios rectángulos apoyados en piezas más pequeñas de Q, sin mezclar información perteneciente a dos piezas distintas de P. La relación $\preceq$ posee además las propiedades esperadas de la inclusión:

- $P \preceq P$;
- si $P \preceq Q$ y $Q \preceq R$, entonces $P \preceq R$;
- si $P \preceq Q$ y $Q \preceq P$, entonces $P = Q$.

Así, sobre las particiones de un intervalo fijo, el refinamiento es un orden parcial. No todas las particiones son comparables: dos de ellas pueden contener cortes diferentes sin que una contenga a la otra.

Dos particiones incompatibles siempre pueden compararse después. La posibilidad de comparar construcciones provenientes de particiones distintas será esencial. El mecanismo es sencillo pero merece quedar probado antes de utilizarlo.

Proposición 13.1. **Proposición (refinamiento común y control de la malla).** Sean P y Q particiones de un mismo intervalo $[a, b]$.

1. El conjunto
$$P \vee Q := P \cup Q$$
es una partición de $[a, b]$ que refina tanto a P como a Q.
2. Si R es cualquier partición que refina simultáneamente a P y a Q, entonces
$$P \vee Q \subseteq R.$$
Por tanto, $P \vee Q$ es el refinamiento común mínimo de ambas respecto de la inclusión.
3. Si $P \preceq Q$, entonces
$$\boxed{\|Q\| \leq \|P\|.}$$

La recíproca de la tercera afirmación es falsa: una malla menor no implica que haya refinamiento.

Demostración. Como P y Q son finitos, $P \cup Q$ también lo es. Los dos contienen a y b; por tanto su unión contiene ambos extremos. Al ordenar estrictamente los elementos distintos de $P \cup Q$ obtenemos una partición de $[a, b]$. Las inclusiones

$$P \subseteq P \cup Q, \qquad Q \subseteq P \cup Q$$

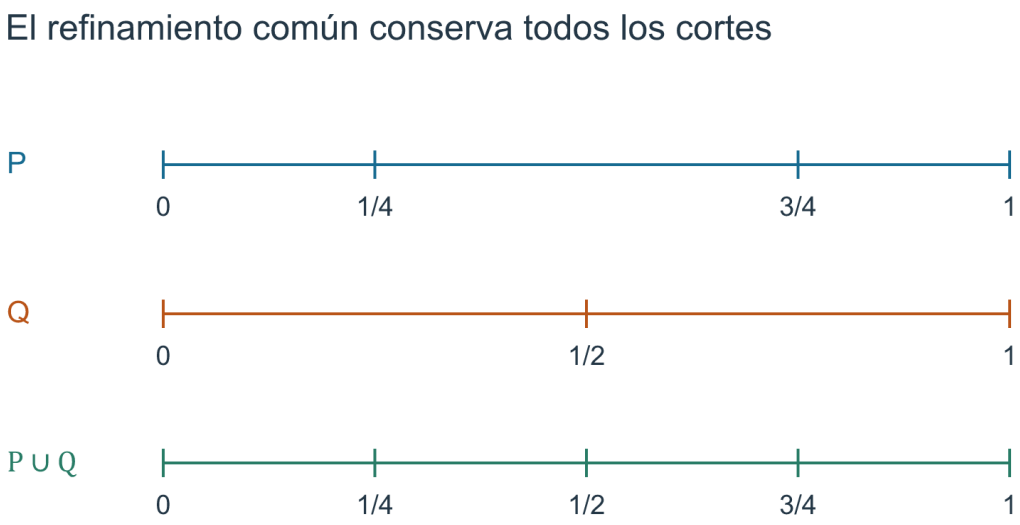


FIGURA 13.5. El refinamiento común conserva los cortes de ambas particiones, incluso cuando estas no se refinan entre sí.

prueban que refina a las dos. Si R refina simultáneamente a P y a Q, entonces $P \subseteq R$ y $Q \subseteq R$. En consecuencia, $P \cup Q \subseteq R$, lo que prueba la segunda afirmación. Supongamos ahora $P \preceq Q$. Sean $y_{j-1} < y_j$ dos puntos consecutivos de Q. Como todos los puntos de P pertenecen a Q, ningún punto de P puede quedar estrictamente entre y_{j-1} y y_j; de lo contrario también sería un punto de Q entre dos elementos consecutivos de Q. Por ello existe un subintervalo $[x_{k-1}, x_k]$ de P que contiene $[y_{j-1}, y_j]$. Entonces $y_j - y_{j-1} \leq x_k - x_{k-1} \leq \|P\|$. Esto vale para cada subintervalo de Q. Tomando el máximo de sus longitudes, $\|Q\| \leq \|P\|$. La falsedad de la recíproca aparecerá en el ejemplo siguiente. □

La primera parte proporciona una herramienta de comparación que usaremos repetidamente. Si dos sumas fueron construidas sobre P y Q, no necesitaremos decidir cuál partición es «mejor». Podremos trasladar ambas construcciones al refinamiento común $P \vee Q$ y compararlas allí.

Refinamiento y malla no dicen lo mismo.

EJEMPLO 13.2. **Tres particiones de $[0,1]$ y dos nociones que no deben confundirse.** Consideremos

$$P = \left\{0, \frac{1}{2}, 1\right\},$$

$$Q = \left\{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\right\},$$

y

$$R = \left\{0, \frac{1}{3}, \frac{2}{3}, 1\right\}.$$

Tenemos $P \subset Q$, luego $P \preceq Q$.
Sus mallas son

$$\|P\| = \frac{1}{2}, \qquad \|Q\| = \frac{1}{4}.$$

En cambio,

$$\|R\| = \frac{1}{3} < \frac{1}{2} = \|P\|,$$

pero R **no** refina a P, porque $1/2 \notin R$. La desigualdad de mallas solo afirma que la pieza más larga de R es menor que la pieza más larga de P; no afirma que R conserve los cortes de P. El refinamiento común de P y R es

$$P \vee R = \left\{0, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, 1\right\}.$$

Sus longitudes consecutivas son

$$\frac{1}{3}, \quad \frac{1}{6}, \quad \frac{1}{6}, \quad \frac{1}{3},$$

de modo que $\|P \vee R\| = \frac{1}{3}$.
Como predice Proposición 13.1, $\|P \vee R\| \leq \min\{\|P\|, \|R\|\} = \frac{1}{3}$. Este ejemplo muestra tres hechos diferentes: Q es un refinamiento genuino de P; R tiene malla menor que P sin refinarla; y $P \vee R$ conserva simultáneamente todos los cortes de ambas.

La última desigualdad del ejemplo puede generalizarse inmediatamente: como $P \preceq P \vee Q$ y $Q \preceq P \vee Q$, $\|P \vee Q\| \leq \min\{\|P\|, \|Q\|\}$. No hace falta numerar esta consecuencia como un nuevo resultado: es una aplicación doble de Proposición 13.1.

Añadir un punto: la operación elemental de refinamiento. Todo refinamiento finito puede construirse añadiendo puntos uno por uno. Conviene examinar qué sucede en un solo paso. Supongamos que P contiene el subintervalo $[u, v]$ y añadimos un punto c con $u < c < v$. La pieza de longitud $v - u$ queda sustituida por dos piezas de longitudes

$$c - u, \qquad v - c,$$

y $(c - u) + (v - c) = v - u$.
La longitud total no cambia, pero ninguna de las dos nuevas piezas es más larga que la antigua. Esta es la versión local de la desigualdad de mallas demostrada arriba. Más adelante, el mismo paso tendrá dos efectos distintos según el objeto que estemos estudiando:

- para una **función escalonada** que ya sea constante en $[u, v]$ salvo quizá en los puntos de corte, subdividir esa franja no deberá cambiar su suma altura por anchura;
- para las **cotas** de una función general, disponer de piezas menores podrá mejorar la información inferior y superior.

La compatibilidad entre esas dos conductas es una de las razones por las que el refinamiento ocupa un lugar estructural en la teoría.

Una malla pequeña no especifica una partición. Fijar un número $\delta > 0$ y exigir $\|P\| < \delta$ no determina una partición única. Hay muchas maneras de cortar $[a, b]$ respetando ese umbral. Esta multiplicidad no es un defecto: la futura definición de integral deberá ser robusta frente a todas las particiones suficientemente finas, no depender de una secuencia especialmente escogida porque produzca el valor que esperamos. Todavía no formularemos esa definición. Por ahora basta retener la distinción:

$$\boxed{\text{refinamiento} = \text{relación entre dos particiones,}}$$

mientras que

$$\boxed{\text{malla} = \text{un número asociado a una partición.}}$$

El refinamiento recuerda **qué cortes** se conservaron; la malla informa únicamente **qué tan grande es la pieza más grande**.

No confundir «más fina» con «malla menor
En lenguaje informal se dice a menudo que una partición es «más fina» que otra. En este libro reservaremos **refinamiento** para la relación exacta $P \subseteq Q$. Una partición puede tener malla menor sin contener los puntos de la otra. Cuando necesitemos comparar dos particiones no comparables, utilizaremos su refinamiento común $P \vee Q$.

Laboratorio de lectura de particiones. Falta un extremo. El conjunto $\{a, (a+b)/2\}$ no es una partición de $[a, b]$: falta b. Añadirlo produce tres puntos y dos subintervalos.

Hay un punto repetido en la lista. Escribir $a = x_0 < x_1 \leq x_2 = b$ no cumple la definición si $x_1 = x_2$. La desigualdad estricta es parte del objeto; una repetición no crea una nueva pieza.

Se compara solo el número de subintervalos. Tener más piezas no implica por sí mismo refinar otra partición. Dos particiones con cuatro subintervalos pueden usar cortes completamente distintos. Para probar refinamiento hay que comprobar inclusión de conjuntos de puntos.

Se compara solo la malla. De $\|Q\| < \|P\|$ no se deduce $P \preceq Q$. Ejemplo 13.2 proporciona un contraejemplo exacto.

Se añaden cortes sucesivamente. Si $P = P_0 \preceq P_1 \preceq \cdots \preceq P_m = Q$, entonces la transitividad da $P \preceq Q$, y Proposición 13.1 produce

$$\|Q\| \leq \cdots \leq \|P_1\| \leq \|P\|.$$

No hemos usado límites ni integrales para obtener esta cadena; es una propiedad finita del orden de los cortes.

Lo que queda preparado para las funciones escalonadas. Ya disponemos de la infraestructura geométrica y combinatoria de las sumas futuras:

$$\boxed{\begin{array}{c} [a,b] \longrightarrow P = \{x_0, \ldots, x_n\} \\ \longrightarrow I_k = [x_{k-1}, x_k],\ \Delta x_k > 0 \\ \longrightarrow \|P\| \text{ y refinamientos.} \end{array}}$$

La siguiente pregunta será qué funciones pueden describirse exactamente mediante una altura constante en cada una de estas franjas. Esa será la función de las **funciones escalonadas** de §13.3. Allí aparecerá una dificultad que el refinamiento común ya está preparado para resolver: la

misma función puede ser constante respecto de varias particiones diferentes. Antes de asociarle una suma definitiva tendremos que demostrar que cambiar la representación no cambia el número calculado.

Antes de seguir. ¿Por qué la finitud de una partición permite hablar de su malla como máximo y no solo como supremo? ¿Qué identidad satisfacen siempre las longitudes Δx_k? ¿Por qué $P \preceq Q$ obliga a $\|Q\| \leq \|P\|$? ¿Por qué la desigualdad estricta de mallas no implica la inclusión de los puntos? Si dos particiones no son comparables, ¿qué construcción permite conservar simultáneamente todos sus cortes?

13.3. Funciones escalonadas y regiones por rectángulos

Las particiones de §13.2 dividen el intervalo, pero todavía no asignan ninguna altura a sus piezas. El siguiente objeto hace exactamente eso: una función que, entre dos cortes consecutivos, mantiene un valor constante. Esta idea parece elemental, pero contiene una dificultad que debemos resolver desde el principio. Dos subintervalos cerrados consecutivos,

$$[x_{k-1}, x_k] \qquad \text{y} \qquad [x_k, x_{k+1}],$$

comparten el punto x_k. Si exigiéramos que la función fuese constante en ambos intervalos **cerrados**, dos escalones de alturas distintas obligarían absurdamente al valor $s(x_k)$ a ser simultáneamente esas dos alturas. Por ello la constancia se formulará sobre los intervalos abiertos determinados por la partición.

Constante entre los cortes.

Definición 13.2. **Función escalonada y partición representante.** Sean $a, b \in \mathbb{R}$ con $a < b$, y sea

$$P = \{x_0, x_1, \ldots, x_n\}, \qquad a = x_0 < x_1 < \cdots < x_n = b,$$

una partición de $[a, b]$.

Una función

$$s : [a, b] \longrightarrow \mathbb{R}$$

se dice **escalonada respecto de** P si para cada $k \in \{1, \ldots, n\}$ existe un número real c_k tal que

$$\boxed{s(x) = c_k \quad \text{para todo } x \in (x_{k-1}, x_k).}$$

En ese caso diremos que P **representa** a s y llamaremos c_k la **altura** de s en el k-ésimo subintervalo. Diremos simplemente que s es una **función escalonada en** $[a, b]$ si existe al menos una partición de $[a, b]$ que la represente.

La definición contiene un cuantificador existencial importante:

$$\boxed{s \text{ es escalonada} \iff \text{existe una partición finita que la representa.}}$$

No se exige, por tanto, que nos entreguen de antemano una partición privilegiada ni que la representación sea única. Los valores $s(x_0), s(x_1), \ldots, s(x_n)$ deben existir porque el dominio de s es todo $[a, b]$, pero **no están obligados** a coincidir con las alturas vecinas. Por ejemplo, si $c_k \neq c_{k+1}$, el valor en el corte x_k puede ser c_k, c_{k+1} o cualquier otro número real sin destruir la constancia en

$$(x_{k-1}, x_k) \quad \text{y} \quad (x_k, x_{k+1}).$$

Esto no significa que el valor puntual sea irrelevante para la función: cambiar $s(x_k)$ produce, en general, **otra función**. Lo que significa es más preciso: el valor aislado en el corte no modifica las alturas de las franjas de anchura positiva.

De los escalones a los rectángulos. Supongamos por un momento que todas las alturas son no negativas: $c_k \geq 0$. Sobre cada intervalo abierto (x_{k-1}, x_k), la gráfica mantiene la altura c_k. Ignorando momentáneamente las líneas verticales situadas en los cortes, la región geométrica correspondiente se descompone en un número finito de franjas rectangulares de anchuras $\Delta x_k = x_k - x_{k-1}$ y alturas c_k. Cada una posee un área elemental $c_k \Delta x_k$. Por eso las funciones escalonadas constituyen el primer caso donde la idea de «sumar contribuciones locales» puede ejecutarse mediante aritmética finita y exacta.

Pero el concepto analítico será más amplio que esta imagen. Permitiremos también alturas negativas. En ese caso ya no estamos describiendo simplemente un área geométrica ordinaria no negativa; estamos preparando una **suma con signo**. Esta separación será esencial más adelante: la integral no será sinónimo de «área positiva bajo una gráfica».

Una representación no es única. Si una función es constante en una franja, podemos insertar puntos de corte adicionales dentro de ella sin perder la constancia. El refinamiento construido en §13.2 formaliza esta observación.

Proposición 13.2. **Proposición (transporte de una representación por refinamiento).** Sea $s : [a, b] \to \mathbb{R}$ una función escalonada respecto de una partición P.

1. Si $P \preceq Q$, entonces s también es escalonada respecto de Q.
2. Si P y Q representan a la misma función s, entonces el refinamiento común
$$P \vee Q = P \cup Q$$
también representa a s.
3. Si $\widetilde{s} : [a, b] \to \mathbb{R}$ satisface
$$\widetilde{s}(x) = s(x) \qquad \text{para todo } x \in [a, b] \setminus P,$$
entonces P también representa a $\widetilde{s}$, cualesquiera que sean los valores reales asignados a $\widetilde{s}$ en los puntos de P.

Demostración. Supongamos primero $P \preceq Q$. Por Proposición 13.1, cada subintervalo determinado por dos puntos consecutivos de Q queda contenido en algún subintervalo de P. Más precisamente, si $y_{j-1} < y_j$ son consecutivos en Q, no puede haber un punto de P estrictamente entre ellos, porque $P \subseteq Q$. Por tanto existe k tal que $(y_{j-1}, y_j) \subseteq (x_{k-1}, x_k)$.

Como s es constante en (x_{k-1}, x_k), también lo es en su subconjunto (y_{j-1}, y_j). Así Q representa a s. La segunda afirmación es inmediata de la primera: Proposición 13.1 demuestra que $P \vee Q$ refina tanto a P como a Q. Para la tercera, los intervalos abiertos (x_{k-1}, x_k) no contienen puntos de P. En todos ellos $\widetilde{s} = s$, de modo que $\widetilde{s}$ conserva exactamente las mismas alturas c_k. Los valores modificados en los cortes no afectan la condición de constancia exigida por Definición 13.2. □

La proposición tiene dos consecuencias conceptuales distintas. Primero, una partición representante puede **refinarse libremente**. La función no cambia; solo cambia la manera finita de describir sus franjas constantes. Segundo, las alturas sobre los intervalos abiertos no determinan por sí solas todos los valores de la función. Dos funciones distintas pueden compartir las mismas alturas entre cortes y diferir únicamente en los puntos de la partición. Estas dos observaciones

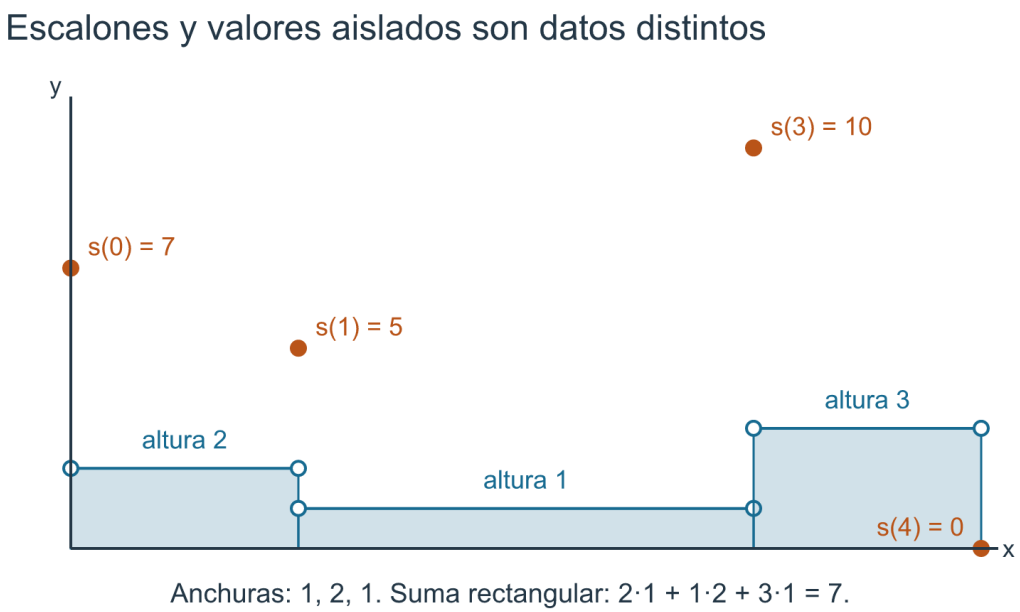


FIGURA 13.6. Los escalones tienen alturas $2, 1, 3$ y anchuras $1, 2, 1$; los valores aislados se indican por separado.

explican qué debe demostrarse en §13.4: si queremos asociar a una función escalonada la suma de sus «altura por anchura», ese número deberá ser independiente de la partición representante y no deberá recibir contribuciones artificiales de puntos aislados.

Un mismo perfil, muchas particiones.

EJEMPLO 13.3. **Una función escalonada con valores excepcionales en los cortes.** Definamos $s : [0, 4] \to \mathbb{R}$ mediante

$$s(x) = \begin{cases} 2, & 0 < x < 1, \\ 1, & 1 < x < 3, \\ 3, & 3 < x < 4, \end{cases}$$

y fijemos, por ejemplo,

$$s(0) = 7, \qquad s(1) = 5, \qquad s(3) = 10, \qquad s(4) = 0.$$

La partición $P = \{0, 1, 3, 4\}$ representa a s, con alturas $2,\ 1,\ 3$ y anchuras $1,\ 2,\ 1$. Los valores $7, 5, 10, 0$ en los puntos de corte no son las alturas de las franjas; son valores puntuales adicionales de la función.

Ahora consideremos

$$Q = \left\{0, \frac{1}{2}, 1, 2, 3, \frac{7}{2}, 4\right\}.$$

Como $P \subseteq Q$, tenemos $P \preceq Q$. Por Proposición 13.2, Q también representa a s. Las alturas correspondientes a las seis franjas son 2, 2, 1, 1, 3, 3. La suma puramente geométrica de las áreas de los rectángulos de anchura positiva es, usando P, $2 \cdot 1 + 1 \cdot 2 + 3 \cdot 1 = 7$. Si subdividimos las mismas franjas usando Q, obtenemos exactamente

$$2\left(\frac{1}{2} + \frac{1}{2}\right) + 1(1+1) + 3\left(\frac{1}{2} + \frac{1}{2}\right) = 7.$$

Este cálculo **no es todavía una definición general de integral**. Es un ejemplo de la identidad finita que la demostración de §13.4 deberá establecer para cualquier refinamiento y, después, para cualesquiera dos particiones representantes.

El ejemplo revela por qué la independencia de representación no puede declararse por intuición. Aquí la igualdad se comprueba a mano porque cada intervalo original fue simplemente subdividido. En general tendremos que demostrar que toda subdivisión conserva la suma altura por anchura y luego pasar a un refinamiento común.

Los puntos aislados pueden ser escalones de anchura cero solo en sentido informal. Consideremos una función $r : [0,1] \to \mathbb{R}$ dada por

$$r(x) = 0 \quad \text{si } x \notin \left\{\frac{1}{3}, \frac{2}{3}\right\},$$

con

$$r\left(\frac{1}{3}\right) = 5, \qquad r\left(\frac{2}{3}\right) = -4.$$

La partición

$$P = \left\{0, \frac{1}{3}, \frac{2}{3}, 1\right\}$$

representa a r, porque la función vale 0 en cada subintervalo abierto de P. Es tentador decir que los valores 5 y -4 forman «rectángulos de anchura cero». Usaremos esa frase, si aparece, solo como intuición. Formalmente no hemos añadido subintervalos degenerados a la partición: Definición 13.1 exige longitudes positivas. Los dos valores excepcionales son simplemente valores de la función en puntos de corte. Esta precisión evita introducir por la puerta trasera una teoría de área de conjuntos de un punto. La futura suma escalonada podrá ignorar esos valores por su propia definición algebraica, sin necesitar todavía demostrar que los puntos tienen área cero.

Tener pocos valores no basta. El nombre «escalonada» no significa únicamente «función con un número finito de valores». La estructura decisiva es la existencia de **finitos cortes** entre los cuales la función sea constante.

Ejemplo 13.4. **Una función con solo dos valores que no es escalonada.** Definamos $d : [0,1] \to \mathbb{R}$ por

$$d(x) = \begin{cases} 1, & x \in \mathbb{Q}, \\ 0, & x \notin \mathbb{Q}. \end{cases}$$

La imagen de d contiene solamente dos números, 0 y 1. Sin embargo, d no es escalonada. En efecto, supongamos que una partición finita $P = \{x_0, \ldots, x_n\}$ representara a d. Como $x_{k-1} < x_k$, cada intervalo abierto (x_{k-1}, x_k) es no vacío. Por la densidad de los racionales y de los irracionales demostrada en Sección 1.8, contiene al menos un racional q y un irracional u. Entonces

$$d(q) = 1, \qquad d(u) = 0,$$

por lo que d no es constante en ese subintervalo abierto. Esto contradice Definición 13.2. Por tanto, una imagen finita no garantiza una descomposición finita del dominio en franjas constantes.

Este contraejemplo marca una frontera que será importante más adelante. Las funciones escalonadas son muy simples desde el punto de vista de su estructura horizontal, pero la teoría de integración deberá enfrentarse a funciones que no admiten ninguna representación de este tipo. No extraeremos todavía de Ejemplo 13.4 ninguna conclusión sobre integrabilidad. Esa pregunta pertenece a C15.

Qué información conserva una representación. Si P representa a s, la información necesaria para describir sus franjas constantes consta de

$$\boxed{P = \{x_0, \dots, x_n\}, \qquad (c_1, \dots, c_n).}$$

A partir de ella conocemos, para cada k,

$$\text{anchura } \Delta x_k = x_k - x_{k-1} \qquad \text{y altura } c_k.$$

Los valores $s(x_k)$ en los cortes pertenecen también a la función completa, pero no son parte de la lista de alturas de los intervalos abiertos. Esta distinción permitirá que la próxima sección formule el candidato natural

$$\sum_{k=1}^{n} c_k \Delta x_k$$

sin confundir tres objetos diferentes:

1. la función s completa;
2. una partición concreta que la representa;
3. la lista de alturas correspondiente a esa representación.

La dificultad será demostrar que, aunque los objetos 2 y 3 pueden cambiar al refinar o elegir otra representación, el número calculado no cambia.

Lo que todavía no podemos escribir sin prueba. A partir de Ejemplo 13.3 resulta muy natural querer escribir inmediatamente

$$\int_a^b s(x)\,dx = \sum_{k=1}^{n} c_k \Delta x_k.$$

El número asignado debe depender de s, con independencia de la partición utilizada para representarla. Supongamos que la misma función admite otra partición representante $Q = \{y_0, \dots, y_m\}$ con otra lista de alturas $d_1, \dots, d_m$. Antes de utilizar el símbolo integral debemos demostrar que

$$\sum_{k=1}^{n} c_k (x_k - x_{k-1}) = \sum_{j=1}^{m} d_j (y_j - y_{j-1}).$$

El refinamiento común Proposición 13.1 y el transporte Proposición 13.2 contienen ya toda la infraestructura necesaria para esa prueba. Por eso el siguiente paso no será «integrar» una función más complicada. Será probar que la primera integral que definamos está **bien definida**.

Antes de seguir. ¿Por qué la constancia se exige en los intervalos abiertos y no en los cerrados? ¿Qué parte de una función escalonada puede cambiar sin alterar sus alturas entre cortes? ¿Por qué un refinamiento de una partición representante sigue representando la misma función? ¿Por qué tener imagen finita no basta para ser escalonada? ¿Qué igualdad debe demostrarse antes de que la expresión $\sum c_k \Delta x_k$ pueda recibir legítimamente el nombre de integral?

13.4. La integral elemental de una función escalonada

La sección anterior terminó con una expresión natural pero todavía sospechosa. Si una función escalonada s está representada por

$$P = \{x_0, x_1, \dots, x_n\}, \qquad a = x_0 < \cdots < x_n = b,$$

y toma el valor constante c_k en cada intervalo abierto (x_{k-1}, x_k), la suma

$$\sum_{k=1}^{n} c_k (x_k - x_{k-1})$$

es exactamente lo que sugiere el cálculo de rectángulos: altura por anchura, sumado sobre todas las franjas. Pero P no forma parte de la identidad de la función. La misma s puede admitir otra partición representante con otro número de subintervalos. Por ello no definiremos todavía la integral mediante esa fórmula. Primero debemos demostrar que el número no depende de la representación elegida.

La suma rectangular de una representación. Fijemos temporalmente una partición representante P de s. Si

$$s(x) = c_k \qquad \text{para } x \in (x_{k-1}, x_k),$$

definimos **solo para esta representación** su suma rectangular por

$$\mathcal{R}(s; P) := \sum_{k=1}^{n} c_k \Delta x_k = \sum_{k=1}^{n} c_k (x_k - x_{k-1}).$$

La notación conserva deliberadamente a P. En este momento todavía no hemos demostrado que $\mathcal{R}(s; P)$ sea una propiedad de s sola. Los valores $s(x_k)$ en los puntos de corte no aparecen en esta suma. No diremos que esos puntos «tienen área cero» como premisa: simplemente, la definición algebraica de $\mathcal{R}(s; P)$ utiliza las alturas de los intervalos abiertos y sus anchuras positivas.

Subdividir una franja no cambia su contribución total. El caso fundamental es insertar nuevos cortes sin cambiar la función.

Lema 13.1. **Lema (invariancia de la suma rectangular bajo refinamiento).** Sea $s : [a, b] \to \mathbb{R}$ escalonada respecto de una partición

$$P = \{x_0, \dots, x_n\},$$

y sea Q un refinamiento de P. Entonces Q también representa a s por Proposición 13.2 y

$$\boxed{\mathcal{R}(s;Q) = \mathcal{R}(s;P).}$$

Demostración. Escribamos $Q = \{y_0, \dots, y_m\}$, con los puntos ordenados crecientemente. Como $P \subseteq Q$, para cada $i \in \{0, \dots, n\}$ existe un índice r_i tal que $x_i = y_{r_i}$, y

$$0 = r_0 < r_1 < \cdots < r_n = m.$$

Fijemos un subintervalo original (x_{i-1}, x_i), donde s vale constantemente c_i. Los puntos de Q que subdividen esta franja son

$$y_{r_{i-1}} = x_{i-1} < y_{r_{i-1}+1} < \cdots < y_{r_i} = x_i.$$

Cada intervalo abierto

$$(y_{j-1}, y_j), \qquad r_{i-1} < j \le r_i,$$

está contenido en (x_{i-1}, x_i). Por tanto la altura inducida por Q en todas esas piezas sigue siendo c_i. La contribución de las piezas refinadas dentro de la franja original es entonces

$$\begin{aligned}\sum_{j=r_{i-1}+1}^{r_i} c_i(y_j - y_{j-1}) &= c_i \sum_{j=r_{i-1}+1}^{r_i} (y_j - y_{j-1}) \\ &= c_i(y_{r_i} - y_{r_{i-1}}) \\ &= c_i(x_i - x_{i-1}).\end{aligned}$$

La igualdad central es telescópica. Así, subdividir la i-ésima franja no cambia su contribución total. Sumando esta identidad para $i = 1, \dots, n$ obtenemos

$$\mathcal{R}(s;Q) = \sum_{i=1}^{n} c_i(x_i - x_{i-1}) = \mathcal{R}(s;P).$$

Esto prueba el lema. □

La prueba contiene la identidad aritmética esencial de toda la construcción:

$$\boxed{c(v-u) = c(t_1-u) + c(t_2-t_1) + \cdots + c(v-t_r)}$$

cuando los puntos intermedios subdividen $[u, v]$. El refinamiento no añade ni quita anchura; solo reparte la misma longitud entre más sumandos que conservan la misma altura.

Dos representaciones arbitrarias dan el mismo número. El lema resuelve el caso en que una partición refina a la otra. Para dos representaciones no comparables utilizamos exactamente la herramienta construida en §13.2: su refinamiento común.

Teorema 13.1. **Teorema (independencia de la partición representante).** Sea $s : [a, b] \to \mathbb{R}$ una función escalonada. Si P y Q son dos particiones que representan a s, entonces

$$\boxed{\mathcal{R}(s;P) = \mathcal{R}(s;Q).}$$

En consecuencia, la suma altura por anchura depende de la función escalonada s y del intervalo $[a, b]$, pero no de la partición escogida para representarla.

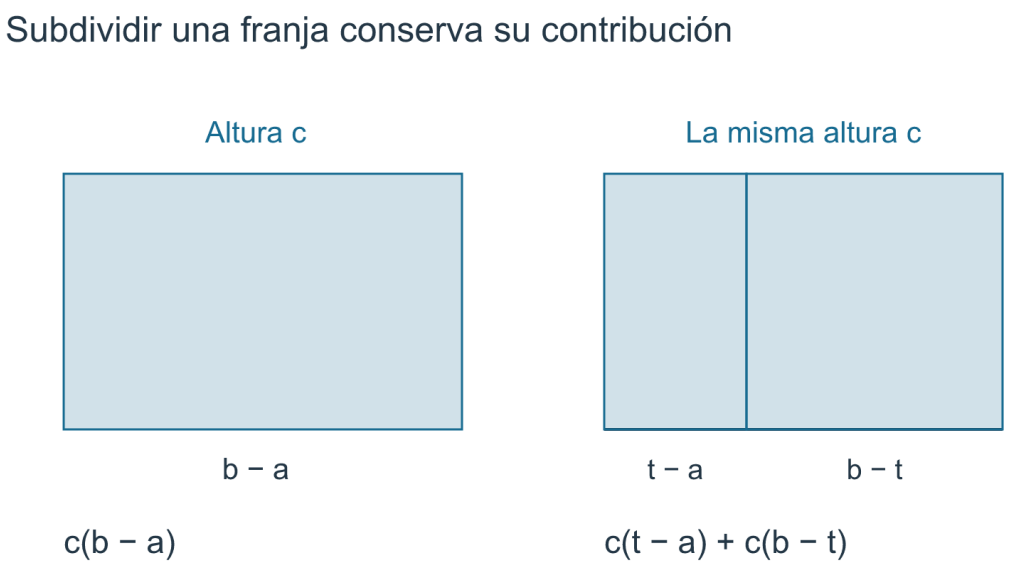


FIGURA 13.7. La subdivisión de una franja mantiene la altura y descompone su anchura: la contribución total permanece igual.

Demostración. Por Proposición 13.1, $R := P \vee Q = P \cup Q$ es una partición que refina simultáneamente a P y a Q. Por Proposición 13.2, R también representa a s.

Aplicando Lema 13.1 primero a $P \preceq R$ y luego a $Q \preceq R$, $\mathcal{R}(s;P) = \mathcal{R}(s;R)$ y $\mathcal{R}(s;Q) = \mathcal{R}(s;R)$. Por transitividad de la igualdad, $\mathcal{R}(s;P) = \mathcal{R}(s;Q)$. Eso era lo que faltaba demostrar. □

El argumento es corto porque las dos secciones anteriores prepararon su infraestructura. La lógica completa es

$$\boxed{\begin{array}{c}\text{dos representaciones}\\ \longrightarrow \text{ refinamiento común}\\ \longrightarrow \text{ invariancia bajo refinamiento}\\ \longrightarrow \text{ un único número.}\end{array}}$$

Ahora sí: la integral elemental. Una vez demostrada la independencia, podemos retirar de la notación la partición auxiliar.

DEFINICIÓN 13.3. **Integral elemental de una función escalonada.** Sean $a, b \in \mathbb{R}$ con $a < b$ y sea $s : [a, b] \to \mathbb{R}$ una función escalonada. Si

$$P = \{x_0, \dots, x_n\}$$

es cualquier partición que representa a s y

$$s(x) = c_k \qquad \text{para } x \in (x_{k-1}, x_k),$$

definimos

$$\boxed{\int_a^b s(x)\,dx := \sum_{k=1}^{n} c_k(x_k - x_{k-1}).}$$

Por Teorema 13.1, el valor del miembro derecho es independiente de la partición representante. Por ello la definición está bien definida.

El adjetivo **elemental** recuerda el alcance actual: el símbolo integral acaba de ser definido para funciones escalonadas y nada más. Todavía no hemos definido $\int_a^b f$ para una función acotada general, continua o no. Esa extensión pertenece a C15.

La letra x en

$$\int_a^b s(x)\,dx$$

funciona aquí como variable ligada de la notación. El número se obtiene de las alturas constantes de s y de las longitudes de los intervalos de una representación; no estamos introduciendo todavía una operación inversa de la derivación ni usando el teorema fundamental del cálculo.

El ejemplo anterior ya puede recibir legítimamente el símbolo integral.

Ejemplo 13.5. **La misma función, dos cálculos y una sola integral.** Retomemos la función de Ejemplo 13.3, representada por $P = \{0, 1, 3, 4\}$, con alturas 2, 1 y 3 en los intervalos abiertos correspondientes. Entonces

$$\begin{aligned}\int_0^4 s(x)\,dx &= 2(1-0) + 1(3-1) + 3(4-3)\\ &= 2 + 2 + 3\\ &= 7.\end{aligned}$$

Si usamos el refinamiento

$$Q = \left\{0, \frac{1}{2}, 1, 2, 3, \frac{7}{2}, 4\right\},$$

obtenemos

$$\begin{aligned}\int_0^4 s(x)\,dx &= 2\left(\frac{1}{2} + \frac{1}{2}\right) + 1(1+1) + 3\left(\frac{1}{2} + \frac{1}{2}\right)\\ &= 7.\end{aligned}$$

La igualdad de ambos cálculos ya no es una coincidencia verificada a mano: es una instancia de Teorema 13.1. Además, los valores excepcionales que s pueda tomar en los puntos de corte no aparecen en ninguna de las dos sumas. La integral elemental depende de las franjas constantes, no de esos valores aislados.

Cambiar finitísimos valores no altera la integral escalonada. La libertad en los puntos de corte puede expresarse de una manera más general.

Corolario 13.1. **Corolario (invariancia ante modificaciones finitas).** Sean $s, t : [a, b] \to \mathbb{R}$ funciones escalonadas. Supongamos que existe un conjunto finito $F \subseteq [a, b]$ tal que

$$s(x) = t(x) \qquad \text{para todo } x \in [a,b] \setminus F.$$

Entonces

$$\boxed{\int_a^b s(x)\,dx = \int_a^b t(x)\,dx.}$$

Demostración. Sean P_s y P_t particiones representantes de s y t. Consideremos $R := P_s \cup P_t \cup F$.

Al eliminar repeticiones y ordenar sus elementos, R es una partición de $[a,b]$: es finita y contiene los extremos porque ya los contienen P_s y P_t. Por Proposición 13.2, R representa tanto a s como a t. Además, ningún punto de F queda en el interior de un subintervalo abierto de R. En cada uno de esos intervalos abiertos se cumple $s = t$ y ambas funciones son constantes; por tanto sus alturas allí coinciden. Las dos integrales, calculadas mediante R, son entonces la misma suma altura por anchura. □

Este corolario no afirma que «cambiar un conjunto pequeño» sea siempre irrelevante: solo se ha demostrado el caso **finito** dentro de la clase de funciones escalonadas. Resultados más amplios requerirían una teoría posterior que no pertenece a este capítulo.

Qué se ha ganado con la buena definición. El símbolo integral no apareció porque la fórmula pareciera razonable, sino porque se resolvieron dos posibles ambigüedades.

Ambigüedad de partición. Una función escalonada puede tener muchas particiones representantes. Teorema 13.1 demuestra que todas producen el mismo número.

Ambigüedad de los puntos de corte. La función completa incluye sus valores en los puntos de subdivisión, pero esos valores no forman parte de las alturas de los intervalos abiertos. Corolario 13.1 muestra, en particular, que finitísimas modificaciones puntuales no alteran la integral elemental.

Así hemos construido una primera operación integral completamente algebraica y finita:

$$\boxed{\text{función escalonada} \longmapsto \text{número bien definido.}}$$

No ha aparecido ningún proceso límite en la **definición** de esta integral. El límite será necesario más adelante para extender la construcción a funciones que no son escalonadas.

Lo que todavía no se ha demostrado. De Definición 13.3 no debemos inferir automáticamente propiedades que aún no han sido probadas. En particular, todavía falta justificar de manera sistemática que

- la integral respeta sumas y múltiplos escalares;
- una función escalonada no negativa tiene integral no negativa;
- $s \le t$ implica una desigualdad entre sus integrales;
- la integral se comporta aditivamente cuando se corta el intervalo de integración.

Esas propiedades constituyen §13.5. Algunas parecerán evidentes al mirar rectángulos, pero las demostraremos a partir de particiones comunes y sumas finitas, no a partir del dibujo.

Antes de seguir. ¿Por qué no era legítimo definir $\int s$ antes de Teorema 13.1? ¿Dónde se usa que un refinamiento conserva todos los cortes anteriores? ¿Qué papel cumple el refinamiento común en la prueba de independencia? ¿Por qué los valores de una función escalonada en los puntos de corte no aparecen en su integral? ¿Qué parte de Corolario 13.1 dejaría de estar justificada si el conjunto excepcional no fuera finito?

13.5. Álgebra, orden y aditividad de la integral escalonada

La definición de §13.4 asigna un número a cada función escalonada. El paso siguiente es comprobar que esa asignación respeta las operaciones y el orden que ya conocemos en las funciones. No basta con que las fórmulas resulten plausibles por una figura: debemos demostrar que todas ellas sobreviven a la posible elección de particiones diferentes. La idea técnica será siempre la misma. Si dos funciones escalonadas vienen dadas inicialmente sobre particiones distintas, las trasladamos a un **refinamiento común**. Una vez que ambas están descritas sobre las mismas franjas, las propiedades de la integral se reducen a propiedades elementales de sumas finitas.

Sumar funciones exige primero una partición común. Sean $s, t : [a, b] \to \mathbb{R}$ escalonadas. Puede ocurrir que P represente a s y Q represente a t, pero que ninguna de las dos particiones refine a la otra. En ese caso no tendría sentido sumar directamente las listas de alturas asociadas a P y Q: ni siquiera tienen por qué poseer el mismo número de términos. El refinamiento común $R = P \vee Q = P \cup Q$ resuelve el problema. Por Proposición 13.2, R representa simultáneamente a s y a t. Si $R = \{z_0, z_1, \dots, z_N\}$ y en cada intervalo abierto (z_{j-1}, z_j) tenemos

$$s(x) = c_j, \qquad t(x) = d_j,$$

entonces las operaciones puntuales se convierten en operaciones sobre alturas: $(\alpha s + \beta t)(x) = \alpha c_j + \beta d_j$. Ésta es toda la estructura necesaria para la linealidad.

Proposición 13.3. **Proposición (álgebra y orden de la integral escalonada).** Sean $a < b$ y sean $s, t : [a, b] \to \mathbb{R}$ funciones escalonadas. Sean además $\alpha, \beta \in \mathbb{R}$.

1. La función $\alpha s + \beta t$ es escalonada y
$$\boxed{\int_a^b (\alpha s(x) + \beta t(x))\, dx = \alpha \int_a^b s(x)\, dx + \beta \int_a^b t(x)\, dx.}$$
2. Para toda constante $c \in \mathbb{R}$,
$$\boxed{\int_a^b c\, dx = c(b - a).}$$
3. Si $s(x) \geq 0$ para todo $x \in [a, b]$, entonces
$$\boxed{\int_a^b s(x)\, dx \geq 0.}$$
4. Si
$$s(x) \leq t(x) \qquad \text{para todo } x \in [a, b],$$
entonces
$$\boxed{\int_a^b s(x)\, dx \leq \int_a^b t(x)\, dx.}$$

Demostración. Elijamos particiones P y Q que representen respectivamente a s y a t, y consideremos su refinamiento común $R = P \vee Q = \{z_0, \dots, z_N\}$.

Por Proposición 13.2, R representa a ambas funciones. Escribamos

$$s(x) = c_j, \qquad t(x) = d_j \qquad \text{si } x \in (z_{j-1}, z_j).$$

Entonces $\alpha s + \beta t$ es constante en ese mismo intervalo abierto, con altura $\alpha c_j + \beta d_j$. Por Definición 13.3,

$$\begin{aligned}\int_a^b (\alpha s + \beta t) &= \sum_{j=1}^{N} (\alpha c_j + \beta d_j)(z_j - z_{j-1}) \\ &= \alpha \sum_{j=1}^{N} c_j (z_j - z_{j-1}) + \beta \sum_{j=1}^{N} d_j (z_j - z_{j-1}) \\ &= \alpha \int_a^b s + \beta \int_a^b t.\end{aligned}$$

Esto prueba la primera afirmación. Para la segunda basta representar la función constante $s(x) = c$ mediante la partición $\{a, b\}$. Su única altura es c, luego

$$\int_a^b c\,dx = c(b-a).$$

Supongamos ahora $s \geq 0$ en todo $[a, b]$. En cualquier partición representante, cada altura c_j es no negativa, porque corresponde a los valores de s en un intervalo abierto no vacío. Como también $z_j - z_{j-1} > 0$, cada sumando $c_j(z_j - z_{j-1})$ es no negativo. Su suma también lo es, lo que prueba la tercera afirmación.

Finalmente, si $s \leq t$, entonces $t - s$ es escalonada por la primera parte y satisface $t - s \geq 0$. Por positividad,

$$0 \leq \int_a^b (t - s).$$

Usando linealidad,

$$0 \leq \int_a^b t - \int_a^b s,$$

y por tanto

$$\int_a^b s \leq \int_a^b t.$$

Quedan demostradas las cuatro afirmaciones. □

La prueba muestra por qué el refinamiento común no es un detalle técnico. La linealidad no se obtiene sumando dos representaciones arbitrarias término a término; se obtiene **después** de expresar ambas funciones sobre una misma partición.

Una desigualdad útil: el valor absoluto controla la integral. Si s es escalonada, también lo es $|s|$: en cada franja donde s tiene altura c_k, la función $|s|$ tiene altura $|c_k|$. Como $-|s(x)| \leq s(x) \leq |s(x)|$, la monotonía de Proposición 13.3 proporciona

$$-\int_a^b |s(x)|\,dx \leq \int_a^b s(x)\,dx \leq \int_a^b |s(x)|\,dx.$$

Por la definición de valor absoluto,

$$\boxed{\left|\int_a^b s(x)\,dx\right| \leq \int_a^b |s(x)|\,dx.}$$

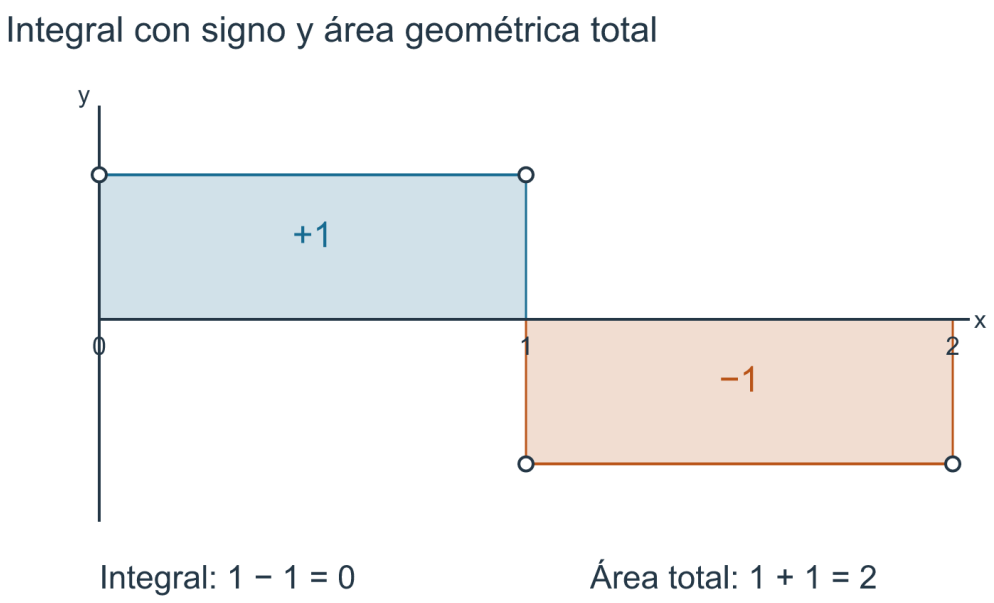


FIGURA 13.8. Las contribuciones $+1$ y -1 se cancelan en la integral, mientras que las dos áreas geométricas se suman.

No necesitamos numerar esta consecuencia como un nuevo teorema: es una combinación directa de linealidad y monotonía.

Integral con signo no significa área geométrica total.

EJEMPLO 13.6. **Cancelación entre franjas positivas y negativas.** Definamos $s : [0, 2] \to \mathbb{R}$ por

$$s(x) = \begin{cases} 1, & 0 < x < 1, \\ -1, & 1 < x < 2, \end{cases}$$

y asignemos valores arbitrarios a $s(0)$, $s(1)$ y $s(2)$.

La partición $P = \{0, 1, 2\}$ representa a s. Entonces

$$\int_0^2 s(x)\,dx = 1 \cdot 1 + (-1) \cdot 1 = 0.$$

Sin embargo, $|s(x)| = 1$ en los dos intervalos abiertos, de modo que

$$\int_0^2 |s(x)|\,dx = 1 \cdot 1 + 1 \cdot 1 = 2.$$

Así, la integral de s es cero por **cancelación de contribuciones con signo**. No expresa el área geométrica total comprendida entre la gráfica y el eje horizontal; esa cantidad corresponde aquí a la integral de $|s|$.

Este ejemplo separa dos conceptos que conviene mantener distintos desde el comienzo:

integral con signo $\neq$ área geométrica total, en general.

Para una función escalonada no negativa ambas interpretaciones coinciden con la suma de áreas rectangulares. Cuando aparecen valores negativos, la definición algebraica sigue siendo perfectamente válida, pero las contribuciones pueden cancelarse.

Cortar el intervalo sin cambiar la suma total. Las propiedades anteriores mantienen fijo el intervalo $[a, b]$. Ahora debemos comparar integrales de la misma función sobre intervalos distintos. Sea $c \in (a, b)$. Aunque c no sea un punto de una partición representante de s, podemos **insertarlo**. El conjunto $Q = P \cup \{c\}$ es un refinamiento de P, y por Proposición 13.2 representa todavía a s. Una vez que c aparece como corte, la suma que define la integral puede separarse exactamente en los términos situados a su izquierda y a su derecha.

Proposición 13.4. **Proposición (aditividad respecto del intervalo).** Sean $a < c < b$ y sea $s : [a, b] \to \mathbb{R}$ una función escalonada. Entonces las restricciones

$$s|_{[a,c]} \qquad \text{y} \qquad s|_{[c,b]}$$

son funciones escalonadas y

$$\boxed{\int_a^b s(x)\,dx = \int_a^c s(x)\,dx + \int_c^b s(x)\,dx.}$$

Demostración. Sea P una partición de $[a, b]$ que representa a s. Formemos $Q = P \cup \{c\}$.

Por Proposición 13.1, Q es una partición de $[a, b]$ y refina a P; por Proposición 13.2, también representa a s. Escribamos sus puntos en orden como

$$a = z_0 < z_1 < \cdots < z_r = c < \cdots < z_N = b.$$

Sean h_j las alturas de s en (z_{j-1}, z_j). Entonces

$$\int_a^b s(x)\,dx = \sum_{j=1}^{N} h_j(z_j - z_{j-1}).$$

Los puntos $\{z_0, \ldots, z_r\}$ forman una partición de $[a, c]$ que representa a $s|_{[a,c]}$, mientras que $\{z_r, \ldots, z_N\}$ forman una partición de $[c, b]$ que representa a $s|_{[c,b]}$. Por Definición 13.3,

$$\int_a^c s(x)\,dx = \sum_{j=1}^{r} h_j(z_j - z_{j-1})$$

y

$$\int_c^b s(x)\,dx = \sum_{j=r+1}^{N} h_j(z_j - z_{j-1}).$$

Sumando las dos expresiones recuperamos exactamente la suma de $j = 1$ a N. Por tanto,

$$\int_a^b s = \int_a^c s + \int_c^b s.$$

□

El valor puntual $s(c)$ no crea un término adicional. En ambas restricciones aparece como valor de borde o punto de corte, pero la integral elemental se construye con franjas de anchura positiva.

Ejemplo 13.7. **Cortar dentro de una franja constante.** Retomemos la función de Ejemplo 13.5, con alturas 2, 1 y 3 sobre

$$(0,1), \qquad (1,3), \qquad (3,4),$$

cuya integral total es 7. Cortemos ahora en $c = \frac{5}{2}$, que no pertenecía a la partición original $\{0,1,3,4\}$. Al insertar $5/2$, la franja de altura 1 sobre $(1,3)$ se divide en dos. A la izquierda,

$$\int_0^{5/2} s(x)\,dx = 2 \cdot 1 + 1\left(\frac{5}{2} - 1\right) = 2 + \frac{3}{2} = \frac{7}{2}.$$

A la derecha,

$$\int_{5/2}^4 s(x)\,dx = 1\left(3 - \frac{5}{2}\right) + 3 \cdot 1 = \frac{1}{2} + 3 = \frac{7}{2}.$$

Por tanto,

$$\int_0^{5/2} s + \int_{5/2}^4 s = \frac{7}{2} + \frac{7}{2} = 7 = \int_0^4 s.$$

El corte adicional no altera la función ni la integral total; solo reparte la misma contribución entre dos intervalos de integración.

Qué no estamos definiendo todavía. La proposición anterior se formuló deliberadamente para $a < c < b$. No hemos fijado todavía convenciones para

$$\int_a^a s \qquad \text{o} \qquad \int_b^a s.$$

Esas convenciones son útiles más adelante, pero no son necesarias para establecer la estructura elemental sobre intervalos no degenerados. Introducirlas aquí mezclaría la propiedad que acabamos de demostrar con una extensión de notación que puede esperar. Tampoco hemos demostrado aún propiedades para funciones generales. Todas las afirmaciones de esta sección pertenecen a la clase de **funciones escalonadas** y descansan en sumas finitas.

Orden y signo: dos advertencias
De $s \le t$ se deduce $\int s \le \int t$, pero una desigualdad estricta en un solo punto no fuerza una desigualdad estricta de integrales: por Corolario 13.1, finitísimas modificaciones puntuales pueden dejar la integral intacta. Asimismo, $\int s = 0$ no implica $s = 0$. Ejemplo 13.6 muestra una cancelación exacta entre una franja positiva y otra negativa.

El mapa estructural construido hasta ahora. Las primeras cinco secciones de la Parte III forman ya una cadena cerrada:

$$\boxed{\begin{array}{c}\text{partición} \longrightarrow \text{función escalonada} \longrightarrow \text{integral elemental bien definida} \\ \longrightarrow \text{linealidad, orden y aditividad por intervalos.}\end{array}}$$

Todo esto se ha obtenido sin definir la integral de Riemann de una función general y sin recurrir a un proceso límite en las propiedades demostradas. Para una función **acotada** $f : [a,b] \to \mathbb{R}$

podemos buscar cotas inferiores y superiores en cada pieza, aunque no sea escalonada. Esas cotas proporcionarán las sumas que la aproximan por abajo y por arriba.

Antes de seguir. ¿Por qué la prueba de linealidad necesita una partición común? ¿Cómo se deduce monotonía de linealidad y positividad? ¿Por qué $\int s = 0$ no implica que s sea la función nula? ¿Qué papel cumple insertar el punto c en la prueba de aditividad por intervalos? ¿Por qué el valor puntual $s(c)$ no aporta un término separado a ninguna de las dos integrales?

13.6. Cotas por abajo y por arriba en una partición

Hasta ahora la función que integrábamos era escalonada: una partición podía capturarla exactamente mediante una altura constante en cada franja. A partir de esta sección cambia el problema. Sea $f : [a, b] \to \mathbb{R}$ una función **acotada**, pero no necesariamente escalonada, continua ni monótona. Fijada una partición, ya no esperamos una única altura exacta en cada subintervalo. En su lugar buscaremos dos números que controlen todos los valores de la función en esa pieza: uno por abajo y otro por arriba. La idea geométrica sigue siendo la de §13.1, pero ahora debemos formularla sin depender de que la función alcance un mínimo o un máximo.

Por qué aparecen ínfimos y supremos. Sea $P = \{x_0, x_1, \ldots, x_n\}$ una partición de $[a, b]$, y escribamos $I_k = [x_{k-1}, x_k]$. Como f está acotada en $[a, b]$, existe algún par de números reales A, B tales que

$$A \le f(x) \le B \qquad \text{para todo } x \in [a, b].$$

Entonces cada conjunto de valores $f(I_k) = \{f(x) : x \in I_k\}$ es no vacío y está acotado inferior y superiormente. La definición de supremo e ínfimo de Definición 1.7, junto con el axioma de completitud de Sección 1.5.2 y la existencia de ínfimos de Proposición 1.10, garantiza entonces la existencia de

$$\text{ínf } f(I_k) \qquad \text{y} \qquad \sup f(I_k).$$

No necesitamos que esos números sean valores efectivamente alcanzados por f.

Definición 13.4. **Sumas inferior y superior respecto de una partición.** Sean $a, b \in \mathbb{R}$ con $a < b$, sea $f : [a, b] \to \mathbb{R}$ una función acotada y sea

$$P = \{x_0, x_1, \ldots, x_n\}$$

una partición de $[a, b]$. Para cada $k \in \{1, \ldots, n\}$ definimos

$$m_k(f, P) := \text{ínf}\{f(x) : x \in [x_{k-1}, x_k]\},$$

y

$$M_k(f, P) := \sup\{f(x) : x \in [x_{k-1}, x_k]\}.$$

La **suma inferior** de f respecto de P es

$$\boxed{L(f, P) := \sum_{k=1}^{n} m_k(f, P)\, \Delta x_k,}$$

mientras que la **suma superior** es

$$\boxed{U(f,P) := \sum_{k=1}^{n} M_k(f,P)\,\Delta x_k.}$$

Estas cantidades dependen de la función **y** de la partición. En este capítulo todavía no las llamaremos integrales inferior o superior globales de f.

La última frase es una frontera importante. $L(f,P)$ y $U(f,P)$ pertenecen a una partición concreta. En C15 estudiaremos qué ocurre al considerar **todas** las particiones y cuándo esas aproximaciones determinan un único número.

Un supremo puede existir aunque no haya máximo.

Ejemplo 13.8. **Una cota superior que no es un valor de la función.** Definamos $f : [0,1] \to \mathbb{R}$ por

$$f(x) = \begin{cases} x, & 0 \le x < 1, \\ 0, & x = 1. \end{cases}$$

Tomemos la partición mínima $P = \{0,1\}$.

En el único subintervalo $I_1 = [0,1]$ tenemos $f(I_1) = [0,1)$, por lo que

$$m_1(f,P) = 0, \qquad M_1(f,P) = 1.$$

El ínfimo sí se alcanza, pero el supremo 1 **no** es valor de f en $[0,1]$. Aun así, las dos sumas están perfectamente definidas:

$$L(f,P) = 0, \qquad U(f,P) = 1.$$

Si hubiéramos exigido un mínimo y un máximo en cada pieza, la construcción fallaría para esta función acotada. El uso de ínfimo y supremo evita esa dependencia de la continuidad o del alcanzamiento de extremos.

Este ejemplo muestra por qué la definición no debe formularse con las palabras «altura mínima» y «altura máxima» sin más. Para funciones continuas en subintervalos cerrados esas alturas sí se alcanzarán por el teorema de valores extremos, pero la presente construcción se aplica a toda función acotada.

Cada partición produce dos funciones escalonadas auxiliares. Las cantidades $m_k(f,P)$ y $M_k(f,P)$ pueden verse como alturas de dos funciones escalonadas. Definamos $\ell_{f,P}, u_{f,P} : [a,b] \to \mathbb{R}$ de la siguiente manera:

- si $x \in (x_{k-1}, x_k)$, ponemos
$$\ell_{f,P}(x) = m_k(f,P), \qquad u_{f,P}(x) = M_k(f,P);$$
- si x es un punto de la partición, ponemos $\ell_{f,P}(x) = f(x) = u_{f,P}(x)$.

Por Definición 13.2, ambas funciones son escalonadas respecto de P. Además, para todo $x \in [a,b]$,

$$\boxed{\ell_{f,P}(x) \le f(x) \le u_{f,P}(x).}$$

En los puntos de corte hay igualdad por definición; en el interior de cada pieza la desigualdad es precisamente la propiedad de ínfimo y supremo. Por Definición 13.3, sus integrales elementales son

$$\int_a^b \ell_{f,P}(x)\,dx = \sum_{k=1}^{n} m_k(f,P)\Delta x_k = L(f,P),$$

y

$$\int_a^b u_{f,P}(x)\,dx = \sum_{k=1}^{n} M_k(f,P)\Delta x_k = U(f,P).$$

Así, las nuevas sumas no aparecen desconectadas de la teoría anterior: son integrales elementales de dos escalonadas construidas a partir de cotas locales de f.

PROPOSICIÓN 13.5. **Proposición (orden básico de las sumas inferior y superior).** Sean $f, g : [a,b] \to \mathbb{R}$ funciones acotadas y sea P una partición de $[a,b]$.

1. Siempre se cumple
$$\boxed{L(f,P) \le U(f,P).}$$
2. Si existen $A, B \in \mathbb{R}$ tales que
$$A \le f(x) \le B \qquad \text{para todo } x \in [a,b],$$
entonces
$$\boxed{A(b-a) \le L(f,P) \le U(f,P) \le B(b-a).}$$
3. Si
$$f(x) \le g(x) \qquad \text{para todo } x \in [a,b],$$
entonces
$$\boxed{L(f,P) \le L(g,P), \qquad U(f,P) \le U(g,P).}$$

Demostración. Para cada k, $m_k(f,P) \le M_k(f,P)$.

Como $\Delta x_k > 0$, multiplicar por la anchura conserva el orden. Sumando sobre k obtenemos $L(f,P) \le U(f,P)$. Si $A \le f \le B$ en todo $[a,b]$, entonces para cada subintervalo

$$A \le m_k(f,P) \le M_k(f,P) \le B.$$

Multiplicando por Δx_k y sumando,

$$A\sum_{k=1}^{n}\Delta x_k \le L(f,P) \le U(f,P) \le B\sum_{k=1}^{n}\Delta x_k.$$

La identidad telescópica de §13.2 da

$$\sum_{k=1}^{n}\Delta x_k = b-a,$$

lo que prueba la segunda afirmación. Por último, si $f \le g$, entonces en cada I_k $\inf f(I_k) \le \inf g(I_k)$ y $\sup f(I_k) \le \sup g(I_k)$. En efecto, todo valor de f está por debajo del valor correspondiente de g en el mismo punto; por tanto cualquier cota inferior de $f(I_k)$ no necesita ser cota inferior de $g(I_k)$, pero el número $\inf f(I_k)$ sí satisface

$$\text{ínf}\, f(I_k) \le f(x) \le g(x) \qquad \text{para todo } x \in I_k,$$

de modo que ínf $f(I_k)$ es una cota inferior de $g(I_k)$ y, por maximalidad del ínfimo, ínf $f(I_k) \le$ ínf $g(I_k)$. Análogamente, $f(x) \le g(x) \le \sup g(I_k)$ muestra que $\sup g(I_k)$ es una cota superior de $f(I_k)$; por minimalidad del supremo, $\sup f(I_k) \le \sup g(I_k)$. Multiplicando por anchuras positivas y sumando se obtienen las dos desigualdades restantes. □

La tercera parte es una versión de orden **a partición fija**. Todavía no afirma nada sobre cómo cambian $L(f,P)$ y $U(f,P)$ al reemplazar P por otra partición. Esa será exactamente la cuestión de §13.7.

Un cálculo explícito en una partición no uniforme. Volvamos a la función

$$f(x) = x, \qquad 0 \le x \le 1,$$

pero tomemos ahora

$$P = \left\{0, \frac{1}{4}, \frac{3}{4}, 1\right\}.$$

Como f es creciente, en cada pieza el ínfimo es el valor del extremo izquierdo y el supremo el del extremo derecho. Así,

$$\begin{aligned} L(f,P) &= 0 \cdot \frac{1}{4} + \frac{1}{4} \cdot \frac{1}{2} + \frac{3}{4} \cdot \frac{1}{4} \\ &= \frac{1}{8} + \frac{3}{16} = \frac{5}{16}, \end{aligned}$$

mientras que

$$\begin{aligned} U(f,P) &= \frac{1}{4} \cdot \frac{1}{4} + \frac{3}{4} \cdot \frac{1}{2} + 1 \cdot \frac{1}{4} \\ &= \frac{1}{16} + \frac{3}{8} + \frac{1}{4} = \frac{11}{16}. \end{aligned}$$

Por tanto,

$$\boxed{L(f,P) = \frac{5}{16} < \frac{1}{2} < \frac{11}{16} = U(f,P).}$$

El valor $1/2$ aparece aquí solo porque conocemos previamente el área del triángulo y podemos usarla como referencia, exactamente como en §13.1. Las sumas $L(f,P)$ y $U(f,P)$ se han obtenido sin usar ninguna integral de f.

Los valores en los cortes vuelven a importar para las cotas de una partición. En §13.4 demostramos que modificar finitísimos valores de una función escalonada no cambia su **integral elemental**. No debe concluirse que tales valores sean invisibles para toda construcción posterior.

Ejemplo 13.9. **Una punta aislada que sí afecta una suma superior fija.** Definamos $s : [0,1] \to \mathbb{R}$ por

$$s(x) = \begin{cases} 1, & x = \frac{1}{2}, \\ 0, & x \neq \frac{1}{2}. \end{cases}$$

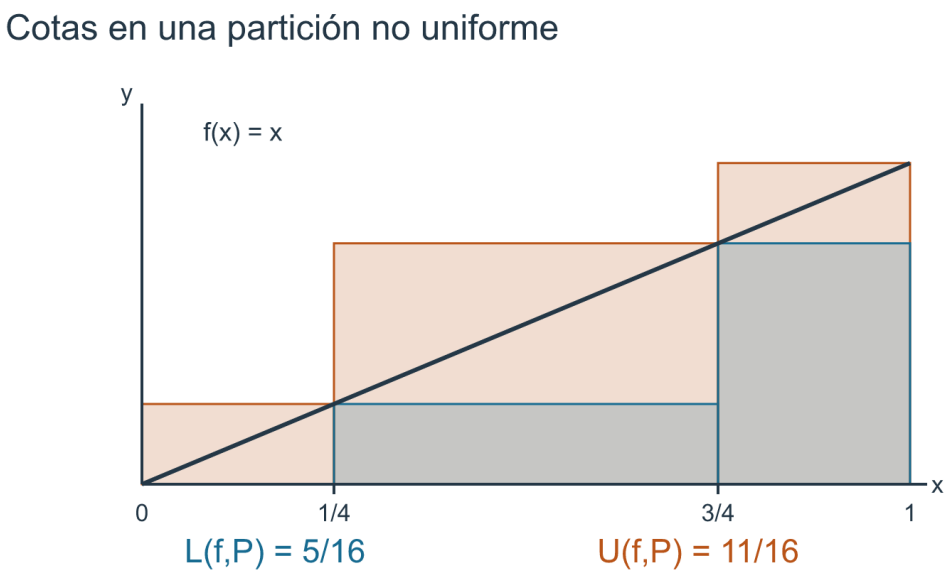


FIGURA 13.9. Para $P = \{0, 1/4, 3/4, 1\}$, las cotas son $L(f, P) = 5/16$ y $U(f, P) = 11/16$.

La función es escalonada y, por Corolario 13.1,

$$\int_0^1 s(x)\,dx = 0.$$

Tomemos, sin embargo, la partición

$$P = \left\{0, \frac{1}{2}, 1\right\}.$$

En ambos subintervalos cerrados,

$$\left[0, \frac{1}{2}\right] \qquad \text{y} \qquad \left[\frac{1}{2}, 1\right],$$

el ínfimo vale 0 y el supremo vale 1. Por tanto, $L(s, P) = 0$ y $U(s, P) = 1 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = 1$. La integral elemental de s es 0, pero la suma superior respecto de esta partición es 1. Si $0 < \delta < 1/2$ y añadimos los cortes $1/2 - \delta$ y $1/2 + \delta$, obtenemos

$$Q_\delta = \left\{0, \frac{1}{2} - \delta, \frac{1}{2}, \frac{1}{2} + \delta, 1\right\}.$$

Los únicos subintervalos donde el supremo sigue siendo 1 son los dos adyacentes a $1/2$, ambos de anchura δ. Así, $U(s, Q_\delta) = 2\delta$. Este cálculo anticipa el efecto que formalizaremos en §13.7: al refinar una partición podemos mejorar las cotas. Aquí la mejora puede hacerse tan pronunciada como queramos alrededor de un único valor excepcional.

La diferencia con §13.4 es exacta. La integral elemental de una escalonada utiliza sus **alturas en intervalos abiertos** y no asigna anchura a un punto aislado. En cambio, $m_k(f, P)$ y $M_k(f, P)$

se calculan sobre los **subintervalos cerrados** $[x_{k-1}, x_k]$, de modo que un valor situado en un corte puede influir en las cotas de las piezas vecinas. No hay contradicción: son construcciones distintas con objetivos distintos.

Cotas locales frente a información global. Una partición produce una lista finita de intervalos $I_1, \dots, I_n$, y, sobre cada uno, un par $m_k(f,P) \le M_k(f,P)$. La diferencia $M_k(f,P) - m_k(f,P)$ mide cuánto espacio vertical queda sin resolver en esa pieza. Todavía no sumaremos sistemáticamente estas diferencias; la identidad global correspondiente se reservará para §13.9. Por ahora basta observar que $L(f,P)$ usa todas las cotas inferiores locales, mientras que $U(f,P)$ usa todas las superiores. La información de la partición es, pues, un **encierro finito** de la función por dos escalonadas.

Una suma inferior o superior no es todavía la integral
Para una partición fija P, los números $L(f,P)$ y $U(f,P)$ son dos aproximaciones construidas con ínfimos y supremos locales. No hemos definido aún un número $\int_a^b f$ para una función acotada general. Tampoco basta que una partición particular produzca una brecha pequeña. Antes debemos comprender cómo se comportan estas cantidades al **refinar** la partición y cómo se comparan sumas provenientes de particiones distintas.

Lo que queda preparado para el refinamiento. La cadena conceptual disponible es ahora

$$\boxed{\begin{array}{c} f \text{ acotada} + P \\ \Downarrow \\ m_k = \inf f(I_k), \quad M_k = \sup f(I_k) \\ \Downarrow \\ L(f,P) \le U(f,P). \end{array}}$$

La siguiente pregunta es inevitable: si añadimos cortes y reemplazamos P por un refinamiento Q, ¿las cotas mejoran en el sentido esperado?

En §13.7 demostraremos que sí:

$$L(f,P) \le L(f,Q) \qquad \text{y} \qquad U(f,Q) \le U(f,P).$$

Esa comparación no será una consecuencia de tener simplemente una malla menor. Necesitará la relación estructural exacta $P \preceq Q$ desarrollada en §13.2.

Antes de seguir. ¿Por qué la acotación global de f garantiza que todos los m_k y M_k existen como números reales? ¿Por qué se usan ínfimos y supremos en lugar de mínimos y máximos? ¿Cómo se construyen las escalonadas $\ell_{f,P}$ y $u_{f,P}$? ¿Por qué $L(f,P) \le U(f,P)$? ¿Cómo puede un único valor excepcional alterar $U(f,P)$ y, sin embargo, no alterar la integral elemental de una escalonada?

13.7. Refinar mejora las cotas

En §13.6 fijamos una partición P y construimos dos números, $L(f,P) \le U(f,P)$, que encierran a la función mediante dos escalonadas auxiliares. Falta todavía una condición de coherencia decisiva: si añadimos puntos de corte, la información no debería empeorar. Una descomposición más fina debe permitir cotas inferiores al menos tan altas y cotas superiores al menos tan

bajas como antes. Esta afirmación parece evidente en un dibujo, pero la prueba debe salir exclusivamente del orden de los conjuntos de valores. El mecanismo elemental será el siguiente:

$$J \subseteq I \quad \Longrightarrow \quad \text{ínf}\, f(I) \leq \text{ínf}\, f(J) \leq \sup f(J) \leq \sup f(I).$$

Al pasar de una pieza grande I a piezas más pequeñas J, el conjunto sobre el que tomamos el ínfimo o el supremo se reduce. Por eso el ínfimo puede subir y el supremo puede bajar. El refinamiento convierte esta observación local en una desigualdad entre sumas completas.

La comparación local dentro de una pieza. Sean $P = \{x_0, x_1, \ldots, x_n\}$ y Q una partición que refina a P. Fijemos un subintervalo $I_k = [x_{k-1}, x_k]$ de P. Los puntos de Q que pertenecen a I_k pueden escribirse como

$$x_{k-1} = y_{k,0} < y_{k,1} < \cdots < y_{k,r_k} = x_k.$$

Así, Q descompone I_k en los intervalos

$$J_{k,j} = [y_{k,j-1}, y_{k,j}], \qquad 1 \leq j \leq r_k.$$

Como $J_{k,j} \subseteq I_k$, se tiene

$$m_k(f, P) \leq \text{ínf}\, f(J_{k,j}) \leq \sup f(J_{k,j}) \leq M_k(f, P).$$

Además, las anchuras de las piezas nuevas recuperan exactamente la anchura antigua:

$$\sum_{j=1}^{r_k} (y_{k,j} - y_{k,j-1}) = x_k - x_{k-1} = \Delta x_k.$$

Estas dos observaciones contienen toda la prueba del resultado fundamental de la sección.

Proposición 13.6. **Proposición (monotonía de las sumas bajo refinamiento).** Sea $f : [a, b] \to \mathbb{R}$ una función acotada y sean P, Q particiones de $[a, b]$. Si

$$P \preceq Q,$$

entonces

$$\boxed{L(f, P) \leq L(f, Q) \leq U(f, Q) \leq U(f, P).}$$

En particular, al refinar una partición la suma inferior no disminuye y la suma superior no aumenta.

Demostración. Conservemos la notación anterior. Para cada pieza I_k de P y cada subpieza $J_{k,j}$ de Q contenida en ella, escribamos

$$m_{k,j}^{Q} := \text{ínf}\, f(J_{k,j}), \qquad M_{k,j}^{Q} := \sup f(J_{k,j}).$$

Entonces

$$m_k(f, P) \leq m_{k,j}^{Q} \leq M_{k,j}^{Q} \leq M_k(f, P).$$

Como todas las anchuras $y_{k,j} - y_{k,j-1}$ son positivas, podemos multiplicar sin alterar el sentido de las desigualdades y sumar sobre j:

$$\begin{aligned} m_k(f,P)\sum_{j=1}^{r_k}(y_{k,j}-y_{k,j-1}) &\le \sum_{j=1}^{r_k} m_{k,j}^{Q}(y_{k,j}-y_{k,j-1}) \\ &\le \sum_{j=1}^{r_k} M_{k,j}^{Q}(y_{k,j}-y_{k,j-1}) \\ &\le M_k(f,P)\sum_{j=1}^{r_k}(y_{k,j}-y_{k,j-1}). \end{aligned}$$

El telescopado de las anchuras convierte los extremos en $m_k(f,P)\Delta x_k$ y $M_k(f,P)\Delta x_k$. Por tanto, dentro de cada pieza de P,

$$\begin{aligned} m_k(f,P)\Delta x_k &\le \sum_{j=1}^{r_k} m_{k,j}^{Q}\,(y_{k,j}-y_{k,j-1}) \\ &\le \sum_{j=1}^{r_k} M_{k,j}^{Q}\,(y_{k,j}-y_{k,j-1}) \\ &\le M_k(f,P)\Delta x_k. \end{aligned}$$

Finalmente sumamos estas desigualdades para $k=1,\dots,n$. Todas las subpiezas de Q aparecen exactamente una vez, de modo que obtenemos

$$L(f,P)\le L(f,Q)\le U(f,Q)\le U(f,P).$$

Esto prueba la proposición. □

La palabra «mejora» debe interpretarse en sentido no estricto. Un refinamiento puede dejar una de las sumas sin cambios. Si f ya es constante sobre una pieza de P, subdividir esa pieza no aporta nueva información vertical y la contribución correspondiente permanece igual. Más generalmente, para una cadena finita de refinamientos $P_0 \preceq P_1 \preceq \cdots \preceq P_r$, la proposición produce dos cadenas ordenadas en sentidos opuestos:

$$L(f,P_0)\le L(f,P_1)\le \cdots \le L(f,P_r),$$

y

$$U(f,P_r)\le \cdots \le U(f,P_1)\le U(f,P_0).$$

El refinamiento comprime, por tanto, el intervalo numérico de incertidumbre entre la cota inferior y la superior.

Ejemplo 13.10. **Refinar una partición para $f(x)=x$.** En $[0,1]$ consideremos $f(x)=x$ y las particiones

$$P=\left\{0,\frac{1}{2},1\right\},$$

$$Q=\left\{0,\frac{1}{4},\frac{1}{2},\frac{3}{4},1\right\}.$$

Como $P\subset Q$, tenemos $P\preceq Q$.

Para P, $L(f,P)=0\cdot\frac{1}{2}+\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{4}$, mientras que $U(f,P)=\frac{1}{2}\cdot\frac{1}{2}+1\cdot\frac{1}{2}=\frac{3}{4}$.

Para Q,

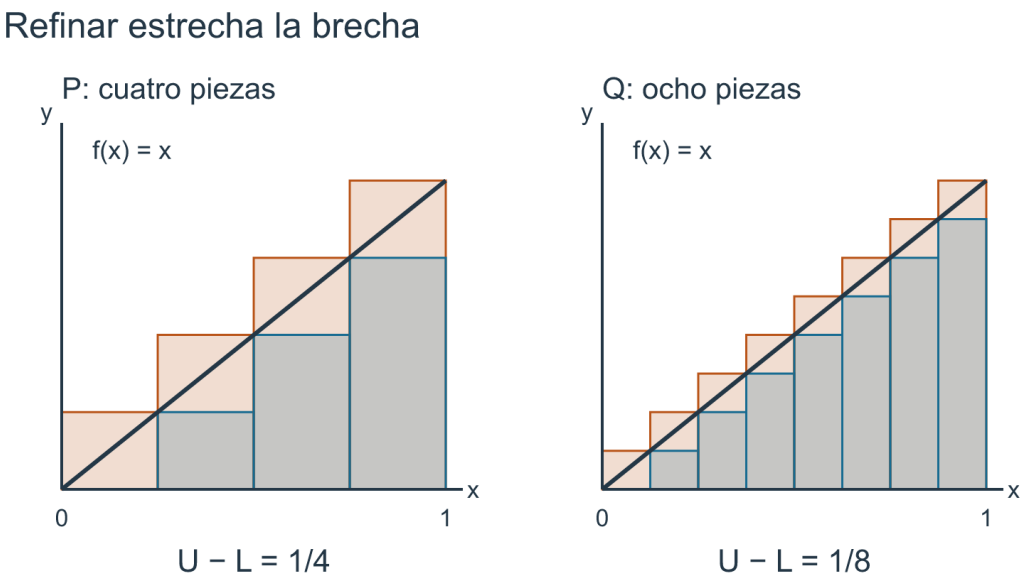


FIGURA 13.10. Al pasar de cuatro a ocho piezas uniformes, la brecha de las sumas de la identidad disminuye de 1/4 a 1/8.

$$\begin{aligned} L(f,Q) &= 0\cdot\frac{1}{4}+\frac{1}{4}\cdot\frac{1}{4}+\frac{1}{2}\cdot\frac{1}{4}+\frac{3}{4}\cdot\frac{1}{4} \\ &= \frac{6}{16}=\frac{3}{8}, \end{aligned}$$

y

$$\begin{aligned} U(f,Q) &= \frac{1}{4}\cdot\frac{1}{4}+\frac{1}{2}\cdot\frac{1}{4}+\frac{3}{4}\cdot\frac{1}{4}+1\cdot\frac{1}{4} \\ &= \frac{10}{16}=\frac{5}{8}. \end{aligned}$$

Por tanto,

$$\boxed{\frac{1}{4}<\frac{3}{8}<\frac{5}{8}<\frac{3}{4}.}$$

La cota inferior sube de 1/4 a 3/8 y la superior baja de 3/4 a 5/8. La mejora no depende de una intuición gráfica: es una instancia concreta de Proposición 13.6.

Dos particiones no necesitan ser comparables. La proposición anterior compara directamente P y Q solo cuando una refina a la otra. Pero en §13.2 vimos que dos particiones cualesquiera admiten el refinamiento común $P \vee Q = P \cup Q$. Esa construcción permite comparar una suma inferior construida con una partición y una suma superior construida con otra completamente distinta.

COROLARIO 13.2. **Corolario (toda suma inferior queda bajo toda suma superior).** Sea $f : [a, b] \to \mathbb{R}$ una función acotada. Si P y Q son dos particiones cualesquiera de $[a, b]$, entonces

$$\boxed{L(f, P) \le U(f, Q).}$$

Intercambiando los papeles de P y Q también se obtiene

$$L(f, Q) \le U(f, P).$$

Demostración. Sea $R = P \vee Q$.

Por Proposición 13.1, R refina tanto a P como a Q. Aplicando Proposición 13.6 dos veces, $L(f, P) \le L(f, R)$ y $U(f, R) \le U(f, Q)$. Además, por la desigualdad básica de Proposición 13.5, $L(f, R) \le U(f, R)$.

Encadenando,

$$\boxed{L(f, P) \le L(f, R) \le U(f, R) \le U(f, Q).}$$

Esto demuestra la primera afirmación; la segunda se obtiene intercambiando P y Q. □

Este corolario es más fuerte que $L(f, P) \le U(f, P)$. No afirma solamente que una partición produzca cotas compatibles consigo misma, sino que **ninguna suma inferior, cualquiera sea la partición que la origine, puede sobrepasar ninguna suma superior obtenida de otra partición**. Esta compatibilidad global es exactamente lo que necesitábamos antes de poder hablar, en el capítulo siguiente, de familias completas de sumas. Aquí nos detenemos antes de tomar supremos o ínfimos sobre todas las particiones.

Una malla menor no sustituye al refinamiento. Proposición 13.6 depende de la inclusión $P \subseteq Q$, no de la desigualdad entre mallas. Ya vimos en §13.2 que $\|Q\| < \|P\|$ no implica $P \preceq Q$. Para las sumas inferior y superior, esta distinción tiene consecuencias cuantitativas reales.

EJEMPLO 13.11. **Malla menor, pero peor cota superior.** Consideremos nuevamente la función con una punta aislada

$$s(x) = \begin{cases} 1, & x = \frac{1}{2}, \\ 0, & x \neq \frac{1}{2}, \end{cases} \qquad 0 \le x \le 1.$$

Tomemos

$$P = \left\{0, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, 1\right\}.$$

Su malla es $\|P\| = \frac{2}{5}$.

El punto $1/2$ pertenece a los dos subintervalos cerrados adyacentes

$$\left[\frac{2}{5}, \frac{1}{2}\right], \qquad \left[\frac{1}{2}, \frac{3}{5}\right],$$

cada uno de anchura $1/10$. En ellos el supremo de s es 1; en las otras piezas es 0. Por tanto, $U(s, P) = \frac{1}{10} + \frac{1}{10} = \frac{1}{5}$.

Ahora consideremos

$$R = \left\{0, \frac{1}{3}, \frac{2}{3}, 1\right\}.$$

Entonces

$$\|R\| = \frac{1}{3} < \frac{2}{5} = \|P\|,$$

pero R no refina a P. El punto $1/2$ queda en el interior de la pieza $[1/3, 2/3]$, cuyo supremo es 1, mientras que las otras dos piezas tienen supremo 0. Así, $U(s, R) = \frac{1}{3}$.

En consecuencia,

$$\boxed{\|R\| < \|P\| \qquad \text{pero} \qquad U(s, R) = \frac{1}{3} > \frac{1}{5} = U(s, P).}$$

Una malla menor puede coincidir con una peor cota superior si los nuevos cortes no conservan la estructura de la partición anterior. La monotonía demostrada en Proposición 13.6 pertenece al **refinamiento**, no al mero tamaño de la malla.

Qué hemos ganado sin definir todavía la integral. Disponemos ahora de tres niveles de comparación:

1. para una partición fija, $L(f, P) \le U(f, P)$;
2. si $P \preceq Q$,
$$L(f, P) \le L(f, Q) \le U(f, Q) \le U(f, P);$$
3. para dos particiones arbitrarias P y Q, $L(f, P) \le U(f, Q)$.

El segundo nivel muestra que refinar comprime las cotas. El tercero muestra que las construcciones provenientes de particiones incompatibles siguen siendo mutuamente coherentes gracias al refinamiento común.

La comparación global no es todavía una definición de integral
Hemos probado que toda suma inferior queda por debajo de toda suma superior. Todavía no hemos definido el supremo de todas las sumas inferiores, el ínfimo de todas las superiores ni una condición de igualdad entre ambos números. Esas operaciones pertenecen a C15, donde se definirá la integrabilidad de Riemann. En este capítulo seguimos trabajando únicamente con sumas finitas asociadas a particiones concretas.

Al elegir un punto de cada subintervalo y evaluar allí la función obtenemos una **partición etiquetada** y su suma de Riemann. La comparación con ínfimos y supremos encerrará esa suma entre las cotas correspondientes a la misma partición.

Antes de seguir. ¿Por qué $J \subseteq I$ hace subir el ínfimo y bajar el supremo? ¿Dónde se usa la positividad de las anchuras en la demostración de Proposición 13.6? ¿Por qué el refinamiento común permite comparar sumas construidas con particiones no comparables? ¿Qué demuestra Ejemplo 13.11 sobre la diferencia entre malla pequeña y refinamiento? ¿Qué paso sigue faltando antes de poder definir una integral para una función acotada general?

13.8. Particiones etiquetadas y sumas de Riemann

Las sumas inferior y superior de §§13.6–13.7 utilizan, en cada subintervalo, las mejores cotas disponibles por abajo y por arriba. Hay otra construcción igualmente natural. En vez de preguntar por **todos** los valores de f en una pieza, podemos escoger un punto concreto de esa pieza, evaluar allí la función y utilizar ese valor como altura de un rectángulo.

Esta elección recupera una idea familiar de los primeros ejemplos de integración: rectángulos con altura tomada en el extremo izquierdo, en el extremo derecho, en el punto medio o en

cualquier otro punto permitido. La diferencia es que ahora la elección será parte explícita de los datos y no una convención oculta.

Una etiqueta en cada subintervalo. Sea $P = \{x_0, x_1, \dots, x_n\}$ una partición de $[a, b]$. Para cada pieza $I_k = [x_{k-1}, x_k]$ elegimos un punto $\xi_k \in I_k$. El punto ξ_k será la **etiqueta** de ese subintervalo. Las etiquetas se eligen independientemente: pueden ser extremos, puntos interiores e incluso un mismo punto de corte puede servir como etiqueta de las dos piezas adyacentes.

Definición 13.5. **Partición etiquetada y suma de Riemann.** Sean $a, b \in \mathbb{R}$ con $a < b$ y sea

$$P = \{x_0, x_1, \dots, x_n\}$$

una partición de $[a, b]$. Una **elección de etiquetas** para P es una n-upla

$$\xi = (\xi_1, \dots, \xi_n)$$

tal que

$$\xi_k \in [x_{k-1}, x_k] \qquad (1 \le k \le n).$$

El par (P, ξ) se llama **partición etiquetada** de $[a, b]$. Su malla es, por definición, la malla de la partición subyacente:

$$\|(P, \xi)\| := \|P\|.$$

Si $f : [a, b] \to \mathbb{R}$, la **suma de Riemann** de f asociada a (P, ξ) es

$$\boxed{R(f; P, \xi) := \sum_{k=1}^{n} f(\xi_k)\, \Delta x_k.}$$

La suma anterior es finita. Por ello, para **definirla** no necesitamos suponer que f sea acotada: cada $f(\xi_k)$ es un número real y aparecen solamente finitísimos términos. La acotación sí será necesaria cuando queramos compararla con $L(f, P)$ y $U(f, P)$, porque esas dos cantidades fueron definidas en §13.6 para funciones acotadas. La malla tampoco depende de las etiquetas. Dos particiones etiquetadas pueden tener exactamente los mismos puntos de corte y, por tanto, la misma malla, aunque sus sumas de Riemann sean diferentes. La geometría horizontal está codificada por P; las alturas elegidas están codificadas por ξ.

Qué representa cada término. El término $f(\xi_k)\Delta x_k$ es el producto de una altura efectivamente tomada por la función en el subintervalo y la anchura de esa pieza. Si $f \ge 0$, puede interpretarse como el área de un rectángulo. Si $f(\xi_k) < 0$, el término es negativo y la lectura correcta es algebraica: la suma es una combinación **con signo** de contribuciones rectangulares. Esta observación evita una identificación prematura. Una suma de Riemann no es, por definición, un «área aproximada» no negativa. Es un número construido mediante valores de la función y longitudes de subintervalos.

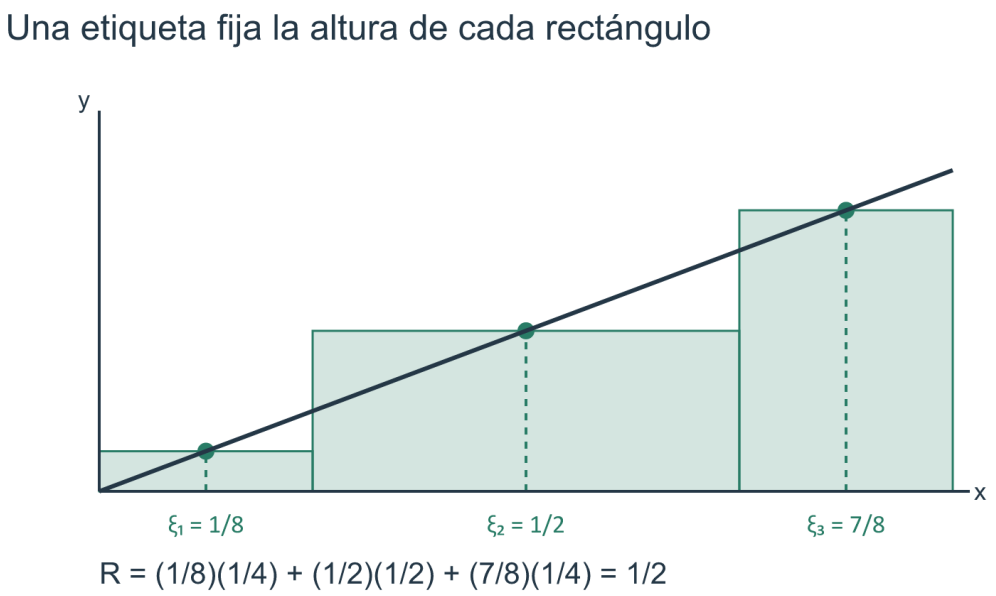


FIGURA 13.11. Los puntos medios $1/8, 1/2, 7/8$ determinan las alturas de tres rectángulos sobre la misma partición no uniforme.

Las sumas inferior y superior encierran cualquier elección de etiquetas. Para una función acotada ya conocemos, sobre cada pieza I_k,

$$m_k(f,P) = \text{ínf}\, f(I_k), \qquad M_k(f,P) = \sup f(I_k).$$

Como la etiqueta ξ_k pertenece precisamente a I_k, su valor queda atrapado entre esas dos cotas: $m_k(f,P) \le f(\xi_k) \le M_k(f,P)$. La positividad de las anchuras transforma inmediatamente este control local en un encierro para la suma completa.

PROPOSICIÓN 13.7. **Proposición (encierro de toda suma de Riemann).** Sea $f : [a,b] \to \mathbb{R}$ una función acotada y sea (P,ξ) una partición etiquetada de $[a,b]$. Entonces

$$\boxed{L(f,P) \le R(f;P,\xi) \le U(f,P).}$$

Demostración. Para cada k tenemos $m_k(f,P) \le f(\xi_k) \le M_k(f,P)$.
Como $\Delta x_k > 0$, podemos multiplicar por la anchura sin invertir el orden:

$$m_k(f,P)\Delta x_k \le f(\xi_k)\Delta x_k \le M_k(f,P)\Delta x_k.$$

Sumando para $k = 1, \dots, n$ obtenemos

$$\sum_{k=1}^{n} m_k(f,P)\Delta x_k \le \sum_{k=1}^{n} f(\xi_k)\Delta x_k \le \sum_{k=1}^{n} M_k(f,P)\Delta x_k.$$

Por Definición 13.4 y Definición 13.5, los tres miembros son, respectivamente,

$$L(f,P), \qquad R(f;P,\xi), \qquad U(f,P).$$

Esto prueba la afirmación. □

El resultado tiene una lectura importante. Una suma de Riemann utiliza una muestra finita de valores reales de f; las sumas inferior y superior utilizan información de orden sobre **todos** los valores de cada pieza. Por eso las segundas funcionan como barreras para cualquier elección posible de etiquetas sobre la misma partición.

Extremos izquierdos, puntos medios y extremos derechos son casos de una misma construcción.

Ejemplo 13.12. **Tres elecciones de etiquetas para una misma partición.** Consideremos

$$f(x) = x, \qquad 0 \le x \le 1,$$

y la partición no uniforme

$$P = \left\{0, \frac{1}{4}, \frac{3}{4}, 1\right\}.$$

En §13.6 calculamos

$$L(f,P) = \frac{5}{16}, \qquad U(f,P) = \frac{11}{16}.$$

Primero elijamos los extremos izquierdos:

$$\xi^- = \left(0, \frac{1}{4}, \frac{3}{4}\right).$$

Entonces

$$\begin{aligned} R(f;P,\xi^-) &= 0 \cdot \frac{1}{4} + \frac{1}{4} \cdot \frac{1}{2} + \frac{3}{4} \cdot \frac{1}{4} \\ &= \frac{5}{16}. \end{aligned}$$

Para esta función creciente, los extremos izquierdos realizan los ínfimos de cada pieza; por eso aquí la suma coincide con $L(f,P)$. Tomemos ahora los puntos medios de los tres subintervalos:

$$\xi^0 = \left(\frac{1}{8}, \frac{1}{2}, \frac{7}{8}\right).$$

La suma correspondiente es

$$\begin{aligned} R(f;P,\xi^0) &= \frac{1}{8} \cdot \frac{1}{4} + \frac{1}{2} \cdot \frac{1}{2} + \frac{7}{8} \cdot \frac{1}{4} \\ &= \frac{1}{32} + \frac{8}{32} + \frac{7}{32} \\ &= \frac{1}{2}. \end{aligned}$$

Finalmente, con los extremos derechos

$$\xi^+ = \left(\frac{1}{4}, \frac{3}{4}, 1\right),$$

obtenemos

$$\begin{aligned} R(f;P,\xi^{+}) &= \frac{1}{4}\cdot\frac{1}{4}+\frac{3}{4}\cdot\frac{1}{2}+1\cdot\frac{1}{4} \\ &= \frac{11}{16}. \end{aligned}$$

Así aparecen, sobre una misma partición, tres sumas distintas:

$$\boxed{\frac{5}{16}=R(f;P,\xi^{-})<R(f;P,\xi^{0})=\frac{1}{2}<R(f;P,\xi^{+})=\frac{11}{16}.}$$

En particular, $L(f,P)\le R(f;P,\xi)\le U(f,P)$ se hace visible en tres elecciones concretas. Las coincidencias con las sumas inferior y superior ocurren aquí porque $f(x)=x$ alcanza sus extremos locales en los extremos de cada subintervalo; no debemos suponer que esto ocurra para toda función acotada.

La última advertencia enlaza con Ejemplo 13.8. Allí el supremo de f sobre $[0,1]$ era 1, pero ese valor no era alcanzado. Con la partición $P=\{0,1\}$, la suma superior vale 1, mientras que para toda etiqueta $\xi_1\in[0,1]$ se cumple

$$R(f;P,\xi)=f(\xi_1)<1=U(f,P).$$

Por tanto, una suma superior no tiene por qué ser ella misma una suma de Riemann obtenida mediante una elección de etiquetas. Es una **cota** construida con supremos.

La misma partición puede producir sumas muy diferentes. El hecho de fijar P no fija la suma. Las etiquetas forman parte esencial de la construcción.

Ejemplo 13.13. **Un mismo corte, dos elecciones extremas de etiquetas.** Recuperemos la función

$$s(x)=\begin{cases}1, & x=\frac{1}{2},\\ 0, & x\neq\frac{1}{2},\end{cases}\qquad 0\le x\le 1,$$

y la partición

$$P=\left\{0,\frac{1}{2},1\right\}.$$

En §13.6 vimos que

$$L(s,P)=0,\qquad U(s,P)=1.$$

Elijamos primero etiquetas que eviten el punto excepcional:

$$\eta=\left(\frac{1}{4},\frac{3}{4}\right).$$

Como $s(1/4)=s(3/4)=0$, $R(s;P,\eta)=0$.

Ahora elijamos

$$\xi=\left(\frac{1}{2},\frac{1}{2}\right).$$

Esta elección es válida: $1/2$ pertenece tanto a $[0, 1/2]$ como a $[1/2, 1]$, y las etiquetas de piezas distintas no tienen que ser distintas. Entonces

$$\begin{aligned} R(s; P, \xi) &= s\left(\frac{1}{2}\right)\frac{1}{2} + s\left(\frac{1}{2}\right)\frac{1}{2} \\ &= 1. \end{aligned}$$

Por tanto, con **la misma partición y la misma malla** obtenemos

$$\boxed{R(s; P, \eta) = 0, \qquad R(s; P, \xi) = 1.}$$

Ambos valores respetan Proposición 13.7:

$$0 = L(s, P) \leq R(s; P, \eta), R(s; P, \xi) \leq U(s, P) = 1.$$

El ejemplo muestra que la malla controla el tamaño horizontal de las piezas, pero no decide qué valores de la función serán muestreados dentro de ellas.

No hay contradicción con la integral elemental de la función escalonada s, que en §13.6 recordamos que vale 0. La integral elemental y una suma de Riemann asociada a etiquetas concretas son construcciones diferentes. Una puede ignorar cambios en puntos de corte porque esos puntos no tienen anchura en la representación escalonada; la otra evalúa deliberadamente la función en puntos escogidos, y por eso puede detectar esos valores excepcionales.

Una muestra finita no debe confundirse con una definición integral. Para una partición etiquetada concreta, $R(f; P, \xi)$ es solamente un número obtenido mediante un cálculo finito. Incluso una larga colección de tales números no define todavía la integral de una función general. Hay dos fuentes independientes de variación:

- podemos cambiar los puntos de corte de P;
- aun con P fija, podemos cambiar las etiquetas ξ_k.

Proposición 13.7 muestra que, cuando f es acotada, **todas** las elecciones de etiquetas para una partición fija quedan encerradas en el mismo intervalo numérico $[L(f, P), U(f, P)]$.

§13.7 demostró, además, que el refinamiento hace que esas barreras se acerquen en sentido de orden. Estas dos observaciones explican por qué las sumas etiquetadas son compatibles con la teoría construida hasta ahora, pero todavía no especifican qué cuantificadores deben exigirse para que exista un único valor integral.

Una suma de Riemann no es todavía la integral de Riemann
El número $R(f; P, \xi)$ pertenece a una partición y a una elección concreta de etiquetas. No definiremos en este capítulo la integral diciendo simplemente «tomamos una partición muy fina» ni declararemos suficiente la convergencia de una familia particular de sumas. En C15 formularemos la condición que hace independiente el valor límite de las particiones y etiquetas admisibles —y demostraremos su relación con las cotas inferiores y superiores—. Aquí solo necesitamos la estructura finita.

Lo que queda preparado para medir la incertidumbre. Disponemos ahora de tres alturas posibles sobre cada pieza:

$$m_k(f, P), \qquad f(\xi_k), \qquad M_k(f, P),$$

ordenadas por $m_k(f,P) \le f(\xi_k) \le M_k(f,P)$.

Al sumar sus contribuciones obtenemos

$$\boxed{L(f,P) \le R(f;P,\xi) \le U(f,P).}$$

La diferencia $U(f,P) - L(f,P)$ mide la incertidumbre que deja el encierro de una partición. Podemos expresarla mediante las oscilaciones de f en sus subintervalos.

Antes de seguir. ¿Qué información adicional aporta una etiqueta respecto de una partición sin etiquetas? ¿Por qué la malla no cambia al modificar ξ? ¿Dónde se usa que $\xi_k \in [x_{k-1}, x_k]$ en la prueba de Proposición 13.7? ¿Por qué una suma superior puede no ser una suma de Riemann efectivamente alcanzable? ¿Qué demuestra Ejemplo 13.13 acerca de dos elecciones de etiquetas con la misma malla? ¿Qué condición falta todavía para pasar de sumas finitas a una integral de una función general?

13.9. La brecha de una partición y la oscilación

En §13.8 vimos que, para una función acotada y una partición fija P, toda suma de Riemann queda encerrada entre dos barreras: $L(f,P) \le R(f;P,\xi) \le U(f,P)$. Conocer el encierro no basta. También queremos medir **cuán ancho es**. Si las dos barreras están muy separadas, las etiquetas todavía pueden producir resultados muy distintos; si están próximas, todas las sumas etiquetadas construidas sobre esa partición quedan obligadas a concentrarse en un intervalo numérico estrecho.

La cantidad natural es $U(f,P) - L(f,P)$.

Esta diferencia se descompone exactamente en contribuciones locales. La pieza que corresponde al subintervalo $I_k = [x_{k-1}, x_k]$ depende de cuánto puede variar f dentro de él.

Oscilación sobre un conjunto. Sea $f : [a,b] \to \mathbb{R}$ acotada y sea $I \subseteq [a,b]$ no vacío. En esta sección llamaremos **oscilación de f sobre I** al número

$$\omega(f;I) := \sup\{f(x) : x \in I\} - \text{ínf}\{f(x) : x \in I\}.$$

Como f es acotada, ambos extremos existen como números reales. Además, $\omega(f;I) \ge 0$. La oscilación mide la amplitud vertical de los valores sobre el conjunto considerado. Una función constante tiene oscilación cero cualquiera que sea su altura; una función con valores positivos y negativos puede tener oscilación grande aunque sus sumas presenten cancelaciones. Para una partición $P = \{x_0, x_1, \dots, x_n\}$ escribiremos abreviadamente $\omega_k(f,P) := \omega(f;[x_{k-1},x_k])$. Por Definición 13.4, $\omega_k(f,P) = M_k(f,P) - m_k(f,P)$. La oscilación depende, pues, de la función **y del subintervalo que estamos mirando**. Al refinar una partición, las nuevas piezas pueden tener oscilaciones menores porque se examina a f sobre conjuntos más pequeños.

La brecha es la suma ponderada de las oscilaciones.

Proposición 13.8. **Proposición (brecha de una partición y control de las sumas etiquetadas).** Sea $f : [a,b] \to \mathbb{R}$ una función acotada y sea

$$P = \{x_0, x_1, \dots, x_n\}$$

una partición de $[a,b]$. Entonces:

1. la brecha entre la suma superior y la inferior satisface

$$\boxed{U(f,P) - L(f,P) = \sum_{k=1}^{n} \omega_k(f,P)\,\Delta x_k;}$$

2. si ξ y η son dos elecciones cualesquiera de etiquetas para la misma partición P, entonces

$$\boxed{|R(f;P,\xi) - R(f;P,\eta)| \le U(f,P) - L(f,P);}$$

3. si $P \preceq Q$, entonces

$$\boxed{0 \le U(f,Q) - L(f,Q) \le U(f,P) - L(f,P).}$$

Así, la brecha cuantifica simultáneamente la incertidumbre local acumulada de la partición y un límite superior para la discrepancia entre cualesquiera dos sumas etiquetadas construidas sobre ella.

Demostración. Por Definición 13.4,

$$U(f,P) = \sum_{k=1}^{n} M_k(f,P)\Delta x_k$$

y

$$L(f,P) = \sum_{k=1}^{n} m_k(f,P)\Delta x_k.$$

Restando término a término,

$$\begin{aligned} U(f,P) - L(f,P) &= \sum_{k=1}^{n} \big(M_k(f,P) - m_k(f,P)\big)\Delta x_k \\ &= \sum_{k=1}^{n} \omega_k(f,P)\Delta x_k. \end{aligned}$$

Esto prueba la primera afirmación. Para la segunda, Proposición 13.7 da $L(f,P) \le R(f;P,\xi) \le U(f,P)$ y $L(f,P) \le R(f;P,\eta) \le U(f,P)$.

Por tanto,

$$R(f;P,\xi) - R(f;P,\eta) \le U(f,P) - L(f,P),$$

y, al intercambiar ξ y η,

$$R(f;P,\eta) - R(f;P,\xi) \le U(f,P) - L(f,P).$$

Las dos desigualdades equivalen a

$$|R(f;P,\xi) - R(f;P,\eta)| \le U(f,P) - L(f,P).$$

Finalmente, si $P \preceq Q$, Proposición 13.6 proporciona

$$L(f,P) \le L(f,Q) \le U(f,Q) \le U(f,P).$$

Como $L(f,Q) \ge L(f,P)$ y $U(f,Q) \le U(f,P)$,

$$U(f,Q) - L(f,Q) \le U(f,P) - L(f,P).$$

La no negatividad de ambas brechas sigue de Proposición 13.5. □

La primera identidad contiene una lectura que conviene conservar. Cada término $\omega_k(f,P)\Delta x_k$ combina dos fuentes de incertidumbre:

- una **amplitud vertical**, $\omega_k(f,P)$;
- una **anchura horizontal**, Δx_k.

Una oscilación grande puede tener una contribución pequeña si está confinada a una franja muy estrecha. Del mismo modo, una pieza ancha puede ser inocua si la función varía muy poco en ella.

Brecha pequeña significa acuerdo forzado, no integral ya definida. La segunda parte de Proposición 13.8 permite interpretar el número $U(f,P) - L(f,P)$ sin hablar todavía de integrabilidad. Si, por ejemplo, $U(f,P) - L(f,P) < 10^{-4}$, entonces **cualesquiera dos** elecciones de etiquetas sobre esa misma partición producen sumas que difieren en menos de 10^{-4}. Esto es más fuerte que calcular dos elecciones favorables y observar que están cerca. El encierro controla simultáneamente todas las etiquetas posibles para P. Pero la afirmación sigue siendo local a una partición concreta. No hemos demostrado que exista una partición con brecha arbitrariamente pequeña para toda función acotada; tampoco hemos definido qué conclusión general debería extraerse si tales particiones existen. Esos cuantificadores pertenecen a C15.

Una función suave en un ejemplo concreto: x^2. La continuidad y la monotonía pueden hacer que las oscilaciones locales se reduzcan al estrechar las piezas. En lugar de formular todavía un teorema general, veamos el mecanismo en una función que podemos calcular exactamente.

Ejemplo 13.14. **La brecha de $f(x) = x^2$ sobre particiones uniformes.** Sea

$$f(x) = x^2, \qquad 0 \le x \le 1,$$

y, para $n \ge 1$, consideremos la partición uniforme

$$P_n = \left\{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1\right\}.$$

Su malla es $\|P_n\| = \frac{1}{n}$.

Como x^2 es creciente en $[0,1]$, sobre

$$I_k = \left[\frac{k-1}{n}, \frac{k}{n}\right]$$

tenemos

$$m_k = \left(\frac{k-1}{n}\right)^2, \qquad M_k = \left(\frac{k}{n}\right)^2.$$

Por tanto,

$$\omega_k = M_k - m_k = \frac{k^2 - (k-1)^2}{n^2} = \frac{2k-1}{n^2}.$$

Como $\Delta x_k = 1/n$, Proposición 13.8 da

$$U(f,P_n)-L(f,P_n)=\sum_{k=1}^{n}\frac{2k-1}{n^2}\frac{1}{n}$$
$$=\frac{1}{n^3}\sum_{k=1}^{n}(2k-1).$$

La suma de los primeros n números impares es n^2, así que

$$\boxed{U(f,P_n)-L(f,P_n)=\frac{1}{n}.}$$

En consecuencia, para cualesquiera dos elecciones de etiquetas ξ y η sobre P_n, $|R(f;P_n,\xi)-R(f;P_n,\eta)|\le\frac{1}{n}$. La malla y la brecha coinciden numéricamente en este ejemplo: $\|P_n\|=U(f,P_n)-L(f,P_n)=\frac{1}{n}$. No convertiremos esta igualdad particular en una regla universal. Procede de la forma concreta de x^2 y de la partición uniforme elegida. Tampoco concluimos aquí, mediante una definición todavía no introducida, que x^2 sea integrable. Lo que hemos demostrado es exactamente esto: existe una familia explícita de particiones para la cual el intervalo que encierra todas las sumas etiquetadas tiene anchura $1/n$.

El cálculo permite ver el papel de la variación local. Cerca de 0, las diferencias entre cuadrados consecutivos son pequeñas; hacia 1, son mayores. La suma pondera todas esas oscilaciones por el ancho $1/n$ y el telescopado algebraico final produce una brecha controlada.

Una discontinuidad aislada también puede quedar confinada. La punta de Ejemplo 13.9 ofrece un contraste útil. Para

$$s(x)=\begin{cases}1, & x=\frac{1}{2},\\ 0, & x\neq\frac{1}{2},\end{cases}$$

y la partición

$$Q_\delta=\left\{0,\frac{1}{2}-\delta,\frac{1}{2},\frac{1}{2}+\delta,1\right\},\qquad 0<\delta<\frac{1}{2},$$

la oscilación es 1 únicamente en los dos intervalos adyacentes a $1/2$ y es 0 en las demás piezas. Por tanto,

$$U(s,Q_\delta)-L(s,Q_\delta)=1\cdot\delta+1\cdot\delta=2\delta.$$

Aquí no hemos reducido la amplitud vertical del salto puntual: sigue siendo 1. Lo que hemos reducido es la **anchura total de las piezas en las que esa oscilación aparece**. Esta distinción es importante. Para disminuir la brecha no hay una única estrategia local: puede hacerse pequeña la oscilación en muchas piezas, o puede confinarse una oscilación grande a piezas de anchura total pequeña.

Una función cuya oscilación no mejora con ningún corte finito. El caso opuesto muestra por qué una malla pequeña, por sí sola, no obliga a que la brecha sea pequeña.

Ejemplo 13.15. **Oscilación máxima en cada subintervalo.** Consideremos nuevamente la función de Ejemplo 13.4,

$$d(x) = \begin{cases} 1, & x \in \mathbb{Q}, \\ 0, & x \notin \mathbb{Q}, \end{cases} \qquad 0 \le x \le 1.$$

Sea $P = \{x_0, \dots, x_n\}$ una partición **arbitraria** de $[0,1]$. Cada subintervalo $I_k = [x_{k-1}, x_k]$ tiene longitud positiva. Su interior contiene racionales e irracionales por la densidad establecida en Sección 1.8. En consecuencia, d toma allí los dos valores 0 y 1, y por tanto

$$m_k(d,P) = 0, \qquad M_k(d,P) = 1, \qquad \omega_k(d,P) = 1$$

para todo k.

La identidad de Proposición 13.8 produce entonces

$$\begin{aligned} U(d,P) - L(d,P) &= \sum_{k=1}^{n} 1 \cdot \Delta x_k \\ &= \sum_{k=1}^{n} \Delta x_k \\ &= 1. \end{aligned}$$

Por tanto,

$$\boxed{U(d,P) - L(d,P) = 1 \qquad \text{para toda partición } P \text{ de } [0,1].}$$

Podemos hacer la malla tan pequeña como queramos y añadir tantos cortes finitos como deseemos; la oscilación dentro de **cada** pieza no disminuye. La partición cambia, pero la brecha permanece igual a la longitud total del intervalo. No traduciremos todavía este hecho a un dictamen sobre integrabilidad. En C15, una vez fijada la definición general, podremos decir exactamente qué consecuencia tiene esta persistencia de la brecha.

Los ejemplos Ejemplo 13.14 y Ejemplo 13.15 muestran que la expresión

$$\sum_{k} \omega_k \Delta x_k$$

contiene más información que la malla aislada. Dos particiones pueden tener piezas horizontalmente muy pequeñas y, sin embargo, comportarse de manera radicalmente distinta según cuánto varíe la función dentro de ellas.

El refinamiento redistribuye la incertidumbre. Cuando $P \preceq Q$, cada pieza de Q está contenida en una pieza de P. La oscilación sobre un subconjunto no puede exceder la oscilación sobre el conjunto original:

$$J \subseteq I \quad \Longrightarrow \quad \omega(f;J) \le \omega(f;I).$$

Sin embargo, no conviene intentar probar la disminución global de la brecha comparando solamente estas oscilaciones una por una: al refinar también cambian el número y las anchuras de las piezas. Proposición 13.6 ya resolvió correctamente el problema global y Proposición 13.8 lo traduce en

$$U(f,Q) - L(f,Q) \le U(f,P) - L(f,P).$$

La interpretación mediante oscilaciones explica **por qué** esa mejora tiene sentido; la prueba rigurosa sigue descansando en las desigualdades de refinamiento ya establecidas.

Brecha pequeña: una información finita, no todavía un criterio de integrabilidad
En este capítulo podemos afirmar rigurosamente:

- qué brecha produce una partición concreta;
- cómo se descompone en oscilaciones locales;
- que el refinamiento no aumenta esa brecha;
- que la brecha controla la diferencia entre todas las sumas etiquetadas sobre la misma partición.

Todavía no formulamos la condición «para todo $\varepsilon > 0$ existe una partición…», ni la relacionamos con una integral superior, una integral inferior o un valor integral único. Ese paso es exactamente uno de los problemas de C15.

Hacia el laboratorio final. La arquitectura finita del capítulo está casi completa. Para una función acotada disponemos ahora de

$$\boxed{\begin{array}{c}\text{partición } P\\ \Downarrow\\ m_k,\ M_k,\ \omega_k\\ \Downarrow\\ L(f,P) \le R(f;P,\xi) \le U(f,P)\\ \Downarrow\\ U(f,P) - L(f,P) = \displaystyle\sum_k \omega_k \Delta x_k.\end{array}}$$

§13.10 reunirá estas construcciones en un solo laboratorio: funciones escalonadas, cotas de Darboux y sumas etiquetadas. El objetivo no será definir anticipadamente la integral general, sino identificar con precisión qué problema lógico queda todavía abierto antes de que podamos hablar de un candidato único independiente del modo de aproximación.

Antes de seguir. ¿Por qué $\omega(f;I)$ puede ser cero aunque f tome valores grandes? ¿Cómo se obtiene la identidad de la brecha a partir de Definición 13.4? ¿Por qué esa brecha controla la diferencia entre dos sumas etiquetadas sobre la misma partición? En Ejemplo 13.14, ¿qué propiedad de x^2 permite calcular exactamente cada oscilación? ¿Por qué en Ejemplo 13.15 podemos hacer la malla tan pequeña como queramos sin que ninguna oscilación local disminuya? ¿Qué cuantificador deliberadamente seguimos sin introducir antes de C15?

13.10. Laboratorio: de rectángulos a un candidato único

Las secciones anteriores introdujeron varios objetos que pueden parecer variantes de una misma idea. No lo son. Cada uno responde a una pregunta diferente y conserva una cantidad distinta de información. Este laboratorio tiene un objetivo preciso: **reconstruir la arquitectura finita de C14 sin adelantar la definición general de integral**. En particular, no buscaremos todavía un criterio de integrabilidad. Nos limitaremos a comparar los objetos disponibles, detectar inferencias inválidas y aislar el problema que queda abierto para C15.

Estación 1. Cuatro objetos, cuatro papeles. Sea $f : [a, b] \to \mathbb{R}$ acotada y sea $P = \{x_0, \dots, x_n\}$ una partición. Conviene separar los siguientes objetos.

1. Una función escalonada. Si una función s es constante entre los cortes de alguna partición, su integral elemental

$$\int_a^b s(x)\,dx$$

ya está definida por Definición 13.3. No depende de la partición representante, por Teorema 13.1. Aquí disponemos de un número exacto porque la función pertenece a una clase finita especial.

2. La suma inferior. Para una función acotada general,

$$L(f, P) = \sum_{k=1}^{n} m_k(f, P)\Delta x_k$$

usa los ínfimos locales. Es una cota construida con toda la información de orden disponible en cada pieza.

3. La suma superior. Del mismo modo,

$$U(f, P) = \sum_{k=1}^{n} M_k(f, P)\Delta x_k$$

usa los supremos locales. Junto con la suma inferior forma un encierro asociado a una partición concreta.

4. Una suma de Riemann. Una elección de etiquetas $\xi_k \in [x_{k-1}, x_k]$ produce

$$R(f; P, \xi) = \sum_{k=1}^{n} f(\xi_k)\Delta x_k.$$

Esta suma utiliza una muestra finita de valores efectivamente tomados por f. Las relaciones demostradas hasta ahora pueden condensarse en

$$\boxed{L(f, P) \le R(f; P, \xi) \le U(f, P).}$$

Pero las tres cantidades no tienen el mismo estatuto. L y U dependen de P; R depende además de ξ; la integral elemental solo está definida, por ahora, para funciones escalonadas. La expresión «suma de Darboux» se usará en este laboratorio como nombre colectivo para $L(f, P)$ y $U(f, P)$. No introduce un objeto nuevo: son exactamente las sumas inferior y superior de Definición 13.4.

Estación 2. Una misma función vista por tres lentes. Consideremos

$$f(x) = x^2, \qquad 0 \le x \le 1,$$

y la partición

$$P = \left\{0, \frac{1}{2}, 1\right\}.$$

Como x^2 es creciente en $[0, 1]$,

$$m_1 = 0, \qquad M_1 = \frac{1}{4},$$

y

$$m_2 = \frac{1}{4}, \qquad M_2 = 1.$$

Por tanto, $L(f,P) = 0 \cdot \frac{1}{2} + \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8}$, mientras que $U(f,P) = \frac{1}{4} \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{5}{8}$.
La brecha es $U(f,P) - L(f,P) = \frac{1}{2}$.
Ahora tomemos las etiquetas de punto medio

$$\xi = \left(\frac{1}{4}, \frac{3}{4}\right).$$

Entonces

$$\begin{aligned} R(f;P,\xi) &= \left(\frac{1}{4}\right)^2 \frac{1}{2} + \left(\frac{3}{4}\right)^2 \frac{1}{2} \\ &= \frac{1}{32} + \frac{9}{32} \\ &= \frac{5}{16}. \end{aligned}$$

Y efectivamente, $\frac{1}{8} < \frac{5}{16} < \frac{5}{8}$.
Nada de esto define todavía $\int_0^1 x^2\,dx$ como integral de Riemann. Lo que sí demuestra es que una partición concreta produce un intervalo de incertidumbre

$$\left[\frac{1}{8}, \frac{5}{8}\right]$$

que contiene toda suma etiquetada sobre P.
Refinemos ahora a

$$Q = \left\{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\right\}.$$

Como $P \preceq Q$, Proposición 13.6 garantiza que el encierro mejora. El cálculo directo da

$$L(f,Q) = \left(0 + \frac{1}{16} + \frac{1}{4} + \frac{9}{16}\right) \frac{1}{4} = \frac{7}{32},$$

y

$$U(f,Q) = \left(\frac{1}{16} + \frac{1}{4} + \frac{9}{16} + 1\right) \frac{1}{4} = \frac{15}{32}.$$

Por tanto,

$$\boxed{\left[\frac{7}{32}, \frac{15}{32}\right] \subset \left[\frac{1}{8}, \frac{5}{8}\right].}$$

La anchura del nuevo intervalo es $\frac{15}{32} - \frac{7}{32} = \frac{1}{4}$, la mitad de la anterior. Si usamos los puntos medios de las cuatro piezas,

$$\eta = \left(\frac{1}{8}, \frac{3}{8}, \frac{5}{8}, \frac{7}{8}\right),$$

obtenemos

$$\begin{aligned} R(f;Q,\eta) &= \left(\frac{1}{64}+\frac{9}{64}+\frac{25}{64}+\frac{49}{64}\right)\frac{1}{4} \\ &= \frac{84}{256} \\ &= \frac{21}{64}. \end{aligned}$$

Y nuevamente

$$\frac{7}{32} = \frac{14}{64} < \frac{21}{64} < \frac{30}{64} = \frac{15}{32}.$$

El punto central no es el valor numérico obtenido, sino la estructura:

$$\boxed{P \preceq Q \quad \Longrightarrow \quad [L(f,Q),U(f,Q)] \subseteq [L(f,P),U(f,P)].}$$

Esta inclusión es solo una reescritura de Proposición 13.6.

Estación 3. Una escalonada exacta puede tener una brecha de Darboux positiva. Conviene comparar ahora dos construcciones que se parecen mucho pero no coinciden.

Definamos $s : [0,2] \to \mathbb{R}$ por

$$s(x) = \begin{cases} 1, & 0 \le x < 1, \\ 2, & x = 1, \\ 3, & 1 < x \le 2. \end{cases}$$

La función es escalonada respecto de $P = \{0,1,2\}$. Su integral elemental está completamente definida y vale

$$\int_0^2 s(x)\,dx = 1 \cdot 1 + 3 \cdot 1 = 4.$$

El valor aislado $s(1) = 2$ no aparece en esta suma, porque la integral elemental usa las alturas constantes de los intervalos abiertos. Pero las sumas de Darboux se calculan sobre los subintervalos **cerrados**. En $[0,1]$ los valores relevantes son 1 y 2, mientras que en $[1,2]$ aparecen 2 y 3. Por tanto, $L(s,P) = 1 \cdot 1 + 2 \cdot 1 = 3$, y $U(s,P) = 2 \cdot 1 + 3 \cdot 1 = 5$.

Así,

$$\boxed{L(s,P) = 3 < \int_0^2 s(x)\,dx = 4 < 5 = U(s,P).}$$

No hay contradicción. La integral elemental de una escalonada es un número exacto ya bien definido por §13.4; las sumas de Darboux pertenecen a una partición fija y son cotas construidas con valores sobre intervalos cerrados. Incluso sobre esta misma partición las sumas etiquetadas pueden tomar distintos valores. Si elegimos

$$\xi = \left(\frac{1}{2}, \frac{3}{2}\right),$$

entonces $R(s;P,\xi) = 1 + 3 = 4$.

Si elegimos $\eta = (1,1)$, entonces $R(s;P,\eta) = 2 + 2 = 4$.

Pero también son válidas, por ejemplo,

$$\alpha = \left(\frac{1}{2}, 1\right), \qquad \beta = \left(1, \frac{3}{2}\right),$$

que producen $R(s; P, \alpha) = 1 + 2 = 3$, y $R(s; P, \beta) = 2 + 3 = 5$.

Toda la familia queda encerrada, como exige Proposición 13.7, entre 3 y 5. Ahora añadamos cortes a distancia δ del salto:

$$Q_\delta = \{0, 1 - \delta, 1, 1 + \delta, 2\}, \qquad 0 < \delta < 1.$$

La oscilación es cero en las piezas exteriores y vale 1 en los dos intervalos adyacentes a 1. Por Proposición 13.8, $U(s, Q_\delta) - L(s, Q_\delta) = \delta + \delta = 2\delta$. La integral elemental continúa siendo exactamente 4, mientras que el encierro de Darboux se estrecha al confinar la irregularidad puntual. Este ejemplo separa cuatro afirmaciones:

- «s es escalonada»;
- «$\int s$ está definida elementalmente»;
- «$L(s, P)$ y $U(s, P)$ dependen de P»;
- «las sumas etiquetadas pueden variar aun cuando la función tenga ya una integral elemental».

Confundir estas afirmaciones produce varios de los argumentos circulares que debemos evitar.

Estación 4. Intervalos de incertidumbre y refinamiento común. Para una función acotada f y una partición P, escribamos de manera puramente auxiliar $J(f, P) := [L(f, P), U(f, P)]$. No estamos definiendo una integral ni una nueva noción matemática sustantiva; solo damos nombre al intervalo numérico que ya existe. Por Proposición 13.7, toda suma etiquetada sobre P pertenece a $J(f, P)$: $R(f; P, \xi) \in J(f, P)$. Si $P \preceq Q$, entonces Proposición 13.6 implica

$$\boxed{J(f, Q) \subseteq J(f, P).}$$

Y si P y Q no son comparables, el refinamiento común $R = P \vee Q$ satisface

$$\boxed{J(f, R) \subseteq J(f, P) \cap J(f, Q).}$$

En efecto,

$$L(f, P) \le L(f, R) \le U(f, R) \le U(f, P),$$

y lo mismo respecto de Q.

Esta propiedad es central. Las particiones no producen intervalos de incertidumbre arbitrariamente incompatibles. Frente a dos de ellos siempre podemos construir un tercero, obtenido por refinamiento común, que queda contenido en ambos. La arquitectura disponible puede resumirse así:

$$\boxed{\begin{array}{c} P,\ Q \\ \Downarrow \\ R = P \vee Q \\ \Downarrow \\ J(f, R) \subseteq J(f, P) \cap J(f, Q). \end{array}}$$

Todavía no tomamos la intersección de **todos** esos intervalos ni sus extremos globales. Esa operación pertenece a C15.

Estación 5. Reparar seis argumentos incorrectos. El objetivo de este bloque es distinguir una intuición plausible de una demostración válida.

Argumento incorrecto A. «Las sumas de puntos medios para f parecen converger a un número, luego ese número es la integral.»

La observación solo controla una familia particular de particiones y una regla particular de etiquetas. No demuestra independencia respecto de otras particiones o etiquetas.

Reparación. En C14 solo podemos afirmar lo demostrado para esa familia. La condición que permitirá pasar de una familia especial a un valor independiente del procedimiento pertenece a C15.

Argumento incorrecto B. «Si $\|P\|$ es pequeña, entonces $U(f,P) - L(f,P)$ es pequeña.»

Ejemplo 13.15 muestra que esto es falso para una función que oscila entre 0 y 1 en todo subintervalo no degenerado: la malla puede hacerse arbitrariamente pequeña y la brecha seguir siendo 1.

Reparación. La brecha depende de la combinación

$$\sum_k \omega_k \Delta x_k,$$

no de la malla aislada.

Argumento incorrecto C. «Si encuentro una partición con una brecha muy pequeña, ya he definido la integral.»

Una partición con brecha pequeña obliga a que todas sus sumas etiquetadas estén próximas entre sí, por Proposición 13.8. Pero C14 todavía no ha definido la integral de una función acotada general.

Reparación. Registrar exactamente la conclusión disponible: esa partición produce un encierro estrecho. El salto desde encierros finitos a un valor integral general exige una definición adicional.

Argumento incorrecto D. «La suma superior siempre se obtiene escogiendo en cada pieza una etiqueta donde f alcanza el máximo.»

Esto falla cuando el supremo no se alcanza, como en Ejemplo 13.8.

Reparación. $U(f,P)$ usa supremos, no necesariamente máximos. Es una cota y no tiene por qué ser una suma etiquetada realizable.

Argumento incorrecto E. «Una partición con malla menor es automáticamente un refinamiento.»

Ejemplo 13.2 y Ejemplo 13.11 refutan esta identificación.

Reparación. Verificar la inclusión de los puntos: $P \preceq Q \iff P \subseteq Q$.

La desigualdad de mallas es solo una consecuencia del refinamiento, no una caracterización.

Argumento incorrecto F. «Si s es escalonada, cualquier suma inferior o superior respecto de una partición representante coincide con su integral elemental.»

La estación 3 acaba de mostrar que los valores en los cortes pueden modificar ínfimos y supremos sobre subintervalos cerrados, aunque no cambien la integral elemental.

Reparación. Separar la representación escalonada —constancia en intervalos abiertos— de las cotas de Darboux —ínfimos y supremos sobre intervalos cerrados—.

Estación 6. Qué significa realmente «un candidato único». Las construcciones de C14 proporcionan una red de información finita.

Para cada partición P obtenemos $J(f,P) = [L(f,P), U(f,P)]$.

Dentro de ese intervalo viven todas las sumas etiquetadas de la misma partición: $R(f;P,\xi) \in J(f,P)$. Al refinar, los intervalos no se ensanchan:

$$P \preceq Q \quad \Longrightarrow \quad J(f,Q) \subseteq J(f,P).$$

Y para dos particiones arbitrarias existe un refinamiento común cuyo intervalo está contenido en ambos. Por tanto, la pregunta pendiente ya puede formularse con precisión conceptual, aun sin responderla:

> ¿Cuándo esta familia compatible de encierros finitos obliga a que todas las aproximaciones admisibles señalen un mismo número real, independientemente de cómo se elijan los cortes y las etiquetas?

Para responder necesitaremos una definición que reúna todas las aproximaciones compatibles y un criterio que garantice que determinan un solo valor. No basta con afirmar que «los rectángulos se hacen pequeños». Tampoco basta con observar una secuencia convergente. Lo que falta es un mecanismo que convierta la **compatibilidad de todas las aproximaciones** en la existencia de un valor único.

Estación 7. El mapa completo de C14. Podemos reconstruir todo el capítulo mediante una cadena de dependencias:

área y exhaución como motivación

$\Downarrow$

particiones, malla y refinamiento

$\Downarrow$

funciones escalonadas

$\Downarrow$

integral elemental bien definida

$\Downarrow$

sumas inferior y superior

$\Downarrow$

monotonía bajo refinamiento

$\Downarrow$

particiones etiquetadas y sumas de Riemann

$\Downarrow$

$$L(f,P) \le R(f;P,\xi) \le U(f,P)$$

$\Downarrow$

$$U(f,P) - L(f,P) = \sum_k \omega_k(f,P)\Delta x_k$$

$\Downarrow$

pregunta por un valor único independiente de la aproximación.

La última línea es una **pregunta**, no un teorema de este capítulo.

Frontera exacta entre C14 y C15
C14 termina habiendo construido y comparado todos los objetos finitos necesarios: funciones escalonadas, sumas inferior y superior, refinamientos, sumas etiquetadas y brechas. C15 comenzará cuando se formule qué significa que estas aproximaciones determinen un único número para una función acotada general. Allí aparecerán los cuantificadores globales, las integrales inferior y superior y la definición de integrabilidad de Riemann. No trasladamos esa definición hacia atrás.

Antes del banco de ejercicios. Los ejercicios ponen en relación los objetos construidos: sumas, particiones, refinamientos y brechas. Habrá que calcular y comparar encierros, reconocer qué información reduce la incertidumbre y detectar los argumentos que presuponen el valor integral que aún se busca definir. La integral elemental de una escalonada y las sumas de Darboux asociadas a una partición concreta son objetos distintos. Refinar coordina los encierros, mientras que reducir la malla exige otro tipo de control. La convergencia de una sola familia etiquetada deja abierta la compatibilidad con todas las demás; esa será la dificultad que deberá resolver la definición de integrabilidad.

13.11. Ejercicios y soluciones

El banco recorre la arquitectura completa del capítulo sin cruzar la frontera de C15. Los cuarenta ejercicios son originales y siguen la distribución fijada en el diseño fino:

$$6A + 6B + 6C + 6D + 6E + 6F + 4G = 40.$$

Cada enunciado tiene una solución desarrollada con el mismo número. Solo se utilizan definiciones y resultados disponibles hasta §13.10. En particular, ningún ejercicio presupone la definición general de integrabilidad de Riemann, integrales inferior o superior globales, el criterio global de la brecha, el teorema fundamental del cálculo ni técnicas de integración.

Nivel A — Particiones, refinamientos y malla. Ejercicio A1. Leer una partición completa. Sea

$$P = \left\{0, \frac{1}{5}, \frac{2}{5}, 1\right\}.$$

1. Escribe sus subintervalos y sus anchuras.
2. Calcula $\|P\|$.
3. Verifica directamente que la suma de las anchuras es la longitud de $[0, 1]$.
4. Explica por qué la lista $\{0, 1/5, 2/5\}$ no es una partición de $[0, 1]$.

Ejercicio A2. Refinamiento, malla y no comparabilidad. Considera

$$P = \left\{0, \frac{1}{2}, 1\right\}, \qquad Q = \left\{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\right\},$$

y

$$R = \left\{0, \frac{1}{3}, \frac{2}{3}, 1\right\}.$$

Determina cuáles de las relaciones $P \preceq Q$, $P \preceq R$, $R \preceq Q$ son verdaderas. Calcula las tres mallas y los refinamientos comunes $P \vee R$ y $Q \vee R$.

Ejercicio A3. Un refinamiento común no uniforme. En $[0, 2]$ sean

$$P = \left\{0, \frac{1}{2}, \frac{3}{2}, 2\right\}, \qquad Q = \left\{0, 1, \frac{7}{4}, 2\right\}.$$

Construye $P \vee Q$, calcula las anchuras de todas sus piezas y su malla. Comprueba que $\|P \vee Q\| \leq \min\{\|P\|, \|Q\|\}$.

Ejercicio A4. Refinar no siempre reduce estrictamente la malla. Sea

$$P = \left\{0, \frac{1}{2}, \frac{3}{4}, 1\right\}.$$

1. Añade el punto $1/8$ y calcula la malla de la nueva partición.
2. Añade, en cambio, el punto $7/8$ a P y calcula la nueva malla.
3. Explica por qué ambos son refinamientos estrictos de P, aunque solo uno reduce estrictamente la malla.

Ejercicio A5. La malla no puede ser menor que la anchura media. Sea $P = \{x_0, \ldots, x_n\}$ una partición de $[a, b]$ con exactamente n subintervalos. Demuestra que

$$\boxed{\|P\| \geq \frac{b-a}{n}.}$$

Demuestra además que hay igualdad si y solo si todos los subintervalos tienen la misma longitud.

Ejercicio A6. Tener más piezas no significa refinar. Considera

$$P = \left\{0, \frac{1}{2}, 1\right\}$$

y

$$Q = \left\{0, \frac{1}{3}, \frac{2}{3}, \frac{5}{6}, 1\right\}.$$

1. Comprueba que Q tiene más subintervalos y menor malla que P.
2. Demuestra que Q no refina a P.
3. Construye $P \vee Q$ y calcula su malla.
4. Identifica exactamente qué información aporta «tener más piezas», cuál aporta «tener menor malla» y cuál aporta «ser refinamiento».

Nivel B — Funciones escalonadas e integral elemental. Ejercicio B1. Una suma rectangular con signos. Sea $s : [0, 4] \to \mathbb{R}$ escalonada con alturas

$$2, \quad -1, \quad 3$$

respectivamente sobre

$$(0, 1), \qquad (1, 3), \qquad (3, 4),$$

y valores arbitrarios en $0, 1, 3, 4$.

1. Calcula $\int_0^4 s(x)\,dx$.
2. Repite el cálculo usando el refinamiento

$$Q = \left\{0, \frac{1}{2}, 1, 2, 3, \frac{7}{2}, 4\right\}.$$

3. Explica qué resultado general garantiza que ambos cálculos coincidan.

Ejercicio B2. Los valores de corte pertenecen a la función, pero no a la suma rectangular. Sea s la función del ejercicio B1. Define t igual a s fuera de $F = \{0, 1, 3, 4\}$, pero asigna a t valores distintos de los de s en los cuatro puntos de F.

1. Justifica que t sigue siendo escalonada.
2. Demuestra que

$$\int_0^4 t(x)\,dx = \int_0^4 s(x)\,dx.$$

3. Explica por qué esto no afirma que $s = t$.

Ejercicio B3. Reconocer una función escalonada. Define $u : [0, 2] \to \mathbb{R}$ por

$$u(x) = \begin{cases} -2, & 0 \le x < 1, \\ 7, & x = 1, \\ 1, & 1 < x \le 2. \end{cases}$$

1. Exhibe una partición que represente a u.
2. Calcula su integral elemental.
3. Da un refinamiento que también la represente y verifica la misma integral mediante ese refinamiento.

Ejercicio B4. Dos representaciones y un refinamiento común. Sea $s : [0, 3] \to \mathbb{R}$ una función que vale -2 en $(0, 1)$ y 1 en $(1, 3)$, con valores arbitrarios en los cortes. Considera

$$P = \{0, 1, 3\}, \qquad Q = \left\{0, \frac{1}{3}, 1, 2, 3\right\}.$$

1. Comprueba que ambas particiones representan a s.
2. Calcula $\mathcal{R}(s; P)$ y $\mathcal{R}(s; Q)$.
3. Explica cómo Lema 13.1 organiza la igualdad obtenida.

Ejercicio B5. Integral nula sin función nula. Sea $s : [0,3] \to \mathbb{R}$ escalonada con altura 2 en $(0,1)$ y altura -1 en $(1,3)$.

1. Calcula $\int_0^3 s$.
2. Calcula $\int_0^3 |s|$.
3. Explica por qué el primer resultado no implica que s sea la función nula.

Ejercicio B6. Dos valores y ningún número finito de cortes. Define $h : [0,1] \to \mathbb{R}$ por

$$h(x) = \begin{cases} 1, & x = 1/n \text{ para algún entero } n \geq 1, \\ 0, & \text{en otro caso.} \end{cases}$$

Demuestra que h tiene imagen finita pero no es una función escalonada. Puedes usar la propiedad arquimediana y la densidad de los irracionales.

Nivel C — Álgebra, orden y aditividad. Ejercicio C1. Linealidad sobre un refinamiento común. En $[0,2]$ sean s y t escalonadas dadas por

$$s(x) = \begin{cases} 1, & 0 < x < 1, \\ -1, & 1 < x < 2, \end{cases}$$

y

$$t(x) = \begin{cases} 2, & 0 < x < \frac{1}{2}, \\ 0, & \frac{1}{2} < x < 2, \end{cases}$$

con valores en los cortes elegidos arbitrariamente.

1. Calcula $\int s$ y $\int t$.
2. Usa el refinamiento común $\{0, 1/2, 1, 2\}$ para calcular directamente $\int (2s - 3t)$.
3. Verifica la identidad de linealidad.

Ejercicio C2. Monotonía en un ejemplo exacto. Sean $s, t : [0,2] \to \mathbb{R}$ escalonadas tales que, fuera de los cortes,

$$s = \begin{cases} 0, & 0 < x < 1, \\ 1, & 1 < x < 2, \end{cases} \qquad t = \begin{cases} 1, & 0 < x < 1, \\ 3, & 1 < x < 2. \end{cases}$$

Supón además que los valores en los cortes se eligen de modo que $s \leq t$ en todo $[0,2]$. Calcula ambas integrales y verifica numéricamente la conclusión de monotonía de Proposición 13.3.

Ejercicio C3. La desigualdad del valor absoluto puede ser estricta. Sea $s : [0,2] \to \mathbb{R}$ escalonada con altura 2 en $(0,1)$ y altura -1 en $(1,2)$.

Calcula

$$\left| \int_0^2 s \right| \qquad \text{y} \qquad \int_0^2 |s|.$$

Comprueba que la desigualdad

$$\left|\int s\right| \le \int |s|$$

es estricta y explica la razón geométrica-algebraica.

Ejercicio C4. Cortar dentro de una franja. Recupera la función s del ejercicio B1 y toma $c = \frac{5}{2}$.

Calcula por separado

$$\int_0^{5/2} s(x)\,dx \qquad \text{y} \qquad \int_{5/2}^{4} s(x)\,dx,$$

y verifica la aditividad respecto del intervalo.

Ejercicio C5. Dos funciones distintas con la misma integral. Sea

$$s(x) = 1 \qquad (0 \le x \le 1),$$

y define $t : [0,1] \to \mathbb{R}$ por $t(x) = 1$ salvo en

$$t\left(\frac{1}{3}\right) = 7, \qquad t\left(\frac{2}{3}\right) = -4.$$

1. Demuestra que t es escalonada.
2. Calcula $\int_0^1 s$ y $\int_0^1 t$.
3. Explica qué resultado del capítulo permite anticipar la igualdad.

Ejercicio C6. Igualdad de integrales no implica igualdad de funciones. En $[0,1]$ sea $s(x) = 1$ y sea t escalonada con altura 2 en $(0, 1/2)$ y altura 0 en $(1/2, 1)$.

Elige valores no negativos en los cortes. Demuestra que s y t son distintas, no negativas y satisfacen

$$\int_0^1 s = \int_0^1 t.$$

¿Qué conversa falsa de la monotonía podría sugerir erróneamente este ejemplo?

Nivel D — Sumas inferiores, superiores y refinamiento. Ejercicio D1. Sumas inferior y superior para x^2. Para

$$f(x) = x^2, \qquad P = \left\{0, \frac{1}{2}, 1\right\},$$

calcula m_k, M_k, $L(f,P)$ y $U(f,P)$. Determina también la brecha $U - L$.

Ejercicio D2. Una partición no uniforme. Para $f(x) = x^2$ en $[0,1]$ y

$$P = \left\{0, \frac{1}{4}, 1\right\},$$

calcula exactamente $L(f,P)$, $U(f,P)$ y $U(f,P) - L(f,P)$.

Ejercicio D3. Un supremo no alcanzado entra igualmente en la suma superior. Sea

$$f(x) = \begin{cases} x, & 0 \le x < 1, \\ 0, & x = 1, \end{cases}$$

y

$$P = \left\{0, \frac{1}{2}, 1\right\}.$$

Calcula los ínfimos y supremos locales, indica cuáles se alcanzan y cuáles no, y obtiene $L(f, P)$ y $U(f, P)$.

Ejercicio D4. Verificar la cadena de refinamiento. Para $f(x) = x^2$ en $[0, 1]$, toma

$$P = \left\{0, \frac{1}{2}, 1\right\},$$

y

$$Q = \left\{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\right\}.$$

Calcula las cuatro sumas y verifica

$$L(f, P) \le L(f, Q) \le U(f, Q) \le U(f, P).$$

Compara además las dos brechas.

Ejercicio D5. Comparar particiones no comparables. Para $f(x) = x$ en $[0, 1]$ considera

$$P = \left\{0, \frac{1}{2}, 1\right\}, \qquad Q = \left\{0, \frac{1}{3}, 1\right\}.$$

1. Calcula $L(f, P), U(f, P), L(f, Q), U(f, Q)$.
2. Construye $R = P \vee Q$ y calcula $L(f, R), U(f, R)$.
3. Verifica las dos cadenas que prueban

$$L(f, P) \le U(f, Q), \qquad L(f, Q) \le U(f, P).$$

Ejercicio D6. Menor malla y peor suma superior. Define

$$s(x) = \begin{cases} 1, & x = \frac{1}{4}, \\ 0, & x \ne \frac{1}{4}, \end{cases} \qquad 0 \le x \le 1.$$

Considera

$$P = \left\{0, \frac{1}{5}, \frac{1}{4}, \frac{3}{10}, 1\right\}, \qquad Q = \left\{0, \frac{1}{3}, \frac{2}{3}, 1\right\}.$$

Calcula $\|P\|, \|Q\|, U(s, P)$ y $U(s, Q)$. Usa los resultados para refutar la afirmación «una malla menor siempre produce una suma superior menor».

Nivel E — Sumas etiquetadas, malla y encierros. Ejercicio E1. Tres etiquetados para x^2. Para

$$f(x) = x^2, \qquad P = \left\{0, \frac{1}{2}, 1\right\},$$

calcula las sumas de Riemann obtenidas con extremos izquierdos, puntos medios y extremos derechos. Verifica que las tres quedan entre $L(f, P)$ y $U(f, P)$.

Ejercicio E2. Un punto de corte puede etiquetar dos piezas. Define

$$s(x) = \begin{cases} 1, & x = \frac{1}{3}, \\ 0, & x \neq \frac{1}{3}, \end{cases}$$

y toma

$$P = \left\{0, \frac{1}{3}, \frac{2}{3}, 1\right\}.$$

Calcula las sumas asociadas a

$$\eta = \left(\frac{1}{6}, \frac{1}{2}, \frac{5}{6}\right)$$

y

$$\xi = \left(\frac{1}{3}, \frac{1}{3}, \frac{5}{6}\right).$$

Calcula también $L(s, P)$ y $U(s, P)$.

Ejercicio E3. Una suma superior que ninguna etiqueta realiza. Para la función del ejercicio D3, toma la partición mínima $P = \{0, 1\}$.

Demuestra que $U(f, P) = 1$, pero que para toda etiqueta $\xi_1 \in [0, 1]$, $R(f; P, \xi) < 1$. Explica por qué esto no contradice Proposición 13.7.

Ejercicio E4. Una elección interior en una partición no uniforme. Para $f(x) = x$ y

$$P = \left\{0, \frac{1}{4}, \frac{3}{4}, 1\right\},$$

usa las etiquetas

$$\xi = \left(\frac{1}{8}, \frac{1}{2}, 1\right).$$

Calcula $R(f; P, \xi)$ y verifica $L(f, P) \leq R(f; P, \xi) \leq U(f, P)$ usando los valores de §13.6.

Ejercicio E5. La misma malla, una diferencia de tres unidades. Define

$$s(x) = \begin{cases} 2, & x = \frac{1}{2}, \\ -1, & x \neq \frac{1}{2}, \end{cases}$$

en $[0,1]$, y toma

$$P = \left\{0, \frac{1}{2}, 1\right\}.$$

Encuentra dos elecciones de etiquetas sobre P que produzcan respectivamente sumas -1 y 2. Calcula $L(s,P)$ y $U(s,P)$ y explica qué controla la malla y qué no controla.

Ejercicio E6. Las sumas etiquetadas no son monótonas bajo refinamiento. Sea $f(x) = x$ en $[0,1]$.

1. Con $P = \{0,1\}$ y etiqueta 1, calcula $R(f;P,\xi)$.
2. Refina a $Q = \{0,1/2,1\}$ y usa etiquetas 0 y $1/2$. Comprueba que la suma disminuye.
3. Repite tomando en P la etiqueta 0 y en Q los extremos derechos. Comprueba que ahora la suma aumenta.
4. Explica por qué esto no contradice Proposición 13.6.

Nivel F — Brecha, oscilación y reconstrucción de argumentos. Ejercicio F1. Oscilaciones locales para x^2. Para $f(x) = x^2$ y

$$P = \left\{0, \frac{1}{2}, 1\right\},$$

calcula ω_1, ω_2 y verifica directamente $U(f,P) - L(f,P) = \omega_1 \frac{1}{2} + \omega_2 \frac{1}{2}$.

Ejercicio F2. La brecha de una función lineal. Para $f(x) = x$ en $[0,1]$ y la partición uniforme

$$P_n = \left\{0, \frac{1}{n}, \ldots, 1\right\},$$

demuestra usando oscilaciones que $U(f,P_n) - L(f,P_n) = \frac{1}{n}$.

Ejercicio F3. Amplitud grande confinada en poco ancho. Fija $c = 2/5$ y define

$$s(x) = \begin{cases} 5, & x = c, \\ 0, & x \neq c, \end{cases}$$

en $[0,1]$. Para $0 < \delta < 2/5$ considera $Q_\delta = \{0, c-\delta, c, c+\delta, 1\}$.

Calcula la brecha $U(s,Q_\delta) - L(s,Q_\delta)$ mediante oscilaciones y explica qué parte del cálculo reduce la incertidumbre: la amplitud vertical o la anchura horizontal.

Ejercicio F4. Oscilación persistente de amplitud tres. Define

$$d(x) = \begin{cases} 2, & x \in \mathbb{Q}, \\ -1, & x \notin \mathbb{Q}, \end{cases} \qquad 0 \le x \le 1.$$

Demuestra que para toda partición P,

$$L(d,P) = -1, \qquad U(d,P) = 2, \qquad U(d,P) - L(d,P) = 3.$$

Ejercicio F5. Malla pequeña no implica brecha pequeña. Conserva la función d del ejercicio F4 y considera las particiones uniformes P_n.

Demuestra que $\|P_n\| = \frac{1}{n} \longrightarrow 0$ mientras $U(d,P_n) - L(d,P_n) = 3$ para todo n. Identifica la inferencia incorrecta que este ejemplo refuta.

Ejercicio F6. Dos cotas generales para la brecha. Sea $f : [a,b] \to \mathbb{R}$ acotada y $P = \{x_0, \dots, x_n\}$.

1. Si $A \le f \le B$ en todo $[a,b]$, demuestra

$$0 \le U(f,P) - L(f,P) \le (B-A)(b-a).$$

2. Supón que $S \subseteq \{1, \dots, n\}$ y que $C, W \ge 0$. Si $\omega_k = 0$ fuera de S, $\omega_k \le C$ para $k \in S$, y

$$\sum_{k \in S} \Delta x_k \le W,$$

Demuestra $U(f,P) - L(f,P) \le CW$.

Interpreta la segunda cota.

Nivel G — Desafíos integradores: del área a las sumas. Ejercicio G1. Una función, cuatro construcciones. Define $s : [0,2] \to \mathbb{R}$ por

$$s(x) = \begin{cases} -1, & 0 \le x < 1, \\ 2, & x = 1, \\ 4, & 1 < x \le 2. \end{cases}$$

y toma $P = \{0,1,2\}$.

1. Calcula la integral elemental de s.
2. Calcula $L(s,P)$ y $U(s,P)$.
3. Encuentra etiquetados que produzcan cada uno de los valores $1, 3, 4, 6$.
4. Para

$$Q_\delta = \{0, 1-\delta, 1, 1+\delta, 2\}, \qquad 0 < \delta < 1,$$

calcula la brecha mediante oscilaciones. 5. Explica por qué los resultados anteriores son compatibles entre sí.

Ejercicio G2. Dos encierros incompatibles en cortes, compatibles en orden. Para $f(x) = x$ en $[0, 1]$ sean

$$P = \left\{0, \frac{1}{2}, 1\right\}, \qquad Q = \left\{0, \frac{1}{3}, 1\right\}.$$

1. Calcula $J(f, P) = [L(f, P), U(f, P)]$ y $J(f, Q)$.
2. Construye $R = P \vee Q$ y calcula $J(f, R)$.
3. Verifica

$J(f, R) \subseteq J(f, P) \cap J(f, Q)$.

4. Elige puntos medios como etiquetas de R, calcula la suma correspondiente y localízala en los tres intervalos.

Ejercicio G3. Forzar acuerdo de sumas sin definir todavía una integral general. Sea

$$s(x) = \begin{cases} 1, & x = \frac{1}{2}, \\ 0, & x \neq \frac{1}{2}, \end{cases}$$

en $[0, 1]$. Dado un número $\eta > 0$:

1. elige $0 < \delta < \min\{1/2, \eta/2\}$;
2. usa

$$Q_\delta = \left\{0, \frac{1}{2} - \delta, \frac{1}{2}, \frac{1}{2} + \delta, 1\right\}$$

para demostrar que $0 \leq R(s; Q_\delta, \xi) < \eta$ para toda elección de etiquetas ξ; 3. explica con precisión qué se ha demostrado y qué definición de C15 deliberadamente no se ha utilizado.

Ejercicio G4. Dos familias de sumas, dos candidatos incompatibles. Sea

$$d(x) = \begin{cases} 1, & x \in \mathbb{Q}, \\ 0, & x \notin \mathbb{Q}, \end{cases}$$

en $[0, 1]$, y sea P_n la partición uniforme en n piezas.

1. Justifica que en cada subintervalo de P_n puede elegirse una etiqueta racional y también una irracional.
2. Construye un etiquetado racional $\xi^{(n)}$ y uno irracional $\eta^{(n)}$ y demuestra

$$R(d; P_n, \xi^{(n)}) = 1, \qquad R(d; P_n, \eta^{(n)}) = 0$$

para todo n. 3. Calcula $L(d, P_n)$ y $U(d, P_n)$. 4. Explica por qué la convergencia de una sola familia particular de sumas etiquetadas no basta para producir un candidato independiente del procedimiento.

Soluciones desarrolladas.

Soluciones del nivel A. **Solución A1.** Los puntos están ordenados

$$0 < \frac{1}{5} < \frac{2}{5} < 1.$$

Por tanto los subintervalos son

$$\left[0, \frac{1}{5}\right], \qquad \left[\frac{1}{5}, \frac{2}{5}\right], \qquad \left[\frac{2}{5}, 1\right].$$

Sus anchuras son

$$\frac{1}{5}, \qquad \frac{1}{5}, \qquad \frac{3}{5}.$$

Así,

$$\boxed{\|P\| = \frac{3}{5}}.$$

La suma de anchuras es $\frac{1}{5} + \frac{1}{5} + \frac{3}{5} = 1$, que coincide con $1 - 0$. La lista $\{0, 1/5, 2/5\}$ no es una partición de $[0, 1]$ porque Definición 13.1 exige que ambos extremos del intervalo pertenezcan al conjunto de cortes y falta el extremo 1.

Solución A2. Como todos los puntos de P aparecen en Q, $P \subseteq Q$, luego $P \preceq Q$. En cambio, $1/2 \notin R$, así que $P \not\preceq R$. Tampoco $R \preceq Q$, porque $1/3$ y $2/3$ no pertenecen a Q.

Las mallas son

$$\|P\| = \frac{1}{2}, \qquad \|Q\| = \frac{1}{4}, \qquad \|R\| = \frac{1}{3}.$$

El refinamiento común de P y R es

$$P \vee R = \left\{0, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, 1\right\}.$$

El de Q y R es la unión de todos sus cortes:

$$Q \vee R = \left\{0, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, 1\right\}.$$

La primera comparación ilustra que una malla menor no basta para certificar refinamiento: R tiene malla menor que P, pero no conserva su corte $1/2$.

Solución A3. La unión ordenada es

$$P \vee Q = \left\{0, \frac{1}{2}, 1, \frac{3}{2}, \frac{7}{4}, 2\right\}.$$

Las anchuras consecutivas son

$$\frac{1}{2}, \quad \frac{1}{2}, \quad \frac{1}{2}, \quad \frac{1}{4}, \quad \frac{1}{4}.$$

Por tanto

$$\boxed{\|P \vee Q\| = \frac{1}{2}}.$$

Para las particiones originales, $\|P\| = 1$ porque la pieza $[1/2, 3/2]$ tiene longitud 1, mientras $\|Q\| = 1$.

Así,

$$\|P \vee Q\| = \frac{1}{2} \leq 1 = \min\{\|P\|, \|Q\|\},$$

como predice Proposición 13.1. El refinamiento común conserva todos los cortes y puede reducir la pieza máxima de ambas particiones a la vez.

Solución A4. En P las anchuras son

$$\frac{1}{2}, \quad \frac{1}{4}, \quad \frac{1}{4},$$

de modo que $\|P\| = \frac{1}{2}$.

Si añadimos $1/8$, la primera pieza se divide en longitudes

$$\frac{1}{8} \qquad \text{y} \qquad \frac{3}{8}.$$

Las demás conservan longitud $1/4$, de modo que la nueva malla es

$$\boxed{\frac{3}{8}}.$$

Si añadimos $7/8$, solo se divide la última pieza $[3/4, 1]$ en dos piezas de longitud $1/8$. La primera pieza $[0, 1/2]$ permanece intacta, por lo que la malla sigue siendo

$$\boxed{\frac{1}{2}}.$$

En los dos casos se añadió un punto nuevo sin eliminar ninguno anterior; ambos refinamientos son estrictos. Proposición 13.1 garantiza que la malla no aumente, pero no que disminuya estrictamente.

Solución A5. Como cada anchura satisface $\Delta x_k \leq \|P\|$, al sumar obtenemos

$$b - a = \sum_{k=1}^{n} \Delta x_k \leq \sum_{k=1}^{n} \|P\| = n\|P\|.$$

Como $n > 0$,

$$\boxed{\|P\| \geq \frac{b-a}{n}}.$$

Si todos los subintervalos tienen longitud $(b-a)/n$, entonces claramente la malla es ese número.

Recíprocamente, supongamos $\|P\| = \frac{b-a}{n}$.

Todas las anchuras son a lo sumo $\|P\|$ y su suma es exactamente $b-a = n\|P\|$. Si alguna anchura fuese estrictamente menor que $\|P\|$, como ninguna puede exceder $\|P\|$, la suma total sería estrictamente menor que $n\|P\|$, contradicción. Por tanto todas satisfacen $\Delta x_k = \frac{b-a}{n}$. La igualdad caracteriza precisamente a la partición uniforme.

Solución A6. P tiene dos subintervalos, ambos de longitud $1/2$, así que $\|P\| = \frac{1}{2}$.

Q tiene cuatro subintervalos, de longitudes

$$\frac{1}{3}, \quad \frac{1}{3}, \quad \frac{1}{6}, \quad \frac{1}{6},$$

por lo que $\|Q\| = \frac{1}{3} < \frac{1}{2}$.

Sin embargo, $\frac{1}{2} \notin Q$, de modo que $P \not\subseteq Q$ y Q no refina a P.

El refinamiento común es

$$P \vee Q = \left\{0, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{5}{6}, 1\right\}.$$

Sus anchuras son

$$\frac{1}{3}, \quad \frac{1}{6}, \quad \frac{1}{6}, \quad \frac{1}{6}, \quad \frac{1}{6},$$

y por tanto $\|P \vee Q\| = \frac{1}{3}$.

Tener más piezas cuenta solamente cuántos subintervalos hay; tener menor malla controla la longitud de la pieza más larga; ser refinamiento exige la condición estructural exacta de inclusión de cortes.

Soluciones del nivel B. **Solución B1.** Usando la partición natural $\{0, 1, 3, 4\}$,

$$\begin{aligned}\int_0^4 s(x)\,dx &= 2(1-0) + (-1)(3-1) + 3(4-3) \\ &= 2 - 2 + 3 \\ &= \boxed{3}.\end{aligned}$$

Con el refinamiento Q, las seis alturas son $2,\ 2,\ -1,\ -1,\ 3,\ 3$, y las anchuras son $\frac{1}{2},\ \frac{1}{2},\ 1,\ 1,\ \frac{1}{2},\ \frac{1}{2}$.

Entonces

$$2\frac{1}{2} + 2\frac{1}{2} - 1 - 1 + 3\frac{1}{2} + 3\frac{1}{2} = 1 + 1 - 1 - 1 + \frac{3}{2} + \frac{3}{2} = 3.$$

La igualdad no depende del cálculo particular: Lema 13.1 demuestra que la suma rectangular de una escalonada es invariante bajo refinamiento, y Teorema 13.1 permite retirar la partición de la definición de integral elemental.

Solución B2. Los puntos de F son precisamente cortes de la partición $P = \{0, 1, 3, 4\}$.

En cada intervalo abierto $(0, 1)$, $(1, 3)$ y $(3, 4)$ las funciones s y t coinciden, por lo que t conserva las mismas alturas $2, -1, 3$. Así P representa también a t. Como $s = t$ fuera del conjunto finito F, Corolario 13.1 da

$$\int_0^4 t = \int_0^4 s = 3.$$

No se concluye $s = t$: por construcción sus valores difieren en puntos de F. La integral elemental es insensible a esas modificaciones finitas dentro de la clase escalonada, pero la igualdad de integrales no convierte dos funciones en una sola.

Solución B3. La partición $P = \{0, 1, 2\}$ representa a u: la altura es -2 en $(0, 1)$ y 1 en $(1, 2)$. El valor excepcional $u(1) = 7$ no altera esas alturas.

Por tanto

$$\int_0^2 u(x)\,dx = (-2)(1) + 1(1) = \boxed{-1}.$$

Por ejemplo,

$$Q = \left\{0, \frac{1}{2}, 1, \frac{3}{2}, 2\right\}$$

refina a P y también representa a u. Con alturas $-2, -2, 1, 1$ y anchuras $1/2$,

$$\int_0^2 u = -2\frac{1}{2} - 2\frac{1}{2} + 1\frac{1}{2} + 1\frac{1}{2} = -1.$$

La coincidencia es una instancia de la invariancia bajo refinamiento.

Solución B4. P representa a s por definición. Como $P \subseteq Q$,

Proposición 13.2 garantiza que Q también la representa.

Con P, $\mathcal{R}(s; P) = (-2)(1) + 1(2) = 0$.

Con Q,

$$\mathcal{R}(s; Q) = (-2)\frac{1}{3} + (-2)\frac{2}{3} + 1(1) + 1(1) = -\frac{2}{3} - \frac{4}{3} + 2 = 0.$$

Lema 13.1 explica la igualdad sin necesidad de recalcular: cada franja original se subdivide, pero la altura permanece constante y las nuevas anchuras telescopan a la anchura original.

Solución B5. La integral con signo es

$$\int_0^3 s = 2(1) + (-1)(2) = 0.$$

En cambio $|s|$ tiene alturas 2 y 1, de modo que

$$\int_0^3 |s| = 2(1) + 1(2) = \boxed{4}.$$

s no es la función nula: por ejemplo, vale 2 en todo $(0, 1)$. El valor cero de la integral procede de cancelación exacta entre la contribución positiva 2 y la negativa -2.

Solución B6. La imagen de h está contenida en $\{0,1\}$ y ambos valores aparecen, así que es finita.

Supongamos, para obtener una contradicción, que una partición finita

$$P = \{0 = x_0 < x_1 < \cdots < x_n = 1\}$$

representa a h. El primer intervalo abierto $(0, x_1)$ es no vacío. Por la propiedad arquimediana existe N suficientemente grande para que $0 < \frac{1}{N} < x_1$. Entonces $h(1/N) = 1$. Por densidad de los irracionales, existe un irracional $u \in (0, x_1)$. Todo número $1/n$ es racional, de modo que u no pertenece a la lista $\{1/n : n \geq 1\}$ y $h(u) = 0$. Así h toma dos valores distintos dentro del mismo intervalo abierto $(0, x_1)$, contradiciendo Definición 13.2. Por tanto h no es escalonada. Tener imagen finita no basta: se necesita una descomposición finita del dominio en franjas constantes.

Soluciones del nivel C. **Solución C1.** Para s,

$$\int_0^2 s = 1(1) + (-1)(1) = 0.$$

Para t,

$$\int_0^2 t = 2\left(\frac{1}{2}\right) + 0\left(\frac{3}{2}\right) = 1.$$

Sobre el refinamiento común

$$R = \left\{0, \frac{1}{2}, 1, 2\right\},$$

las alturas de $2s - 3t$ son

$$2(1) - 3(2) = -4, \qquad 2(1) - 3(0) = 2, \qquad 2(-1) - 3(0) = -2.$$

Por tanto,

$$\begin{aligned} \int_0^2 (2s - 3t) &= (-4)\frac{1}{2} + 2\frac{1}{2} + (-2)(1) \\ &= -2 + 1 - 2 \\ &= \boxed{-3}. \end{aligned}$$

La linealidad de Proposición 13.3 predice exactamente

$$2\int s - 3\int t = 2(0) - 3(1) = -3.$$

Solución C2. Las alturas de s son 0 y 1, ambas sobre intervalos de longitud 1, así que

$$\int_0^2 s = 0 + 1 = 1.$$

Las de t son 1 y 3, por lo que

$$\int_0^2 t = 1 + 3 = 4.$$

Así

$$\boxed{\int_0^2 s = 1 \le 4 = \int_0^2 t}.$$

La desigualdad puntual $s \le t$, incluidos los cortes por hipótesis, activa la monotonía de Proposición 13.3. El cálculo ilustra la proposición, pero la prueba general no depende de esta representación concreta.

Solución C3. Tenemos

$$\int_0^2 s = 2(1) + (-1)(1) = 1,$$

de modo que

$$\left|\int_0^2 s\right| = 1.$$

Por otra parte,

$$\int_0^2 |s| = 2(1) + 1(1) = 3.$$

Por tanto

$$\boxed{1 < 3}.$$

La desigualdad es estricta porque la integral con signo cancela parte de las contribuciones positiva y negativa, mientras que $|s|$ convierte ambas en contribuciones no negativas antes de sumar.

Solución C4. En $[0, 5/2]$ aparecen la franja completa $(0, 1)$ de altura 2 y una porción de longitud $3/2$ de la franja de altura -1:

$$\int_0^{5/2} s = 2(1) + (-1)\left(\frac{3}{2}\right) = \frac{1}{2}.$$

En $[5/2, 4]$ queda media unidad de la franja de altura -1 y la franja completa de altura 3:

$$\int_{5/2}^4 s = (-1)\left(\frac{1}{2}\right) + 3(1) = \frac{5}{2}.$$

Sumando,

$$\frac{1}{2} + \frac{5}{2} = 3 = \int_0^4 s.$$

El punto $5/2$ no era un corte original, pero puede insertarse como refinamiento sin cambiar la integral elemental.

Solución C5. La partición

$$P = \left\{0, \frac{1}{3}, \frac{2}{3}, 1\right\}$$

representa a t, porque en cada intervalo abierto t vale constantemente 1. Por tanto t es escalonada. La función s es constante de altura 1, así que

$$\int_0^1 s = 1.$$

t difiere de s solo en el conjunto finito $\{1/3, 2/3\}$. Por Corolario 13.1,

$$\boxed{\int_0^1 t = \int_0^1 s = 1}.$$

El corolario permite anticipar la igualdad antes de calcular. Los valores 7 y -4 siguen siendo valores auténticos de t, pero no aportan anchura a la suma escalonada.

Solución C6. s vale 1 en todo el intervalo, mientras t vale 2 en $(0, 1/2)$ y 0 en $(1/2, 1)$; por tanto son distintas. Con valores no negativos en los cortes, ambas son no negativas.

Sus integrales son

$$\int_0^1 s = 1(1) = 1$$

y

$$\int_0^1 t = 2\left(\frac{1}{2}\right) + 0\left(\frac{1}{2}\right) = 1.$$

Así dos funciones no negativas distintas pueden tener la misma integral. La conversa falsa sería: de

$$\int s \le \int t$$

inferir $s \le t$ punto a punto, o de igualdad de integrales inferir igualdad de funciones. Proposición 13.3 solo afirma la dirección desde el orden puntual hacia el orden de las integrales.

Soluciones del nivel D. **Solución D1.** En $[0, 1/2]$, como x^2 es creciente,

$$m_1 = 0, \qquad M_1 = \frac{1}{4}.$$

En $[1/2, 1]$,

$$m_2 = \frac{1}{4}, \qquad M_2 = 1.$$

Como ambas anchuras son $1/2$,

$$L(f, P) = 0\frac{1}{2} + \frac{1}{4}\frac{1}{2} = \boxed{\frac{1}{8}},$$

y

$$U(f,P) = \frac{1}{4}\frac{1}{2} + 1\frac{1}{2} = \boxed{\frac{5}{8}}.$$

La brecha es

$$\boxed{U - L = \frac{1}{2}}.$$

Solución D2. En $[0, 1/4]$,

$$m_1 = 0, \qquad M_1 = \frac{1}{16}, \qquad \Delta x_1 = \frac{1}{4}.$$

En $[1/4, 1]$,

$$m_2 = \frac{1}{16}, \qquad M_2 = 1, \qquad \Delta x_2 = \frac{3}{4}.$$

Por tanto,

$$L(f,P) = 0 + \frac{1}{16}\frac{3}{4} = \boxed{\frac{3}{64}},$$

mientras

$$U(f,P) = \frac{1}{16}\frac{1}{4} + 1\frac{3}{4} = \frac{1}{64} + \frac{48}{64} = \boxed{\frac{49}{64}}.$$

La diferencia es

$$\boxed{\frac{49}{64} - \frac{3}{64} = \frac{46}{64} = \frac{23}{32}}.$$

Solución D3. En

$$I_1 = \left[0, \frac{1}{2}\right],$$

$f(x) = x$, así que

$$m_1 = 0, \qquad M_1 = \frac{1}{2},$$

y ambos valores se alcanzan.
En

$$I_2 = \left[\frac{1}{2}, 1\right],$$

los valores son

$$\{0\} \cup \left[\frac{1}{2}, 1\right).$$

Por tanto $m_2 = 0$ se alcanza en $x = 1$, mientras $M_2 = 1$ es el supremo pero no se alcanza. Las dos piezas tienen longitud $1/2$, de modo que $L(f, P) = 0$ y

$$U(f, P) = \frac{1}{2}\frac{1}{2} + 1\frac{1}{2} = \boxed{\frac{3}{4}}.$$

La construcción funciona porque Definición 13.4 usa ínfimo y supremo, no mínimo y máximo.

Solución D4. Para P, el ejercicio D1 dio

$$L(f, P) = \frac{1}{8}, \qquad U(f, P) = \frac{5}{8}.$$

Para Q, los extremos izquierdos producen

$$L(f, Q) = \left(0 + \frac{1}{16} + \frac{1}{4} + \frac{9}{16}\right)\frac{1}{4} = \frac{14}{64} = \boxed{\frac{7}{32}},$$

y los derechos

$$U(f, Q) = \left(\frac{1}{16} + \frac{1}{4} + \frac{9}{16} + 1\right)\frac{1}{4} = \frac{30}{64} = \boxed{\frac{15}{32}}.$$

Escribiendo con denominador 32,

$$\frac{1}{8} = \frac{4}{32} \le \frac{7}{32} \le \frac{15}{32} \le \frac{20}{32} = \frac{5}{8}.$$

La brecha de P es $\frac{1}{2}$, y la de Q es $\frac{15}{32} - \frac{7}{32} = \frac{1}{4}$. El refinamiento reduce la brecha a la mitad.

Solución D5. Para $f(x) = x$ y

$$P = \left\{0, \frac{1}{2}, 1\right\},$$

tenemos $L(f, P) = 0\frac{1}{2} + \frac{1}{2}\frac{1}{2} = \frac{1}{4}$, y $U(f, P) = \frac{1}{2}\frac{1}{2} + 1\frac{1}{2} = \frac{3}{4}$.
Para

$$Q = \left\{0, \frac{1}{3}, 1\right\},$$

las anchuras son $1/3$ y $2/3$. Entonces $L(f, Q) = 0\frac{1}{3} + \frac{1}{3}\frac{2}{3} = \frac{2}{9}$, y $U(f, Q) = \frac{1}{3}\frac{1}{3} + 1\frac{2}{3} = \frac{7}{9}$.
El refinamiento común es

$$R = \left\{0, \frac{1}{3}, \frac{1}{2}, 1\right\}.$$

Sus anchuras son $1/3, 1/6, 1/2$, de donde

$$L(f, R) = 0 + \frac{1}{3}\frac{1}{6} + \frac{1}{2}\frac{1}{2} = \frac{1}{18} + \frac{1}{4} = \boxed{\frac{11}{36}},$$

y

$$U(f, R) = \frac{1}{3}\frac{1}{3} + \frac{1}{2}\frac{1}{6} + 1\frac{1}{2} = \frac{1}{9} + \frac{1}{12} + \frac{1}{2} = \boxed{\frac{25}{36}}.$$

La primera cadena es

$$\frac{1}{4} = \frac{9}{36} \le \frac{11}{36} \le \frac{25}{36} \le \frac{28}{36} = \frac{7}{9},$$

y la segunda,

$$\frac{2}{9} = \frac{8}{36} \le \frac{11}{36} \le \frac{25}{36} \le \frac{27}{36} = \frac{3}{4}.$$

Ambas son instancias de la prueba de Corolario 13.2 mediante un refinamiento común.

Solución D6. En P las anchuras son

$$\frac{1}{5}, \quad \frac{1}{20}, \quad \frac{1}{20}, \quad \frac{7}{10},$$

por lo que $\|P\| = \frac{7}{10}$.

El punto excepcional $1/4$ pertenece a los dos subintervalos cerrados adyacentes, ambos de longitud $1/20$. En ellos el supremo es 1; en los otros es 0. Así,

$$\boxed{U(s,P) = \frac{1}{20} + \frac{1}{20} = \frac{1}{10}}.$$

Para Q, la malla es $\|Q\| = \frac{1}{3} < \frac{7}{10}$.

El punto $1/4$ está en el interior de $[0, 1/3]$, y solo esa pieza tiene supremo 1. Por tanto,

$$\boxed{U(s,Q) = \frac{1}{3}}.$$

Luego

$$\|Q\| < \|P\| \qquad \text{pero} \qquad U(s,Q) > U(s,P).$$

La afirmación propuesta es falsa. La monotonía de Proposición 13.6 requiere refinamiento, no solo una malla menor.

Soluciones del nivel E. **Solución E1.** Las etiquetas izquierdas son 0 y $1/2$:

$$R_{\text{izq}} = 0^2\frac{1}{2} + \left(\frac{1}{2}\right)^2\frac{1}{2} = \boxed{\frac{1}{8}}.$$

Los puntos medios son $1/4$ y $3/4$:

$$R_{\text{med}} = \left(\frac{1}{4}\right)^2\frac{1}{2} + \left(\frac{3}{4}\right)^2\frac{1}{2} = \frac{1}{32} + \frac{9}{32} = \boxed{\frac{5}{16}}.$$

Los extremos derechos son $1/2$ y 1:

$$R_{\text{der}} = \left(\frac{1}{2}\right)^2\frac{1}{2} + 1^2\frac{1}{2} = \boxed{\frac{5}{8}}.$$

Por D1,

$$L(f,P)=\frac{1}{8},\qquad U(f,P)=\frac{5}{8}.$$

Así

$$L(f,P)=R_{\text{izq}}<R_{\text{med}}<R_{\text{der}}=U(f,P).$$

Solución E2. Con

$$\eta=\left(\frac{1}{6},\frac{1}{2},\frac{5}{6}\right),$$

ninguna etiqueta es $1/3$, luego los tres valores de s son cero y

$$\boxed{R(s;P,\eta)=0}.$$

En

$$\xi=\left(\frac{1}{3},\frac{1}{3},\frac{5}{6}\right),$$

el punto $1/3$ se usa como etiqueta de las dos primeras piezas, lo cual es válido porque pertenece a ambas. Como todas las anchuras son $1/3$,

$$R(s;P,\xi)=1\frac{1}{3}+1\frac{1}{3}+0\frac{1}{3}=\boxed{\frac{2}{3}}.$$

El ínfimo es 0 en todas las piezas. El supremo es 1 en las dos piezas que contienen $1/3$ y 0 en la tercera. Por tanto,

$$\boxed{L(s,P)=0,\qquad U(s,P)=\frac{2}{3}}.$$

Las dos sumas etiquetadas realizan aquí los extremos del encierro.

Solución E3. En el único subintervalo $[0,1]$, $\sup f([0,1])=1$, aunque f nunca toma ese valor. Como la anchura es 1, $U(f,P)=1$.

Para cualquier etiqueta $\xi_1\in[0,1]$ hay dos casos. Si $\xi_1<1$, entonces $f(\xi_1)=\xi_1<1$. Si $\xi_1=1$, entonces $f(1)=0<1$. En ambos casos $R(f;P,\xi)=f(\xi_1)<1$. No hay contradicción con Proposición 13.7: la proposición solo afirma $R(f;P,\xi)\le U(f,P)$, no que la cota superior tenga que ser alcanzada por alguna etiqueta.

Solución E4. Las anchuras son $1/4, 1/2, 1/4$. Por tanto,

$$\begin{aligned}R(f;P,\xi)&=\frac{1}{8}\frac{1}{4}+\frac{1}{2}\frac{1}{2}+1\frac{1}{4}\\&=\frac{1}{32}+\frac{1}{4}+\frac{1}{4}\\&=\boxed{\frac{17}{32}}.\end{aligned}$$

En §13.6 se obtuvo

$$L(f,P)=\frac{5}{16}=\frac{10}{32},\qquad U(f,P)=\frac{11}{16}=\frac{22}{32}.$$

Así,

$$\boxed{\frac{10}{32}\le\frac{17}{32}\le\frac{22}{32}}.$$

El cálculo es una instancia directa del encierro de Proposición 13.7.

Solución E5. Si elegimos, por ejemplo,

$$\eta=\left(\frac{1}{4},\frac{3}{4}\right),$$

ambas etiquetas evitan $1/2$, de modo que

$$R(s;P,\eta)=(-1)\frac{1}{2}+(-1)\frac{1}{2}=\boxed{-1}.$$

Si elegimos

$$\xi=\left(\frac{1}{2},\frac{1}{2}\right),$$

la misma etiqueta puede usarse en las dos piezas y

$$R(s;P,\xi)=2\frac{1}{2}+2\frac{1}{2}=\boxed{2}.$$

En cada uno de los dos subintervalos cerrados el ínfimo es -1 y el supremo es 2. Por ello

$$L(s,P)=-1,\qquad U(s,P)=2.$$

La malla es $1/2$ en ambos etiquetados porque depende solo de P. No controla qué valores de la función se muestrean; por eso dos elecciones con la misma malla pueden diferir en tres unidades.

Solución E6. Con $P=\{0,1\}$ y etiqueta 1, $R(f;P,\xi)=1$.

Refinamos a $Q=\{0,1/2,1\}$ y elegimos las etiquetas 0 y $1/2$. Entonces $R(f;Q,\eta)=0\frac{1}{2}+\frac{1}{2}\frac{1}{2}=\frac{1}{4}$.

La suma disminuyó de 1 a $1/4$.

Ahora tomemos en P la etiqueta 0. La suma es $R(f;P,\xi')=0$.

En Q, usando extremos derechos $1/2$ y 1, $R(f;Q,\eta')=\frac{1}{2}\frac{1}{2}+1\frac{1}{2}=\frac{3}{4}$.

Esta vez la suma aumentó.

Proposición 13.6 no habla de sumas etiquetadas arbitrarias: afirma monotonía de las sumas **inferior y superior** bajo refinamiento. Cambiar simultáneamente la partición y las etiquetas no produce una relación de orden fija entre las sumas de Riemann.

Soluciones del nivel F. **Solución F1.** En $[0, 1/2]$, $\omega_1 = \frac{1}{4} - 0 = \frac{1}{4}$.
En $[1/2, 1]$, $\omega_2 = 1 - \frac{1}{4} = \frac{3}{4}$.
Como ambas anchuras son $1/2$, $\omega_1 \frac{1}{2} + \omega_2 \frac{1}{2} = \frac{1}{8} + \frac{3}{8} = \frac{1}{2}$.
Por D1, $U(f, P) - L(f, P) = \frac{5}{8} - \frac{1}{8} = \frac{1}{2}$.
Por tanto se verifica exactamente la identidad de Proposición 13.8.

Solución F2. Cada subintervalo de P_n es

$$I_k = \left[\frac{k-1}{n}, \frac{k}{n}\right].$$

Como $f(x) = x$ es creciente, $\omega_k = \frac{k}{n} - \frac{k-1}{n} = \frac{1}{n}$.
También $\Delta x_k = \frac{1}{n}$.
La identidad de la brecha da

$$U(f, P_n) - L(f, P_n) = \sum_{k=1}^{n} \frac{1}{n}\frac{1}{n} = n\frac{1}{n^2} = \boxed{\frac{1}{n}}.$$

Aquí cada pieza contribuye exactamente $1/n^2$.

Solución F3. En las piezas exteriores la función es idénticamente cero, por lo que la oscilación es 0. En los dos intervalos adyacentes a c, el punto c pertenece al subintervalo cerrado y la función toma allí 5, mientras toma 0 en los demás puntos. La oscilación es, por tanto, 5 en cada una de esas dos piezas.

Ambas tienen anchura δ. Así,

$$U(s, Q_\delta) - L(s, Q_\delta) = 5\delta + 5\delta = \boxed{10\delta}.$$

La amplitud vertical no se ha reducido: sigue siendo 5. Lo que disminuye con δ es la anchura horizontal total, 2δ, sobre la que aparece esa oscilación.

Solución F4. Cada subintervalo de una partición tiene interior no vacío. Por densidad de racionales e irracionales contiene puntos de ambos tipos. Por ello, en toda pieza,

$$m_k = -1, \qquad M_k = 2, \qquad \omega_k = 3.$$

Entonces

$$L(d, P) = \sum_k (-1)\Delta x_k = -(1 - 0) = \boxed{-1},$$

y

$$U(d, P) = \sum_k 2\Delta x_k = 2(1 - 0) = \boxed{2}.$$

La brecha es

$$\boxed{3}.$$

Ningún corte finito reduce la oscilación vertical de una pieza.

Solución F5. Para la partición uniforme, $\|P_n\| = \frac{1}{n} \to 0$.

Sin embargo, por F4 cada subintervalo contiene racionales e irracionales y tiene oscilación 3. Por tanto $U(d, P_n) - L(d, P_n) = 3$ para todo n. El ejemplo refuta la inferencia

«si la malla es pequeña, entonces la brecha de Darboux es pequeña».

La malla controla solo el ancho horizontal máximo. La brecha depende además de la oscilación vertical dentro de cada pieza.

Solución F6. Si $A \le f(x) \le B$, entonces en cada subintervalo

$$A \le m_k \le M_k \le B,$$

y por tanto

$$0 \le \omega_k = M_k - m_k \le B - A.$$

Usando Proposición 13.8,

$$\begin{aligned} 0 \le U - L &= \sum_k \omega_k \Delta x_k \\ &\le \sum_k (B - A)\Delta x_k \\ &= (B - A)(b - a). \end{aligned}$$

Para la segunda parte, los términos fuera de S son cero. Así,

$$U - L = \sum_{k \in S} \omega_k \Delta x_k \le \sum_{k \in S} C \Delta x_k = C \sum_{k \in S} \Delta x_k \le \boxed{CW}.$$

La interpretación es la misma que en los ejemplos de puntas aisladas: una oscilación potencialmente grande solo afecta la brecha en la medida en que ocupa anchura horizontal.

Soluciones del nivel G. **Solución G1.** La integral elemental usa las alturas de los intervalos abiertos:

$$\int_0^2 s = (-1)(1) + 4(1) = \boxed{3}.$$

Para $P = \{0, 1, 2\}$, en $[0, 1]$ la función toma los valores -1 y 2, de modo que el ínfimo es -1 y el supremo 2. En $[1, 2]$ toma 2 y 4, así que el ínfimo es 2 y el supremo 4. Por tanto,

$$\boxed{L(s, P) = (-1)(1) + 2(1) = 1},$$

y

$$\boxed{U(s, P) = 2(1) + 4(1) = 6}.$$

Los cuatro valores pedidos se realizan así:

- etiquetas $(1/2, 1)$: $R = -1 + 2 = 1$;
- etiquetas $(1/2, 3/2)$: $R = -1 + 4 = 3$;
- etiquetas $(1, 1)$: $R = 2 + 2 = 4$;
- etiquetas $(1, 3/2)$: $R = 2 + 4 = 6$.

Para Q_δ, las piezas exteriores tienen oscilación cero. En $[1-\delta, 1]$ los valores relevantes son -1 y 2, por lo que la oscilación es 3; en $[1, 1+\delta]$ es $4-2=2$. Ambas anchuras son δ. Por tanto,

$$\boxed{U(s, Q_\delta) - L(s, Q_\delta) = 3\delta + 2\delta = 5\delta.}$$

No hay contradicción. La integral elemental ya está bien definida para la escalonada y vale 3. L, U y las sumas etiquetadas son construcciones dependientes de una partición y, en el caso etiquetado, también de las etiquetas. El refinamiento puede estrechar sus cotas sin cambiar la integral elemental.

Solución G2. De D5,

$$J(f, P) = \left[\frac{1}{4}, \frac{3}{4}\right],$$

y

$$J(f, Q) = \left[\frac{2}{9}, \frac{7}{9}\right].$$

El refinamiento común es

$$R = \left\{0, \frac{1}{3}, \frac{1}{2}, 1\right\},$$

y también por D5,

$$J(f, R) = \left[\frac{11}{36}, \frac{25}{36}\right].$$

En denominador 36,

$$J(f, P) = \left[\frac{9}{36}, \frac{27}{36}\right],$$

y

$$J(f, Q) = \left[\frac{8}{36}, \frac{28}{36}\right].$$

Por tanto,

$$J(f, P) \cap J(f, Q) = \left[\frac{9}{36}, \frac{27}{36}\right],$$

y claramente

$$\boxed{\left[\frac{11}{36}, \frac{25}{36}\right] \subseteq \left[\frac{9}{36}, \frac{27}{36}\right].}$$

Los puntos medios de las tres piezas de R son

$$\frac{1}{6}, \qquad \frac{5}{12}, \qquad \frac{3}{4}.$$

La suma etiquetada es

$$\begin{aligned} R(f;R,\xi) &= \frac{1}{6}\frac{1}{3} + \frac{5}{12}\frac{1}{6} + \frac{3}{4}\frac{1}{2} \\ &= \frac{1}{18} + \frac{5}{72} + \frac{3}{8} \\ &= \frac{4}{72} + \frac{5}{72} + \frac{27}{72} \\ &= \boxed{\frac{1}{2}}. \end{aligned}$$

El número $1/2$ pertenece a $J(f,R)$ y, por las inclusiones demostradas, también a $J(f,P)$ y $J(f,Q)$.

Solución G3. Dado $\eta > 0$, podemos elegir

$$0 < \delta < \text{mín}\left\{\frac{1}{2}, \frac{\eta}{2}\right\}.$$

En Q_δ, el ínfimo de s es 0 en todas las piezas. El supremo es 1 solo en los dos subintervalos cerrados adyacentes a $1/2$, ambos de anchura δ. Por tanto, $L(s,Q_\delta) = 0$, y $U(s,Q_\delta) = 2\delta < \eta$. Por Proposición 13.7, toda elección de etiquetas satisface

$$0 = L(s,Q_\delta) \le R(s;Q_\delta,\xi) \le U(s,Q_\delta) = 2\delta < \eta.$$

Así,

$$\boxed{0 \le R(s;Q_\delta,\xi) < \eta}$$

para todas las etiquetas de esa partición. Lo demostrado es una afirmación finita y específica sobre esta función y una partición construida a partir de η: todas sus sumas etiquetadas quedan encerradas en un intervalo de ancho menor que η. No hemos usado ni formulado la definición general de integrabilidad de Riemann de C15, ni integrales inferior/superior globales.

Solución G4. Cada subintervalo de P_n tiene interior no vacío. Por densidad contiene al menos un racional y un irracional. Podemos, por tanto, escoger en cada pieza una etiqueta racional $\xi_k^{(n)}$ y una irracional $\eta_k^{(n)}$.

Con el etiquetado racional, $d(\xi_k^{(n)}) = 1$ para todo k. Como las anchuras suman 1,

$$R(d;P_n,\xi^{(n)}) = \sum_{k=1}^{n} 1 \cdot \Delta x_k = \boxed{1}.$$

Con el etiquetado irracional, $d(\eta_k^{(n)}) = 0$ para todo k, de modo que

$$R(d;P_n,\eta^{(n)}) = \boxed{0}.$$

En cada pieza el ínfimo es 0 y el supremo es 1. Por tanto,

$$L(d,P_n) = 0, \qquad U(d,P_n) = 1.$$

Así obtenemos dos familias de sumas sobre las **mismas particiones uniformes**: una es constantemente 1 y la otra constantemente 0. Cada familia converge, pero a un número distinto. Por ello la convergencia de una sola elección sistemática de etiquetas no produce un candidato

independiente del procedimiento. C14 llega hasta este diagnóstico; C15 formulará la condición global que falta.

Capítulo 14

Integral de Riemann: definición, integrabilidad y propiedades

Para una función acotada $f : [a, b] \to \mathbb{R}$, cada partición P proporciona una suma inferior $L(f, P)$, una superior $U(f, P)$ y sumas etiquetadas $R(f; P, \xi)$ comprendidas entre ambas. El refinamiento estrecha ese encierro, pero todavía debemos relacionar los resultados de todas las particiones. Pero ninguno de esos números pertenece todavía a la función **sin referencia a una partición**. Ese es el problema de este capítulo. Buscamos una cantidad que dependa únicamente de f y del intervalo. Para eliminar la elección auxiliar de una partición consideraremos las familias completas de sumas inferiores y superiores y preguntaremos si sus extremos coinciden.

La completitud de $\mathbb{R}$ permite tomar el supremo de un conjunto no vacío acotado superiormente y el ínfimo de uno no vacío acotado inferiormente. Aplicaremos esa propiedad a las familias de sumas; primero tendremos que demostrar que satisfacen las condiciones necesarias.

La arquitectura será

sumas finitas $\longrightarrow$ familias de sumas $\longrightarrow$ supremo e ínfimo $\longrightarrow$ pregunta por su coincidencia.

La definición de integrabilidad debe establecerse antes de relacionar la integral con las derivadas o las primitivas. Por eso partiremos de las sumas y de sus extremos globales; los teoremas de valor medio y el teorema fundamental se apoyarán después en esta construcción.

14.1. De todas las particiones a dos extremos globales

En C14 cada partición produjo información localmente organizada pero globalmente finita. Si $P = \{x_0, \dots, x_n\}$, entonces Definición 13.4 asocia a una función acotada $f : [a, b] \to \mathbb{R}$ los números

$$L(f, P) \qquad \text{y} \qquad U(f, P).$$

La primera tentación podría ser escoger una partición especialmente conveniente y declarar que una de esas cantidades, o alguna suma etiquetada intermedia, es «la» integral. Eso haría depender el objeto buscado de una elección auxiliar. El cambio decisivo consiste en no privilegiar ninguna partición.

La colección de todas las particiones. Fijemos $a < b$. Denotaremos por $\mathcal{P}[a, b]$ la colección de todas las particiones finitas de $[a, b]$ en el sentido de Definición 13.1. Esta colección no es vacía: al menos contiene la partición más sencilla, $\{a, b\}$. Si $f : [a, b] \to \mathbb{R}$ es acotada, cada $P \in \mathcal{P}[a, b]$ produce dos números reales. Podemos reunirlos en dos conjuntos: $\mathcal{L}(f) := \{L(f, P) : P \in \mathcal{P}[a, b]\}$, y $\mathcal{U}(f) := \{U(f, P) : P \in \mathcal{P}[a, b]\}$. La notación debe leerse con cuidado.

- $\mathcal{L}(f)$ no es una suma inferior: es el **conjunto de todas las sumas inferiores** de f.
- $\mathcal{U}(f)$ no es una suma superior: es el **conjunto de todas las sumas superiores** de f.
- cambiar P cambia un elemento de la familia, pero no cambia la familia que estamos considerando.

El problema ya no es, por tanto, comparar dos números asociados a una misma partición. Es comprender la posición relativa de dos conjuntos de números reales.

La compatibilidad finita produce acotación global. En Corolario 13.2 demostramos un hecho más fuerte que la desigualdad $L(f,P) \le U(f,P)$. Para **dos particiones arbitrarias** $P, Q \in \mathcal{P}[a,b]$ se cumple

$$\boxed{L(f,P) \le U(f,Q).}$$

La demostración utilizó un refinamiento común. Ahora esa desigualdad adquiere una función nueva. Fijemos temporalmente una partición Q. Para toda partición P tenemos $L(f,P) \le U(f,Q)$. Por tanto, el único número $U(f,Q)$ es una cota superior de **todo** el conjunto $\mathcal{L}(f)$. Como Q era arbitraria, obtenemos algo todavía más fuerte:

> **cada suma superior es una cota superior del conjunto de todas las sumas inferiores.**

De manera dual, si fijamos una partición P, la misma desigualdad muestra que $L(f,P) \le U(f,Q)$ para toda Q. Luego $L(f,P)$ es una cota inferior de todo el conjunto $\mathcal{U}(f)$.

Así,

> **cada suma inferior es una cota inferior del conjunto de todas las sumas superiores.**

La situación puede resumirse como

$$\boxed{\mathcal{L}(f) \quad \text{queda completamente a la izquierda de} \quad \mathcal{U}(f).}$$

Más precisamente,

$$\ell \le u \qquad \text{para todo } \ell \in \mathcal{L}(f) \text{ y todo } u \in \mathcal{U}(f).$$

Este es el punto en el que la construcción finita de C14 se vuelve una afirmación global.

Aquí entra la completitud. Sabemos ahora cuatro cosas:

1. $\mathcal{L}(f) \neq \varnothing$;
2. $\mathcal{U}(f) \neq \varnothing$;
3. $\mathcal{L}(f)$ está acotado superiormente;
4. $\mathcal{U}(f)$ está acotado inferiormente.

La definición de supremo e ínfimo ya fue fijada en Definición 1.7. Lo nuevo es que podemos **garantizar su existencia**. Por el axioma de completitud formulado en Sección 1.5.2, el conjunto no vacío y acotado superiormente $\mathcal{L}(f)$ posee un supremo real. Por Proposición 1.10, el conjunto no vacío y acotado inferiormente $\mathcal{U}(f)$ posee un ínfimo real. La completitud asegura que esos supremos e ínfimos son números reales. Escribir los símbolos no bastaría para justificar su existencia sin las cotas y la no vacuidad de las familias a las que los aplicamos. Ya podemos dar nombre a esos dos extremos.

Definición 14.1. **Integral inferior global e integral superior global.** Sean $a < b$ y sea $f : [a,b] \to \mathbb{R}$ una función acotada.

Definimos la **integral inferior** de f en $[a,b]$ por

$$\boxed{\underline{\int_a^b} f(x)\,dx := \sup \mathcal{L}(f) = \sup_{P\in\mathcal{P}[a,b]} L(f,P).}$$

Definimos la **integral superior** de f en $[a,b]$ por

$$\boxed{\overline{\int_a^b} f(x)\,dx := \operatorname{ínf} \mathcal{U}(f) = \operatorname*{ínf}_{P\in\mathcal{P}[a,b]} U(f,P).}$$

Ambos números existen por la completitud de $\mathbb{R}$ y por la compatibilidad entre sumas inferiores y superiores demostrada en C14.

La palabra **global** sirve para recordar el cambio de nivel:

$$L(f,P),\ U(f,P) \quad \text{dependen de una partición,}$$

mientras que

$$\underline{\int_a^b} f, \qquad \overline{\int_a^b} f$$

se obtienen después de considerar **todas** las particiones. No debe confundirse tampoco «supremo» con «máximo» ni «ínfimo» con «mínimo». Puede ocurrir que ninguna partición produzca exactamente la integral inferior o la integral superior. La definición no exige una «mejor partición»: exige solamente los extremos de las dos familias de números.

Todo encierro finito contiene al encierro global. La posición relativa de las dos familias proporciona inmediatamente una cadena fundamental.

Proposición 14.1. **Proposición (encierro global por sumas de Darboux).** Sea $f : [a,b] \to \mathbb{R}$ acotada. Para cualesquiera particiones $P, Q \in \mathcal{P}[a,b]$,

$$\boxed{L(f,P) \le \underline{\int_a^b} f(x)\,dx \le \overline{\int_a^b} f(x)\,dx \le U(f,Q).}$$

Demostración. Como $L(f,P) \in \mathcal{L}(f)$, la definición de supremo implica

$$L(f,P) \le \sup \mathcal{L}(f) = \underline{\int_a^b} f(x)\,dx.$$

Análogamente, $U(f,Q) \in \mathcal{U}(f)$, y la definición de ínfimo da

$$\overline{\int_a^b} f(x)\,dx = \operatorname{ínf} \mathcal{U}(f) \le U(f,Q).$$

Falta comparar los dos extremos globales. Fijemos una partición Q. Por Corolario 13.2, $L(f,P) \le U(f,Q)$ para toda partición P. Así, $U(f,Q)$ es una cota superior de $\mathcal{L}(f)$. Como la integral inferior es la **menor** cota superior,

$$\underline{\int_a^b} f(x)\,dx \le U(f,Q).$$

Extremos globales de Darboux

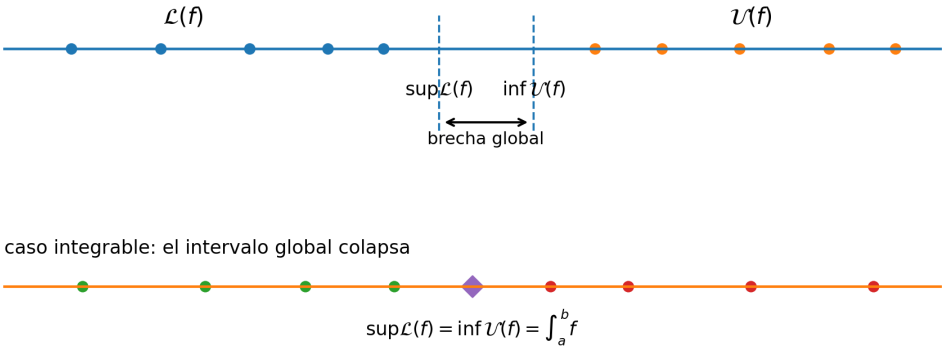


FIGURA 14.1. Familias de sumas inferiores y superiores y sus extremos globales. La integrabilidad corresponde al colapso de la brecha global.

Esto vale para toda partición Q. Por tanto, la integral inferior es una cota inferior de $\mathcal{U}(f)$. Como la integral superior es la **mayor** cota inferior de ese conjunto,

$$\underline{\int_a^b} f(x)\,dx \le \overline{\int_a^b} f(x)\,dx.$$

Uniendo las tres desigualdades obtenemos la cadena anunciada. □

La proposición dice algo más fuerte que

$$\underline{\int_a^b} f \le \overline{\int_a^b} f.$$

Afirma que **cada** suma inferior queda por debajo del extremo inferior global y que **cada** suma superior queda por encima del extremo superior global. Así, cualquier encierro finito $[L(f,P), U(f,Q)]$ contiene al intervalo global

$$\left[\underline{\int_a^b} f(x)\,dx,\ \overline{\int_a^b} f(x)\,dx\right].$$

No hemos afirmado todavía que ese intervalo tenga anchura cero.

Dos extremos pueden coincidir o permanecer separados. Los casos más simples muestran ambas posibilidades sin introducir aún ninguna definición nueva.

Una función constante. Si

$$f(x) = c \qquad \text{para todo } x \in [a,b],$$

entonces en toda partición P los ínfimos y supremos locales son iguales a c. Por tanto, $L(f,P) = U(f,P) = c(b-a)$. Las dos familias son el mismo conjunto unitario: $\mathcal{L}(f) = \mathcal{U}(f) = \{c(b-a)\}$.

Luego

$$\underline{\int_a^b} f(x)\,dx = \overline{\int_a^b} f(x)\,dx = c(b-a).$$

Aquí los dos extremos globales coinciden. El fenómeno opuesto también puede persistir incluso después de considerar todas las particiones.

EJEMPLO 14.1. **Dos extremos globales separados en cada escala finita.** Consideremos la función

$$d(x) = \begin{cases} 1, & x \in \mathbb{Q}, \\ 0, & x \notin \mathbb{Q}, \end{cases} \qquad 0 \le x \le 1.$$

En Ejemplo 13.15 vimos que todo subintervalo no degenerado contiene racionales e irracionales. Por tanto, para **cualquier** partición P de $[0,1]$,

$$L(d,P) = 0, \qquad U(d,P) = 1.$$

En consecuencia,

$$\mathcal{L}(d) = \{0\}, \qquad \mathcal{U}(d) = \{1\},$$

y de Definición 14.1 obtenemos

$$\boxed{\underline{\int_0^1} d(x)\,dx = 0, \qquad \overline{\int_0^1} d(x)\,dx = 1.}$$

Ningún cambio de partición modifica la separación entre ambos extremos.

Este ejemplo muestra por qué la existencia de la integral inferior y de la integral superior **no resuelve por sí sola** el problema de la integración. La completitud garantiza dos números, pero no garantiza que sean el mismo.

Lo que hemos construido y lo que todavía falta. C14 trabajaba con una partición cada vez: $P \longmapsto [L(f,P), U(f,P)]$. C15 ha comenzado considerando todas ellas simultáneamente:

$$\{L(f,P)\}_P \qquad \text{y} \qquad \{U(f,P)\}_P.$$

La completitud transforma esas dos familias en dos números bien definidos:

$$\boxed{\underline{\int_a^b} f \le \overline{\int_a^b} f.}$$

Pero todavía no hemos definido

$$\int_a^b f(x)\,dx$$

para una función acotada general. Ese símbolo sin barra inferior ni superior aparecerá solo después de contestar la pregunta que ahora queda aislada con precisión:

$$\boxed{\text{¿cuándo coinciden la integral inferior y la integral superior?}}$$

La respuesta será la definición de integrabilidad de Riemann de §14.2.

No adelantemos la definición
En esta sección, toda función acotada posee una integral inferior global y una integral superior global. No hemos dicho todavía que toda función acotada sea integrable, ni hemos definido la integral de Riemann general. La función de Ejemplo 14.1 muestra exactamente por qué esas afirmaciones serían distintas.

La colección $\mathcal{L}(f)$ es no vacía porque existen particiones de $[a,b]$. La comparación mediante refinamiento común permite acotar todas sus sumas inferiores por cualquier suma superior. La completitud produce entonces el supremo, aunque ninguna partición lo alcance. Quedará por decidir si ese extremo coincide con el ínfimo de las sumas superiores.

14.2. Definición de integrabilidad de Riemann

La sección anterior dejó aislada una única obstrucción. Para toda función acotada $f : [a,b] \to \mathbb{R}$ existen dos números reales,

$$\underline{\int_a^b} f(x)\,dx \le \overline{\int_a^b} f(x)\,dx,$$

pero no hay razón automática para que sean iguales. La función de Ejemplo 14.1 muestra que la separación puede persistir aunque consideremos **todas** las particiones. La definición de integral de Riemann consiste precisamente en exigir que ese intervalo global se reduzca a un solo punto.

Definición 14.2. **Integrabilidad de Riemann e integral de Riemann.** Sean $a < b$ y sea $f : [a,b] \to \mathbb{R}$ una función acotada. Diremos que f es **Riemann integrable** en $[a,b]$ si

$$\boxed{\underline{\int_a^b} f(x)\,dx = \overline{\int_a^b} f(x)\,dx.}$$

Cuando esta igualdad se cumple, llamamos **integral de Riemann** de f en $[a,b]$ al valor común y escribimos

$$\boxed{\int_a^b f(x)\,dx := \underline{\int_a^b} f(x)\,dx = \overline{\int_a^b} f(x)\,dx.}$$

Denotaremos por $\mathcal{R}[a,b]$ la clase de las funciones Riemann integrables en $[a,b]$.

Hay dos partes de la definición que conviene mantener separadas. Primero, la **acotación** no es una conclusión de esta formulación: es una hipótesis previa, necesaria para que las sumas de Darboux de C14 y los extremos globales de Definición 14.1 estén definidos como números reales. Segundo, la palabra **integrable** designa una propiedad de la función; el símbolo

$$\int_a^b f(x)\,dx$$

se introduce solo cuando esa propiedad ya ha sido verificada. No estamos asignando por convención un número a toda función acotada.

El valor integral es único. La unicidad no requiere un nuevo teorema. Las integrales inferior y superior globales son números determinados por f y por $[a,b]$. Si coinciden, su valor común queda determinado de manera única. En otras palabras, no hay una elección escondida en la notación

$$\int_a^b f(x)\,dx.$$

No elegimos una partición, ni una sucesión de particiones, ni etiquetas particulares. El número integral aparece únicamente cuando los dos extremos construidos sobre **todas** las particiones son el mismo. Esta observación evita una confusión frecuente: una sucesión concreta de aproximaciones puede converger a cierto número y, sin embargo, eso por sí solo no convierte ese número en la integral de Riemann. La definición exige una afirmación global sobre las familias completas de sumas inferiores y superiores.

Toda suma inferior y superior encierra la integral. Si $f \in \mathcal{R}[a,b]$, Proposición 14.1 se simplifica inmediatamente. Para toda partición $P \in \mathcal{P}[a,b]$,

$$L(f,P) \le \underline{\int_a^b} f(x)\,dx = \int_a^b f(x)\,dx = \overline{\int_a^b} f(x)\,dx \le U(f,P).$$

Por tanto,

$$\boxed{L(f,P) \le \int_a^b f(x)\,dx \le U(f,P) \qquad \text{para toda partición } P.}$$

Y si ξ es una elección de etiquetas sobre la misma partición, el encierro de C14 da además $L(f,P) \le R(f;P,\xi) \le U(f,P)$. Obsérvese lo que **no** se sigue: en general no podemos ordenar directamente una suma etiquetada y la integral,

$$R(f;P,\xi) \quad \text{y} \quad \int_a^b f.$$

Ambas están encerradas entre los mismos extremos, pero cualquiera de las dos puede quedar por encima de la otra. Lo que importa para la teoría será poder hacer tan estrecho como queramos el intervalo $[L(f,P), U(f,P)]$. Ese será exactamente el contenido de §14.3.

Dos decisiones inmediatas: una función integrable y otra que no lo es. Las funciones constantes ya quedaron resueltas en §14.1. Si $f(x) = c$ en $[a,b]$, entonces

$$\underline{\int_a^b} f(x)\,dx = \overline{\int_a^b} f(x)\,dx = c(b-a),$$

por lo que $f \in \mathcal{R}[a,b]$ y

$$\int_a^b c\,dx = c(b-a).$$

Por el contrario, para la función racional/irracional d de Ejemplo 14.1,

$$\underline{\int_0^1} d(x)\,dx = 0, \qquad \overline{\int_0^1} d(x)\,dx = 1.$$

Luego $d \notin \mathcal{R}[0,1]$.

No hace falta buscar una partición más ingeniosa: la separación ya ha sido calculada después de considerar **todas** las particiones. El siguiente ejemplo muestra un tercer comportamiento. Una familia concreta de particiones no define la integral, pero puede servir para **certificar** que los dos extremos globales tienen que coincidir.

EJEMPLO 14.2. **La función identidad en $[0,1]$: una familia que atrapa los extremos globales.** Sea

$$f(x) = x, \qquad 0 \le x \le 1.$$

Para $n \ge 1$ consideremos la partición uniforme

$$P_n = \left\{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1\right\}.$$

En el subintervalo

$$\left[\frac{k-1}{n}, \frac{k}{n}\right]$$

la función identidad tiene ínfimo $(k-1)/n$ y supremo k/n. Como todos los subintervalos tienen longitud $1/n$,

$$L(f, P_n) = \frac{1}{n}\sum_{k=1}^{n} \frac{k-1}{n} = \frac{n-1}{2n} = \frac{1}{2} - \frac{1}{2n},$$

y

$$U(f, P_n) = \frac{1}{n}\sum_{k=1}^{n} \frac{k}{n} = \frac{n+1}{2n} = \frac{1}{2} + \frac{1}{2n}.$$

Por Proposición 14.1, para todo n,

$$\frac{1}{2} - \frac{1}{2n} \le \underline{\int_0^1} f(x)\,dx \le \overline{\int_0^1} f(x)\,dx \le \frac{1}{2} + \frac{1}{2n}.$$

Mostremos que los dos números centrales deben ser $1/2$.
Si

$$\underline{\int_0^1} f(x)\,dx < \frac{1}{2},$$

la diferencia

$$\delta := \frac{1}{2} - \underline{\int_0^1} f(x)\,dx$$

sería positiva. Eligiendo n suficientemente grande para que $1/(2n) < \delta$, obtendríamos

$$\frac{1}{2} - \frac{1}{2n} > \underline{\int_0^1} f(x)\,dx,$$

contradicción. Por tanto,

$$\frac{1}{2} \leq \underline{\int_0^1} f(x)\,dx.$$

De manera análoga, si la integral superior fuese mayor que $1/2$, un n suficientemente grande produciría

$$\overline{\int_0^1} f(x)\,dx > \frac{1}{2} + \frac{1}{2n},$$

contradiciendo el encierro anterior. Luego

$$\overline{\int_0^1} f(x)\,dx \leq \frac{1}{2}.$$

Como siempre tenemos

$$\underline{\int_0^1} f(x)\,dx \leq \overline{\int_0^1} f(x)\,dx,$$

concluimos

$$\boxed{\underline{\int_0^1} f(x)\,dx = \overline{\int_0^1} f(x)\,dx = \frac{1}{2}.}$$

Así, $f \in \mathcal{R}[0,1]$ y

$$\boxed{\int_0^1 x\,dx = \frac{1}{2}.}$$

El ejemplo es importante por su lógica. No hemos dicho «las sumas de esta familia convergen, luego la función es integrable». Lo que hicimos fue usar, para **cada** n, el encierro global de Proposición 14.1. Las cotas producidas por P_n obligaron a las integrales inferior y superior —que ya estaban definidas sobre todas las particiones— a caer en intervalos cada vez más estrechos. Ese mecanismo sugiere una caracterización mucho más útil que la igualdad abstracta de dos extremos:

poder encontrar particiones cuya brecha $\quad U(f,P) - L(f,P) \quad$ sea arbitrariamente pequeña.

Pero «arbitrariamente pequeña» contiene un cuantificador que todavía debemos formular y demostrar correctamente. No basta con exhibir una sola partición buena, ni con observar numéricamente que varias brechas parecen decrecer.

Definición primero, criterio después
En §14.2, la integrabilidad de Riemann significa exclusivamente

$$\underline{\int_a^b} f = \overline{\int_a^b} f.$$

La condición $\forall \varepsilon > 0\ \exists P : U(f,P) - L(f,P) < \varepsilon$ aún **no** ha sido declarada equivalente. Esa equivalencia será el teorema de §14.3.

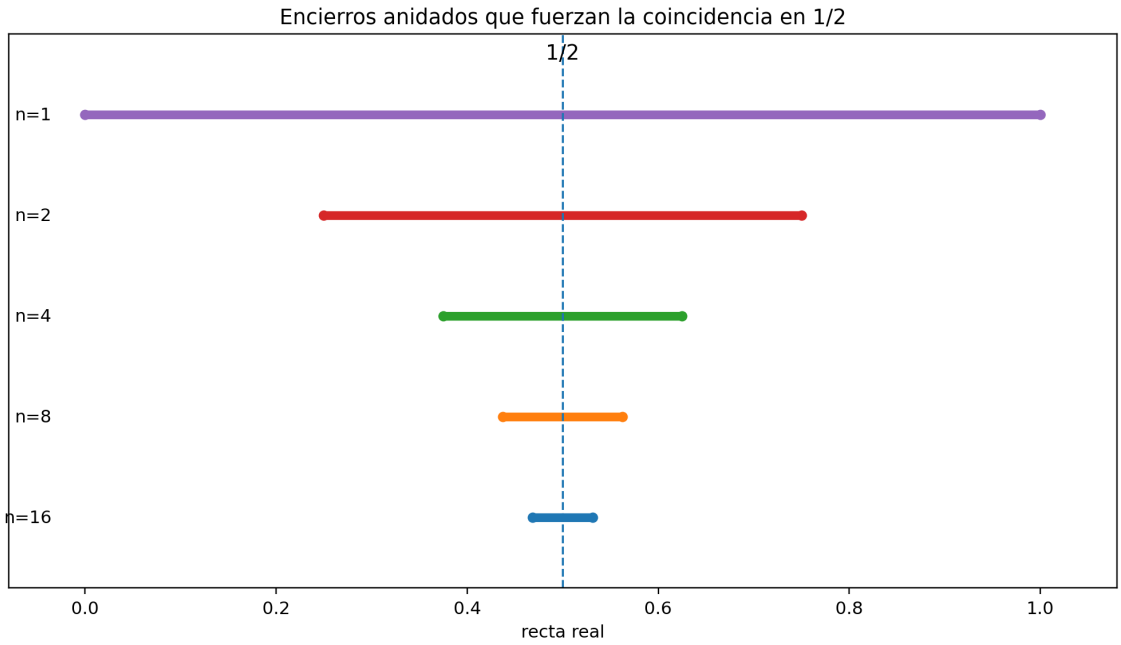


FIGURA 14.2. Los encierros producidos por las particiones uniformes de la identidad se estrechan alrededor de 1/2.

La integrabilidad afirma la existencia de un valor común a los extremos globales; calcularlo es una tarea adicional. Una familia particular de sumas puede sugerir el número buscado, pero para certificarlo debemos enlazarla con la definición, por ejemplo demostrando que las brechas de Darboux se hacen arbitrariamente pequeñas. La proposición Proposición 14.1 garantiza el encierro por cualquier partición cuando el valor integral existe; en Ejemplo 14.2 se justifica su existencia mediante el control de la brecha.

14.3. El criterio de Darboux: brechas arbitrariamente pequeñas

La definición de §14.2 es conceptualmente limpia, pero no siempre es la forma más cómoda de demostrar integrabilidad. Para usarla directamente habría que identificar dos extremos globales,

$$\underline{\int_a^b} f \qquad \text{y} \qquad \overline{\int_a^b} f,$$

y después probar que coinciden.

El ejemplo Ejemplo 14.2 sugiere una estrategia más local. No calculamos primero los dos extremos: construimos particiones cuyos encierros $[L(f,P), U(f,P)]$ son cada vez más estrechos. Si podemos hacer su anchura menor que cualquier número positivo prefijado, entonces no queda espacio para que las integrales inferior y superior permanezcan separadas. La palabra decisiva es **cualquier**.

Del intervalo global a una partición suficientemente buena. Recordemos que, por Proposición 14.1, para toda partición P,

$$L(f,P) \le \underline{\int_a^b} f \le \overline{\int_a^b} f \le U(f,P).$$

Por consiguiente,

$$0 \le \overline{\int_a^b} f - \underline{\int_a^b} f \le U(f,P) - L(f,P).$$

Toda brecha de una partición domina la separación global. La pregunta es si la recíproca lógica también puede obtenerse: cuando la separación global es cero, ¿podemos encontrar una sola partición cuya brecha sea tan pequeña como queramos?

Sí, pero hace falta usar correctamente las propiedades de supremo e ínfimo y después combinar dos particiones distintas.

TEOREMA 14.1. **Teorema (criterio de Darboux).** Sean $a < b$ y sea $f : [a,b] \to \mathbb{R}$ acotada. Entonces son equivalentes:

1. f es Riemann integrable en $[a,b]$;
2. para todo $\varepsilon > 0$ existe una partición $P \in \mathcal{P}[a,b]$ tal que

$$\boxed{U(f,P) - L(f,P) < \varepsilon.}$$

Demostración. Supongamos primero que f es Riemann integrable y escribamos

$$I := \int_a^b f(x)\,dx.$$

Entonces, por Definición 14.2,

$$I = \sup_{P\in\mathcal{P}[a,b]} L(f,P) = \inf_{P\in\mathcal{P}[a,b]} U(f,P).$$

Sea $\varepsilon > 0$.

Como I es el supremo de todas las sumas inferiores, $I - \varepsilon/2$ no puede ser una cota superior de $\mathcal{L}(f)$. Por tanto existe una partición P_- tal que $L(f,P_-) > I - \frac{\varepsilon}{2}$. Análogamente, como I es el ínfimo de todas las sumas superiores, $I + \varepsilon/2$ no puede ser una cota inferior de $\mathcal{U}(f)$. Existe entonces una partición P_+ con $U(f,P_+) < I + \frac{\varepsilon}{2}$. No hay razón para que P_- y P_+ sean la misma partición. Tomemos su refinamiento común $P := P_- \vee P_+$. Por la monotonía bajo refinamiento de Proposición 13.6, $L(f,P) \ge L(f,P_-) > I - \frac{\varepsilon}{2}$, y $U(f,P) \le U(f,P_+) < I + \frac{\varepsilon}{2}$.

Restando,

$$U(f,P) - L(f,P) < \left(I + \frac{\varepsilon}{2}\right) - \left(I - \frac{\varepsilon}{2}\right) = \varepsilon.$$

Esto demuestra la segunda condición. Recíprocamente, supongamos que para todo $\varepsilon > 0$ existe una partición P con $U(f,P) - L(f,P) < \varepsilon$.

Escribamos

$$\underline{I} := \underline{\int_a^b} f(x)\,dx, \qquad \overline{I} := \overline{\int_a^b} f(x)\,dx.$$

Por Proposición 14.1, para cualquier partición P,

$$L(f,P) \le \underline{I} \le \overline{I} \le U(f,P).$$

Luego $0 \le \overline{I} - \underline{I} \le U(f,P) - L(f,P)$.

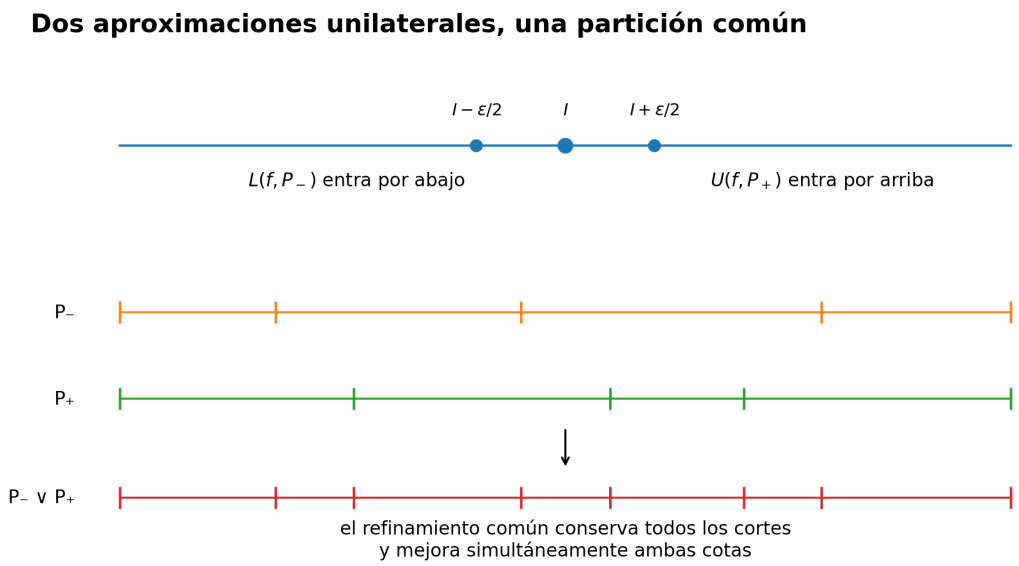


FIGURA 14.3. El refinamiento común combina la aproximación por abajo y la aproximación por arriba en una sola partición.

Dado $\varepsilon > 0$, elegimos una partición con brecha menor que ε y obtenemos $0 \le \overline{I} - \underline{I} < \varepsilon$. Si $\overline{I} - \underline{I}$ fuese positivo, podríamos tomar $\varepsilon = \frac{\overline{I}-\underline{I}}{2} > 0$, lo que daría la contradicción $\overline{I} - \underline{I} < \frac{\overline{I}-\underline{I}}{2}$.

Por tanto, $\underline{I} = \overline{I}$.

Según Definición 14.2, f es Riemann integrable. □

Qué hace realmente el refinamiento común. La primera mitad de la prueba contiene un detalle fácil de ocultar si se escribe demasiado rápido. De la igualdad $\sup_P L(f,P) = \inf_P U(f,P) = I$ obtenemos inicialmente **dos** datos distintos:

- una partición P_- cuya suma inferior está cerca de I por abajo;
- una partición P_+ cuya suma superior está cerca de I por arriba.

Nada garantiza que una sola de ellas satisfaga simultáneamente ambas aproximaciones. El refinamiento común resuelve exactamente ese problema: $P_- \vee P_+$ mejora la suma inferior de P_- y, al mismo tiempo, mejora la suma superior de P_+. Ese paso convierte dos aproximaciones unilaterales en un único encierro estrecho.

El orden de los cuantificadores importa. El criterio no dice que exista una partición con una brecha «pequeña» en algún sentido informal. Dice

$$\boxed{\forall \varepsilon > 0 \ \exists P \in \mathcal{P}[a,b] \text{ tal que } U(f,P) - L(f,P) < \varepsilon.}$$

Primero se fija la precisión exigida; después puede elegirse una partición que dependa de esa precisión. Una sola brecha positiva, aunque numéricamente pequeña, no basta. Por ejemplo, para un número fijo $c > 0$ consideremos

$$d_c(x) = \begin{cases} c, & x \in \mathbb{Q}, \\ 0, & x \notin \mathbb{Q}, \end{cases} \qquad 0 \le x \le 1.$$

En toda partición P,

$$L(d_c, P) = 0, \qquad U(d_c, P) = c,$$

así que la brecha es siempre c. Si $c = 10^{-6}$, por ejemplo, disponemos de una brecha muy pequeña en escala decimal; sin embargo, al pedir $\varepsilon = \frac{c}{2}$, ninguna partición satisface el criterio. La función sigue sin ser integrable. Por eso «pequeña» no es una noción matemática suficiente aquí. Lo relevante es **arbitrariamente pequeña**.

Una sucesión de brechas sí puede certificar el criterio. El teorema permite una reformulación útil. Si podemos construir particiones P_n tales que $U(f, P_n) - L(f, P_n) \longrightarrow 0$, entonces, dado $\varepsilon > 0$, para n suficientemente grande la brecha será menor que ε. Por Teorema 14.1, f es integrable. La conversa también es cierta: si f es integrable, para cada $n \ge 1$ el criterio permite elegir una partición P_n con $U(f, P_n) - L(f, P_n) < \frac{1}{n}$.

Así,

$$f \in \mathcal{R}[a, b] \quad \Longleftrightarrow \quad \text{existe una sucesión de particiones } (P_n) \text{ con } U(f, P_n) - L(f, P_n) \to 0.$$

La afirmación es correcta porque controla **la brecha completa de Darboux**. Es muy distinta de observar que una sucesión particular de sumas etiquetadas converge: ese último hecho, por sí solo, no controla todas las elecciones posibles y no es todavía un criterio de integrabilidad.

El criterio puede probar existencia sin calcular el valor.

Ejemplo 14.3. **Integrabilidad de x^2 sin calcular todavía su integral.** Sea

$$f(x) = x^2, \qquad 0 \le x \le 1.$$

En Ejemplo 13.14 se calculó, para la partición uniforme

$$P_n = \left\{0, \frac{1}{n}, \dots, 1\right\},$$

la brecha exacta $U(f, P_n) - L(f, P_n) = \frac{1}{n}$.

Sea $\varepsilon > 0$. Por la propiedad arquimediana podemos elegir n con $n > \frac{1}{\varepsilon}$.

Entonces $U(f, P_n) - L(f, P_n) = \frac{1}{n} < \varepsilon$.

Por Teorema 14.1,

$$\boxed{f \in \mathcal{R}[0, 1].}$$

El argumento demuestra existencia de la integral, pero no ha calculado todavía su valor. Esa distinción será importante repetidas veces: **probar integrabilidad** y **evaluar la integral** son tareas diferentes.

Refinar una partición buena conserva la certificación. Si una partición P satisface $U(f,P) - L(f,P) < \varepsilon$ y Q refina a P, entonces Proposición 13.8 implica

$$U(f,Q) - L(f,Q) \leq U(f,P) - L(f,P) < \varepsilon.$$

Por tanto, una vez encontrada una partición suficientemente buena, **todo refinamiento suyo sigue siendo suficientemente bueno** para la misma tolerancia. Esto no debe confundirse con afirmar que toda partición de malla menor es buena. C14 ya mostró que «malla menor» y «refinamiento» son relaciones diferentes. El criterio de Darboux habla de existencia de particiones con brecha pequeña; el criterio uniforme por malla, mucho más fuerte en su forma aparente, será objeto de §14.4 al introducir la equivalencia con sumas etiquetadas.

La definición no ha cambiado
El criterio de Darboux es un **teorema equivalente** a la definición de §14.2. No reemplaza retroactivamente la definición primaria de CPM. La ruta lógica sigue siendo

$$\text{extremos globales} \longrightarrow \text{definición por coincidencia} \longrightarrow \text{criterio } U - L < \varepsilon.$$

Antes de seguir. En la primera dirección de Teorema 14.1, ¿por qué aparecen inicialmente dos particiones diferentes? ¿Qué propiedad del refinamiento común permite fundirlas en una sola? ¿Por qué una brecha fija de 10^{-6} no basta para probar integrabilidad? ¿Qué diferencia hay entre una sucesión de **brechas de Darboux** que tiende a cero y una sucesión particular de sumas etiquetadas que converge? En Ejemplo 14.3, ¿qué hemos demostrado sobre $\int_0^1 x^2\,dx$ y qué sigue todavía sin determinar?

14.4. Sumas etiquetadas: equivalencia con la definición por malla

Hasta ahora la teoría de C15 ha utilizado como objeto primario las sumas inferior y superior. La definición de integrabilidad compara sus extremos globales y Teorema 14.1 permite reconocer la integrabilidad buscando particiones cuya brecha de Darboux sea arbitrariamente pequeña. Pero las sumas etiquetadas de C14 plantean una formulación que se parece más a la idea informal de «sumar rectángulos cada vez más finos». Si (P,ξ) es una partición etiquetada, recordemos que

$$R(f;P,\xi) = \sum_{k=1}^{n} f(\xi_k)\Delta x_k.$$

La pregunta correcta no es si **alguna** sucesión especialmente escogida de estas sumas converge. Debemos preguntar si, una vez que la malla es suficientemente pequeña, **todas** las sumas etiquetadas quedan cerca del mismo número, sin importar cómo se hayan distribuido los cortes ni cómo se hayan elegido las etiquetas. El orden de los cuantificadores será, por tanto,

$$\boxed{\forall \varepsilon > 0\ \exists \delta > 0\ \forall (P,\xi) : \|P\| < \delta \Longrightarrow |R(f;P,\xi) - I| < \varepsilon.}$$

Esta afirmación es considerablemente más fuerte que la convergencia de una familia particular. Antes de demostrar que equivale a la integrabilidad de Darboux necesitamos resolver una dificultad: una partición de malla pequeña no tiene por qué refinar la partición buena proporcionada por Teorema 14.1.

Insertar finitísimos cortes solo puede afectar a pocas piezas. Sea Q una partición fija. Una partición muy fina P puede omitir todos los cortes interiores de Q. Sin embargo, cada corte omitido puede caer en el interior de **a lo sumo una** pieza de P. Si f está acotada, la contribución total de esas piezas problemáticas puede controlarse mediante la malla.

LEMA 14.1. **Lema (transferencia de una partición de Darboux a particiones de malla pequeña).** Sea $f : [a, b] \to \mathbb{R}$ acotada y sea $B > 0$ tal que

$$|f(x)| \le B \qquad (x \in [a, b]).$$

Fijemos una partición

$$Q = \{q_0, q_1, \dots, q_m\}, \qquad q_0 = a < q_1 < \cdots < q_m = b,$$

y sea $r = m - 1$ el número de sus puntos interiores. Entonces, para toda partición P de $[a, b]$,

$$\boxed{U(f, P) - L(f, P) \le U(f, Q) - L(f, Q) + 2Br\,\|P\|.}$$

Demostración. Formemos el refinamiento común $S := P \vee Q$.

Como S refina a Q, Proposición 13.8 da

$$U(f, S) - L(f, S) \le U(f, Q) - L(f, Q).$$

Comparemos ahora las brechas de P y S. Un subintervalo de P que no contenga en su interior ningún punto q_j de Q no se divide al pasar de P a S; por tanto, su contribución a la brecha permanece exactamente igual. Solo pueden cambiar los subintervalos de P que contienen algún punto interior de Q. Llamémoslos **subintervalos afectados**. Hay a lo sumo r de ellos: distintos puntos de Q pueden incluso caer en una misma pieza de P, pero cada pieza afectada contiene por lo menos uno de los r cortes interiores. Si $I = [x_{k-1}, x_k]$ es una pieza afectada, la acotación $|f| \le B$ implica $\sup_I f - \inf_I f \le 2B$.

Su contribución original a la brecha de P es, por tanto, a lo sumo $2B\,\Delta x_k \le 2B\,\|P\|$. La suma de las contribuciones originales de todas las piezas afectadas es a lo sumo $2Br\,\|P\|$. En las piezas no afectadas las contribuciones de P y S coinciden, mientras que las nuevas contribuciones creadas dentro de las piezas afectadas son no negativas. En consecuencia,

$$U(f, P) - L(f, P) \le U(f, S) - L(f, S) + 2Br\,\|P\|.$$

Combinando esta desigualdad con la cota obtenida desde Q llegamos a

$$U(f, P) - L(f, P) \le U(f, Q) - L(f, Q) + 2Br\,\|P\|.$$

Esto prueba el lema. □

La estimación explica la diferencia entre **refinamiento** y **malla pequeña**. Si P refinara a Q, el término adicional sería innecesario. Cuando no la refina, la única pérdida posible está concentrada en las pocas piezas que atraviesan los cortes fijos de Q; al hacer pequeña la malla, la longitud total de esas piezas se hace pequeña.

La formulación uniforme mediante sumas etiquetadas. Ya podemos demostrar la equivalencia completa.

TEOREMA 14.2. **Teorema (equivalencia entre Darboux y sumas etiquetadas de malla pequeña).** Sean $a < b$, sea $f : [a, b] \to \mathbb{R}$ acotada y sea $I \in \mathbb{R}$. Son equivalentes:

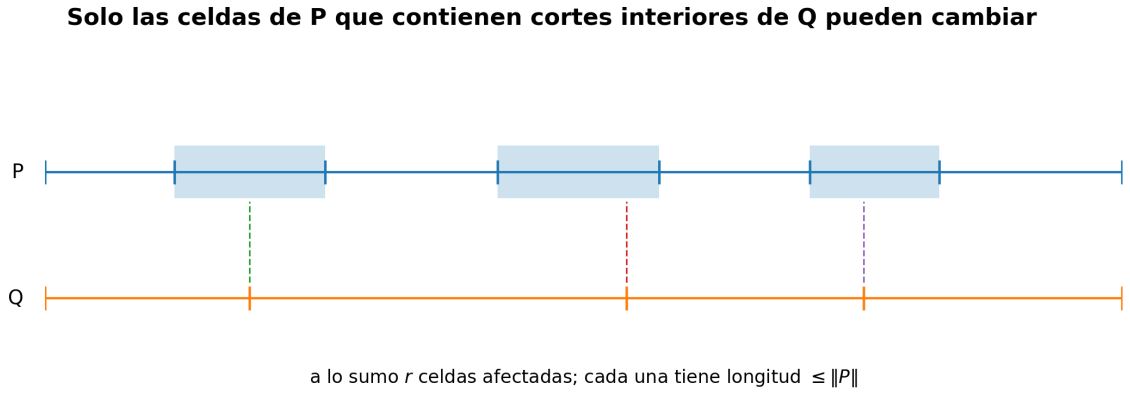


FIGURA 14.4. Solo las celdas de P atravesadas por cortes interiores de Q pueden cambiar al formar el refinamiento común.

1. f es Riemann integrable en $[a,b]$ y
$$I = \int_a^b f(x)\,dx;$$
2. para todo $\varepsilon > 0$ existe $\delta > 0$ tal que, para **toda** partición etiquetada (P,ξ) de $[a,b]$,
$$\|P\| < \delta \quad\Longrightarrow\quad \boxed{|R(f;P,\xi) - I| < \varepsilon.}$$

Así, el valor integral queda caracterizado por un control uniforme de todas las sumas etiquetadas cuando la malla es suficientemente pequeña.

Demostración. Demostraremos las dos implicaciones por separado.
(1 ⇒ 2). Supongamos que f es integrable y
$$I = \int_a^b f(x)\,dx.$$
Como f es acotada, existe $B > 0$ tal que $|f(x)| \le B$ en todo $[a,b]$. Sea $\varepsilon > 0$. Por el criterio de Darboux Teorema 14.1 existe una partición fija $Q = \{q_0, \dots, q_m\}$ tal que $U(f,Q) - L(f,Q) < \frac{\varepsilon}{2}$. Sea $r = m - 1$ el número de cortes interiores de Q y definamos $\delta := \frac{\varepsilon}{4B\,\max\{1,r\}}$. Tomemos ahora **cualquier** partición etiquetada (P,ξ) con $\|P\| < \delta$. Por Lema 14.1,
$$\begin{aligned} U(f,P) - L(f,P) &\le U(f,Q) - L(f,Q) + 2Br\,\|P\| \\ &< \frac{\varepsilon}{2} + 2Br\,\delta \\ &\le \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$
Como f es integrable, §14.2 nos da $L(f,P) \le I \le U(f,P)$, y Proposición 13.7 da $L(f,P) \le R(f;P,\xi) \le U(f,P)$. Dos números contenidos en el mismo intervalo de longitud $U(f,P) - L(f,P)$ difieren en valor absoluto a lo sumo esa longitud. Por tanto,
$$|R(f;P,\xi) - I| \le U(f,P) - L(f,P) < \varepsilon.$$
La elección de δ dependió de ε, de f y de la partición auxiliar fija Q, pero **no** dependió de la partición etiquetada (P,ξ) que apareció después. Queda demostrada la uniformidad exigida.

(2 ⇒ 1). Supongamos ahora que existe $I \in \mathbb{R}$ con la propiedad uniforme indicada. Queremos recuperar el criterio de Darboux.

Sea $\varepsilon > 0$. Aplicando la hipótesis con tolerancia $\varepsilon/4$, existe $\delta > 0$ tal que

$$\|P\| < \delta \quad \Longrightarrow \quad |R(f; P, \xi) - I| < \frac{\varepsilon}{4}$$

para toda elección de etiquetas ξ. Elegimos una partición $P = \{x_0, \ldots, x_n\}$ con $\|P\| < \delta$; por ejemplo, una partición uniforme con número de piezas suficientemente grande. Para cada subintervalo $I_k = [x_{k-1}, x_k]$ escribamos

$$m_k := \inf_{I_k} f, \qquad M_k := \sup_{I_k} f.$$

Los extremos m_k y M_k no tienen por qué alcanzarse. Sin embargo, por la definición de ínfimo y supremo, dado cualquier $\eta > 0$ podemos escoger etiquetas

$$\xi_k^- \in I_k, \qquad \xi_k^+ \in I_k$$

tales que

$$f(\xi_k^-) < m_k + \eta, \qquad f(\xi_k^+) > M_k - \eta.$$

Tomemos $\eta := \frac{\varepsilon}{4(b-a)}$.

Las correspondientes sumas etiquetadas satisfacen

$$0 \le R(f; P, \xi^-) - L(f, P) < \eta \sum_{k=1}^{n} \Delta x_k = \frac{\varepsilon}{4},$$

y $0 \le U(f, P) - R(f; P, \xi^+) < \frac{\varepsilon}{4}$.

Como $\|P\| < \delta$, la hipótesis uniforme también proporciona

$$|R(f; P, \xi^-) - I| < \frac{\varepsilon}{4}, \qquad |R(f; P, \xi^+) - I| < \frac{\varepsilon}{4}.$$

Por tanto,

$$\begin{aligned} U(f, P) - L(f, P) &= \big(U(f, P) - R(f; P, \xi^+)\big) \\ &\quad + \big(R(f; P, \xi^+) - I\big) \\ &\quad + \big(I - R(f; P, \xi^-)\big) \\ &\quad + \big(R(f; P, \xi^-) - L(f, P)\big) \\ &< \varepsilon. \end{aligned}$$

Así, para todo $\varepsilon > 0$ existe una partición con brecha menor que ε. Por Teorema 14.1, f es Riemann integrable.

Sea

$$J := \int_a^b f(x)\, dx.$$

Falta demostrar que el número I de la hipótesis es precisamente J. Ya probada la primera implicación, para toda tolerancia $\rho > 0$ existe $\delta_1 > 0$ tal que todas las sumas etiquetadas con malla menor que δ_1 satisfacen $|R(f; P, \xi) - J| < \frac{\rho}{2}$. La hipótesis original proporciona otro

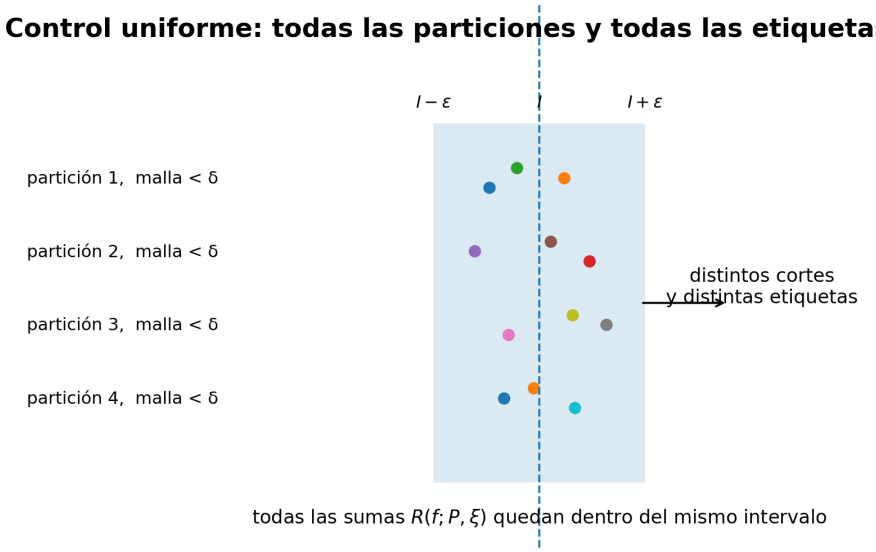


FIGURA 14.5. El criterio por sumas etiquetadas cuantifica sobre todas las particiones y todas las etiquetas de malla suficientemente pequeña.

$\delta_2 > 0$ para el cual $|R(f;P,\xi) - I| < \frac{\rho}{2}$. Elegimos una partición etiquetada con malla menor que $\min\{\delta_1, \delta_2\}$. Entonces

$$|I - J| \leq |I - R(f;P,\xi)| + |R(f;P,\xi) - J| < \rho.$$

Como esto vale para todo $\rho > 0$, necesariamente $I = J$. Queda demostrada la equivalencia. □

El teorema cuantifica sobre todas las etiquetas. La segunda condición de Teorema 14.2 puede leerse como una afirmación de independencia progresiva:

> cuando la malla es suficientemente pequeña,
>
> ni los cortes ni las etiquetas pueden cambiar mucho la suma.

No significa que todas las sumas sean iguales. Significa que, dada una tolerancia, existe una escala de malla por debajo de la cual **todas** caen dentro del mismo intervalo de radio ε alrededor de I. Esto permite distinguir tres afirmaciones:

1. existe una sucesión particular de sumas etiquetadas que converge a I;
2. para cada partición de cierta familia elegimos etiquetas especiales y obtenemos sumas cercanas a I;
3. existe $\delta > 0$ tal que **toda** partición etiquetada de malla menor que δ produce una suma cercana a I.

Solo la tercera es la condición equivalente del teorema. Las dos primeras pueden ser verdaderas sin controlar las demás elecciones.

El supremo y el ínfimo no necesitan ser etiquetas. La conversa del teorema también explica por qué C14 insistió en distinguir supremum de máximo. Para reconstruir una brecha de Darboux desde sumas etiquetadas no necesitamos que exista una etiqueta donde f alcance exactamente M_k o m_k. La propiedad del supremo y del ínfimo permite elegir valores de la función

arbitrariamente próximos a esos extremos. Como la partición tiene un número finito de piezas, podemos repartir un error total entre ellas y controlar la suma de los errores.

Por eso las desigualdades

$$f(\xi_k^+) > M_k - \eta, \qquad f(\xi_k^-) < m_k + \eta$$

son suficientes. Exigir máximos y mínimos locales habría introducido una hipótesis de continuidad que la teoría general no necesita.

Un control explícito para x^2.

EJEMPLO 14.4. **Todas las sumas etiquetadas de x^2 son cercanas cuando la malla es pequeña.** Sea

$$f(x) = x^2, \qquad 0 \le x \le 1.$$

En Ejemplo 14.3 ya demostramos que f es Riemann integrable. Sea

$$I := \int_0^1 x^2\,dx,$$

sin calcular todavía el valor de I.

Tomemos una partición arbitraria

$$P = \{0 = x_0 < x_1 < \cdots < x_n = 1\}.$$

Como x^2 es creciente en $[0,1]$, en $[x_{k-1}, x_k]$ tenemos

$$M_k - m_k = x_k^2 - x_{k-1}^2 = (x_k + x_{k-1})\Delta x_k \le 2\Delta x_k.$$

Por la identidad de la brecha de Proposición 13.8,

$$\begin{aligned} U(f,P) - L(f,P) &= \sum_{k=1}^{n} (M_k - m_k)\Delta x_k \\ &\le 2\sum_{k=1}^{n} (\Delta x_k)^2 \\ &\le 2\|P\| \sum_{k=1}^{n} \Delta x_k \\ &= 2\|P\|. \end{aligned}$$

Para cualquier elección de etiquetas ξ, tanto $R(f;P,\xi)$ como I están entre $L(f,P)$ y $U(f,P)$. Por tanto,

$$\boxed{|R(f;P,\xi) - I| \le 2\|P\|.}$$

Dado $\varepsilon > 0$, basta exigir $\|P\| < \frac{\varepsilon}{2}$ para obtener $|R(f;P,\xi) - I| < \varepsilon$

para todas las etiquetas. El cálculo no ha utilizado una partición uniforme ni una regla especial para escoger ξ_k.

El ejemplo muestra una ventaja adicional de la formulación por malla: cuando logramos una estimación explícita, podemos producir directamente un radio δ que controla todas las sumas etiquetadas.

Una familia convergente no basta
Si alguien exhibe únicamente particiones etiquetadas $(P_n, \xi^{(n)})$ tales que

$$\|P_n\| \to 0 \qquad \text{y} \qquad R(f; P_n, \xi^{(n)}) \to I,$$

ha demostrado el comportamiento de **esa familia**. Para aplicar Teorema 14.2 debe probar algo más fuerte: que, dado $\varepsilon > 0$, existe una sola escala δ que funciona simultáneamente para **todas** las particiones y **todas** las etiquetas con malla menor que δ.

La equivalencia entre Darboux y sumas etiquetadas queda así cerrada. La definición primaria de CPM sigue siendo Definición 14.2; Teorema 14.2 demuestra que la formulación clásica mediante sumas de Riemann suficientemente finas selecciona exactamente las mismas funciones y el mismo valor integral.

Antes de seguir. ¿Por qué una partición de malla pequeña no tiene por qué refinar la partición buena obtenida por Teorema 14.1? ¿Qué subintervalos pueden cambiar al insertar los cortes de esa partición fija? ¿Dónde utiliza Lema 14.1 la acotación de f? En la conversa de Teorema 14.2, ¿por qué escogemos etiquetas que solo se aproximan al supremo y al ínfimo en vez de exigir que los alcancen? ¿Qué diferencia lógica hay entre una sucesión concreta de sumas convergentes y el cuantificador uniforme sobre todas las particiones etiquetadas de malla suficientemente pequeña?

14.5. Compatibilidad con la integral elemental y cambios finitos

La integral elemental de C14 y la integral de Riemann de C15 se han construido por caminos distintos. Para una función escalonada, Definición 13.3 suma exactamente **altura por anchura** sobre una partición representante. Para una función acotada general, Definición 14.2 compara los extremos globales de todas las sumas inferiores y superiores. Que ambas construcciones utilicen el mismo símbolo no autoriza todavía a identificarlas. Debemos demostrar dos cosas:

1. toda función escalonada es Riemann integrable;
2. cuando lo es, su integral de Riemann coincide exactamente con la integral elemental ya definida en C14.

Después podremos ampliar una propiedad que en C14 solo conocíamos dentro de la clase de escalonadas: cambiar finitísimos valores de una función integrable no altera ni su integrabilidad ni el valor de su integral.

Los puntos de corte son la única posible fuente de brecha. Sea $s : [a, b] \to \mathbb{R}$ una función escalonada y sea

$$Q = \{q_0, q_1, \dots, q_m\}, \qquad a = q_0 < q_1 < \cdots < q_m = b,$$

una partición que la representa. Existen alturas $c_1, \dots, c_m$ tales que

$$s(x) = c_j \qquad \text{si } x \in (q_{j-1}, q_j).$$

Dentro de cada uno de esos intervalos abiertos la oscilación de s es cero. La única dificultad procede de los finitísimos puntos $q_0, q_1, \dots, q_m$, donde la definición de función escalonada permite valores excepcionales. Como s es acotada, fijemos $B > 0$ con

$$|s(x)| \le B \qquad (x \in [a, b]).$$

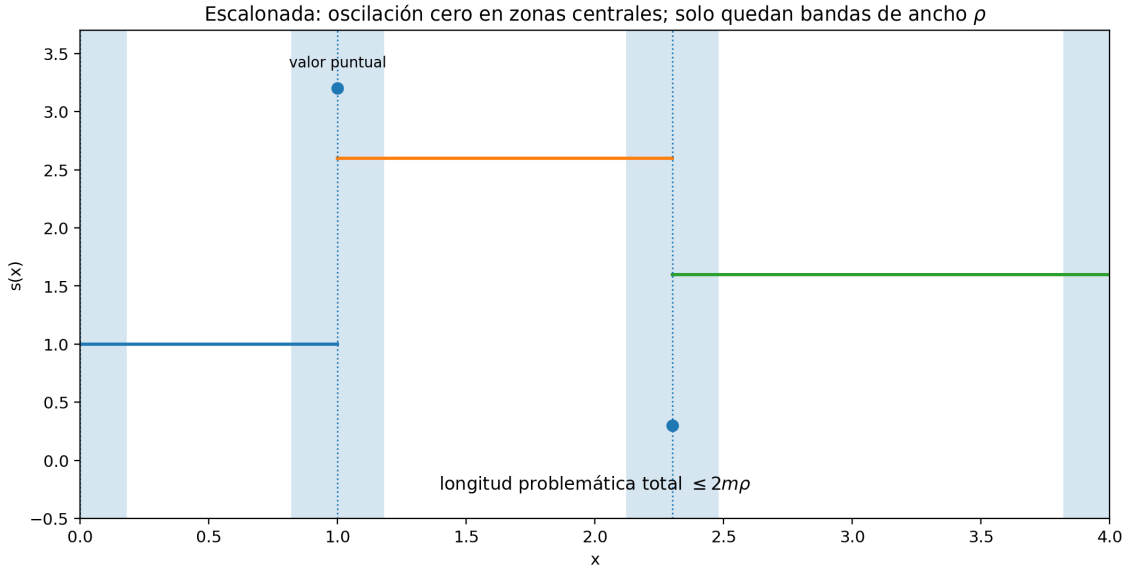


FIGURA 14.6. Para una función escalonada, la oscilación queda confinada a bandas pequeñas alrededor de los cortes; en las zonas centrales es nula.

Sea además

$$d := \min_{1\le j\le m} (q_j - q_{j-1}) > 0.$$

Dado ρ con $0 < \rho < \frac{d}{2}$, refinemos Q insertando, dentro de cada intervalo (q_{j-1}, q_j), los dos puntos

$$q_{j-1} + \rho, \qquad q_j - \rho.$$

Si $m = 1$ la construcción se interpreta del mismo modo; la condición $2\rho < d$ garantiza que los puntos quedan en el orden correcto. Denotemos por P_ρ la partición resultante. En cada intervalo central $[q_{j-1} + \rho, q_j - \rho]$ la función vale constantemente c_j, incluso en sus extremos, porque todos esos puntos pertenecen al intervalo abierto (q_{j-1}, q_j). Su contribución a la brecha es, por tanto, cero. Solo quedan las dos franjas de anchura ρ adyacentes a cada extremo de cada escalón. La suma de sus longitudes es $2m\rho$. En cualquiera de esas piezas la oscilación es a lo sumo $2B$. Por la identidad de la brecha de Proposición 13.8,

$$\boxed{U(s, P_\rho) - L(s, P_\rho) \le 4Bm\rho.}$$

Así, los valores excepcionales de los cortes pueden influir en una partición concreta, pero su influencia queda confinada a una longitud total que podemos hacer arbitrariamente pequeña.

Las dos integrales coinciden.

TEOREMA 14.3. **Teorema (compatibilidad entre la integral elemental y la integral de Riemann).** Sean $a < b$ y sea $s : [a, b] \to \mathbb{R}$ una función escalonada. Entonces

$$\boxed{s \in \mathcal{R}[a, b].}$$

Además, la integral de Riemann de s coincide con la integral elemental definida en Definición 13.3:

$$\boxed{\int_a^b s(x)\,dx\Big|_{\text{Riemann}} = \int_a^b s(x)\,dx\Big|_{\text{elemental}}.}$$

Por tanto, desde este punto podremos usar el símbolo $\int_a^b s(x)\,dx$ sin distinguir ambas construcciones cuando s sea escalonada.

Demostración. Sea $Q = \{q_0, \dots, q_m\}$ una partición representante de s y elijamos $B > 0$ y $d > 0$ como arriba.

Primero probemos la integrabilidad. Sea $\varepsilon > 0$. Elegimos

$$0 < \rho < \min\left\{\frac{d}{2}, \frac{\varepsilon}{4Bm}\right\}.$$

Para la partición P_ρ recién construida, $U(s, P_\rho) - L(s, P_\rho) \le 4Bm\rho < \varepsilon$. Por el criterio de Darboux Teorema 14.1, $s \in \mathcal{R}[a, b]$.

Escribamos ahora

$$I_R := \int_a^b s(x)\,dx$$

para la integral de Riemann y

$$I_E := \int_a^b s(x)\,dx\Big|_{\text{elemental}}$$

para el número definido en C14. Fijemos una partición arbitraria P de $[a, b]$. En §13.6 se construyeron las funciones escalonadas auxiliares $\ell_{s,P}$ y $u_{s,P}$, que satisfacen $\ell_{s,P} \le s \le u_{s,P}$ punto a punto y cuyas integrales elementales son

$$\int_a^b \ell_{s,P} = L(s, P), \qquad \int_a^b u_{s,P} = U(s, P).$$

Como las tres funciones son escalonadas, la monotonía de la integral elemental demostrada en Proposición 13.3 da $L(s, P) \le I_E \le U(s, P)$. Esto vale para **toda** partición P. Por consiguiente, I_E es una cota superior de todas las sumas inferiores y una cota inferior de todas las sumas superiores. De Definición 14.1 obtenemos

$$\underline{\int_a^b} s \le I_E \le \overline{\int_a^b} s.$$

Pero ya demostramos que s es integrable, de modo que ambos extremos globales coinciden con I_R. Por tanto, $I_R \le I_E \le I_R$, y necesariamente $I_E = I_R$.

Queda demostrada la compatibilidad. $\square$

El teorema cierra una deuda lógica de C14. Allí el adjetivo «elemental» delimitaba una operación definida solo para escalonadas. Ahora sabemos que esa operación no compite con la integral general: es exactamente su restricción a la clase de funciones escalonadas. Por ejemplo, la función de Ejemplo 13.5 tenía integral elemental 7. Desde ahora podemos afirmar, sin introducir un segundo valor,

$$\int_0^4 s(x)\,dx = 7$$

en el sentido general de Riemann.

Cambiar finitísimos valores: ahora para cualquier función integrable. En Corolario 13.1 probamos que dos funciones **escalonadas** que difieren solo en un conjunto finito tienen la misma integral elemental. La equivalencia por sumas etiquetadas de §14.4 permite eliminar ahora la hipótesis de que ambas sean escalonadas.

COROLARIO 14.1. **Corolario (invariancia de la integral de Riemann ante modificaciones finitas).** Sean $f, g : [a, b] \to \mathbb{R}$ funciones que coinciden salvo, a lo sumo, en un conjunto finito

$$F \subseteq [a, b].$$

Entonces

$$\boxed{f \in \mathcal{R}[a, b] \iff g \in \mathcal{R}[a, b].}$$

Cuando estas condiciones se cumplen,

$$\boxed{\int_a^b f(x)\,dx = \int_a^b g(x)\,dx.}$$

Demostración. Por simetría basta demostrar que la integrabilidad de f implica la de g con el mismo valor integral.

Supongamos $f \in \mathcal{R}[a, b]$ y escribamos

$$I := \int_a^b f(x)\,dx.$$

Como f es acotada y g solo puede diferir de ella en finitísimos puntos, g también es acotada.

Escribamos $F = \{c_1, \ldots, c_r\}$ sin repeticiones y pongamos

$$D := \sum_{j=1}^{r} |g(c_j) - f(c_j)|.$$

Si $F = \varnothing$ o $D = 0$, entonces $f = g$ y no hay nada que demostrar. Supongamos $D > 0$. Sea $\varepsilon > 0$. Por Teorema 14.2 aplicado a f, existe $\delta_1 > 0$ tal que toda partición etiquetada (P, ξ) con $\|P\| < \delta_1$ satisface $|R(f; P, \xi) - I| < \frac{\varepsilon}{2}$.

Definamos $\delta := \min\left\{\delta_1, \frac{\varepsilon}{4D}\right\}$.

Tomemos una partición etiquetada arbitraria (P, ξ) con $\|P\| < \delta$. Las dos sumas solo difieren en aquellos subintervalos cuya etiqueta pertenece a F. Para un punto fijo c_j, la etiqueta c_j puede aparecer en **a lo sumo dos** subintervalos de una partición: los dos adyacentes si c_j es un punto de corte, o uno solo si queda en el interior de una pieza. Por tanto, la suma de las anchuras de los términos etiquetados por c_j es a lo sumo $2\|P\|$.

Aplicando la desigualdad triangular,

$$\begin{aligned}|R(g;P,\xi)-R(f;P,\xi)| &\le 2\|P\|\sum_{j=1}^{r}|g(c_j)-f(c_j)|\\ &= 2D\|P\|\\ &< 2D\delta\\ &\le \frac{\varepsilon}{2}.\end{aligned}$$

En consecuencia,

$$\begin{aligned}|R(g;P,\xi)-I| &\le |R(g;P,\xi)-R(f;P,\xi)|+|R(f;P,\xi)-I|\\ &< \frac{\varepsilon}{2}+\frac{\varepsilon}{2}\\ &= \varepsilon.\end{aligned}$$

La misma δ funciona para **todas** las particiones y todas las etiquetas con malla menor que δ. Por Teorema 14.2, g es Riemann integrable y

$$\int_a^b g(x)\,dx = I = \int_a^b f(x)\,dx.$$

Intercambiando los papeles de f y g obtenemos la equivalencia anunciada. □

La finitud de F entra en un lugar preciso: permite reducir toda la perturbación a una suma finita de contribuciones, cada una multiplicada por una longitud controlada por la malla. El corolario **no** autoriza a reemplazar «finito» por «infinito pero pequeño» sin una teoría adicional.

Un valor enorme en unos pocos puntos no cambia la integral.

Ejemplo 14.5. **Dos alteraciones puntuales de una función conocida.** Definamos $g:[0,1]\to\mathbb{R}$ por

$$g(x)=\begin{cases}100, & x=\dfrac{1}{3},\\ -7, & x=1,\\ x, & \text{en otro caso.}\end{cases}$$

La función g difiere de la identidad $f(x)=x$ solamente en el conjunto finito

$$F=\left\{\frac{1}{3},1\right\}.$$

En Ejemplo 14.2 demostramos que

$$f\in\mathcal{R}[0,1] \qquad \text{y} \qquad \int_0^1 x\,dx=\frac{1}{2}.$$

Por Corolario 14.1,

$$\boxed{g\in\mathcal{R}[0,1]}$$

y

$$\boxed{\int_0^1 g(x)\,dx = \frac{1}{2}.}$$

Los valores 100 y -7 pueden afectar sumas particulares si una etiqueta cae exactamente en esos puntos, pero su efecto total queda forzado a cero cuando la malla tiende uniformemente a cero.

Lo que este resultado no dice. La invariancia finita es fuerte, pero tiene una frontera exacta.

- No afirma que toda modificación sobre un conjunto infinito preserve la integral.
- No afirma que dos funciones que coinciden «casi en todas partes» tengan la misma integral: esa expresión todavía no tiene aquí un significado medido.
- No utiliza el criterio de Lebesgue ni ninguna teoría de medida.
- No depende del teorema fundamental del cálculo.

Lo que sí queda demostrado es enteramente riemanniano:

una cantidad finita de valores puntuales no puede aportar anchura total positiva.

Con esto quedan reconciliadas las dos capas construidas desde C14: las escalonadas conservan exactamente su integral elemental y la teoría general ignora modificaciones finitas del mismo modo en que aquella construcción finita ya sugería. La siguiente pregunta será mucho más importante para producir nuevas clases de funciones integrables: ¿cómo transformar la **continuidad uniforme** en una partición con brecha de Darboux arbitrariamente pequeña? Esa será la tarea exclusiva de §14.6.

Antes de seguir. ¿Por qué los valores de una escalonada en sus puntos de corte pueden producir una brecha positiva en una partición fija y, sin embargo, no impedir la integrabilidad? ¿Cómo se usa la monotonía de la integral elemental para encerrar su valor entre $L(s,P)$ y $U(s,P)$ para toda partición? En Corolario 14.1, ¿por qué un punto excepcional puede aparecer como etiqueta en a lo sumo dos subintervalos? ¿Dónde se usa la finitud del conjunto excepcional? ¿Qué afirmación adicional sería necesaria antes de extender el corolario a conjuntos infinitos de puntos?

14.6. Las funciones continuas son integrables

Hasta ahora hemos demostrado integrabilidad para funciones escalonadas y para algunas funciones concretas cuyas brechas podían calcularse explícitamente. La continuidad permite dar un salto mucho mayor: **toda función continua sobre un intervalo cerrado es Riemann integrable**. La idea no será aproximar la función mediante primitivas ni utilizar derivadas. Volveremos directamente al criterio de Darboux.

Para que

$$U(f,P) - L(f,P) = \sum_{k=1}^{n} \omega_k(f,P)\,\Delta x_k$$

sea pequeño, basta controlar uniformemente las oscilaciones locales $\omega_k(f,P)$. La continuidad ordinaria proporciona controles alrededor de cada punto. Lo que necesitamos para una partición entera es algo más fuerte: **un único tamaño de intervalo que funcione en todas partes**. Ese es exactamente el contenido de Heine–Cantor.

Dos consecuencias distintas de la compactitud. Sea $f : [a, b] \to \mathbb{R}$ continua, con $a < b$.

Antes de hablar de sumas de Darboux necesitamos saber que f es acotada. Esto ya se sigue de Teorema 5.5: una función continua sobre $[a, b]$ alcanza un mínimo y un máximo global. En particular, existen números reales m, M tales que

$$m \leq f(x) \leq M \qquad (x \in [a, b]).$$

Por tanto, las sumas inferiores y superiores de C14 están bien definidas. Pero la acotación no basta para demostrar integrabilidad. La segunda consecuencia de la compactitud es la decisiva para controlar la brecha: por Teorema 5.6, f es **uniformemente continua** en $[a, b]$. Esto significa que, dada una tolerancia vertical $\eta > 0$, existe un único número $\delta > 0$ tal que

$$x, y \in [a, b], \qquad |x - y| < \delta \quad \Longrightarrow \quad |f(x) - f(y)| < \eta.$$

La palabra «único» aquí no significa que δ sea numéricamente único. Significa que podemos escoger **una misma** δ que funcione simultáneamente para todos los pares de puntos del intervalo. Esa uniformidad es exactamente la que una partición podrá explotar.

Teorema 14.4. **Teorema (integrabilidad de las funciones continuas).** Sean $a, b \in \mathbb{R}$ con $a < b$ y sea

$$f : [a, b] \to \mathbb{R}$$

continua. Entonces

$$\boxed{f \in \mathcal{R}[a, b].}$$

Demostración. Como acabamos de observar, Teorema 5.5 garantiza que f es acotada, así que podemos aplicar el criterio de Darboux.

Sea $\varepsilon > 0$. Definamos $\eta := \frac{\varepsilon}{b-a} > 0$.

Por Heine–Cantor, Teorema 5.6, existe $\delta > 0$ tal que para cualesquiera $x, y \in [a, b]$,

$$|x - y| < \delta \quad \Longrightarrow \quad |f(x) - f(y)| < \eta.$$

Elegimos una partición $P = \{x_0, \ldots, x_n\}$ de $[a, b]$ cuya malla satisfaga $\|P\| < \delta$. Por ejemplo, basta tomar una partición uniforme con un número suficientemente grande de subintervalos.

Fijemos una pieza $I_k = [x_{k-1}, x_k]$.

La restricción de f a I_k es continua. Como I_k es un intervalo cerrado, Teorema 5.5 aplicado a esa restricción garantiza puntos $u_k, v_k \in I_k$ donde se alcanzan, respectivamente, el máximo y el mínimo locales:

$$M_k = f(u_k), \qquad m_k = f(v_k).$$

Además,

$$|u_k - v_k| \leq \Delta x_k \leq \|P\| < \delta.$$

La continuidad uniforme da entonces $M_k - m_k = |f(u_k) - f(v_k)| < \eta$. Por tanto, todas las oscilaciones locales satisfacen $\omega_k(f, P) < \eta$. Usando la identidad de la brecha de Proposición 13.8,

$$\Delta x_k < \delta \implies \omega_k < \eta$$

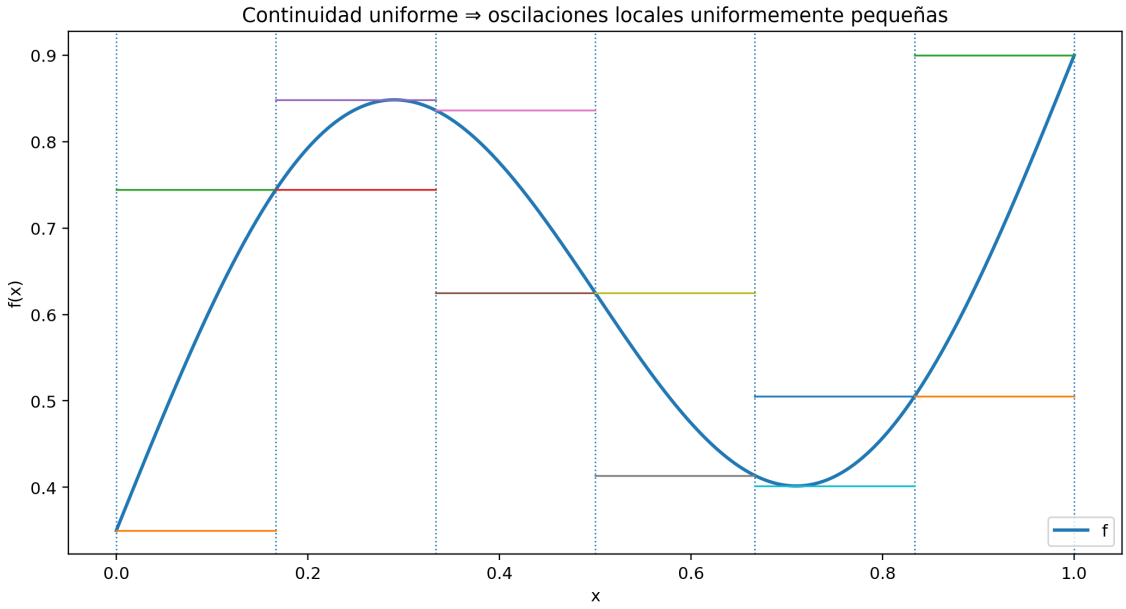


FIGURA 14.7. La continuidad uniforme transforma piezas horizontalmente cortas en oscilaciones verticales uniformemente pequeñas.

$$\begin{aligned} U(f,P) - L(f,P) &= \sum_{k=1}^{n} \omega_k(f,P)\,\Delta x_k \\ &< \eta \sum_{k=1}^{n} \Delta x_k \\ &= \eta(b-a) \\ &= \varepsilon. \end{aligned}$$

Hemos probado que para todo $\varepsilon > 0$ existe una partición P con $U(f,P) - L(f,P) < \varepsilon$. Por el criterio de Darboux Teorema 14.1, $f \in \mathcal{R}[a,b]$. □

Dónde trabaja realmente la continuidad uniforme. La demostración contiene una traducción exacta entre dos lenguajes.

Heine–Cantor dice:

$$\boxed{\text{intervalos horizontalmente cortos} \implies \text{variación vertical pequeña}}$$

con un mismo umbral horizontal válido en todo $[a,b]$. La identidad de la brecha convierte después esa información local en una estimación global:

$$\boxed{\text{oscilación pequeña en cada pieza} \implies \text{brecha total pequeña.}}$$

La suma de las anchuras no crece al aumentar el número de piezas:

$$\sum_{k=1}^{n} \Delta x_k = b - a.$$

Por eso el error local uniforme η se multiplica por una longitud total fija.

Por qué la continuidad punto a punto no basta directamente. Podríamos intentar comenzar solamente con la definición ordinaria de continuidad. Para cada punto $x \in [a, b]$ y cada $\eta > 0$ obtendríamos algún radio $\delta_x > 0$. Pero una partición contiene muchas piezas distribuidas por todo el intervalo. Para imponer mediante la malla que **cada** pieza sea suficientemente corta necesitamos una sola escala que controle todas las zonas a la vez. No está justificado escribir simplemente $\delta = \inf_{x\in[a,b]} \delta_x$ y suponer que ese número es positivo: una familia infinita de radios positivos puede tener ínfimo cero. Precisamente este obstáculo fue resuelto en C06 por la compactitud de $[a, b]$, que conduce a Teorema 5.6.

El papel exacto de la compactitud
En esta prueba la compactitud actúa dos veces, a través de resultados ya demostrados:

1. Teorema 5.5 proporciona acotación global, necesaria para que las sumas de Darboux estén definidas;
2. Teorema 5.6 proporciona continuidad uniforme, necesaria para controlar todas las oscilaciones locales con una sola escala.

La integrabilidad no se obtiene de «continuidad + dibujo», sino de esta cadena de resultados.

Un certificado cuantitativo sin usar derivadas. El teorema anterior es cualitativo: garantiza que existe una partición suficientemente fina. Algunas funciones permiten además una cota explícita de la brecha.

Ejemplo 14.6. **La función valor absoluto: integrable sin ser diferenciable en el origen.** Sea

$$f(x) = |x|, \qquad -1 \le x \le 1.$$

La desigualdad triangular inversa da, para cualesquiera $x, y \in [-1, 1]$, $||x| - |y|| \le |x - y|$.

Tomemos una partición arbitraria

$$P = \{-1 = x_0 < x_1 < \cdots < x_n = 1\}.$$

En cada pieza $I_k = [x_{k-1}, x_k]$, para cualesquiera $x, y \in I_k$, $||x| - |y|| \le |x - y| \le \Delta x_k$. Por consiguiente, la oscilación local satisface $\omega_k(f, P) \le \Delta x_k$. Aplicando Proposición 13.8,

$$\begin{aligned} U(f, P) - L(f, P) &\le \sum_{k=1}^{n} (\Delta x_k)^2 \\ &\le \|P\| \sum_{k=1}^{n} \Delta x_k \\ &= 2\|P\|. \end{aligned}$$

Dado $\varepsilon > 0$, cualquier partición con $\|P\| < \frac{\varepsilon}{2}$ satisface $U(f, P) - L(f, P) < \varepsilon$. Así, Teorema 14.1 vuelve a dar

$$\boxed{|x| \in \mathcal{R}[-1,1].}$$

La demostración no utiliza la derivada. En particular, la falta de diferenciabilidad de $|x|$ en 0 no constituye ninguna obstrucción para su integrabilidad.

El ejemplo muestra que una estimación del tipo $|f(x)-f(y)| \le C|x-y|$ produce directamente un control cuantitativo sobre la brecha. Sin embargo, el teorema general no necesita disponer de una constante C: la continuidad uniforme basta.

Qué hemos demostrado y qué no. El resultado establece una inclusión fundamental:

$$\boxed{C([a,b]) \subseteq \mathcal{R}[a,b],}$$

entendiendo $C([a,b])$ como la clase de funciones reales continuas sobre el intervalo. No hemos demostrado la conversa. Existen funciones Riemann integrables que no son continuas: las funciones escalonadas con saltos de §14.5 ya proporcionan ejemplos, y Corolario 14.1 permite incluso introducir finitísimas discontinuidades puntuales sin alterar la integral. Tampoco hemos calculado las integrales de todas las funciones continuas. El teorema prueba **existencia**, no una fórmula de evaluación.

Finalmente, no hemos utilizado el teorema de valor medio integral ni el teorema fundamental del cálculo. Ambos siguen fuera de este capítulo en las fronteras ya fijadas. Para una función monótona construiremos la cota de la brecha a partir de la variación entre sus valores extremos. Esa estimación dará integrabilidad incluso cuando haya discontinuidades, sin exigir continuidad uniforme.

Antes de seguir. ¿Qué dos consecuencias de la compactitud se usan en Teorema 14.4 y para qué sirve cada una? ¿Por qué una colección de radios δ_x dependientes del punto no proporciona automáticamente una malla global positiva? ¿Cómo convierte Proposición 13.8 el control uniforme de las oscilaciones en una brecha global? En Ejemplo 14.6, ¿por qué la no diferenciabilidad en 0 no interviene en absoluto? ¿Qué diferencia hay entre demostrar que una función continua es integrable y calcular el valor de su integral?

14.7. Las funciones monótonas son integrables

La sección anterior obtuvo integrabilidad a partir de una propiedad local que, gracias a la compactitud, se volvía uniforme: la continuidad impedía oscilaciones grandes en intervalos suficientemente cortos. La monotonía ofrece un mecanismo completamente diferente. Una función monótona puede tener saltos y no necesita ser continua. Lo que sí posee es un control global muy rígido: al recorrer el intervalo en una sola dirección, toda la variación local se acumula sin cancelaciones y queda dominada por la diferencia entre los valores extremos. Ese hecho permite estimar directamente la brecha de Darboux.

La variación total aparece por telescopado. Supongamos primero que $f : [a,b] \to \mathbb{R}$ es no decreciente. Entonces

$$f(a) \le f(x) \le f(b) \qquad (x \in [a,b]),$$

de modo que f es acotada y las sumas de Darboux están bien definidas. Para una partición arbitraria

$$P = \{a = x_0 < x_1 < \cdots < x_n = b\},$$

en cada subintervalo $I_k = [x_{k-1}, x_k]$ la monotonía da

$$\inf_{I_k} f = f(x_{k-1}), \qquad \sup_{I_k} f = f(x_k).$$

Por tanto, la oscilación local es exactamente $\omega_k(f, P) = f(x_k) - f(x_{k-1})$. La identidad de la brecha de Proposición 13.8 produce entonces

$$U(f, P) - L(f, P) = \sum_{k=1}^{n} (f(x_k) - f(x_{k-1}))\Delta x_k.$$

Como cada anchura satisface $\Delta x_k \leq \|P\|$, tenemos

$$\begin{aligned} U(f, P) - L(f, P) &\leq \|P\| \sum_{k=1}^{n} (f(x_k) - f(x_{k-1})) \\ &= \|P\|(f(b) - f(a)). \end{aligned}$$

La última igualdad es telescópica. Para una función no creciente ocurre lo mismo con los signos invertidos. En este caso,

$$f(b) \leq f(x) \leq f(a) \qquad (x \in [a, b]),$$

así que f también es acotada. Además, $\omega_k(f, P) = f(x_{k-1}) - f(x_k)$, y por tanto

$$U(f, P) - L(f, P) \leq \|P\|(f(a) - f(b)).$$

Ambos casos quedan condensados en un único resultado.

TEOREMA 14.5. **Teorema (integrabilidad de las funciones monótonas).** Sean $a < b$ y sea

$$f : [a, b] \to \mathbb{R}$$

monótona, es decir, no decreciente o no creciente. Entonces f es Riemann integrable en $[a, b]$. Más precisamente, para toda partición P de $[a, b]$,

$$\boxed{U(f, P) - L(f, P) \leq \|P\|\, |f(b) - f(a)|.}$$

Demostración. La desigualdad ya quedó demostrada arriba en los dos sentidos posibles de monotonía. Falta convertirla en el criterio de Darboux.

Escribamos $V := |f(b) - f(a)|$.

Hay dos casos.

Caso 1: $V = 0$**.** Si f es no decreciente y $f(a) = f(b)$, entonces para todo $x \in [a, b]$,

$$f(a) \leq f(x) \leq f(b) = f(a),$$

por lo que f es constante. Lo mismo ocurre si f es no creciente. Así, $U(f, P) - L(f, P) = 0$ para toda partición P, y f es integrable por Teorema 14.1.

Caso 2: $V > 0$**.** Sea $\varepsilon > 0$. Elegimos una partición uniforme P_n con $\|P_n\| = \frac{b-a}{n} < \frac{\varepsilon}{V}$, lo cual es posible por la propiedad arquimediana. Entonces la cota cuantitativa anterior da $U(f, P_n) - L(f, P_n) \leq \|P_n\|V < \varepsilon$.

Por Teorema 14.1, f es Riemann integrable. □

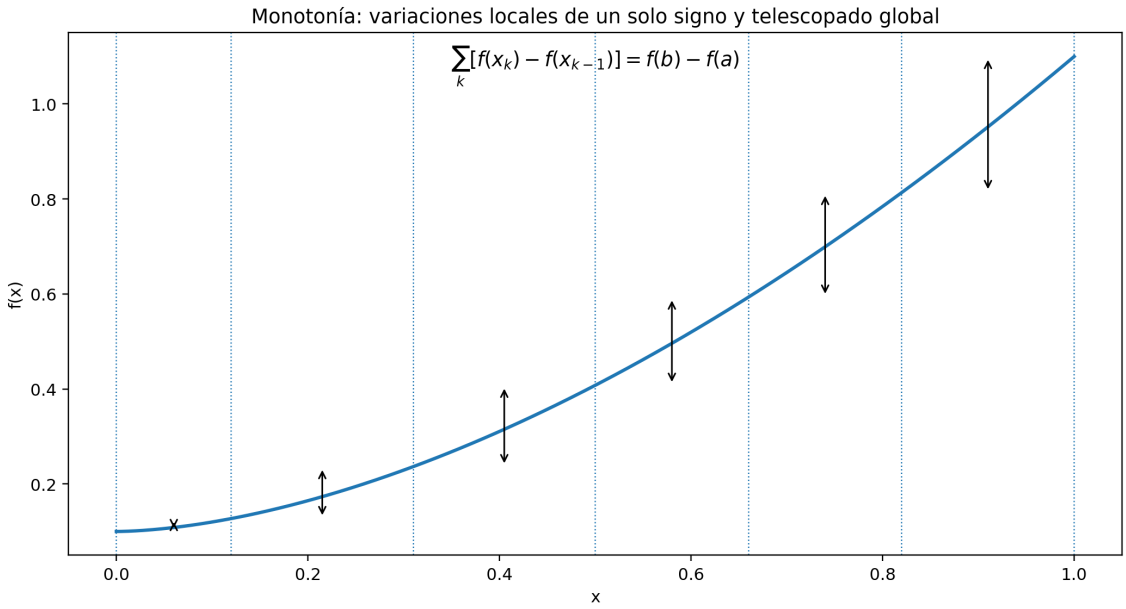


FIGURA 14.8. En una función monótona, las variaciones locales tienen un solo signo y su suma telescopa hasta la variación entre los extremos.

La prueba no necesita continuidad. La diferencia con Teorema 14.4 es estructural. Para una función continua, el argumento fue

$$\text{continuidad uniforme} \longrightarrow \text{oscilaciones locales pequeñas} \longrightarrow U - L < \varepsilon.$$

Para una función monótona, en cambio,

$$\boxed{\text{monotonía} \longrightarrow \text{variaciones locales con un solo signo} \longrightarrow \text{telescopado global.}}$$

No necesitamos demostrar que cada oscilación sea pequeña por continuidad. Basta multiplicar cada variación local por una anchura menor o igual que la malla y sumar. La suma de todas las variaciones locales es exactamente la variación entre los extremos. Esta es también la razón por la cual la prueba tolera saltos.

El caso constante no debe esconderse en una división. La estimación $U(f,P) - L(f,P) \le \|P\|V$ podría tentar a escoger directamente $\|P\| < \frac{\varepsilon}{V}$. Pero esa expresión solo tiene sentido cuando $V > 0$. Si $V = 0$, no hay que dividir por cero: la monotonía fuerza que la función sea constante y la brecha ya es exactamente cero. Este pequeño caso límite es matemáticamente simple, pero lógicamente necesario.

Una función monótona con salto.

EJEMPLO 14.7. **Integrabilidad sin continuidad: un salto.** Definamos $s : [0,1] \to \mathbb{R}$ por

$$s(x) = \begin{cases} 0, & 0 \le x < \frac{1}{2}, \\ 1, & \frac{1}{2} \le x \le 1. \end{cases}$$

La función es no decreciente y satisface $s(1) - s(0) = 1$. Por Teorema 14.5, para toda partición P de $[0,1]$,

$$\boxed{U(s,P) - L(s,P) \leq \|P\|.}$$

Así, si $\|P\| < \varepsilon$, entonces $U(s,P) - L(s,P) < \varepsilon$, y Teorema 14.1 certifica nuevamente que $s \in \mathcal{R}[0,1]$. Sin embargo, s no es continua en $1/2$: sus valores a la izquierda permanecen en 0, mientras que

$$s\left(\frac{1}{2}\right) = 1.$$

La integrabilidad no requiere continuidad. En este ejemplo también podríamos haber invocado §14.5 porque s es escalonada. La utilidad de la prueba actual es distinta: exhibe una estimación que depende únicamente de la monotonía y de la malla, sin utilizar la estructura escalonada particular de s.

Una partición arbitraria ya posee una cota útil. El teorema es ligeramente más fuerte que una mera prueba de existencia. No dice solamente que, dado ε, podemos fabricar una partición conveniente. Afirma que **toda** partición satisface

$$U(f,P) - L(f,P) \leq \|P\|\,|f(b) - f(a)|.$$

Por tanto, si $V = |f(b) - f(a)| > 0$, cualquier partición con $\|P\| < \frac{\varepsilon}{V}$ es automáticamente una partición de Darboux buena para la tolerancia ε. Aquí la malla sí basta porque está acompañada por una propiedad estructural concreta de la función: la monotonía transforma la suma de oscilaciones en una cantidad telescópica global. Para una función acotada arbitraria no existe tal estimación.

Continuidad y monotonía son criterios diferentes. Tenemos ahora dos inclusiones importantes: $C([a,b]) \subseteq \mathcal{R}[a,b]$ por Teorema 14.4, y

$$\boxed{\{f : [a,b] \to \mathbb{R} :\ f \text{ monótona}\} \subseteq \mathcal{R}[a,b]}$$

por Teorema 14.5. Ninguna de estas dos clases contiene a la otra en general.

- Una función monótona puede tener saltos, como Ejemplo 14.7.
- Una función continua puede no ser monótona; por ejemplo, $|x|$ en $[-1,1]$, estudiada en Ejemplo 14.6, decrece en una parte del intervalo y crece en otra.

Por tanto, continuidad y monotonía son **dos mecanismos suficientes distintos** para obtener integrabilidad.

Qué hemos demostrado y qué queda abierto. La teoría dispone ya de varias clases integrables construidas por rutas independientes:

$$\boxed{\text{escalonadas} \quad \cup \quad \text{continuas} \quad \cup \quad \text{monótonas} \subseteq \mathcal{R}[a,b].}$$

Además, §14.5 mostró que las modificaciones finitas de una función integrable preservan integrabilidad y valor. Todavía no hemos demostrado que la clase $\mathcal{R}[a,b]$ sea estable bajo suma, multiplicación por escalares, valor absoluto o producto. Esas propiedades requieren argumentos propios y constituyen exactamente el objeto de §14.8. Tampoco hemos introducido teoremas de valor medio para integrales ni primitivas. La integrabilidad de las funciones monótonas se ha obtenido exclusivamente a partir de sumas de Darboux, malla y telescopado.

Antes de seguir. Para una función no decreciente, ¿por qué los extremos de cada subintervalo determinan exactamente su ínfimo y su supremo? ¿Dónde aparece el telescopado en la prueba de Teorema 14.5? ¿Por qué el caso $f(a) = f(b)$ debe tratarse antes de dividir por $|f(b) - f(a)|$? ¿Qué propiedad adicional hace que una malla pequeña sí controle la brecha de una función monótona, aunque no la controle para una función acotada arbitraria? ¿Qué muestran conjuntamente Ejemplo 14.6 y Ejemplo 14.7 sobre la relación entre continuidad, monotonía e integrabilidad?

14.8. Álgebra y orden de las funciones integrables

Hasta ahora hemos obtenido integrabilidad mediante propiedades especiales de una función: ser escalonada, ser continua, ser monótona o diferir solo en finitísimos puntos de otra función ya integrable. Falta demostrar que la propia clase $\mathcal{R}[a, b]$ posee estructura interna. La pregunta es doble. Si construimos una función nueva a partir de funciones integrables, ¿seguirá siendo integrable? Y, cuando las funciones están ordenadas punto a punto, ¿ese orden se conserva después de integrar?

La primera respuesta será lineal. La segunda mostrará que la integral respeta el orden. Después usaremos ambas para estudiar el valor absoluto, las partes positiva y negativa y el producto.

La integral es lineal. La caracterización uniforme de §14.4 permite demostrar la linealidad sin manipular por separado supremos e ínfimos cuando aparecen escalares negativos.

Teorema 14.6. **Teorema (linealidad y monotonía de la integral de Riemann).** Sean $f, g \in \mathcal{R}[a, b]$ y sean $\alpha, \beta \in \mathbb{R}$. Entonces

$$\alpha f + \beta g \in \mathcal{R}[a, b]$$

y

$$\boxed{\int_a^b (\alpha f(x) + \beta g(x))\, dx = \alpha \int_a^b f(x)\, dx + \beta \int_a^b g(x)\, dx.}$$

Además, si

$$f(x) \leq g(x) \qquad (x \in [a, b]),$$

entonces

$$\boxed{\int_a^b f(x)\, dx \leq \int_a^b g(x)\, dx.}$$

Demostración. Escribamos

$$I := \int_a^b f(x)\, dx, \qquad J := \int_a^b g(x)\, dx.$$

Sea $\varepsilon > 0$. Por Teorema 14.2 aplicado a f, existe $\delta_f > 0$ tal que toda partición etiquetada (P, ξ) con $\|P\| < \delta_f$ satisface $|R(f; P, \xi) - I| < \frac{\varepsilon}{2(1+|\alpha|)}$. Análogamente, existe $\delta_g > 0$ tal que $|R(g; P, \xi) - J| < \frac{\varepsilon}{2(1+|\beta|)}$ si $\|P\| < \delta_g$.

Tomemos $\delta := \min\{\delta_f, \delta_g\}$.

Para cualquier partición etiquetada (P, ξ) con $\|P\| < \delta$, la definición de suma etiquetada da exactamente

$$R(\alpha f + \beta g; P, \xi) = \alpha R(f; P, \xi) + \beta R(g; P, \xi).$$

Por tanto,

$$\begin{aligned}
&|R(\alpha f + \beta g; P, \xi) - (\alpha I + \beta J)| \\
&\quad \le |\alpha|\,|R(f; P, \xi) - I| + |\beta|\,|R(g; P, \xi) - J| \\
&\quad < \frac{\varepsilon}{2}\frac{|\alpha|}{1+|\alpha|} + \frac{\varepsilon}{2}\frac{|\beta|}{1+|\beta|} \\
&\quad < \varepsilon.
\end{aligned}$$

Como f y g son acotadas, también lo es $\alpha f + \beta g$. El número $\alpha I + \beta J$ satisface así la condición uniforme de Teorema 14.2. Luego $\alpha f + \beta g$ es Riemann integrable y su integral es ese número. Probemos ahora la monotonía. Supongamos $f \le g$ y definamos $h := g - f$. Por la linealidad recién demostrada, $h \in \mathcal{R}[a, b]$ y

$$\int_a^b h = \int_a^b g - \int_a^b f.$$

Como $h(x) \ge 0$ en todo el intervalo, para cualquier partición P se tiene $L(h, P) \ge 0$.

El encierro de §14.2 implica

$$0 \le L(h, P) \le \int_a^b h(x)\,dx.$$

Por consiguiente,

$$0 \le \int_a^b g(x)\,dx - \int_a^b f(x)\,dx,$$

que es exactamente la desigualdad anunciada. □

La primera parte afirma que $\mathcal{R}[a, b]$ es estable bajo combinaciones lineales finitas. La segunda convierte desigualdades punto a punto en desigualdades entre números.

Por ejemplo, si

$$m \le f(x) \le M \qquad (x \in [a, b]),$$

entonces, comparando con las funciones constantes m y M,

$$\boxed{m(b-a) \le \int_a^b f(x)\,dx \le M(b-a).}$$

En particular, si $f \ge 0$, entonces

$$\int_a^b f(x)\,dx \ge 0.$$

La conversa es falsa: que una integral sea no negativa no obliga a que la función sea no negativa punto a punto.

El valor absoluto no aumenta la oscilación. Para estudiar $|f|$ volvemos por un momento a las brechas de Darboux. Sean $I \subseteq [a,b]$ un subintervalo y $x, y \in I$. La desigualdad triangular inversa da $\big||f(x)| - |f(y)|\big| \leq |f(x) - f(y)|$.

Si $\omega(f; I) := \sup_I f - \inf_I f$, entonces $|f(x) - f(y)| \leq \omega(f; I)$.

Por tanto, todos los valores de $|f|$ en I difieren entre sí a lo sumo en $\omega(f; I)$, y de aquí

$$\boxed{\omega(|f|; I) \leq \omega(f; I).}$$

Sumando sobre una partición P y usando Proposición 13.8 obtenemos

$$\boxed{U(|f|, P) - L(|f|, P) \leq U(f, P) - L(f, P).}$$

Así, cualquier partición de Darboux buena para f es también buena para $|f|$.

Valor absoluto, partes positiva y negativa, y productos.

Proposición 14.2. **Proposición (estabilidad por valor absoluto, partes positiva/negativa y producto).** Sean $f, g \in \mathcal{R}[a,b]$. Entonces:

1. $|f| \in \mathcal{R}[a,b]$;
2. las funciones
$$f^+(x) := \max\{f(x), 0\} = \frac{|f(x)| + f(x)}{2},$$
y
$$f^-(x) := \max\{-f(x), 0\} = \frac{|f(x)| - f(x)}{2}$$
pertenecen a $\mathcal{R}[a,b]$;
3. $fg \in \mathcal{R}[a,b]$.

Además,

$$f = f^+ - f^-, \qquad |f| = f^+ + f^-.$$

Demostración. Para $|f|$, sea $\varepsilon > 0$. Como f es integrable, Teorema 14.1 proporciona una partición P con $U(f,P) - L(f,P) < \varepsilon$.

La desigualdad de brechas recién obtenida da $U(|f|,P) - L(|f|,P) < \varepsilon$. Por Teorema 14.1, $|f|$ es integrable. Las fórmulas de f^+ y f^- son identidades punto a punto. Como f y $|f|$ son integrables, Teorema 14.6 implica que ambas partes son integrables. Las identidades $f = f^+ - f^-$ y $|f| = f^+ + f^-$ se verifican directamente separando los casos $f(x) \geq 0$ y $f(x) < 0$. Falta el producto. Como f y g son acotadas, existen constantes $A, B > 0$ tales que

$$|f(x)| \leq A, \qquad |g(x)| \leq B$$

para todo $x \in [a,b]$.

En un subintervalo $I \subseteq [a,b]$ y para cualesquiera $x, y \in I$,

$$\begin{aligned} |f(x)g(x) - f(y)g(y)| &\leq |f(x)|\,|g(x) - g(y)| + |g(y)|\,|f(x) - f(y)| \\ &\leq A\,\omega(g; I) + B\,\omega(f; I). \end{aligned}$$

Por tanto, $\omega(fg; I) \leq A\,\omega(g; I) + B\,\omega(f; I)$.

Aplicando esta estimación a cada pieza de una partición P y usando Proposición 13.8,

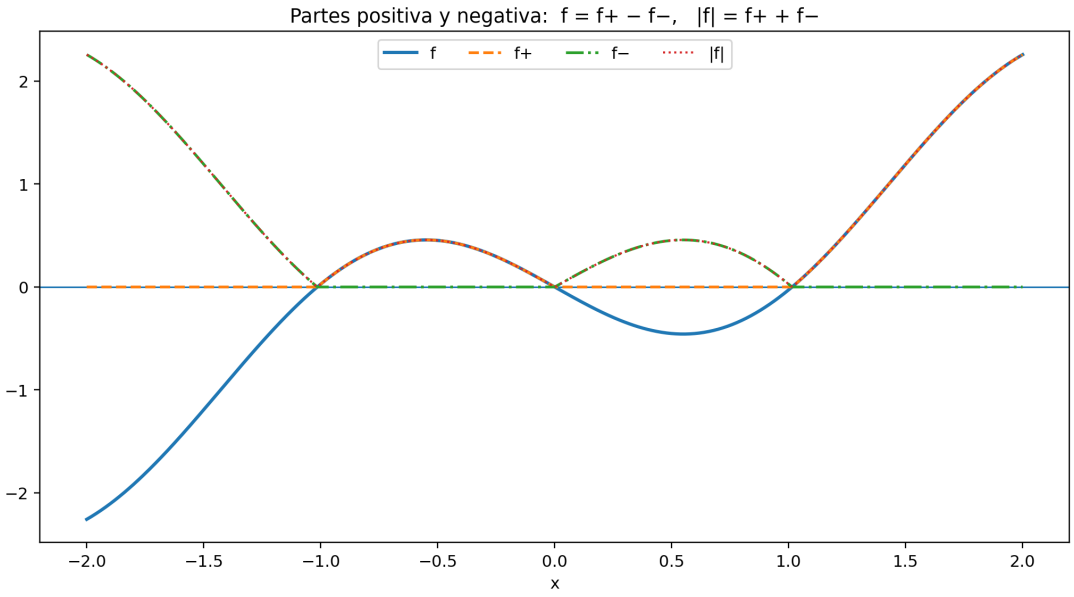


FIGURA 14.9. Las partes positiva y negativa separan las contribuciones de signo de una función: $f = f^+ - f^-$ y $|f| = f^+ + f^-$.

$$\boxed{U(fg,P) - L(fg,P) \le A\big(U(g,P) - L(g,P)\big) + B\big(U(f,P) - L(f,P)\big).}$$

Sea $\varepsilon > 0$. Por Teorema 14.1 existen particiones P_f, P_g tales que $U(f,P_f) - L(f,P_f) < \frac{\varepsilon}{2B}$ y $U(g,P_g) - L(g,P_g) < \frac{\varepsilon}{2A}$.

Tomemos el refinamiento común $P := P_f \vee P_g$.

Por Proposición 13.8, las dos brechas individuales no aumentan al refinar. Por ello,

$$\begin{aligned} U(fg,P) - L(fg,P) &\le A\big(U(g,P) - L(g,P)\big) + B\big(U(f,P) - L(f,P)\big) \\ &< A\frac{\varepsilon}{2A} + B\frac{\varepsilon}{2B} \\ &= \varepsilon. \end{aligned}$$

El criterio de Darboux concluye que fg es integrable. $\square$

La prueba del producto muestra otra vez por qué el refinamiento común es una herramienta estructural. Las particiones que controlan a f y a g pueden ser distintas; al refinarlas simultáneamente obtenemos una sola partición en la cual ambas brechas son pequeñas. No hay, sin embargo, una regla multiplicativa para los valores integrales. En general,

$$\int_a^b f(x)g(x)\,dx \neq \left(\int_a^b f(x)\,dx\right)\left(\int_a^b g(x)\,dx\right).$$

Por ejemplo, en $[0,2]$ la función constante 1 satisface

$$\int_0^2 1\,dx = 2,$$

mientras que

$$\int_0^2 1 \cdot 1 \, dx = 2 \neq 4.$$

La estabilidad del producto significa que fg **es integrable**, no que la integral transforme productos en productos de integrales.

La desigualdad triangular para la integral.

COROLARIO 14.2. **Corolario (desigualdad triangular integral).** Si $f \in \mathcal{R}[a,b]$, entonces

$$\boxed{\left| \int_a^b f(x)\, dx \right| \leq \int_a^b |f(x)|\, dx.}$$

Demostración. Punto a punto, $-|f(x)| \leq f(x) \leq |f(x)|$.

Por Proposición 14.2, $|f|$ es integrable, y por Teorema 14.6 también lo es $-|f|$. Aplicando la monotonía de la integral,

$$-\int_a^b |f(x)|\, dx \leq \int_a^b f(x)\, dx \leq \int_a^b |f(x)|\, dx.$$

La caracterización elemental del valor absoluto de un número real produce la desigualdad anunciada. □

Como consecuencia, las partes positiva y negativa satisfacen

$$\int_a^b f = \int_a^b f^+ - \int_a^b f^-,$$

y

$$\int_a^b |f| = \int_a^b f^+ + \int_a^b f^-.$$

También las funciones $\max\{f,g\} = \frac{f+g+|f-g|}{2}$ y $\min\{f,g\} = \frac{f+g-|f-g|}{2}$ son Riemann integrables.

La clausura genera funciones fuera de las clases iniciales.

EJEMPLO 14.8. **Una función integrable que no es continua ni monótona, obtenida algebraicamente.** En $[0,1]$ consideremos la función identidad $f(x) = x$ y la función de salto de Ejemplo 14.7,

$$s(x) = \begin{cases} 0, & 0 \leq x < \frac{1}{2}, \\ 1, & \frac{1}{2} \leq x \leq 1. \end{cases}$$

Sabemos que f es integrable por Ejemplo 14.2 y que s es integrable por Teorema 14.5. Por Teorema 14.6, $1 - s \in \mathcal{R}[0,1]$, y por Proposición 14.2, $h := f(1-s) \in \mathcal{R}[0,1]$.

Explícitamente,

$$h(x) = \begin{cases} x, & 0 \leq x < \frac{1}{2}, \\ 0, & \frac{1}{2} \leq x \leq 1. \end{cases}$$

La función no es continua en $1/2$: los valores por la izquierda se aproximan a $1/2$, mientras que

$$h\left(\frac{1}{2}\right) = 0.$$

Tampoco es monótona en todo $[0,1]$: aumenta antes de $1/2$ y después cae a 0. Así, la estabilidad algebraica de $\mathcal{R}[0,1]$ produce funciones integrables que no pertenecen a ninguna de las dos clases suficientes estudiadas en §§14.6–14.7.

Qué estructura tenemos ahora. Podemos resumir §14.8 diciendo que

$\mathcal{R}[a,b]$ es un espacio vectorial de funciones, cerrado además bajo $\lvert\cdot\rvert$ y producto.

La integral

$$f \longmapsto \int_a^b f(x)\,dx$$

es lineal y preserva el orden:

$$f \le g \Longrightarrow \int_a^b f \le \int_a^b g.$$

De la interacción entre ambas propiedades surgen las cotas por constantes, la positividad y la desigualdad triangular. Todavía falta una estructura distinta: cómo cambia la integral cuando **cambiamos el intervalo**. En §14.9 estudiaremos restricciones, aditividad respecto de un punto intermedio y, solo después, las convenciones de orientación para límites iguales o invertidos.

Antes de seguir. ¿Por qué la prueba de linealidad mediante Teorema 14.2 evita tener que distinguir el signo de α y β en sumas inferiores y superiores? ¿Cómo se deduce la monotonía de la integral a partir de la positividad de $g - f$? ¿Por qué $\omega(|f|; I) \le \omega(f; I)$ basta para probar integrabilidad de $|f|$? En la prueba del producto, ¿por qué necesitamos un refinamiento común de dos particiones inicialmente distintas? ¿Qué diferencia hay entre afirmar que fg es integrable y afirmar falsamente que la integral es multiplicativa? ¿Qué enseña Ejemplo 14.8 sobre el alcance de los criterios de continuidad y monotonía?

14.9. Restricción, aditividad por intervalos y orientación

Hasta ahora el intervalo $[a,b]$ ha permanecido fijo. Hemos cambiado la función —por combinaciones lineales, valor absoluto, productos o modificaciones finitas—, pero no hemos preguntado qué ocurre cuando cortamos el intervalo o invertimos el orden de sus extremos. La primera cuestión es lógica antes que algebraica. Si c es un punto interior, ¿la integrabilidad sobre $[a,b]$ equivale a la integrabilidad de las dos restricciones sobre $[a,c]$ y $[c,b]$? Solo después de resolver esa pregunta tendrá sentido demostrar la fórmula

$$\int_a^b f = \int_a^c f + \int_c^b f.$$

La prueba volverá a las brechas de Darboux. La razón es simple: cuando una partición contiene el punto c, sus piezas quedan separadas exactamente en dos familias, una a cada lado de c.

Una brecha global se descompone en dos brechas locales. Sea $Q = \{x_0, \dots, x_n\}$ una partición de $[a, b]$ que contiene c. Escribamos

$$Q_- := Q \cap [a, c], \qquad Q_+ := Q \cap [c, b].$$

Entonces Q_- es una partición de $[a, c]$ y Q_+ una partición de $[c, b]$. Si

$$f_- := f|_{[a,c]}, \qquad f_+ := f|_{[c,b]},$$

la definición de las sumas inferior y superior da las identidades exactas $L(f, Q) = L(f_-, Q_-) + L(f_+, Q_+)$, y

$$U(f, Q) = U(f_-, Q_-) + U(f_+, Q_+).$$

Por tanto,

$$\boxed{U(f, Q) - L(f, Q) = \big(U(f_-, Q_-) - L(f_-, Q_-)\big) + \big(U(f_+, Q_+) - L(f_+, Q_+)\big).}$$

Las dos cantidades del lado derecho son no negativas. Así, una brecha global pequeña obliga a que ambas brechas locales sean pequeñas; y, recíprocamente, dos brechas locales pequeñas se pueden ensamblar en una brecha global pequeña.

Integrabilidad local y aditividad.

Teorema 14.7. **Teorema (restricción, recomposición y aditividad de la integral).** Sean $a < c < b$ y sea $f : [a, b] \to \mathbb{R}$. Son equivalentes:

1. $f \in \mathcal{R}[a, b]$;
2. las restricciones
$$f|_{[a,c]} \in \mathcal{R}[a, c], \qquad f|_{[c,b]} \in \mathcal{R}[c, b].$$

Cuando estas condiciones se cumplen,

$$\boxed{\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.}$$

Demostración. Supongamos primero que $f \in \mathcal{R}[a, b]$. Sea $\varepsilon > 0$. Por Teorema 14.1 existe una partición P de $[a, b]$ tal que $U(f, P) - L(f, P) < \varepsilon$.

Insertamos el punto c y formamos $Q := P \cup \{c\}$.

Como Q refina a P, Proposición 13.8 implica

$$U(f, Q) - L(f, Q) \le U(f, P) - L(f, P) < \varepsilon.$$

La descomposición anterior escribe esta brecha como suma de las dos brechas correspondientes a las restricciones. Como ambas son no negativas, $U(f_-, Q_-) - L(f_-, Q_-) < \varepsilon$ y $U(f_+, Q_+) - L(f_+, Q_+) < \varepsilon$. Las restricciones son acotadas porque f lo es. El criterio de Darboux concluye que

$$f_- \in \mathcal{R}[a, c], \qquad f_+ \in \mathcal{R}[c, b].$$

Recíprocamente, supongamos que ambas restricciones son integrables. En particular, existen $B_-, B_+ > 0$ tales que

$$|f(x)| \le B_- \quad (a \le x \le c),$$

y

$$|f(x)| \le B_+ \quad (c \le x \le b).$$

Luego f es acotada en $[a,b]$ por $\max\{B_-, B_+\}$. Sea $\varepsilon > 0$. Por Teorema 14.1 podemos elegir una partición P_- de $[a,c]$ y una partición P_+ de $[c,b]$ tales que $U(f_-,P_-) - L(f_-,P_-) < \frac{\varepsilon}{2}$, y $U(f_+,P_+) - L(f_+,P_+) < \frac{\varepsilon}{2}$. La unión $P := P_- \cup P_+$ es una partición de $[a,b]$ que contiene c. Por la identidad de descomposición,

$$\begin{aligned} U(f,P) - L(f,P) &= \big(U(f_-,P_-) - L(f_-,P_-)\big) \\ &\quad + \big(U(f_+,P_+) - L(f_+,P_+)\big) \\ &< \varepsilon. \end{aligned}$$

De nuevo Teorema 14.1 implica $f \in \mathcal{R}[a,b]$. Falta demostrar la igualdad de integrales. Escribamos

$$I := \int_a^b f(x)\,dx, \qquad I_- := \int_a^c f(x)\,dx, \qquad I_+ := \int_c^b f(x)\,dx.$$

Sea $\varepsilon > 0$. Elegimos, como arriba, particiones P_- y P_+ cuyas brechas sean menores que $\varepsilon/2$, y ponemos $P = P_- \cup P_+$. Entonces $U(f,P) - L(f,P) < \varepsilon$. Por el encierro integral de §14.2, $L(f,P) \le I \le U(f,P)$.

Además, $L(f_-,P_-) \le I_- \le U(f_-,P_-)$ y $L(f_+,P_+) \le I_+ \le U(f_+,P_+)$.

Sumando estas dos últimas cadenas y usando la descomposición exacta de las sumas, $L(f,P) \le I_- + I_+ \le U(f,P)$. Así, tanto I como $I_- + I_+$ pertenecen al mismo intervalo de longitud menor que ε. Por tanto, $|I - (I_- + I_+)| < \varepsilon$. Como esto vale para todo $\varepsilon > 0$, necesariamente $I = I_- + I_+$.

Queda probado el teorema. □

El resultado contiene dos mensajes diferentes. La integrabilidad es **local respecto de una descomposición finita del intervalo**, y el valor integral es **aditivo respecto del intervalo**. Ninguna de las dos afirmaciones utiliza primitivas ni derivadas.

De dos piezas a finitísimas piezas. La forma con un solo punto intermedio se itera inmediatamente.

Corolario 14.3. **Corolario (localidad finita y suma por una partición).** Sea

$$P = \{x_0, x_1, \dots, x_n\}, \qquad x_0 = a < x_1 < \cdots < x_n = b.$$

Para una función $f : [a,b] \to \mathbb{R}$, son equivalentes:

1. $f \in \mathcal{R}[a,b]$;
2. para cada $k = 1, \dots, n$,
$$f|_{[x_{k-1},x_k]} \in \mathcal{R}[x_{k-1}, x_k].$$

Cuando se cumplen,

$$\boxed{\int_a^b f(x)\,dx = \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} f(x)\,dx.}$$

Demostración. Aplicamos Teorema 14.7 repetidamente a los puntos interiores $x_1, \dots, x_{n-1}$. La equivalencia de integrabilidad se obtiene por inducción sobre el número de piezas, y en cada paso la fórmula de aditividad sustituye una integral sobre un intervalo por la suma de las dos integrales sobre los subintervalos que lo componen. □

Este corolario formaliza una práctica que, para funciones escalonadas, ya aparecía de manera elemental en C14. Ahora la descomposición vale para toda función Riemann integrable.

Los límites invertidos son una convención de orientación. Hasta aquí todas las integrales de Riemann se han definido con extremo inferior estrictamente menor que el superior. Las expresiones

$$\int_a^a f(x)\,dx$$

y

$$\int_b^a f(x)\,dx \qquad (a < b)$$

no provienen directamente de Definición 14.2. Las introducimos ahora por convención para que la aditividad conserve una forma uniforme al cambiar la orientación.

Definición 14.3. **Definición (integral orientada).** Sea $J \subseteq \mathbb{R}$ un intervalo y sea $f : J \to \mathbb{R}$. Para $u, v \in J$:

- si $u = v$, definimos
$$\boxed{\int_u^u f(x)\,dx := 0;}$$
- si $u < v$ y f es Riemann integrable en $[u, v]$, conservamos la integral ya definida;
- si $u > v$ y f es Riemann integrable en $[v, u]$, definimos
$$\boxed{\int_u^v f(x)\,dx := -\int_v^u f(x)\,dx.}$$

Estas convenciones extienden la notación; no modifican la definición primaria de integrabilidad de Riemann para intervalos no degenerados.

La igualdad

$$\int_u^v f = -\int_v^u f$$

queda entonces válida también cuando $u = v$, pues ambos miembros son cero.

La aditividad orientada ya no necesita ordenar los tres puntos.

Proposición 14.3. **Proposición (aditividad orientada).** Sean r, s, t puntos de un intervalo J y sea $f : J \to \mathbb{R}$. Supongamos que f es Riemann integrable en el intervalo cerrado

$$[m, M], \qquad m := \min\{r, s, t\}, \quad M := \max\{r, s, t\},$$

cuando $m < M$. Entonces

$$\boxed{\int_r^s f(x)\,dx + \int_s^t f(x)\,dx = \int_r^t f(x)\,dx.}$$

Demostración. Fijemos el extremo mínimo m y, para $x \in [m, M]$, escribamos simplemente

$$A(x) := \int_m^x f(u)\,du,$$

usando la convención $A(m) = 0$.

Si $x \le y$, Teorema 14.7 aplicado al punto intermedio x —o trivialmente la convención si $x = m$ o $x = y$— da

$$A(y) = A(x) + \int_x^y f(u)\,du.$$

Por tanto,

$$\int_x^y f(u)\,du = A(y) - A(x).$$

Si $x > y$, la definición de orientación produce la misma identidad:

$$\int_x^y f(u)\,du = -\int_y^x f(u)\,du = -(A(x) - A(y)) = A(y) - A(x).$$

Así, para cualesquiera $x, y \in [m, M]$,

$$\boxed{\int_x^y f(u)\,du = A(y) - A(x).}$$

Aplicándola a los pares (r, s), (s, t) y (r, t),

$$\begin{aligned}\int_r^s f + \int_s^t f &= [A(s) - A(r)] + [A(t) - A(s)] \\ &= A(t) - A(r) \\ &= \int_r^t f.\end{aligned}$$

La cancelación es puramente algebraica; no se ha introducido ninguna primitiva de f. □

El símbolo $A(x)$ de la prueba es solo una abreviatura para un número integral con extremo variable. No hemos afirmado nada sobre continuidad o diferenciabilidad de A; esas cuestiones pertenecen al teorema fundamental del cálculo de C17.

Un salto visto con ambas orientaciones.

Ejemplo 14.9. **Aditividad y orientación para una función con salto.** Consideremos en $[0, 1]$ la función de Ejemplo 14.7,

$$s(x) = \begin{cases} 0, & 0 \le x < \frac{1}{2}, \\ 1, & \frac{1}{2} \le x \le 1. \end{cases}$$

En $[0, 1/2]$ difiere de la función constante 0 solo en el punto $1/2$. Por Corolario 14.1,

$$\int_0^{1/2} s(x)\,dx = 0.$$

En $[1/2, 1]$ es constantemente igual a 1, así que

Aditividad por intervalos y orientación

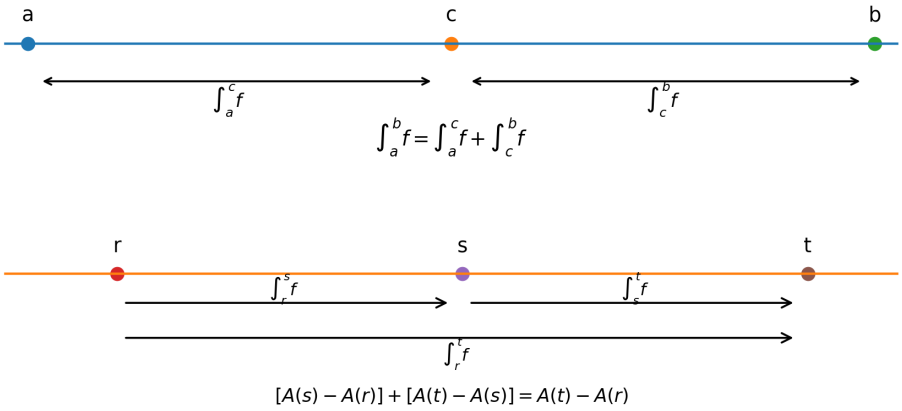


FIGURA 14.10. La aditividad separa un intervalo en piezas; la orientación permite expresar la misma identidad para extremos en cualquier orden.

$$\int_{1/2}^{1} s(x)\,dx = \frac{1}{2}.$$

Por Teorema 14.7,

$$\boxed{\int_{0}^{1} s(x)\,dx = 0 + \frac{1}{2} = \frac{1}{2}.}$$

Al invertir la orientación,

$$\int_{1}^{1/2} s(x)\,dx = -\frac{1}{2}, \qquad \int_{1/2}^{0} s(x)\,dx = 0,$$

y Proposición 14.3 da

$$\int_{1}^{1/2} s + \int_{1/2}^{0} s = \int_{1}^{0} s = -\frac{1}{2}.$$

El punto de salto no genera ningún término adicional: la aditividad se refiere a integrales sobre intervalos, no a una suma de «áreas más un valor en el corte».

Qué queda fijado antes del laboratorio. La teoría dispone ahora de dos tipos de estructura complementarios. Respecto de la **función**, §14.8 estableció linealidad, orden y estabilidad algebraica. Respecto del **intervalo**, esta sección ha establecido

$$\boxed{\int_a^b f = \int_a^c f + \int_c^b f}$$

para $a < c < b$, y después ha extendido esa identidad a cualquier orden de los extremos mediante orientación. Conviene conservar el orden lógico:

$$\text{integrabilidad en subintervalos}$$
$$\longleftrightarrow \text{integrabilidad global}$$
$$\longrightarrow \text{aditividad}$$
$$\longrightarrow \text{convención de orientación.}$$

No hemos usado el teorema de valor medio integral ni el teorema fundamental del cálculo. Tampoco hemos afirmado que la función

$$x \longmapsto \int_a^x f(t)\,dt$$

tenga todavía propiedades analíticas especiales.

Antes de seguir. ¿Por qué insertar c en una partición global permite separar exactamente la brecha en dos términos no negativos? ¿Cómo se demuestra la dirección recíproca cuando las particiones buenas de $[a, c]$ y $[c, b]$ son distintas? En la prueba de aditividad, ¿por qué basta encerrar tanto I como $I_- + I_+$ entre la misma suma inferior y la misma suma superior? ¿Por qué $\int_b^a f = -\int_a^b f$ es una convención posterior y no parte de Definición 14.2? ¿Cómo evita Proposición 14.3 una discusión de los seis posibles órdenes de tres puntos? ¿Por qué la abreviatura $A(x) = \int_m^x f$ no anticipa todavía el teorema fundamental del cálculo?

14.10. Laboratorio de integrabilidad: qué basta y qué no

A esta altura del capítulo disponemos de varios caminos correctos para demostrar integrabilidad y también de varios razonamientos que **parecen** suficientes pero no lo son. Conviene reunirlos antes del banco de ejercicios. El objetivo de este laboratorio no es añadir teoría nueva. Es aprender a reconocer qué tipo de información tenemos en cada problema y qué resultado ya demostrado permite transformarla legítimamente en una conclusión sobre integrabilidad. La pregunta guía será siempre la misma:

¿qué controla realmente todas las aproximaciones relevantes?

No basta con que una aproximación concreta se comporte bien. La definición de Riemann y sus criterios equivalentes son afirmaciones globales sobre particiones, brechas o sumas etiquetadas.

Estación 1. Cuatro puertas suficientes que ya están demostradas. Hasta aquí hemos probado cuatro mecanismos distintos que garantizan integrabilidad.

Funciones escalonadas. Por Teorema 14.3, toda función escalonada es Riemann integrable y su integral general coincide con la integral elemental de C14.

Funciones continuas. Por Teorema 14.4, toda función continua en un intervalo cerrado es integrable. El mecanismo decisivo no fue la derivabilidad, sino la continuidad uniforme producida por la compactitud.

Funciones monótonas. Por Teorema 14.5, toda función monótona en un intervalo cerrado es integrable, incluso si posee saltos. La cota cuantitativa fue

$$U(f, P) - L(f, P) \leq \|P\|\,|f(b) - f(a)|.$$

Modificaciones finitas. Por Corolario 14.1, si una función integrable se cambia en finitísimos puntos, la nueva función sigue siendo integrable y conserva exactamente el mismo valor integral.

Estas cuatro rutas son **criterios suficientes**. Ninguna de ellas caracteriza por sí sola a todas las funciones Riemann integrables. En efecto, Ejemplo 14.8 ya produjo una función integrable obtenida mediante operaciones algebraicas que no pertenece simplemente a las clases «continua» o «monótona». La teoría de §14.8 muestra que la clase $\mathcal{R}[a, b]$ es mucho más amplia que cada una de esas familias iniciales.

Estación 2. Acotación no significa integrabilidad. La definición primaria Definición 14.2 exige que la función sea acotada antes de preguntar si las integrales inferior y superior coinciden. Esa hipótesis es necesaria para la teoría de Darboux desarrollada aquí, pero no es suficiente. La función de Dirichlet de Ejemplo 14.1,

$$d(x) = \begin{cases} 1, & x \in \mathbb{Q}, \\ 0, & x \notin \mathbb{Q}, \end{cases} \qquad 0 \le x \le 1,$$

es acotada y, sin embargo, para toda partición P satisface

$$L(d, P) = 0, \qquad U(d, P) = 1.$$

Por tanto, $U(d, P) - L(d, P) = 1$ para **toda** partición. No existe ninguna escala en la que la brecha comience a cerrarse.

Así,

$$\boxed{\text{acotada} \not\Longrightarrow \text{Riemann integrable.}}$$

La acotación permite formular el problema; no lo resuelve.

Estación 3. Una brecha pequeña no es lo mismo que brechas arbitrariamente pequeñas. Supongamos que encontramos una partición P para la cual $U(f, P) - L(f, P) = 10^{-6}$. Ese dato puede ser excelente como aproximación numérica, pero por sí solo no demuestra integrabilidad. El criterio de Darboux Teorema 14.1 exige

$$\forall \varepsilon > 0 \ \exists P : U(f, P) - L(f, P) < \varepsilon.$$

El orden de los cuantificadores es parte del teorema. La familia escalada de la función de Dirichlet lo muestra con precisión. Para un número fijo $c > 0$, sea

$$d_c(x) = \begin{cases} c, & x \in \mathbb{Q}, \\ 0, & x \notin \mathbb{Q}. \end{cases}$$

En toda partición, $U(d_c, P) - L(d_c, P) = c$.

Si $c = 10^{-6}$, entonces toda partición posee una brecha que puede parecer «muy pequeña», pero al pedir $\varepsilon = \frac{c}{2}$ ninguna sirve. Hay, sin embargo, un caso especial importante: si para alguna partición se obtiene exactamente $U(f, P) - L(f, P) = 0$, entonces esa misma partición satisface el criterio para **todo** $\varepsilon > 0$, y la función sí es integrable. Lo insuficiente es una brecha fija **positiva**, por pequeña que parezca.

Estación 4. Malla pequeña no implica brecha pequeña. Otro error frecuente consiste en suponer que una partición muy fina debe producir automáticamente una buena aproximación. La función d anterior lo refuta de inmediato. Para cada n podemos tomar la partición uniforme

$$P_n = \left\{0, \frac{1}{n}, \frac{2}{n}, \dots, 1\right\},$$

cuya malla satisface $\|P_n\| = \frac{1}{n} \longrightarrow 0$.

Sin embargo, $U(d, P_n) - L(d, P_n) = 1$ para todo n.

Por tanto,

$$\boxed{\|P_n\| \to 0 \not\Longrightarrow U(f, P_n) - L(f, P_n) \to 0}$$

para una función acotada arbitraria. La malla es una medida geométrica de la partición; la brecha depende además de la oscilación de la función dentro de sus piezas. En §§14.6–14.7 pudimos relacionar ambas cantidades porque la continuidad uniforme y la monotonía proporcionaban precisamente ese control adicional. Y en §14.4, Teorema 14.2 demostró algo todavía más fuerte para una función **ya integrable**: existe una escala de malla a partir de la cual todas las sumas etiquetadas quedan cerca de la integral. Ese resultado no autoriza a comenzar la teoría suponiendo que «malla pequeña» implica por sí sola «buena suma».

Estación 5. Una familia convergente de sumas etiquetadas tampoco basta. La misma función d permite aislar un error lógico más sutil. Tomemos las particiones uniformes P_n anteriores. Cada subintervalo no degenerado contiene números racionales e irracionales. Por tanto podemos elegir, para cada pieza de P_n, una etiqueta irracional. Llamemos $\xi^{(n)}$ a esa elección. Entonces $d(\xi_k^{(n)}) = 0$ para toda etiqueta y $R(d; P_n, \xi^{(n)}) = 0$ para todo n. En particular, $R(d; P_n, \xi^{(n)}) \longrightarrow 0$. Si en las mismas particiones escogemos en cambio etiquetas racionales $\eta^{(n)}$, obtenemos $R(d; P_n, \eta^{(n)}) = 1$ para todo n, y por tanto $R(d; P_n, \eta^{(n)}) \longrightarrow 1$. Así, una función no integrable puede poseer una familia perfectamente convergente de sumas etiquetadas —incluso muchas—. Lo que falla es la independencia respecto de las elecciones.

El contraste correcto es

una familia escogida converge	insuficiente,
todas las sumas con malla pequeña se acercan al mismo I	equivalente a integrabilidad.

La segunda afirmación es exactamente Teorema 14.2.

Estación 6. Qué control sí es suficiente. Podemos organizar los certificados correctos en tres niveles.

Nivel A: definición. Si las integrales inferior y superior globales coinciden,

$$\underline{\int_a^b} f = \overline{\int_a^b} f,$$

entonces f es integrable por Definición 14.2.

Nivel B: brechas de Darboux. Si para cada $\varepsilon > 0$ podemos construir una partición con $U(f, P) - L(f, P) < \varepsilon$, entonces f es integrable por Teorema 14.1.

Nivel C: todas las sumas etiquetadas finas. Si existe $I \in \mathbb{R}$ tal que

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall (P, \xi) : \|P\| < \delta \Longrightarrow |R(f; P, \xi) - I| < \varepsilon,$$

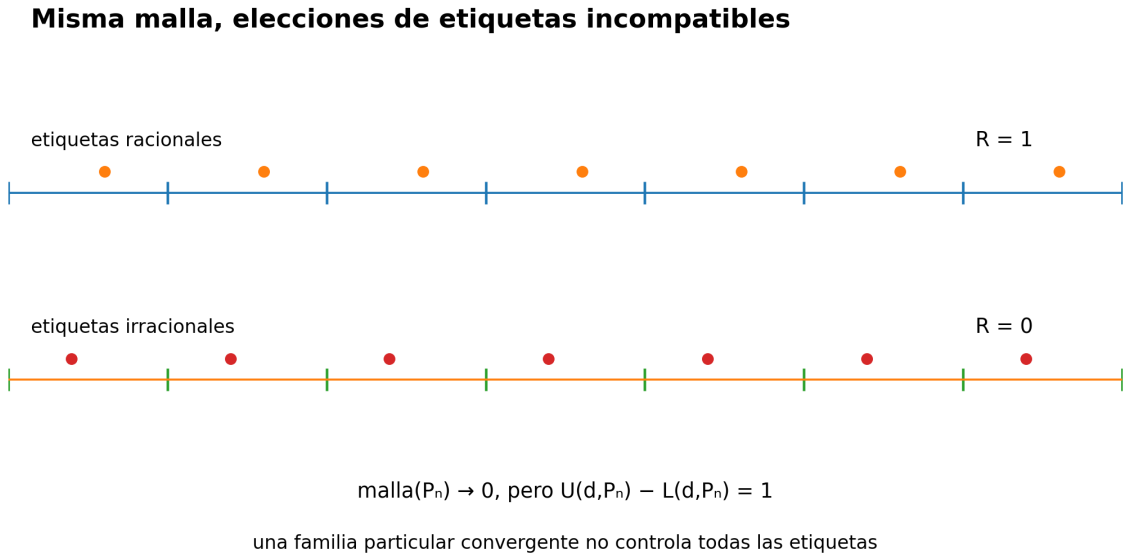


FIGURA 14.11. La función racional/irracional muestra que una misma sucesión de particiones puede producir sumas etiquetadas incompatibles: 0 o 1.

entonces f es integrable y su integral vale I por Teorema 14.2. Los niveles B y C no son definiciones nuevas dentro de CPM: son caracterizaciones equivalentes ya demostradas de la definición primaria del nivel A. Por debajo de estos niveles quedan datos que pueden ser útiles pero que necesitan una pieza lógica adicional: una partición concreta, una brecha positiva fija, una sucesión particular de mallas, una elección especial de etiquetas o una gráfica que «parece tener área».

Estación 7. Reparación de seis razonamientos circulares o incompletos. Razonamiento 1. «f es continua, luego tiene una primitiva, y por el teorema fundamental del cálculo es integrable.»

Utilizar el teorema fundamental aquí introduciría una dependencia circular. La integrabilidad de las funciones continuas ya se obtuvo en Teorema 14.4 mediante continuidad uniforme y el criterio de Darboux; esa es la justificación disponible antes de relacionar integración y derivación.

Razonamiento 2. «f es monótona, así que el área bajo su gráfica existe.»

La frase geométrica no constituye una prueba. La reparación es cuantitativa: Teorema 14.5 demuestra

$$U(f,P) - L(f,P) \leq \|P\|\,|f(b) - f(a)|,$$

y de allí se aplica el criterio de Darboux.

Razonamiento 3. «Una escalonada ya tenía integral en C14, por tanto esa integral es automáticamente la integral de Riemann.»

No. Antes había dos construcciones diferentes. Teorema 14.3 fue necesario para demostrar primero integrabilidad de la escalonada y después igualdad entre la integral elemental y la integral general.

Razonamiento 4. «Cambiar unos pocos puntos no importa porque los puntos tienen anchura cero.»

La intuición es útil, pero no prueba la afirmación para sumas etiquetadas, donde un punto excepcional puede ser elegido como etiqueta. Corolario 14.1 controla exactamente cuánto pueden cambiar las sumas: solo las piezas etiquetadas en esos finitísimos puntos contribuyen, y su longitud total queda controlada por la malla.

Razonamiento 5. «Como la integral es lineal, si f y g son integrables entonces $f+g$ también lo es.»

El orden está invertido. Para poder escribir la integral de $f+g$ hay que demostrar primero que $f+g$ es integrable. Teorema 14.6 establece simultáneamente el cierre de $\mathcal{R}[a,b]$ bajo combinaciones lineales y la fórmula correspondiente para la integral.

Razonamiento 6. «La igualdad $\int_a^b f = \int_a^c f + \int_c^b f$ es evidente porque estamos partiendo el área en dos.»

De nuevo, la imagen geométrica anticipa el resultado pero no sustituye su prueba. Teorema 14.7 demuestra antes la equivalencia entre integrabilidad global y local, y solo después identifica los valores integrales. La orientación de §14.9 se añadió todavía más tarde como convención compatible con esa aditividad. La regla de trabajo que emerge es sencilla:

$$\boxed{\text{antes de usar una propiedad de } \int f, \text{ hay que justificar que la integral existe.}}$$

Estación 8. Tabla de diagnóstico.

Información disponible	¿Basta para integrabilidad?	Razón correcta
f es acotada	No	Es hipótesis previa, no criterio suficiente; Ejemplo 14.1.
Existe una partición con brecha positiva pequeña	No	Falta controlar toda tolerancia $\varepsilon > 0$.
Existe una partición con brecha exactamente 0	Sí	La misma partición sirve para todo $\varepsilon > 0$ en Teorema 14.1.
Existe (P_n) con $U(f,P_n) - L(f,P_n) \to 0$	Sí	Reformulación secuencial del criterio de Darboux.
Existe (P_n) con $\|P_n\| \to 0$	No	La malla no controla por sí sola la oscilación de f.
Una familia particular $R(f;P_n,\xi^{(n)})$ converge	No	Solo controla esas elecciones; la función de Dirichlet da contraejemplo.
Todas las sumas etiquetadas suficientemente finas se acercan al mismo I	Sí	Equivalente por Teorema 14.2; además $I = \int f$.
f es escalonada	Sí	Teorema 14.3.
f es continua en $[a,b]$	Sí	Teorema 14.4.
f es monótona en $[a,b]$	Sí	Teorema 14.5.
f difiere finitamente de una función integrable	Sí	Corolario 14.1.

La tabla no reemplaza las demostraciones. Su función es decidir **qué puerta lógica debe abrirse** en cada problema.

El mapa final antes de ejercitar la teoría. Las ideas centrales del capítulo pueden ahora leerse como una cadena de controles cada vez más robustos:

función acotada
$\Downarrow$
sumas inferior y superior para cada partición
$\Downarrow$
extremos globales
$\Downarrow$
coincidencia $\iff$ brechas arbitrariamente pequeñas
$\iff$
todas las sumas etiquetadas suficientemente finas
se concentran en un único valor.

Sobre esa base se apoyan después las clases suficientes —escalonadas, continuas y monótonas—, la invariancia ante cambios finitos, el álgebra de funciones integrables y la aditividad respecto del intervalo. La construcción y las propiedades anteriores proceden de las sumas y sus estimaciones. La relación con la derivación exigirá teoremas adicionales sobre las funciones de acumulación.

Frontera del capítulo
En este laboratorio no se usa el teorema de valor medio integral, no se usa ninguna primitiva y no se usa el teorema fundamental del cálculo. Tampoco se invoca el criterio de Lebesgue. C15 sigue respondiendo una pregunta previa: **qué significa ser Riemann integrable y cómo puede certificarse esa propiedad desde la definición.**

Una brecha pequeña en una partición solo garantiza el encierro correspondiente a esa precisión. La integrabilidad requiere obtener brechas arbitrariamente pequeñas. De manera equivalente, el criterio uniforme controla todas las etiquetas cuando la malla es suficientemente fina; controlar una familia elegida de etiquetas no basta para esa conclusión.

14.11. Ejercicios y soluciones

Los cuarenta ejercicios siguientes recorren la arquitectura completa del capítulo sin introducir teoría posterior. El banco está organizado en siete niveles. Cada problema tiene una solución desarrollada y puede resolverse únicamente con los resultados disponibles hasta §14.10.

Nivel A — Extremos globales y definición de integrabilidad. Ejercicio A1. Extremos globales de una función constante. Sea $f(x) = c$ en $[a, b]$, con $a < b$.

1. Calcula $L(f, P)$ y $U(f, P)$ para una partición arbitraria P.
2. Determina $\mathcal{L}(f)$ y $\mathcal{U}(f)$.
3. Obtén la integral inferior global, la integral superior global y decide si f es Riemann integrable.

Ejercicio A2. Atrapar los extremos globales de la identidad. Para $f(x) = x$ en $[0,1]$, considera la partición uniforme

$$P_n = \left\{0, \frac{1}{n}, \dots, 1\right\}.$$

Demuestra que

$$L(f, P_n) = \frac{1}{2} - \frac{1}{2n}, \qquad U(f, P_n) = \frac{1}{2} + \frac{1}{2n},$$

y usa únicamente el encierro global de Proposición 14.1 para concluir

$$\underline{\int_0^1} f = \overline{\int_0^1} f = \frac{1}{2}.$$

Ejercicio A3. Extremos globales separados. Define

$$g(x) = \begin{cases} 3, & x \in \mathbb{Q}, \\ -1, & x \notin \mathbb{Q}, \end{cases} \qquad 0 \le x \le 2.$$

Calcula $L(g, P)$ y $U(g, P)$ para una partición arbitraria P, determina las integrales inferior y superior globales y decide si g es Riemann integrable.

Ejercicio A4. Cuando una suma inferior alcanza una suma superior. Sea $f : [a,b] \to \mathbb{R}$ acotada. Supón que existen particiones P y Q tales que $L(f,P) = U(f,Q) = I$.

Demuestra que f es Riemann integrable y que

$$\int_a^b f(x)\,dx = I.$$

No supongas que $P = Q$ ni que una refine a la otra.

Ejercicio A5. Aproximar el valor integral por ambos lados. Sea $f \in \mathcal{R}[a,b]$ y

$$I = \int_a^b f(x)\,dx.$$

Demuestra que para todo $\varepsilon > 0$ existen particiones P_- y P_+ tales que

$$I - \varepsilon < L(f, P_-) \le I \le U(f, P_+) < I + \varepsilon.$$

Explica por qué no hay razón para que $P_- = P_+$.

Ejercicio A6. Orden antes de la integrabilidad. Sean $f, g : [a,b] \to \mathbb{R}$ acotadas y supón que

$$f(x) \le g(x) \qquad \text{para todo } x \in [a,b].$$

Demuestra que

$$\underline{\int_a^b} f \le \underline{\int_a^b} g, \qquad \overline{\int_a^b} f \le \overline{\int_a^b} g.$$

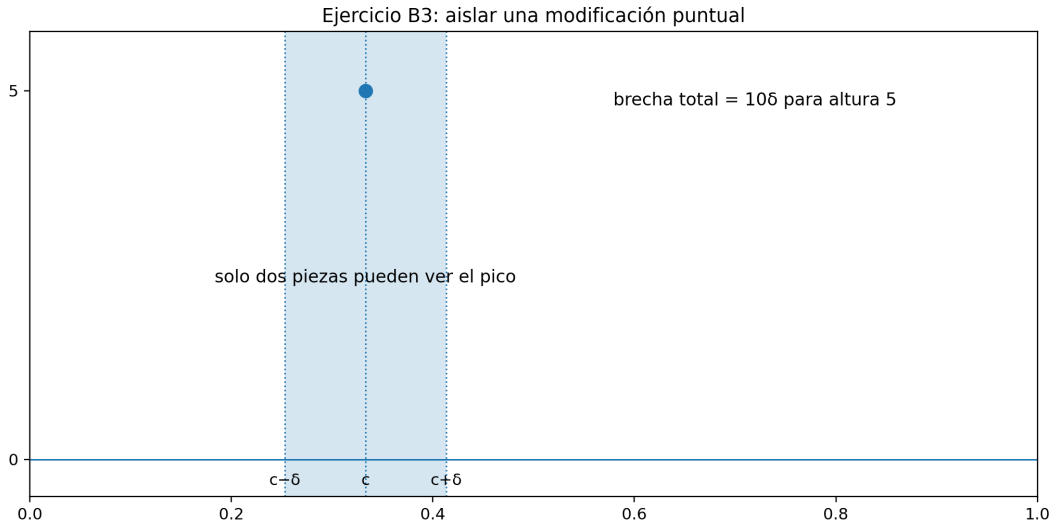


FIGURA 14.12. En el ejercicio B3, el valor excepcional queda confinado a las dos piezas adyacentes a $1/3$, cuya longitud total es 2δ.

Observa que no se supone que ninguna de las dos funciones sea integrable.

Nivel B — Criterio de Darboux y construcción de particiones. Ejercicio B1. Certificar x^2 con una tolerancia dada. Para $f(x) = x^2$ en $[0,1]$ se sabe que, con la partición uniforme P_n, $U(f,P_n) - L(f,P_n) = \frac{1}{n}$.

Dado $\varepsilon > 0$, elige explícitamente un entero n que garantice $U(f,P_n) - L(f,P_n) < \varepsilon$ y concluye la integrabilidad de f sin calcular el valor de su integral.

Ejercicio B2. Una cota explícita para $|x|$. Sea

$$f(x) = |x|, \qquad -1 \le x \le 1.$$

Usa la desigualdad $\big||x| - |y|\big| \le |x-y|$ para demostrar que, para toda partición P de $[-1,1]$, $U(f,P) - L(f,P) \le 2\|P\|$. Deduce directamente del criterio de Darboux que f es Riemann integrable.

Ejercicio B3. Aislar una modificación puntual. Define

$$s(x) = \begin{cases} 5, & x = \frac{1}{3}, \\ 0, & x \neq \frac{1}{3}, \end{cases} \qquad 0 \le x \le 1.$$

Para $0 < \delta < 1/3$ considera

$$P_\delta = \left\{0, \frac{1}{3} - \delta, \frac{1}{3}, \frac{1}{3} + \delta, 1\right\}.$$

Calcula exactamente $U(s,P_\delta) - L(s,P_\delta)$ y usa el resultado para demostrar la integrabilidad de s mediante Teorema 14.1.

Ejercicio B4. Monotonía y variación total. Para $f(x) = x^3$ en $[0,1]$, usa Teorema 14.5 para demostrar que, para toda partición P, $U(f,P) - L(f,P) \le \|P\|$.

Dado $\varepsilon > 0$, construye una partición uniforme que certifique la integrabilidad de f.

Ejercicio B5. Una sucesión de brechas es suficiente. Sea $f : [a,b] \to \mathbb{R}$ acotada. Supón que existe una sucesión de particiones (P_n) tal que $U(f,P_n) - L(f,P_n) \le \frac{1}{n}$ para todo $n \ge 1$.

Demuestra directamente, a partir de Teorema 14.1, que f es Riemann integrable.

Ejercicio B6. Una brecha pequeña pero fija no basta. Para un número $c > 0$ define

$$d_c(x) = \begin{cases} c, & x \in \mathbb{Q}, \\ 0, & x \notin \mathbb{Q}, \end{cases} \qquad 0 \le x \le 1.$$

Demuestra que toda partición tiene brecha exactamente c. Explica por qué el dato «existe una partición con brecha 10^{-9}» no sería suficiente para probar integrabilidad.

Nivel C — Sumas etiquetadas y criterio por malla. Ejercicio C1. Todas las sumas de la identidad son cercanas. Sea $f(x) = x$ en $[0,1]$ y sea (P,ξ) una partición etiquetada arbitraria.

Demuestra que

$$\left| R(f;P,\xi) - \frac{1}{2} \right| \le \|P\|.$$

Concluye que la condición uniforme de Teorema 14.2 se verifica con, por ejemplo, $\delta = \varepsilon$.

Ejercicio C2. Las mismas particiones, dos límites incompatibles. Sea d la función racional/irracional de Ejemplo 14.1 y sea P_n la partición uniforme de $[0,1]$ en n piezas.

Construye una elección de etiquetas racionales y otra de etiquetas irracionales sobre cada P_n, calcula las dos sumas y explica por qué este comportamiento impide satisfacer Teorema 14.2 para cualquier candidato I.

Como apoyo visual para esta comparación, véase Figura 14.11.

Ejercicio C3. Cuánto puede alterar una sola excepción una suma etiquetada. Sean $f, g : [a,b] \to \mathbb{R}$ iguales salvo en un punto c, y sea $D = |g(c) - f(c)|$.

Demuestra que, para toda partición etiquetada (P,ξ), $|R(g;P,\xi) - R(f;P,\xi)| \le 2D\|P\|$.

Explica el factor 2.

Ejercicio C4. El límite uniforme de las sumas etiquetadas es único. Supón que dos números $I, J \in \mathbb{R}$ satisfacen la condición 2 de Teorema 14.2 para una misma función acotada f.

Demuestra que $I = J$ sin usar la definición por integrales inferior y superior.

Ejercicio C5. Todas las sumas finas son mutuamente cercanas. Sea $f \in \mathcal{R}[a,b]$.

Demuestra que para todo $\varepsilon > 0$ existe $\delta > 0$ tal que, si (P,ξ) y (Q,η) son dos particiones etiquetadas cualesquiera con

$$\|P\| < \delta, \qquad \|Q\| < \delta,$$

entonces $|R(f;P,\xi) - R(f;Q,\eta)| < \varepsilon$.

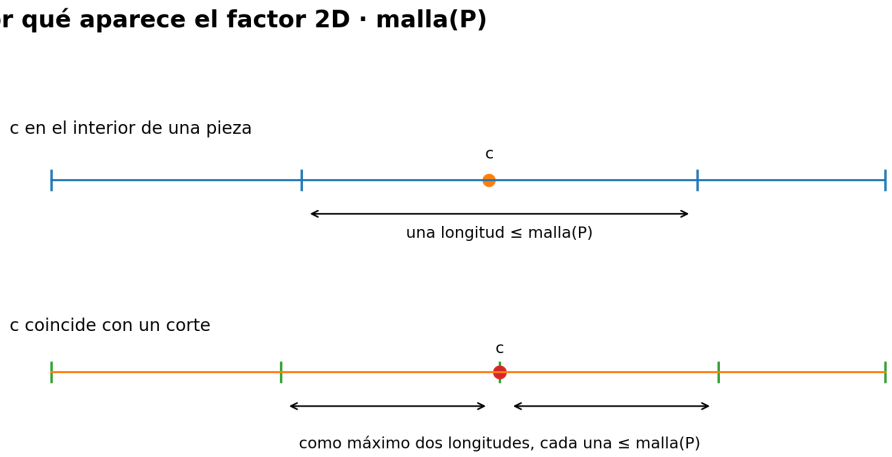


FIGURA 14.13. Un punto excepcional puede ser etiqueta de una sola pieza si está en su interior, o de como máximo dos piezas si coincide con un corte.

Ejercicio C6. Auditar una prueba incompleta. Un estudiante afirma:

«He construido particiones etiquetadas $(P_n, \xi^{(n)})$ con $\|P_n\| \to 0$ y $R(f; P_n, \xi^{(n)}) \to 7$. Por tanto, $\int_a^b f = 7$.»

Indica con precisión qué falta para que el argumento sea válido y escribe una hipótesis adicional, formulada con cuantificadores, que sí permitiría concluir que f es integrable con integral 7.

Nivel D — Escalonadas, continuas, monótonas y modificaciones finitas. Ejercicio D1. Integral elemental e integral de Riemann. Sea

$$s(x) = \begin{cases} 2, & 0 \le x < 1, \\ -1, & 1 < x \le 3, \end{cases}$$

y asigna a $s(1)$ cualquier valor real.

1. Calcula la integral elemental de s.
2. Usa Teorema 14.3 para justificar que el mismo número es su integral de Riemann.
3. Explica por qué el valor escogido en $x = 1$ no cambia la respuesta.

Ejercicio D2. Dos valores enormes no alteran la integral. Define en $[0, 1]$

$$g(x) = \begin{cases} 1000, & x = \frac{1}{4}, \\ -500, & x = \frac{3}{4}, \\ x, & \text{en otro caso.} \end{cases}$$

Demuestra que g es Riemann integrable y calcula su integral usando únicamente resultados ya demostrados en C15.

Ejercicio D3. Continuidad sin monotonía global. Considera $f(x) = |x|$ en $[-1, 1]$.

1. Demuestra que f es continua y por tanto Riemann integrable.
2. Demuestra que f no es monótona en todo $[-1, 1]$.
3. Explica por qué Teorema 14.5 no es necesario para este caso.

Ejercicio D4. Monótona, discontinua e integrable. Define

$$h(x) = \begin{cases} 0, & 0 \le x < \frac{1}{2}, \\ 1, & \frac{1}{2} \le x \le 1. \end{cases}$$

Demuestra que h es no decreciente, que es discontinua en $1/2$ y que es Riemann integrable. Calcula su integral por compatibilidad con la integral elemental.

Ejercicio D5. Modificar finitísimos puntos tampoco repara una función no integrable. Sea d la función racional/irracional de Ejemplo 14.1 y sea $g : [0, 1] \to \mathbb{R}$ una función que coincide con d salvo en un conjunto finito.

Demuestra que g no es Riemann integrable.

Ejercicio D6. Dos criterios suficientes que no se contienen mutuamente. Da:

1. una función continua en un intervalo cerrado que no sea monótona en todo el intervalo;
2. una función monótona Riemann integrable que no sea continua.

Justifica la integrabilidad de cada ejemplo citando el teorema adecuado y explica qué muestra esto sobre §§14.6–14.7.

Nivel E — Álgebra, orden, valor absoluto y productos. Ejercicio E1. Una combinación lineal antes de evaluarla. Sean $f, g \in \mathcal{R}[a, b]$ y sean $\alpha, \beta \in \mathbb{R}$.

Explica por qué $\alpha f + \beta g$ es integrable antes de escribir su integral y demuestra

$$\int_a^b (\alpha f + \beta g) = \alpha \int_a^b f + \beta \int_a^b g.$$

Indica qué sería circular en el argumento «la fórmula es lineal, luego $\alpha f + \beta g$ es integrable».

Ejercicio E2. Producto de funciones integrables. En $[-1, 1]$ define

$$f(x) = x, \qquad g(x) = |x|.$$

Demuestra que f, g y $fg = x|x|$ son Riemann integrables. No calcules la integral del producto.

Ejercicio E3. Máximo y mínimo de dos funciones integrables. Sean $f, g \in \mathcal{R}[a, b]$.

Usa las identidades $\max\{f, g\} = \frac{f+g+|f-g|}{2}$, y $\min\{f, g\} = \frac{f+g-|f-g|}{2}$ para demostrar que ambas funciones son Riemann integrables.

Ejercicio E4. La desigualdad triangular puede ser estricta. Sea

$$s(x) = \begin{cases} 2, & 0 \le x < 1, \\ -1, & 1 \le x \le 3. \end{cases}$$

Calcula

$$\left| \int_0^3 s(x)\,dx \right| \qquad \text{y} \qquad \int_0^3 |s(x)|\,dx,$$

y verifica Corolario 14.2. Decide si hay igualdad o desigualdad estricta.

Ejercicio E5. Partes positiva y negativa. Para $f \in \mathcal{R}[a,b]$ define

$$f^+ = \max\{f, 0\}, \qquad f^- = \max\{-f, 0\}.$$

Demuestra las identidades puntuales

$$f = f^+ - f^-, \qquad |f| = f^+ + f^-,$$

y deduce que

$$\int_a^b f = \int_a^b f^+ - \int_a^b f^-.$$

Ejercicio E6. Acotar una integral con cotas puntuales. Sea $f \in \mathcal{R}[a,b]$ y supón que $m \le f(x) \le M$ para todo $x \in [a,b]$.

Demuestra que

$$m(b-a) \le \int_a^b f(x)\,dx \le M(b-a).$$

Usa solo monotonía de la integral y la integral de funciones constantes.

Nivel F — Aditividad, orientación, modificaciones finitas y contraejemplos. Ejercicio F1. Partir una integral en un punto que no era corte original. Sea

$$s(x) = \begin{cases} 2, & 0 \le x < 1, \\ -1, & 1 < x \le 3, \end{cases}$$

con valor arbitrario en $x = 1$. Calcula

$$\int_0^3 s, \qquad \int_0^{5/2} s, \qquad \int_{5/2}^3 s,$$

y verifica explícitamente Teorema 14.7.

Ejercicio F2. Invertir los límites. Conserva la función s del ejercicio F1.
Calcula

$$\int_3^0 s(x)\,dx \qquad \text{y} \qquad \int_{5/2}^0 s(x)\,dx$$

a partir de Definición 14.3, y comprueba que los signos son coherentes con los valores obtenidos en F1.

Ejercicio F3. Aditividad orientada en un orden no creciente. Sea f Riemann integrable en un intervalo que contiene $0, 1, 2$.
Demuestra, usando Proposición 14.3, que

$$\int_2^0 f = \int_2^1 f + \int_1^0 f.$$

Reescribe la identidad usando únicamente integrales con límites crecientes y verifica que ambas formas son equivalentes.

Ejercicio F4. Una modificación finita con integral conocida. Define en $[0, 1]$

$$g(x) = \begin{cases} 10, & x = \frac{1}{3}, \\ -4, & x = 1, \\ x, & \text{en otro caso.} \end{cases}$$

Demuestra que

$$\int_0^1 g(x)\,dx = \frac{1}{2}.$$

Explica por qué no es necesario calcular sumas de Darboux de g desde cero.

Ejercicio F5. Una mala mitad arruina el intervalo completo. Define

$$f(x) = \begin{cases} 1, & 0 \le x \le \frac{1}{2} \text{ y } x \in \mathbb{Q}, \\ 0, & 0 \le x \le \frac{1}{2} \text{ y } x \notin \mathbb{Q}, \\ 0, & \frac{1}{2} < x \le 1. \end{cases}$$

Demuestra que f no es Riemann integrable en $[0, 1]$ usando la localidad de Teorema 14.7.

Ejercicio F6. Una secuencia de mallas finas no sustituye el criterio de Darboux. Para la función d de Ejemplo 14.1, toma las particiones uniformes P_n.
Demuestra simultáneamente que

$$\|P_n\| \to 0, \qquad U(d, P_n) - L(d, P_n) = 1.$$

Identifica cuál de las dos cantidades es geométrica y cuál incorpora información sobre la oscilación de la función.

Nivel G — Desafíos integradores. Ejercicio G1. Tres puertas para la misma integrabilidad. Sea $f : [a, b] \to \mathbb{R}$ acotada. Supón que existen particiones P_n tales que $U(f, P_n) - L(f, P_n) \leq \frac{1}{n}$.

1. Demuestra que f es Riemann integrable.
2. Si $I = \int_a^b f$, explica por qué todas las sumas etiquetadas de malla suficientemente pequeña se acercan a I.
3. Señala qué resultado justifica cada paso y distingue definición primaria de caracterizaciones equivalentes.

Ejercicio G2. Elegir el criterio adecuado. En cada caso decide si la función es Riemann integrable y justifica la respuesta por la ruta más directa disponible en C15:

1. $f_1(x) = x^2$ en $[0, 1]$;
2. $f_2(x) = 0$ para $x < 1/2$ y $f_2(x) = 1$ para $x \geq 1/2$;
3. $f_3(x) = x$ salvo en $x = 1/3$, donde $f_3(1/3) = 100$;
4. $f_4(x) = 1$ en los racionales y 0 en los irracionales de $[0, 1]$.

No uses FTC, primitivas ni teoría de medida.

Ejercicio G3. Construir una función integrable fuera de las dos clases básicas. En $[-1, 1]$ define

$$s(x) = \begin{cases} 0, & x < 0, \\ 1, & x \geq 0, \end{cases} \qquad h(x) = |x| + s(x).$$

1. Demuestra que s y $|x|$ son Riemann integrables.
2. Deduce que h es Riemann integrable.
3. Demuestra que h es discontinua en 0.
4. Demuestra que h no es monótona en todo $[-1, 1]$.

Explica por qué este ejemplo confirma que continuidad y monotonía son criterios suficientes, no caracterizaciones.

Ejercicio G4. Auditoría final de cuatro argumentos. Para cada afirmación, decide si el argumento es válido. Si no lo es, identifica la laguna y formula una reparación correcta usando resultados de C15.

1. «$\|P_n\| \to 0$, luego f es integrable.»
2. «Una sucesión particular de sumas etiquetadas converge a I, luego $\int f = I$.»
3. «Como $\int (f + g) = \int f + \int g$, la suma de dos funciones integrables es integrable.»
4. «Una función continua es integrable porque el teorema fundamental del cálculo lo garantiza.»

La reparación debe respetar la arquitectura lógica del capítulo.

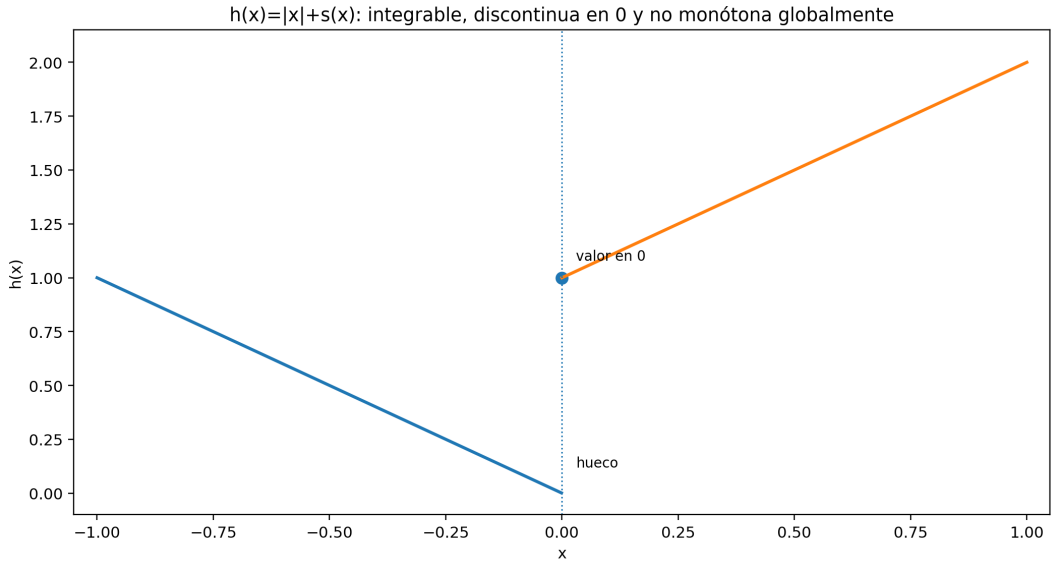


FIGURA 14.14. La función $h(x) = |x| + s(x)$ es integrable por construcción algebraica, pero es discontinua en 0 y no es monótona en todo $[-1, 1]$.

Soluciones desarrolladas.

Soluciones del nivel A. **Solución A1.** En toda pieza $[x_{k-1}, x_k]$ de una partición P, ínf $f =$ sup $f = c$.

Por tanto,

$$L(f, P) = \sum_{k=1}^{n} c\,\Delta x_k = c(b - a),$$

y exactamente lo mismo ocurre con la suma superior: $U(f, P) = c(b - a)$.

Así, $\mathcal{L}(f) = \mathcal{U}(f) = \{c(b - a)\}$.

Su supremo y su ínfimo son el mismo número:

$$\underline{\int_a^b} f = \overline{\int_a^b} f = c(b - a).$$

Por Definición 14.2, f es Riemann integrable y

$$\boxed{\int_a^b c\,dx = c(b - a).}$$

Solución A2. Como $f(x) = x$ es creciente, en

$$\left[\frac{k-1}{n}, \frac{k}{n}\right]$$

el ínfimo es $(k - 1)/n$ y el supremo es k/n. Como todas las piezas tienen longitud $1/n$,

$$L(f,P_n)=\frac{1}{n^2}\sum_{k=1}^{n}(k-1)=\frac{n(n-1)}{2n^2}=\frac{1}{2}-\frac{1}{2n},$$

mientras

$$U(f,P_n)=\frac{1}{n^2}\sum_{k=1}^{n}k=\frac{n(n+1)}{2n^2}=\frac{1}{2}+\frac{1}{2n}.$$

Por Proposición 14.1,

$$\frac{1}{2}-\frac{1}{2n}\le\underline{\int_0^1}f\le\overline{\int_0^1}f\le\frac{1}{2}+\frac{1}{2n}$$

para todo n.

Si la integral inferior fuese menor que $1/2$, la distancia positiva a $1/2$ sería superada por $1/(2n)$ para n suficientemente grande, contradiciendo la desigualdad izquierda. De manera análoga, la integral superior no puede ser mayor que $1/2$. Por tanto,

$$\boxed{\underline{\int_0^1}f=\overline{\int_0^1}f=\frac{1}{2}.}$$

Solución A3. Todo subintervalo no degenerado contiene racionales e irracionales. En cada pieza de cualquier partición P,

$$\operatorname{ínf} g=-1,\qquad \sup g=3.$$

Como la suma de las anchuras es 2, $L(g,P)=-1\cdot 2=-2$, y $U(g,P)=3\cdot 2=6$. Esto vale para toda partición, de modo que

$$\mathcal{L}(g)=\{-2\},\qquad \mathcal{U}(g)=\{6\}.$$

Por tanto,

$$\underline{\int_0^2}g=-2,\qquad \overline{\int_0^2}g=6.$$

Como los dos extremos globales no coinciden,

$$\boxed{g\notin\mathcal{R}[0,2].}$$

Solución A4. Por el encierro global de Proposición 14.1,

$$L(f,P)\le\underline{\int_a^b}f\le\overline{\int_a^b}f\le U(f,Q).$$

La hipótesis dice que los extremos de esta cadena son el mismo número I. Así,

$$I\le\underline{\int_a^b}f\le\overline{\int_a^b}f\le I.$$

Luego necesariamente

$$\underline{\int_a^b} f = \overline{\int_a^b} f = I.$$

Por Definición 14.2,

$$\boxed{f \in \mathcal{R}[a,b], \qquad \int_a^b f = I.}$$

No fue necesario comparar directamente P y Q: Proposición 14.1 ya incorpora la compatibilidad global entre particiones arbitrarias.

Solución A5. Como f es integrable, $I = \sup_P L(f,P) = \inf_P U(f,P)$.

Sea $\varepsilon > 0$. Si ninguna suma inferior superase $I-\varepsilon$, entonces $I-\varepsilon$ sería una cota superior de todas las sumas inferiores, contradiciendo que I es su supremo. Por tanto existe P_- con $L(f,P_-) > I-\varepsilon$. Además, toda suma inferior está por debajo de I, de modo que $I-\varepsilon < L(f,P_-) \le I$. Análogamente, si ninguna suma superior fuese menor que $I+\varepsilon$, ese número sería una cota inferior mayor que el ínfimo I. Existe entonces P_+ tal que $I \le U(f,P_+) < I+\varepsilon$. Las propiedades de supremo e ínfimo producen inicialmente dos particiones independientes; nada obliga a que coincidan. Cuando se necesita una sola partición que herede ambas aproximaciones se pasa al refinamiento común, como en la prueba de Teorema 14.1.

Solución A6. Fijemos una partición P. En cada subintervalo I_k, $\inf_{I_k} f \le \inf_{I_k} g$, porque $f \le g$ punto a punto. Multiplicando por $\Delta x_k > 0$ y sumando, $L(f,P) \le L(g,P)$.

Esto vale para toda partición P. Por tanto, $\sup_P L(f,P) \le \sup_P L(g,P)$, es decir,

$$\boxed{\underline{\int_a^b} f \le \underline{\int_a^b} g.}$$

De manera análoga, $\sup_{I_k} f \le \sup_{I_k} g$, de donde $U(f,P) \le U(g,P)$ para toda P. Tomando ínfimos,

$$\boxed{\overline{\int_a^b} f \le \overline{\int_a^b} g.}$$

La argumentación usa solamente orden y las definiciones globales; no presupone integrabilidad.

Soluciones del nivel B. **Solución B1.** Dado $\varepsilon > 0$, por la propiedad arquimediana existe un entero $n > \frac{1}{\varepsilon}$.

Por ejemplo, puede elegirse cualquier entero mayor que $1/\varepsilon$. Entonces $U(f,P_n) - L(f,P_n) = \frac{1}{n} < \varepsilon$. Como esto puede hacerse para toda tolerancia positiva, Teorema 14.1 implica

$$\boxed{f \in \mathcal{R}[0,1].}$$

El argumento certifica existencia de la integral, pero no calcula su valor.

Solución B2. Sea

$$P = \{-1 = x_0 < x_1 < \cdots < x_n = 1\}.$$

En una pieza $I_k = [x_{k-1}, x_k]$, para cualesquiera $u, v \in I_k$,

$$||u| - |v|| \leq |u - v| \leq \Delta x_k \leq \|P\|.$$

Por tanto la oscilación local satisface $\omega_k \leq \|P\|$. Usando la identidad de la brecha,

$$\begin{aligned} U(f,P) - L(f,P) &= \sum_{k=1}^{n} \omega_k \Delta x_k \\ &\leq \|P\| \sum_{k=1}^{n} \Delta x_k \\ &= 2\|P\|. \end{aligned}$$

Dado $\varepsilon > 0$, basta elegir una partición con $\|P\| < \frac{\varepsilon}{2}$, por ejemplo una partición uniforme suficientemente fina. Entonces la brecha es menor que ε, y Teorema 14.1 da

$$\boxed{|x| \in \mathcal{R}[-1,1].}$$

Solución B3. Fuera de las dos piezas adyacentes a $1/3$, la función es constantemente cero, de modo que la oscilación es nula.

En

$$\left[\frac{1}{3} - \delta, \frac{1}{3}\right] \quad \text{y} \quad \left[\frac{1}{3}, \frac{1}{3} + \delta\right]$$

aparecen los valores 0 y 5, así que la oscilación es 5 en cada una. Ambas tienen longitud δ. Por tanto,

$$U(s, P_\delta) - L(s, P_\delta) = 5\delta + 5\delta = 10\delta.$$

Dado $\varepsilon > 0$, elige

$$0 < \delta < \text{mín}\left\{\frac{1}{3}, \frac{\varepsilon}{10}\right\}.$$

Entonces $U(s, P_\delta) - L(s, P_\delta) < \varepsilon$.

Por Teorema 14.1,

$$\boxed{s \in \mathcal{R}[0,1].}$$

El cálculo cuantifica la idea de que una amplitud vertical fija puede confinarse en una anchura horizontal arbitrariamente pequeña.

Solución B4. $f(x) = x^3$ es creciente en $[0, 1]$. Por Teorema 14.5, toda partición P satisface

$$U(f, P) - L(f, P) \leq \|P\| \, |f(1) - f(0)|.$$

Como $|f(1) - f(0)| = 1$, obtenemos

$$\boxed{U(f, P) - L(f, P) \leq \|P\|.}$$

Dado $\varepsilon > 0$, elige un entero $n > 1/\varepsilon$ y toma la partición uniforme en n piezas. Su malla es $1/n < \varepsilon$, por lo que $U(f, P_n) - L(f, P_n) < \varepsilon$. El criterio de Darboux vuelve a confirmar la integrabilidad.

Solución B5. Sea $\varepsilon > 0$. Elige un entero $n > \frac{1}{\varepsilon}$.

Entonces $U(f, P_n) - L(f, P_n) \leq \frac{1}{n} < \varepsilon$.

Así, para toda tolerancia positiva existe una partición cuya brecha es menor que esa tolerancia. Esta es exactamente la condición de Teorema 14.1. Por tanto,

$$\boxed{f \in \mathcal{R}[a, b].}$$

Solución B6. En cada subintervalo no degenerado hay racionales e irracionales. Por tanto, para cualquier partición P,

$$\text{ínf}\, d_c = 0, \qquad \sup d_c = c$$

en cada pieza. Como la suma de las anchuras es 1,

$$L(d_c, P) = 0, \qquad U(d_c, P) = c.$$

Luego

$$\boxed{U(d_c, P) - L(d_c, P) = c}$$

para toda partición.

Si $c = 10^{-9}$, una brecha de ese tamaño es numéricamente pequeña, pero el criterio exige responder a **toda** tolerancia. Tomando, por ejemplo, $\varepsilon = \frac{c}{2}$, ninguna partición satisface $U - L < \varepsilon$. Por eso una brecha positiva fija, aunque sea diminuta, no prueba integrabilidad.

Soluciones del nivel C. **Solución C1.** Ya sabemos por Ejemplo 14.2 que

$$\int_0^1 x \, dx = \frac{1}{2}.$$

Para una partición arbitraria P, la función identidad es creciente y Teorema 14.5 da

$$U(f, P) - L(f, P) \leq \|P\| \, |f(1) - f(0)| = \|P\|.$$

Tanto la suma etiquetada como la integral están encerradas entre $L(f, P)$ y $U(f, P)$. Por tanto,

$$\left| R(f; P, \xi) - \frac{1}{2} \right| \leq U(f, P) - L(f, P) \leq \|P\|.$$

Dado $\varepsilon > 0$, si $\|P\| < \delta := \varepsilon$, entonces

$$\left| R(f;P,\xi) - \frac{1}{2} \right| < \varepsilon$$

para **toda** elección de etiquetas. Esto verifica la condición uniforme de Teorema 14.2.

Solución C2. Cada pieza de P_n es un intervalo no degenerado y contiene racionales e irracionales.

Si elegimos una etiqueta racional en cada pieza, entonces cada valor de d es 1, de modo que

$$R(d;P_n,\xi^{(n)}) = \sum_{k=1}^{n} 1 \cdot \frac{1}{n} = 1.$$

Si elegimos etiquetas irracionales,

$$R(d;P_n,\eta^{(n)}) = \sum_{k=1}^{n} 0 \cdot \frac{1}{n} = 0.$$

Las dos familias usan las **mismas** particiones, cuyas mallas tienden a cero, pero sus sumas permanecen separadas por 1. Si existiese un número I que satisficiera Teorema 14.2, tomando por ejemplo $\varepsilon = 1/3$ todas las sumas de malla suficientemente pequeña deberían quedar a distancia menor que $1/3$ de I. Dos de ellas no podrían diferir en 1. Contradicción. Por tanto no existe tal I.

Solución C3. Las sumas solo pueden diferir en aquellas piezas cuya etiqueta sea exactamente c.

Si c está en el interior de una pieza, puede ser etiqueta únicamente de esa pieza. Si c es un punto de corte interior, puede pertenecer a las dos piezas adyacentes y, por tanto, puede aparecer como etiqueta en ambas. En cualquier caso intervienen a lo sumo dos subintervalos. Cada uno tiene longitud a lo sumo $\|P\|$. En una pieza etiquetada en c, la diferencia de contribuciones tiene valor absoluto $D\,\Delta x_k$.

Por la desigualdad triangular,

$$|R(g;P,\xi) - R(f;P,\xi)| \le D(\Delta x_{k_1} + \Delta x_{k_2}) \le 2D\|P\|.$$

El factor 2 aparece porque un punto de corte puede ser utilizado como etiqueta de las dos piezas que se encuentran en él.

Solución C4. Sea $\varepsilon > 0$. Aplicando la condición uniforme a I con tolerancia $\varepsilon/2$, existe $\delta_I > 0$ tal que toda suma etiquetada con malla menor que δ_I satisface $|R - I| < \frac{\varepsilon}{2}$.

De manera análoga existe $\delta_J > 0$ con $|R - J| < \frac{\varepsilon}{2}$ para toda suma de malla menor que δ_J. Elige cualquier partición etiquetada con malla menor que $\min\{\delta_I, \delta_J\}$.

Entonces $|I - J| \le |I - R| + |R - J| < \varepsilon$.

Como esto vale para todo $\varepsilon > 0$, necesariamente

$$\boxed{I = J.}$$

La unicidad se obtiene únicamente de la propiedad uniforme y la desigualdad triangular.

Solución C5. Sea

$$I = \int_a^b f.$$

Aplicamos Teorema 14.2 con tolerancia $\varepsilon/2$. Existe $\delta > 0$ tal que, para toda partición etiquetada de malla menor que δ, $|R(f; P, \xi) - I| < \frac{\varepsilon}{2}$. Si (P, ξ) y (Q, η) satisfacen ambas condiciones de malla, entonces

$$\begin{aligned} |R(f; P, \xi) - R(f; Q, \eta)| &\leq |R(f; P, \xi) - I| + |I - R(f; Q, \eta)| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon. \end{aligned}$$

Así, la integrabilidad hace que **todas** las sumas suficientemente finas sean mutuamente cercanas, aunque provengan de particiones y etiquetas diferentes.

Solución C6. El argumento solo controla una familia escogida: $(P_n, \xi^{(n)})$.

No dice nada sobre otras particiones de malla comparable ni sobre otras etiquetas. La función racional/irracional muestra que una familia particular puede converger aunque la función no sea integrable. Una hipótesis suficiente es la condición uniforme completa:

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall (P, \xi) : \|P\| < \delta \Longrightarrow |R(f; P, \xi) - 7| < \varepsilon.$$

Si esta propiedad se cumple, Teorema 14.2 permite concluir

$$\boxed{f \in \mathcal{R}[a, b], \qquad \int_a^b f = 7.}$$

Lo que faltaba era reemplazar el control de **una** secuencia por un control uniforme sobre **todas** las particiones etiquetadas suficientemente finas.

Soluciones del nivel D. **Solución D1.** La partición natural $P = \{0, 1, 3\}$ representa a s. Las alturas sobre los intervalos abiertos son 2 y -1, con anchuras 1 y 2. Por tanto, la integral elemental es $2(1) + (-1)(2) = 0$.

Teorema 14.3 afirma que toda escalonada es Riemann integrable y que ambas construcciones coinciden. Así,

$$\boxed{\int_0^3 s(x)\, dx = 0.}$$

El valor $s(1)$ se encuentra en un punto de corte. Cambiarlo no modifica la integral elemental, y Corolario 14.1 confirma además que una modificación puntual tampoco cambia la integral de Riemann.

Solución D2. La función identidad $f(x) = x$ es Riemann integrable en $[0, 1]$ y Ejemplo 14.2 da

$$\int_0^1 x\, dx = \frac{1}{2}.$$

g difiere de f únicamente en el conjunto finito

$$\left\{\frac{1}{4}, \frac{3}{4}\right\}.$$

Por Corolario 14.1, las dos funciones tienen la misma integrabilidad y el mismo valor integral. Por tanto,

$$\boxed{g \in \mathcal{R}[0,1], \qquad \int_0^1 g(x)\,dx = \frac{1}{2}.}$$

La magnitud de las dos modificaciones puntuales no altera la conclusión; lo decisivo es que son finitísimas.

Solución D3. La función valor absoluto es continua en $\mathbb{R}$ porque $||x| - |y|| \leq |x - y|$.
En particular es continua en el intervalo compacto $[-1, 1]$. Por Teorema 14.4,

$$\boxed{|x| \in \mathcal{R}[-1,1].}$$

No es monótona en todo el intervalo: por ejemplo, $-1 < 0 < 1$, pero

$$|-1| = 1 > |0| = 0 < |1| = 1.$$

En $[-1, 0]$ decrece y en $[0, 1]$ crece. Por tanto Teorema 14.5 no proporciona una única aplicación global sobre $[-1, 1]$; tampoco hace falta, porque la continuidad ya basta.

Solución D4. Si $x < y$, los valores posibles siguen el orden $0 \leq 1$, así que h es no decreciente.
En $x = 1/2$,

$$h\left(\frac{1}{2}\right) = 1,$$

mientras que por la izquierda los valores son constantemente 0. Por tanto la función es discontinua en ese punto. Sin embargo, por Teorema 14.5 toda función monótona en un intervalo cerrado es Riemann integrable. Además h es escalonada: tiene altura 0 en $(0, 1/2)$ y altura 1 en $(1/2, 1)$. Su integral elemental es

$$0\left(\frac{1}{2}\right) + 1\left(\frac{1}{2}\right) = \frac{1}{2}.$$

Por Teorema 14.3,

$$\boxed{\int_0^1 h(x)\,dx = \frac{1}{2}.}$$

Solución D5. Ejemplo 14.1 demuestra que $d \notin \mathcal{R}[0, 1]$.

g coincide con d salvo en un conjunto finito. Corolario 14.1 establece la equivalencia $d \in \mathcal{R}[0, 1] \iff g \in \mathcal{R}[0, 1]$. Como la afirmación de la izquierda es falsa, también lo es la de la derecha:

$$\boxed{g \notin \mathcal{R}[0,1].}$$

La invariancia ante modificaciones finitas funciona en ambas direcciones: no solo preserva la integrabilidad cuando existe, sino también la no integrabilidad dentro de una clase de equivalencia finita.

Solución D6. Para el primer ejemplo podemos tomar

$$f(x) = |x| \qquad (-1 \le x \le 1).$$

Es continua, luego integrable por Teorema 14.4, pero no es monótona en todo $[-1, 1]$. Para el segundo podemos tomar la función escalón

$$h(x) = \begin{cases} 0, & 0 \le x < 1/2, \\ 1, & 1/2 \le x \le 1. \end{cases}$$

Es no decreciente, luego integrable por Teorema 14.5, pero es discontinua en $1/2$. Así, las clases «continua» y «monótona» proporcionan rutas suficientes distintas. Ninguna de las dos contiene a la otra dentro de las funciones del capítulo.

Soluciones del nivel E. **Solución E1.** El orden lógico correcto está contenido en Teorema 14.6.

Primero, el teorema establece que si $f, g \in \mathcal{R}[a, b]$, entonces para cualesquiera escalares α, β también $\alpha f + \beta g \in \mathcal{R}[a, b]$. Solo después de haber garantizado que esa integral existe podemos escribir

$$\boxed{\int_a^b (\alpha f + \beta g) = \alpha \int_a^b f + \beta \int_a^b g.}$$

El argumento «la fórmula es lineal, luego la función es integrable» es circular porque el miembro izquierdo presupone que la integral de $\alpha f + \beta g$ ya está definida.

Solución E2. $f(x) = x$ es continua, por lo que $f \in \mathcal{R}[-1, 1]$ por Teorema 14.4.

$g(x) = |x|$ también es continua, así que $g \in \mathcal{R}[-1, 1]$.

Proposición 14.2 demuestra que el producto de dos funciones Riemann integrables es Riemann integrable. Por tanto,

$$\boxed{x|x| \in \mathcal{R}[-1, 1].}$$

No es necesario disponer todavía de una técnica para evaluar su integral.

Solución E3. Como f, g son integrables, Teorema 14.6 implica que $f - g$ es integrable. Por Proposición 14.2 también lo es $|f - g|$.

Otra aplicación de la linealidad muestra que $f + g + |f - g|$ y $f + g - |f - g|$ son integrables. Multiplicar por $1/2$ preserva la integrabilidad. Usando las identidades puntuales dadas,

$$\boxed{\max\{f, g\},\ \min\{f, g\} \in \mathcal{R}[a, b].}$$

Solución E4. La integral de s es

$$\int_0^3 s = 2(1) + (-1)(2) = 0.$$

Luego

$$\left|\int_0^3 s\right| = 0.$$

En cambio,

$$|s(x)| = \begin{cases} 2, & 0 \le x < 1, \\ 1, & 1 \le x \le 3, \end{cases}$$

salvo el valor de corte irrelevante. Por tanto,

$$\int_0^3 |s| = 2(1) + 1(2) = 4.$$

Así,

$$\boxed{\left|\int_0^3 s\right| = 0 < 4 = \int_0^3 |s|.}$$

La desigualdad de Corolario 14.2 es estricta porque las contribuciones positiva y negativa se cancelan en $\int s$, pero no en $\int |s|$.

Solución E5. Si $f(x) \ge 0$, entonces

$$f^+(x) = f(x), \qquad f^-(x) = 0.$$

Si $f(x) < 0$, entonces

$$f^+(x) = 0, \qquad f^-(x) = -f(x).$$

En ambos casos, $f = f^+ - f^-$.

De manera análoga, $|f| = f^+ + f^-$.

Proposición 14.2 garantiza que f^+ y f^- son integrables. Aplicando linealidad a la primera identidad,

$$\boxed{\int_a^b f = \int_a^b f^+ - \int_a^b f^-.}$$

La segunda identidad también produce

$$\int_a^b |f| = \int_a^b f^+ + \int_a^b f^-.$$

Solución E6. Las funciones constantes

$$x \mapsto m \qquad \text{y} \qquad x \mapsto M$$

son Riemann integrables, con integrales

$$m(b-a) \qquad \text{y} \qquad M(b-a).$$

La hipótesis puntual es $m \le f \le M$.
Por la monotonía de la integral en Teorema 14.6,

$$\int_a^b m\,dx \le \int_a^b f(x)\,dx \le \int_a^b M\,dx.$$

Sustituyendo los valores de las integrales constantes,

$$\boxed{m(b-a) \le \int_a^b f \le M(b-a).}$$

Soluciones del nivel F. **Solución F1.** La integral total es

$$\int_0^3 s = 2(1) + (-1)(2) = 0.$$

En $[0, 5/2]$ aparece la primera franja completa y una longitud $3/2$ de la segunda:

$$\int_0^{5/2} s = 2(1) + (-1)\left(\frac{3}{2}\right) = \frac{1}{2}.$$

En $[5/2, 3]$ solo queda una longitud $1/2$ con altura -1:

$$\int_{5/2}^3 s = -\frac{1}{2}.$$

Por tanto,

$$\boxed{\int_0^3 s = 0 = \frac{1}{2} - \frac{1}{2} = \int_0^{5/2} s + \int_{5/2}^3 s.}$$

Esto verifica concretamente la aditividad de Teorema 14.7, aunque $5/2$ no fuese un corte de la representación original.

Solución F2. Por Definición 14.3,

$$\int_3^0 s = -\int_0^3 s = -0 = 0.$$

También,

$$\int_{5/2}^0 s = -\int_0^{5/2} s = -\frac{1}{2}.$$

Así,

$$\boxed{\int_3^0 s = 0, \qquad \int_{5/2}^0 s = -\frac{1}{2}.}$$

La inversión de límites no cambia la magnitud del valor asociado al mismo tramo; cambia únicamente su signo.

Solución F3. Proposición 14.3 vale para tres puntos en cualquier orden. Tomando

$$r = 2, \qquad s = 1, \qquad t = 0,$$

obtenemos directamente

$$\boxed{\int_2^0 f = \int_2^1 f + \int_1^0 f.}$$

Usando la convención de orientación,

$$\int_2^0 f = -\int_0^2 f, \qquad \int_2^1 f = -\int_1^2 f, \qquad \int_1^0 f = -\int_0^1 f.$$

Multiplicando la identidad por -1 queda

$$\int_0^2 f = \int_0^1 f + \int_1^2 f,$$

que es la forma habitual con límites crecientes. Las dos expresiones son exactamente equivalentes.

Solución F4. La función identidad $f(x) = x$ es Riemann integrable en $[0, 1]$ y

$$\int_0^1 f(x)\,dx = \frac{1}{2}.$$

g difiere de f únicamente en los puntos

$$\frac{1}{3} \qquad \text{y} \qquad 1.$$

Por Corolario 14.1, las modificaciones finitas preservan tanto integrabilidad como valor integral. Así,

$$\boxed{\int_0^1 g(x)\,dx = \frac{1}{2}.}$$

No hace falta recomenzar con sumas inferiores y superiores porque el corolario ya encapsula exactamente el efecto de esas excepciones.

Solución F5. La restricción de f al subintervalo $[0,1/2]$ es la función racional/irracional con valores 1 y 0 sobre ese intervalo.

En toda partición de $[0,1/2]$, cada pieza contiene racionales e irracionales. Por tanto,

$$L = 0, \qquad U = \frac{1}{2},$$

para toda partición, y la restricción no es Riemann integrable.

Teorema 14.7 afirma que si f fuese integrable en $[0,1]$, entonces sus restricciones a $[0,1/2]$ y $[1/2,1]$ también serían integrables. La primera no lo es. Por contraposición,

$$\boxed{f \notin \mathcal{R}[0,1].}$$

Solución F6. La partición uniforme en n piezas tiene malla $\|P_n\| = \frac{1}{n} \longrightarrow 0$.

Pero en cada pieza aparecen racionales e irracionales, de modo que

$$\text{ínf}\, d = 0, \qquad \sup d = 1.$$

Por tanto,

$$L(d, P_n) = 0, \qquad U(d, P_n) = 1,$$

y

$$\boxed{U(d, P_n) - L(d, P_n) = 1}$$

para todo n.

La malla es una propiedad **geométrica de la partición**: mide la longitud de su pieza más grande. La brecha incorpora además la **oscilación de la función** dentro de las piezas. Hacer fina la geometría no obliga a que una función arbitraria oscile menos.

Soluciones del nivel G. **Solución G1.** Para la primera parte, sea $\varepsilon > 0$. Elige un entero $n > \frac{1}{\varepsilon}$.

Entonces $U(f, P_n) - L(f, P_n) \leq \frac{1}{n} < \varepsilon$.

Por Teorema 14.1, $f \in \mathcal{R}[a,b]$. Esta conclusión utiliza la caracterización de Darboux, cuya definición primaria subyacente sigue siendo Definición 14.2.

Escribamos

$$I = \int_a^b f.$$

Como f ya es integrable, Teorema 14.2 garantiza que para todo $\varepsilon > 0$ existe $\delta > 0$ tal que $\|P\| < \delta \Longrightarrow |R(f; P, \xi) - I| < \varepsilon$ para toda elección de partición y etiquetas.

La cadena lógica es

$$\boxed{\text{brechas } 1/n \to 0 \overset{\text{Darboux}}{\Longrightarrow} \text{integrabilidad} \overset{\text{sumas etiquetadas}}{\Longrightarrow} \text{control uniforme de sumas etiquetadas.}}$$

La definición primaria es la coincidencia de integrales inferior y superior; los otros dos pasos son teoremas equivalentes o consecuencias equivalentes ya demostradas.

Solución G2.

1. $f_1(x) = x^2$ es continua en $[0,1]$. Por Teorema 14.4 es Riemann integrable. También podría usarse la brecha explícita de Ejemplo 14.3.
2. f_2 es no decreciente en $[0,1]$, aunque tiene un salto en $1/2$. Por Teorema 14.5 es Riemann integrable; además es escalonada.
3. f_3 difiere de la identidad solo en el punto $1/3$. La identidad es integrable, de modo que Corolario 14.1 implica que f_3 es integrable con el mismo valor.
4. f_4 es la función de Ejemplo 14.1. Para toda partición,

$$L(f_4, P) = 0, \qquad U(f_4, P) = 1.$$

Por tanto no es Riemann integrable.

Cada caso se resuelve con una herramienta interna de C15; no hace falta relacionar la integral con primitivas ni introducir teoría de medida.

Solución G3. s es una función escalonada, por lo que Teorema 14.3 da $s \in \mathcal{R}[-1,1]$.

$|x|$ es continua en $[-1,1]$, así que Teorema 14.4 da $|x| \in \mathcal{R}[-1,1]$. Por linealidad y cierre de Teorema 14.6,

$$\boxed{h = |x| + s \in \mathcal{R}[-1,1].}$$

En 0, $h(0) = 1$, mientras que si $x \to 0^-$, $h(x) = |x| \to 0$.

Luego h es discontinua en 0.

Tampoco es monótona en todo el intervalo. Por ejemplo, $-1 < -\frac{1}{2} < 0$, pero

$$h(-1) = 1, \qquad h\left(-\frac{1}{2}\right) = \frac{1}{2}, \qquad h(0) = 1.$$

Primero disminuye y luego aumenta. Así, h no es monótona globalmente. La integrabilidad se obtuvo por la estructura algebraica de $\mathcal{R}[-1,1]$, mostrando de nuevo que continuidad y monotonía son rutas suficientes, no descripciones exhaustivas de la clase.

Solución G4.

1. «$\|P_n\| \to 0$, luego f es integrable.» Es inválido. La malla por sí sola no controla la oscilación. Ejemplo 14.1 con particiones uniformes da un contraejemplo. Una reparación correcta es demostrar además $U(f, P_n) - L(f, P_n) \to 0$, y aplicar Teorema 14.1.

2. «Una sucesión particular de sumas etiquetadas converge a I, luego $\int f = I$.» También es inválido. Solo se han controlado esas elecciones. La reparación es probar

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall (P, \xi) : \|P\| < \delta \Longrightarrow |R(f; P, \xi) - I| < \varepsilon,$$

y usar Teorema 14.2.

3. «Como $\int (f+g) = \int f + \int g$, la suma de dos funciones integrables es integrable.» El orden lógico es circular: el miembro izquierdo presupone que $f + g$ ya tiene integral. Teorema 14.6 demuestra primero el cierre de $\mathcal{R}[a,b]$ bajo sumas y luego la fórmula de linealidad.

4. «Una función continua es integrable porque el teorema fundamental del cálculo lo garantiza.» Aunque la conclusión sea verdadera, esa ruta viola la arquitectura del capítulo y usaría un resultado posterior. La reparación canónica es Teorema 14.4: continuidad en $[a,b]$

implica continuidad uniforme; las oscilaciones locales se hacen pequeñas, la brecha de Darboux se hace arbitrariamente pequeña y Teorema 14.1 concluye integrabilidad.

En los cuatro casos la corrección consiste en respetar la dirección lógica de las dependencias construidas dentro de C15.

Capítulo 15

Teoremas de valor medio para integrales

15.1. De una integral a una altura media

En el capítulo anterior construimos la integral de Riemann como un número asociado a una función integrable y a un intervalo. Ese número conserva información acumulativa sobre toda la función: incluso cuando conocemos

$$\int_a^b f(x)\,dx,$$

no hemos identificado todavía una altura particular de la gráfica que pueda representar por sí sola ese valor. Si $a < b$, dividir la integral por la longitud $b - a$ produce un número con la escala vertical de los valores de f. Podemos interpretarlo como la altura que, distribuida uniformemente sobre el intervalo, reproduce el valor integral.

Definición 15.1. **Definición (promedio integral).** Sea $f \in \mathcal{R}[a, b]$, con $a < b$. El **promedio integral** de f sobre $[a, b]$ es

$$\boxed{\operatorname{Avg}_{[a,b]}(f) := \frac{1}{b-a}\int_a^b f(x)\,dx.}$$

Para definir el promedio basta la integrabilidad de Riemann. Para asegurar que alguna imagen de la función coincide con ese número necesitaremos una hipótesis adicional; la continuidad permitirá obtenerla mediante el valor intermedio. La identidad que define el promedio puede escribirse también como

$$\int_a^b f(x)\,dx = (b-a)\operatorname{Avg}_{[a,b]}(f).$$

Esta forma explica la expresión *altura media*: el factor $b - a$ aporta la longitud horizontal y $\operatorname{Avg}_{[a,b]}(f)$ aporta una altura. Cuando $f \geq 0$ y la integral admite la lectura geométrica usual, el rectángulo de base $[a, b]$ y altura $\operatorname{Avg}_{[a,b]}(f)$ tiene el mismo contenido acumulado que la región representada por la integral. Para funciones que cambian de signo, la identidad algebraica sigue siendo válida, aunque la interpretación como área geométrica ordinaria ya no debe tomarse literalmente.

La figura Figura 15.1 representa el promedio como una **altura numérica**. La existencia de un punto de la gráfica a esa altura requiere una demostración aparte.

Ejemplo 15.1. **Dos calibraciones básicas del promedio.**

1. Si $f(x) = c$ en todo $[a, b]$, en §14.2 se obtuvo

$$\int_a^b c\,dx = c(b-a).$$

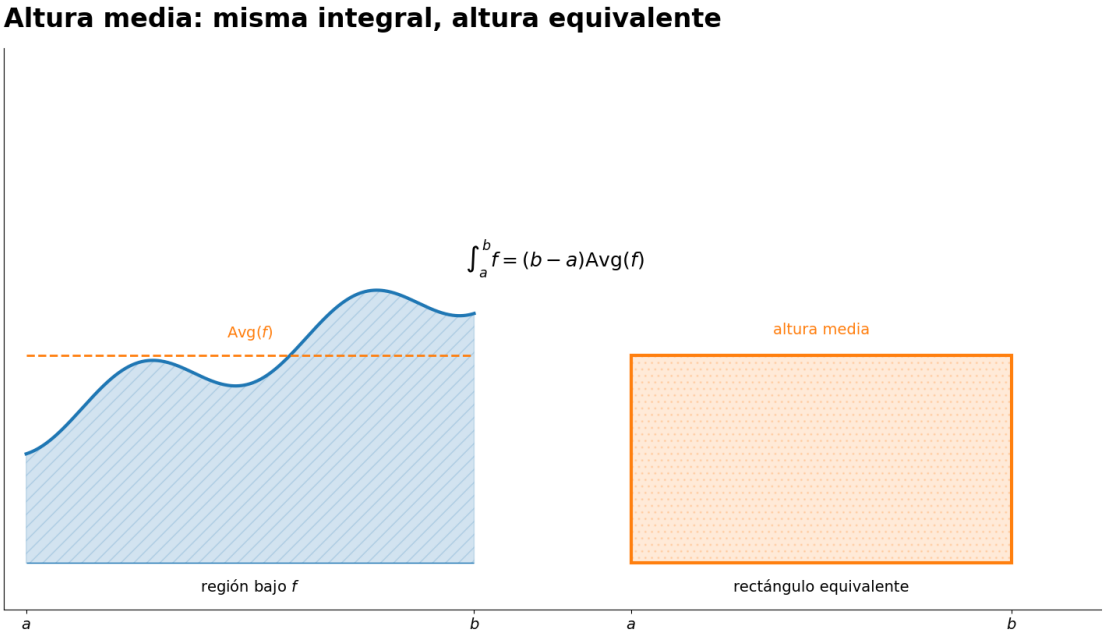


FIGURA 15.1. El promedio integral actúa como una altura media: para una función no negativa, el rectángulo de base $[a,b]$ y altura $\mathrm{Avg}_{[a,b]}(f)$ representa el mismo contenido integral.

Por tanto, $\mathrm{Avg}_{[a,b]}(f) = c$.

Una función constante tiene como promedio exactamente su única altura.

2. Para la identidad $f(x) = x$ en $[0,1]$, Ejemplo 14.2 demostró

$$\int_0^1 x\,dx = \frac{1}{2}.$$

Como el intervalo tiene longitud 1, $\mathrm{Avg}_{[0,1]}(f) = \frac{1}{2}$.

En este caso el promedio coincide con el valor $f(1/2)$, pero esa coincidencia es todavía un **hecho del ejemplo**, no un teorema general.

Un número no es todavía un punto. Conviene separar desde ahora dos preguntas distintas. La primera es puramente numérica: ¿cuánto vale $\mathrm{Avg}_{[a,b]}(f)$?

La segunda es una pregunta de existencia:

$$\text{¿existe } c \in [a,b] \text{ tal que } f(c) = \mathrm{Avg}_{[a,b]}(f)?$$

La primera pregunta tiene respuesta tan pronto como f es Riemann integrable. La segunda requiere hipótesis adicionales: la integrabilidad, por sí sola, no autoriza a afirmar que el promedio pertenece a la imagen de f. Empezaremos por localizar el promedio respecto de las cotas de la función. Esa comparación permitirá distinguir la estimación numérica de la afirmación de que el promedio se alcanza en algún punto.

15.2. El promedio queda entre las cotas de la función

La definición de promedio integral convierte la integral en una altura, pero todavía no indica dónde queda situada esa altura. El orden ya desarrollado para la integral de Riemann proporciona

inmediatamente la primera localización general. Supongamos que $f \in \mathcal{R}[a,b]$, con $a < b$, y que dos números reales m y M satisfacen

$$m \le f(x) \le M \qquad (x \in [a,b]).$$

Las funciones constantes m y M son integrables y, por el cálculo ya establecido en C15,

$$\int_a^b m\,dx = m(b-a), \qquad \int_a^b M\,dx = M(b-a).$$

La monotonía de la integral en Teorema 14.6 conserva las desigualdades punto a punto al integrar. De aquí surge el encierro básico del capítulo.

Proposición 15.1. **Proposición (encierro del promedio integral).** Sea $f \in \mathcal{R}[a,b]$, con $a < b$. Si

$$m \le f(x) \le M \qquad \text{para todo } x \in [a,b],$$

entonces

$$\boxed{m \le \operatorname{Avg}_{[a,b]}(f) \le M.}$$

En particular, como toda función Riemann-integrable es acotada,

$$\text{ínf}\, f([a,b]) \le \operatorname{Avg}_{[a,b]}(f) \le \sup f([a,b]).$$

Demostración. De $m \le f \le M$ y de la monotonía de Teorema 14.6 obtenemos

$$\int_a^b m\,dx \le \int_a^b f(x)\,dx \le \int_a^b M\,dx.$$

Como las integrales de las funciones constantes valen $m(b-a)$ y $M(b-a)$,

$$m(b-a) \le \int_a^b f(x)\,dx \le M(b-a).$$

Finalmente, $a < b$, de modo que $b - a > 0$. Dividir toda la cadena por ese número positivo conserva el sentido de las desigualdades y, usando Definición 15.1, da

$$m \le \frac{1}{b-a}\int_a^b f(x)\,dx \le M,$$

es decir, $m \le \operatorname{Avg}_{[a,b]}(f) \le M$.

Tomando como cotas el ínfimo y el supremo de la imagen se obtiene la última afirmación. □

La distinción entre **encerrar** y **realizar** el promedio queda visible en Figura 15.2.

Qué afirma —y qué no afirma— el encierro. La proposición no necesita continuidad y tampoco exige que las cotas m y M sean alcanzadas. Su contenido es puramente numérico: el promedio integral no puede quedar por debajo de una cota inferior de la función ni por encima de una cota superior. Dos consecuencias inmediatas son útiles como controles de signo:

- si $f \ge 0$, entonces $\operatorname{Avg}_{[a,b]}(f) \ge 0$;
- si $f \le 0$, entonces $\operatorname{Avg}_{[a,b]}(f) \le 0$.

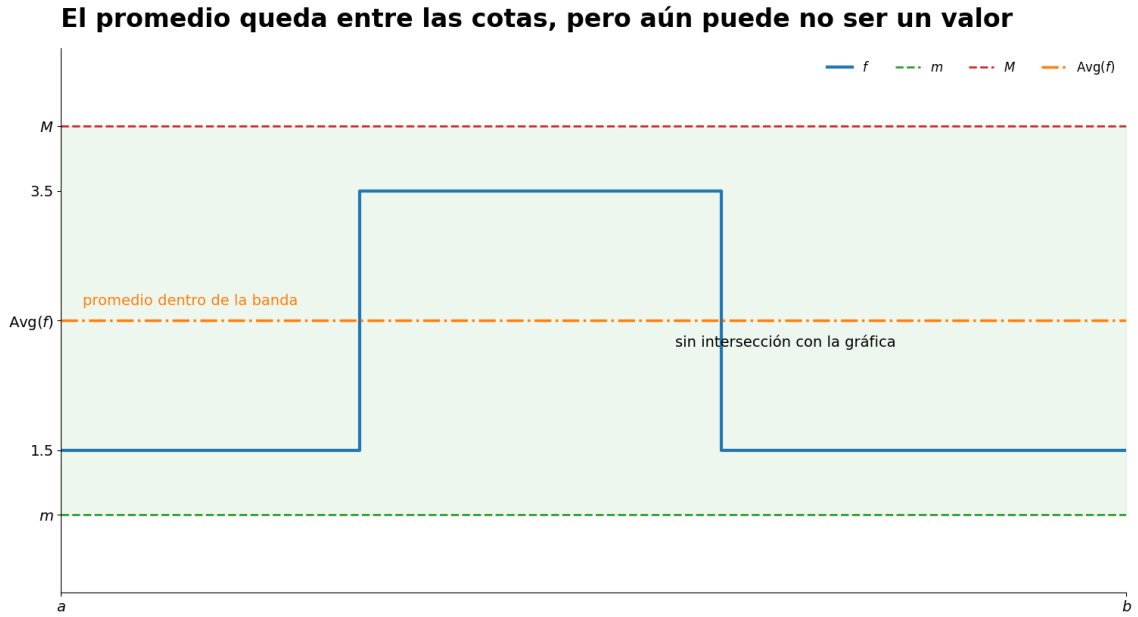


FIGURA 15.2. El encierro $m \le \mathrm{Avg}_{[a,b]}(f) \le M$ localiza el promedio entre dos niveles horizontales, sin garantizar por sí solo que ese nivel sea alcanzado por la función.

Pero todavía no podemos pasar de $\mathrm{Avg}_{[a,b]}(f) \in [m, M]$ a la existencia de un punto c con $f(c) = \mathrm{Avg}_{[a,b]}(f)$. Incluso si usamos las cotas óptimas $\inf f([a,b])$ y $\sup f([a,b])$, el encierro solo localiza un **número**. Para convertir desigualdades no estrictas en una herramienta capaz de forzar un valor intermedio necesitaremos primero una forma de **positividad estricta** de la integral bajo continuidad. Esa es la tarea exclusiva de §15.3.

15.3. Positividad estricta y comparación estricta

La monotonía de la integral solo proporciona desigualdades no estrictas. Si $h \ge 0$, sabemos que

$$\int_a^b h(x)\,dx \ge 0,$$

pero esa información no distingue todavía entre una función nula y una función que sea positiva en alguna parte. La continuidad permite dar ese paso.

LEMA 15.1. **Lema (positividad estricta de la integral).** Sea $h : [a,b] \to \mathbb{R}$ continua, con $a < b$. Si

$$h(x) \ge 0 \qquad (x \in [a,b])$$

y h no es idénticamente nula, entonces

$$\boxed{\int_a^b h(x)\,dx > 0.}$$

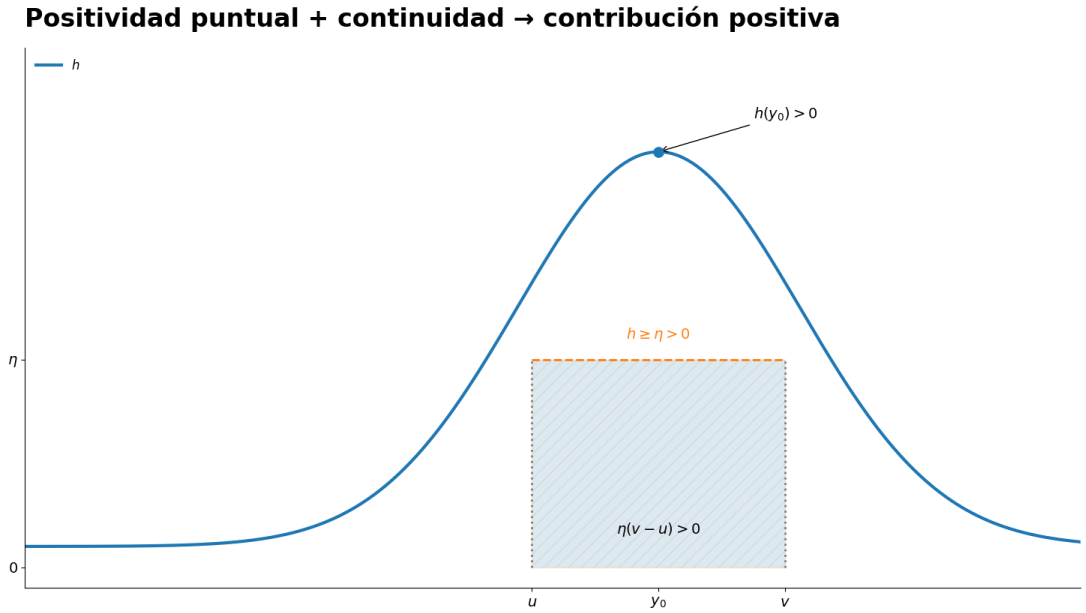


FIGURA 15.3. La continuidad ensancha la positividad puntual $h(y_0) > 0$ a un subintervalo $[u, v]$ donde $h \geq \eta > 0$, lo que produce una contribución integral estrictamente positiva.

Demostración. Como $h \not\equiv 0$ y $h \geq 0$, existe $x_0 \in [a, b]$ tal que $h(x_0) > 0$.

Podemos elegir un punto interior $y_0 \in (a, b)$ con $h(y_0) > 0$. Si x_0 ya es interior, basta tomar $y_0 = x_0$. Si $x_0 = a$ o $x_0 = b$, la continuidad relativa en ese extremo asegura que h sigue siendo positiva en puntos interiores suficientemente próximos.

Fijemos entonces $\eta := \frac{h(y_0)}{2} > 0$.

Por continuidad en y_0, existe $\delta > 0$ tal que

$$|x - y_0| < \delta \quad \Longrightarrow \quad h(x) > \eta.$$

Como $y_0 \in (a, b)$, podemos escoger $\rho > 0$ suficientemente pequeño para que

$$0 < \rho < \delta, \qquad a < y_0 - \rho < y_0 + \rho < b.$$

Definamos

$$u := y_0 - \rho, \qquad v := y_0 + \rho.$$

Entonces

$$h(x) \geq \eta \qquad (x \in [u, v]).$$

El paso geométrico decisivo de la prueba se resume en Figura 15.3. La continuidad de h implica su integrabilidad por Teorema 14.4. En $[u, v]$, la monotonía de Teorema 14.6 da

$$\int_u^v h(x)\,dx \geq \int_u^v \eta\,dx = \eta(v - u) > 0.$$

Por otra parte, como $h \geq 0$,

$$\int_a^u h(x)\,dx \geq 0, \qquad \int_v^b h(x)\,dx \geq 0.$$

Aplicando dos veces la aditividad de Teorema 14.7,

$$\begin{aligned}\int_a^b h(x)\,dx &= \int_a^u h(x)\,dx + \int_u^v h(x)\,dx + \int_v^b h(x)\,dx \\ &> 0.\end{aligned}$$

Esto prueba la afirmación. □

El argumento muestra con precisión dónde entra la continuidad: una desigualdad estricta en **un punto** se ensancha a un subintervalo de longitud positiva. Sin ese ensanchamiento, una modificación aislada puede no alterar la integral.

Corolario 15.1. **Corolario (consecuencias de la positividad estricta).** Sean $a < b$.

1. Si $f, g : [a,b] \to \mathbb{R}$ son continuas,
$$f(x) \leq g(x) \qquad (x \in [a,b]),$$
y existe $x_0 \in [a,b]$ tal que
$$f(x_0) < g(x_0),$$
entonces
$$\boxed{\int_a^b f(x)\,dx < \int_a^b g(x)\,dx.}$$
2. Si $h : [a,b] \to \mathbb{R}$ es continua y $h \geq 0$, entonces
$$\int_a^b h(x)\,dx = 0 \quad \Longleftrightarrow \quad h \equiv 0.$$
3. Si $f : [a,b] \to \mathbb{R}$ es continua y no constante, y
$$m := \min_{[a,b]} f, \qquad M := \max_{[a,b]} f,$$
entonces
$$\boxed{m < \operatorname{Avg}_{[a,b]}(f) < M.}$$

Demostración.

1. Sea $h := g - f$. La función h es continua, satisface $h \geq 0$ y no es idénticamente nula porque $h(x_0) > 0$. Por Lema 15.1,
$$\int_a^b (g - f) > 0.$$
La linealidad de Teorema 14.6 transforma esta desigualdad en
$$\int_a^b g(x)\,dx - \int_a^b f(x)\,dx > 0,$$
que es la comparación estricta.
2. Si $h \equiv 0$, su integral es 0. Recíprocamente, si $h \geq 0$ es continua y no fuera idénticamente nula, Lema 15.1 daría
$$\int_a^b h(x)\,dx > 0,$$
contradicción.

3. Por el teorema de extremos Teorema 5.5, los números m y M están bien definidos y son valores de f. Como f no es constante,

$$m < f(x_1) \quad \text{para algún } x_1,$$

y

$$f(x_2) < M \quad \text{para algún } x_2.$$

Aplicando la primera parte a la función constante m y a f, y luego a f y a la función constante M, obtenemos

$$m(b-a) < \int_a^b f(x)\,dx < M(b-a).$$

Como $b-a>0$, al dividir y usar Definición 15.1 resulta $m < \operatorname{Avg}_{[a,b]}(f) < M$. □

La diferencia con §15.2 es decisiva. Allí bastaba integrabilidad para encerrar el promedio entre cotas no estrictas. Aquí la continuidad permite convertir una desigualdad puntual estricta en una desigualdad **integral** estricta. Para una función continua no constante, el promedio queda en (m, M). El valor intermedio convertirá esa localización estricta en la existencia de un punto cuya imagen sea exactamente el promedio.

15.4. Teorema del valor medio para integrales

El promedio está definido para toda función integrable. Si la función es continua y no constante, ya sabemos que queda estrictamente entre su mínimo y su máximo; podemos aplicar el teorema del valor intermedio para demostrar que se alcanza.

Teorema 15.1. **Teorema (valor medio para integrales).** Sea $f : [a,b] \to \mathbb{R}$ continua, con $a < b$. Entonces existe un punto

$$c \in (a,b)$$

tal que

$$\boxed{\int_a^b f(x)\,dx = f(c)(b-a).}$$

Equivalentemente,

$$\boxed{f(c) = \operatorname{Avg}_{[a,b]}(f).}$$

Demostración. Como f es continua en $[a,b]$, es Riemann-integrable por Teorema 14.4, de modo que su promedio integral está bien definido.

Distinguimos dos casos.

Caso 1: f es constante. Si $f(x) = k$ para todo $x \in [a,b]$, elegimos, por ejemplo, $c := \frac{a+b}{2} \in (a,b)$.

Entonces $f(c) = k$ y, por la integral de una constante,

$$\int_a^b f(x)\,dx = k(b-a) = f(c)(b-a).$$

Caso 2: f no es constante. Por el teorema de extremos Teorema 5.5 existen puntos $x_m, x_M \in [a,b]$ tales que

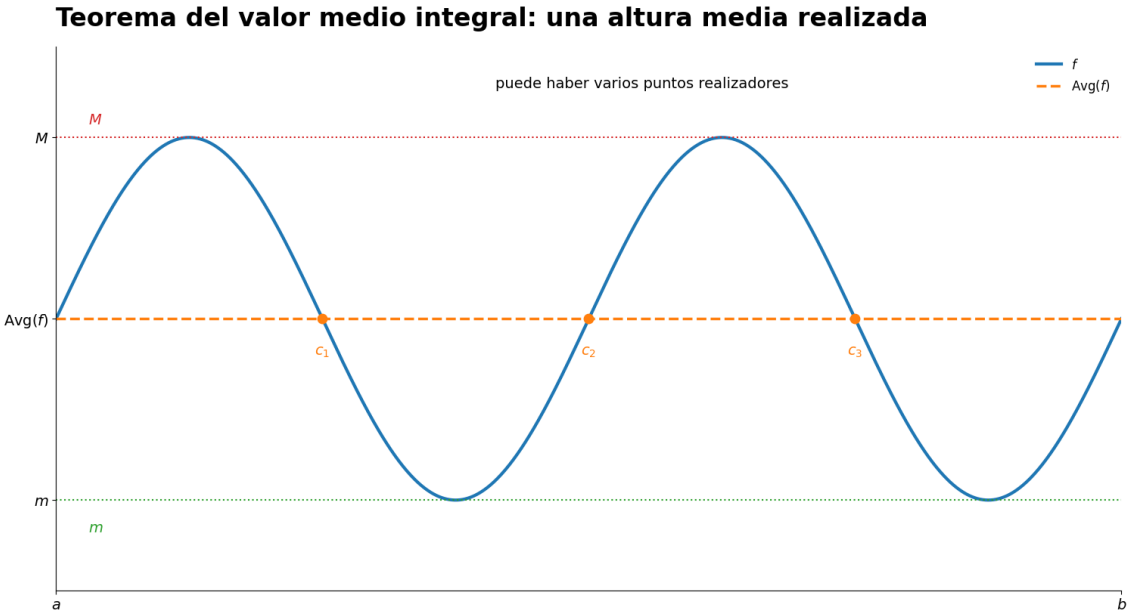


FIGURA 15.4. Para una función continua, el nivel medio es alcanzado por al menos un punto interior; el teorema no exige que ese punto sea único.

$$f(x_m) = m := \min_{[a,b]} f, \qquad f(x_M) = M := \max_{[a,b]} f.$$

Como f no es constante, $m < M$. Además, la tercera parte de Corolario 15.1 da $m < \operatorname{Avg}_{[a,b]}(f) < M$. Los puntos x_m y x_M son distintos. Definamos

$$u := \min\{x_m, x_M\}, \qquad v := \max\{x_m, x_M\}.$$

Entonces $u < v$, la restricción de f a $[u, v]$ es continua y los valores $f(u)$ y $f(v)$ son precisamente m y M en algún orden. Por tanto, $\operatorname{Avg}_{[a,b]}(f)$ está **estrictamente** entre $f(u)$ y $f(v)$. Aplicando el teorema del valor intermedio Teorema 5.2 a la restricción de f sobre $[u, v]$, existe $c \in (u, v)$ tal que $f(c) = \operatorname{Avg}_{[a,b]}(f)$. Como $(u, v) \subset (a, b)$, el punto obtenido satisface realmente $c \in (a, b)$. Finalmente, por Definición 15.1,

$$\int_a^b f(x)\,dx = (b-a)\operatorname{Avg}_{[a,b]}(f) = f(c)(b-a).$$

Esto concluye la demostración. □

Existencia no significa localización. El teorema afirma que **algún** punto interior realiza el promedio, pero no afirma que ese punto sea único ni proporciona por sí solo un procedimiento para calcularlo. Una función puede tomar su valor promedio en varios puntos distintos.

La posible multiplicidad de puntos realizadores se muestra en Figura 15.4. También conviene notar la arquitectura lógica de la prueba:

$$\text{continuidad}$$
$$\Longrightarrow \text{extremos alcanzados}$$
$$\Longrightarrow \text{promedio interior si } f \text{ no es constante}$$
$$\Longrightarrow \text{IVT}$$
$$\Longrightarrow f(c) = \operatorname{Avg}(f).$$

No intervienen primitivas, funciones de acumulación ni el teorema fundamental del cálculo. Las primeras consecuencias directas de este resultado —en particular, qué puede deducirse del signo o de la nulidad de una integral— se estudiarán exclusivamente en §15.5.

15.5. Consecuencias del valor medio

El teorema del valor medio para integrales convierte una afirmación global sobre

$$\int_a^b f(x)\,dx$$

en una afirmación puntual sobre algún valor $f(c)$. Esa conversión permite leer inmediatamente el signo y la nulidad de la integral, siempre bajo la hipótesis de continuidad.

Corolario 15.2. **Corolario (consecuencias puntuales del valor medio).** Sea $f : [a, b] \to \mathbb{R}$ continua, con $a < b$. Entonces:

1. Si
$$\int_a^b f(x)\,dx = 0,$$
existe $c \in (a, b)$ tal que
$$f(c) = 0.$$
2. Si
$$\int_a^b f(x)\,dx > 0,$$
existe $c \in (a, b)$ tal que $f(c) > 0$. Análogamente, si
$$\int_a^b f(x)\,dx < 0,$$
existe $c \in (a, b)$ tal que $f(c) < 0$.
3. Si
$$m := \min_{[a,b]} f, \qquad M := \max_{[a,b]} f,$$
las siguientes afirmaciones son equivalentes:
   - $\operatorname{Avg}_{[a,b]}(f) = m$;
   - $\operatorname{Avg}_{[a,b]}(f) = M$;
   - f es constante en $[a, b]$.

Demostración. Por Teorema 15.1 existe $c \in (a, b)$ tal que

$$\int_a^b f(x)\,dx = f(c)(b - a).$$

Como $b-a>0$, las dos primeras afirmaciones son inmediatas: una integral nula fuerza $f(c)=0$, mientras que el signo estricto de la integral coincide con el signo de ese valor $f(c)$. Para la tercera afirmación, si f es constante, su promedio coincide con su único valor y, por tanto, con su mínimo y su máximo. Recíprocamente, si f no fuera constante, la tercera parte de Corolario 15.1 daría $m<\operatorname{Avg}_{[a,b]}(f)<M$, lo que excluye cualquiera de las dos igualdades extremales. □

Del signo global a un valor puntual. Las dos primeras consecuencias deben leerse con cuidado. Por ejemplo,

$$\int_a^b f(x)\,dx>0$$

no significa que $f(x)>0$ para todo x. Afirma algo más débil pero muy útil: para una función continua, la acumulación neta positiva obliga a que la gráfica alcance al menos una altura positiva en el interior del intervalo. Lo mismo vale, con signos invertidos, para una integral negativa. La hipótesis de continuidad es esencial para la conclusión sobre integral nula. Considérese, en $[-1,1]$,

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

Esta función es Riemann-integrable y

$$\int_{-1}^1 f(x)\,dx=0,$$

pero no existe ningún punto donde $f(x)=0$. La integral puede cancelar contribuciones positivas y negativas sin que el nivel cero sea alcanzado. La continuidad es precisamente la condición que, mediante Teorema 15.1, transforma la información global en un valor puntual efectivo. Hasta aquí seguimos trabajando con la media uniforme determinada por la longitud del intervalo. El siguiente paso será permitir que distintas regiones del intervalo contribuyan con distinto peso; esa generalización comienza exclusivamente en §15.6.

15.6. Promedios ponderados

El promedio integral de §15.1 reparte la contribución de todos los puntos del intervalo con la misma regla de normalización: la longitud total $b-a$. Hay situaciones en las que conviene permitir que distintas regiones intervengan con distinta intensidad. Para ello introducimos una segunda función, el **peso**. Sean $f,w\in\mathcal{R}[a,b]$. Por Proposición 14.2, el producto fw también es Riemann-integrable, de modo que

$$\int_a^b f(x)w(x)\,dx$$

está bien definido. Para normalizar esa integral necesitamos, sin embargo, que la contribución total del peso no sea nula.

Definición 15.2. **Definición (promedio integral ponderado).** Sean $f,w\in\mathcal{R}[a,b]$, con $a<b$, y supongamos

$$\int_a^b w(x)\,dx\neq 0.$$

El **promedio integral de f ponderado por w** sobre $[a,b]$ es

$$\boxed{\operatorname{Avg}_{w,[a,b]}(f) := \frac{\displaystyle\int_a^b f(x)w(x)\,dx}{\displaystyle\int_a^b w(x)\,dx}.}$$

La condición sobre el denominador forma parte de la definición. Si

$$\int_a^b w(x)\,dx = 0,$$

el cociente anterior **no está definido**, aunque las integrales del numerador y del denominador existan por separado.

Dos controles de normalización. La definición recupera inmediatamente la media uniforme. Si el peso es constante,

$$w(x) = \lambda, \qquad \lambda \neq 0,$$

entonces, por linealidad,

$$\operatorname{Avg}_{w,[a,b]}(f) = \frac{\lambda \int_a^b f(x)\,dx}{\lambda(b-a)} = \operatorname{Avg}_{[a,b]}(f).$$

Más generalmente, multiplicar un peso admisible por una constante no nula no modifica el promedio ponderado. Si $\mu \neq 0$, entonces

$$\operatorname{Avg}_{\mu w,[a,b]}(f) = \frac{\mu \int_a^b fw}{\mu \int_a^b w} = \operatorname{Avg}_{w,[a,b]}(f).$$

Por tanto, el promedio depende de las **proporciones relativas** con que el peso distribuye su contribución, no de una escala global elegida para medirla.

Signo puntual y peso total no son la misma condición. Conviene distinguir cuidadosamente dos afirmaciones:

$$w(x) \geq 0 \quad \text{para todo } x$$

y

$$\int_a^b w(x)\,dx > 0.$$

La primera es una condición **punto a punto**; la segunda es una condición **global**. Para pesos continuos no negativos que no sean idénticamente nulos, Lema 15.1 permite pasar de la primera a la segunda. Para un peso meramente Riemann-integrable, ese paso ya no es automático.

Ejemplo 15.2. **Ejemplo (signo puntual frente a peso total).** En $[-1, 1]$ consideremos dos pesos.

1. Definamos
$$w_1(x) = \begin{cases} 1, & x = 0, \\ 0, & x \neq 0. \end{cases}$$
Entonces $w_1 \geq 0$ y $w_1 \not\equiv 0$. Sin embargo, w_1 difiere de la función nula en un solo punto; por Corolario 14.1,
$$\int_{-1}^{1} w_1(x)\,dx = 0.$$
Por tanto, w_1 **no** define un promedio ponderado mediante Definición 15.2.

2. Definamos el peso escalonado
$$w_2(x) = \begin{cases} -1, & -1 \leq x < 0, \\ 2, & 0 \leq x \leq 1. \end{cases}$$
Entonces w_2 cambia de signo, pero
$$\int_{-1}^{1} w_2(x)\,dx = (-1)(1) + 2(1) = 1 > 0.$$
Así, tener integral positiva tampoco implica ser un peso no negativo.

Estos dos ejemplos muestran por qué, en los resultados de valor medio, no conviene reemplazar una hipótesis de **signo fijo** por una simple condición sobre el signo de $\int w$. La definición de promedio ponderado solo exige que el denominador sea distinto de cero; las hipótesis adicionales que permiten garantizar que ese promedio sea realizado por algún valor de f pertenecen exclusivamente a §15.7.

15.7. Teorema del valor medio ponderado

El promedio ponderado de §15.6 solo está definido cuando $\int_a^b w \neq 0$. El teorema de valor medio, en cambio, admite una formulación más básica que sigue teniendo sentido incluso cuando esa integral se anula: primero se factoriza la integral de fw y **solo después**, si el factor $\int w$ es distinto de cero, se divide.

Teorema 15.2. **Teorema (valor medio ponderado para integrales).** Sea $f : [a, b] \to \mathbb{R}$ continua, con $a < b$, y sea $w \in \mathcal{R}[a, b]$ un peso de signo fijo; es decir, supongamos que

$$w(x) \geq 0 \quad \text{para todo } x \in [a, b],$$

o bien

$$w(x) \leq 0 \quad \text{para todo } x \in [a, b].$$

Entonces existe $c \in [a, b]$ tal que

$$\boxed{\int_a^b f(x)w(x)\,dx = f(c)\int_a^b w(x)\,dx.}$$

En particular, si

$$\int_a^b w(x)\,dx \neq 0,$$

entonces

$$\boxed{\operatorname{Avg}_{w,[a,b]}(f) = f(c).}$$

Demostración. Comenzamos con el caso $w \geq 0$.

Como f es continua en $[a, b]$, el teorema de extremos Teorema 5.5 proporciona

$$m := \min_{[a,b]} f, \qquad M := \max_{[a,b]} f.$$

Por tanto,

$$m \leq f(x) \leq M \qquad (x \in [a, b]).$$

Multiplicando por $w(x) \geq 0$ obtenemos $mw(x) \leq f(x)w(x) \leq Mw(x)$.

Las tres funciones son Riemann-integrables: w lo es por hipótesis y fw por Proposición 14.2. Aplicando la monotonía y la linealidad de Teorema 14.6,

$$m \int_a^b w(x)\,dx \leq \int_a^b f(x)w(x)\,dx \leq M \int_a^b w(x)\,dx.$$

Ahora debemos separar dos casos **antes de dividir**.

Caso 1: $\int_a^b w = 0$**.** La desigualdad anterior se reduce a

$$0 \leq \int_a^b f(x)w(x)\,dx \leq 0,$$

por lo que

$$\int_a^b f(x)w(x)\,dx = 0.$$

En consecuencia,

$$\int_a^b f(x)w(x)\,dx = f(c) \int_a^b w(x)\,dx$$

para **cualquier** $c \in [a, b]$. Obsérvese que en este caso no se ha dividido por $\int w$ y que el promedio ponderado de Definición 15.2 sigue sin estar definido.

Caso 2: $\int_a^b w > 0$**.** Podemos dividir la cadena de desigualdades por ese número positivo y obtener

$$m \leq \frac{\displaystyle\int_a^b f(x)w(x)\,dx}{\displaystyle\int_a^b w(x)\,dx} \leq M.$$

El cociente central es $\operatorname{Avg}_{w,[a,b]}(f)$ por Definición 15.2. Como f es continua y toma los valores m y M, el teorema del valor intermedio Teorema 5.2 garantiza un punto $c \in [a, b]$ tal que $f(c) = \operatorname{Avg}_{w,[a,b]}(f)$. Multiplicando por $\int_a^b w$ resulta la identidad requerida.

Finalmente, si $w \leq 0$, definimos $v := -w \geq 0$.

La función v es Riemann-integrable. Aplicando el caso ya demostrado a v, existe $c \in [a, b]$ tal que

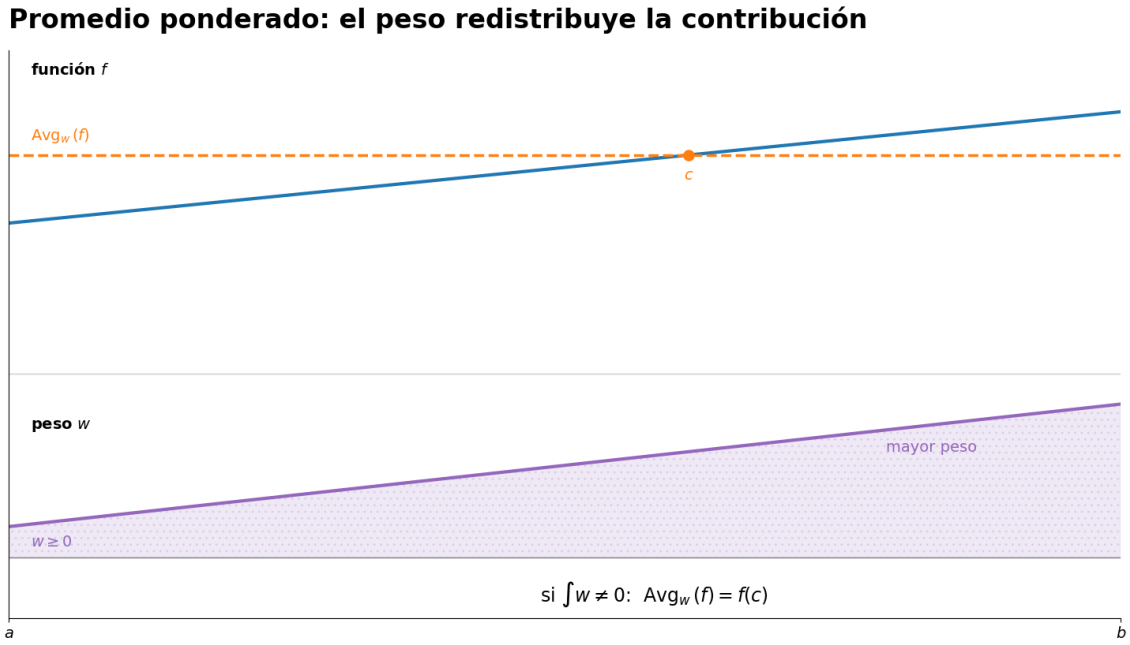


FIGURA 15.5. Un peso de signo fijo redistribuye la contribución del intervalo. Cuando su integral total no es nula, el promedio ponderado queda dentro del rango de f y es realizado por algún punto.

$$\int_a^b f(x)v(x)\,dx = f(c)\int_a^b v(x)\,dx.$$

Como $v = -w$, multiplicar esta igualdad por -1 da

$$\int_a^b f(x)w(x)\,dx = f(c)\int_a^b w(x)\,dx.$$

Esto concluye la demostración. □

La función del peso y la realización del promedio ponderado se sintetizan en Figura 15.5.

Qué se obtiene cuando el peso total no es nulo. Si $\int_a^b w \neq 0$, el teorema afirma exactamente que el promedio ponderado pertenece a la imagen de f: $\mathrm{Avg}_{w,[a,b]}(f) \in f([a,b])$. La diferencia con el caso ordinario de Teorema 15.1 es que aquí no prometemos que el punto realizador sea interior. El peso puede concentrar su contribución efectiva en regiones donde f toma un valor extremo, de modo que la conclusión segura en toda generalidad es $c \in [a,b]$. La hipótesis de signo fijo es la que permite conservar el orden al multiplicar por w. Si el peso cambia de signo, esa cadena puede romperse y el cociente normalizado puede quedar fuera del rango de f. Esa frontera se auditará exclusivamente en §15.8.

15.8. Por qué las hipótesis importan

Los teoremas de §§15.4 y 15.7 son cortos en su formulación, pero cada una de sus hipótesis cumple una función concreta. Conviene someterlas a una auditoría adversarial antes de usarlas de manera automática.

EJEMPLO 15.3. **Cuatro pruebas de estrés para las hipótesis.**

1. **Integrabilidad sin continuidad: el promedio puede no ser un valor de la función.** Consideremos nuevamente la función escalón de Ejemplo 14.9,
$$s(x) = \begin{cases} 0, & 0 \le x < \frac{1}{2}, \\ 1, & \frac{1}{2} \le x \le 1. \end{cases}$$
Allí se obtuvo
$$\int_0^1 s(x)\,dx = \frac{1}{2}.$$
Como el intervalo tiene longitud 1, $\operatorname{Avg}_{[0,1]}(s) = \frac{1}{2}$. Pero $s([0,1]) = \{0,1\}$, de modo que no existe $c \in [0,1]$ con $s(c) = 1/2$. La integrabilidad por sí sola no sustituye la continuidad en Teorema 15.1.

2. **Un peso que cambia de signo puede expulsar el promedio ponderado fuera del rango de f.** En $[0,3]$ definamos la función continua
$$f(x) = \begin{cases} 0, & 0 \le x \le 1, \\ x - 1, & 1 \le x \le 2, \\ 1, & 2 \le x \le 3, \end{cases}$$
y el peso escalonado
$$w(x) = \begin{cases} -1, & 0 \le x < 1, \\ 0, & 1 \le x < 2, \\ 2, & 2 \le x \le 3. \end{cases}$$
Entonces w cambia de signo y
$$\int_0^3 w(x)\,dx = -1 + 0 + 2 = 1.$$
Además, fw vale 0 en $[0,2)$ y 2 en $[2,3]$, salvo elecciones irrelevantes en los puntos de empalme. Por tanto,
$$\int_0^3 f(x)w(x)\,dx = 2.$$
Así, $\operatorname{Avg}_{w,[0,3]}(f) = 2$, mientras que $f([0,3]) = [0,1]$. No existe, pues, ningún c con $f(c) = \operatorname{Avg}_{w,[0,3]}(f)$. La hipótesis de signo fijo en Teorema 15.2 no es decorativa.

Los dos mecanismos de fallo se comparan en Figura 15.6.

3. **No se puede dividir por $\int w$ sin verificar antes que sea distinto de cero.** Tomemos el peso w_1 de Ejemplo 15.2, que satisface $w_1 \ge 0$ pero
$$\int_{-1}^1 w_1(x)\,dx = 0.$$
Para cualquier función continua f, el producto fw_1 difiere de la función nula a lo sumo en el punto 0; por Corolario 14.1,
$$\int_{-1}^1 f(x)w_1(x)\,dx = 0.$$
La identidad de Teorema 15.2 se convierte entonces en $0 = f(c)\,0$, que es válida, pero el cociente
$$\frac{\int fw_1}{\int w_1} = \frac{0}{0}$$

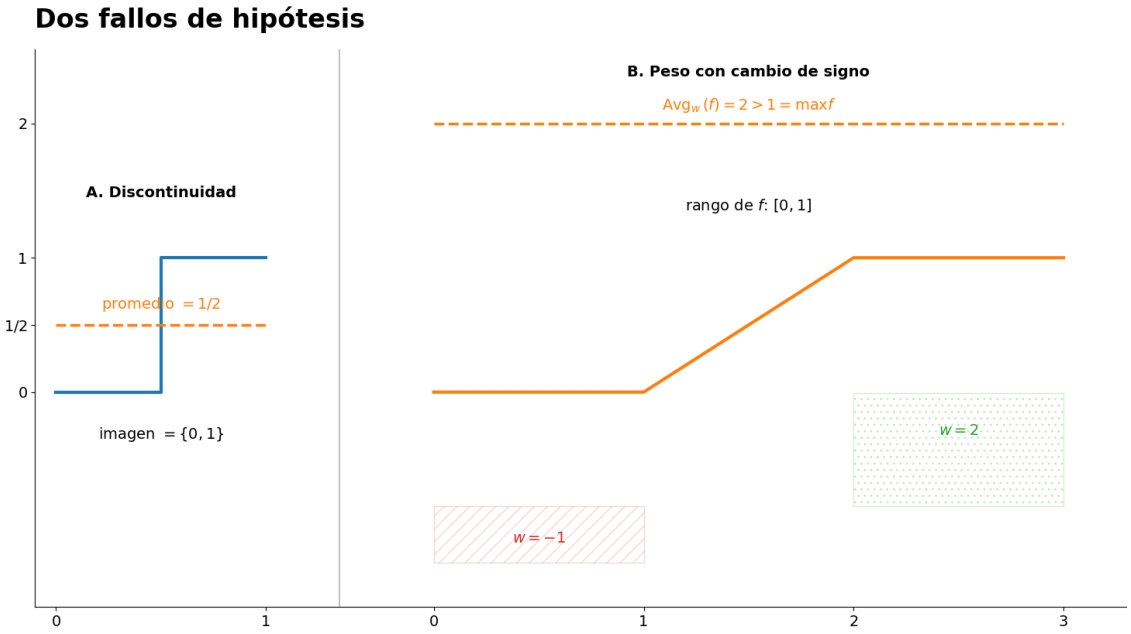


FIGURA 15.6. Dos fallos de hipótesis: una discontinuidad puede dejar el promedio ordinario fuera de la imagen, y un peso que cambia de signo puede expulsar el promedio ponderado fuera del rango de f.

no está definido. El orden lógico correcto es exactamente el usado en §15.7: separar primero el caso $\int w = 0$ y dividir solo en el caso no nulo.

4. **Integral nula no significa función nula sin una hipótesis de signo.** Sea

$$g(x) = 2x - 1, \qquad 0 \le x \le 1.$$

Por linealidad de Teorema 14.6, la integral de la constante y Ejemplo 14.2,

$$\int_0^1 g(x)\,dx = 2\int_0^1 x\,dx - \int_0^1 1\,dx = 2 \cdot \frac{1}{2} - 1 = 0.$$

Sin embargo, g no es idénticamente nula. Lo que permite concluir $h \equiv 0$ a partir de $\int h = 0$ en Corolario 15.1 es la combinación de continuidad con $h \ge 0$, no la nulidad de la integral por sí sola.

Una tabla mental de uso. Los cuatro fallos anteriores se resumen en una regla operativa sencilla:

- para **realizar un promedio ordinario** como $f(c)$, comprobar continuidad de f;
- para **realizar un promedio ponderado**, comprobar además que el peso tenga signo fijo;
- antes de escribir un cociente ponderado, comprobar $\int w \neq 0$;
- antes de deducir que una función es nula a partir de una integral nula, comprobar la hipótesis de signo que impide cancelaciones.

Estas comprobaciones no añaden un teorema nuevo: delimitan el alcance exacto de los resultados ya demostrados. En §15.9 pasaremos de esta auditoría lógica a estimaciones cuantitativas de integrales y promedios.

15.9. Estimaciones y lectura cuantitativa del promedio

Los resultados anteriores no solo garantizan existencia. Las desigualdades que los sostienen permiten obtener información cuantitativa sobre una integral o sobre un promedio **sin calcularlos exactamente**. Esta es la función de las estimaciones: convertir una banda conocida para la función en una banda certificada para la cantidad integral.

PROPOSICIÓN 15.2. **Proposición (cotas cuantitativas para integrales y promedios).** Sea $a < b$.

1. Si $f \in \mathcal{R}[a,b]$ y existen $m, M \in \mathbb{R}$ tales que
$$m \le f(x) \le M \qquad (x \in [a,b]),$$
entonces
$$\boxed{m(b-a) \le \int_a^b f(x)\,dx \le M(b-a).}$$

2. Si $f, g \in \mathcal{R}[a,b]$ y, para algún $\varepsilon \ge 0$,
$$|f(x) - g(x)| \le \varepsilon \qquad (x \in [a,b]),$$
entonces
$$\boxed{\left|\int_a^b f(x)\,dx - \int_a^b g(x)\,dx\right| \le \varepsilon(b-a),}$$
y, por tanto,
$$\boxed{\left|\operatorname{Avg}_{[a,b]}(f) - \operatorname{Avg}_{[a,b]}(g)\right| \le \varepsilon.}$$

3. Si $f, w \in \mathcal{R}[a,b]$, el peso w tiene signo fijo,
$$\int_a^b w(x)\,dx \neq 0,$$
y
$$m \le f(x) \le M \qquad (x \in [a,b]),$$
entonces
$$\boxed{m \le \operatorname{Avg}_{w,[a,b]}(f) \le M.}$$

Demostración.

1. La primera afirmación es la forma no normalizada de Proposición 15.1. En efecto, al multiplicar $m \le \operatorname{Avg}_{[a,b]}(f) \le M$ por $b-a > 0$ y usar Definición 15.1 se obtiene exactamente
$$m(b-a) \le \int_a^b f(x)\,dx \le M(b-a).$$

2. De $|f-g| \le \varepsilon$ se sigue $-\varepsilon \le f - g \le \varepsilon$. Aplicando la primera parte a $f-g$ y usando la linealidad de Teorema 14.6,
$$-\varepsilon(b-a) \le \int_a^b f(x)\,dx - \int_a^b g(x)\,dx \le \varepsilon(b-a).$$
Esta doble desigualdad equivale a
$$\left|\int_a^b f - \int_a^b g\right| \le \varepsilon(b-a).$$
Dividir por $b-a$ y usar Definición 15.1 da la estimación de los promedios.

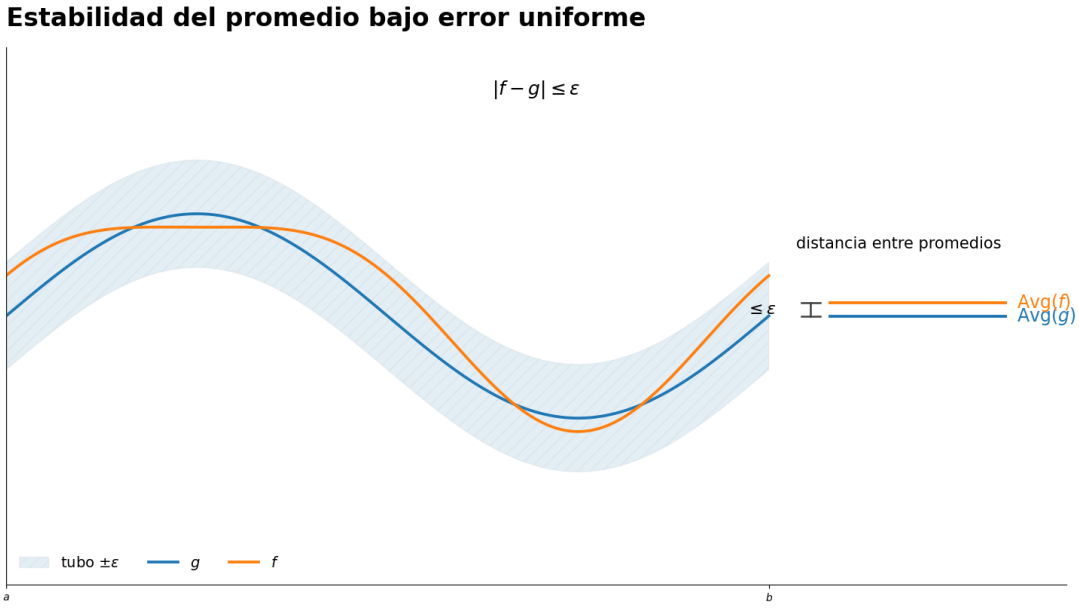


FIGURA 15.7. Si $|f-g| \le \varepsilon$ en todo el intervalo, la banda vertical de radio ε se transmite directamente a $|\operatorname{Avg}(f) - \operatorname{Avg}(g)| \le \varepsilon$.

3. Supongamos primero $w \ge 0$. De $m \le f \le M$ obtenemos $mw \le fw \le Mw$. El producto fw es integrable por Proposición 14.2. Por monotonía y linealidad de Teorema 14.6,

$$m\int_a^b w \le \int_a^b fw \le M\int_a^b w.$$

Como $w \ge 0$, su integral es no negativa; la hipótesis $\int_a^b w \neq 0$ fuerza entonces

$$\int_a^b w > 0.$$

Podemos dividir por ese número positivo y usar Definición 15.2.

Si $w \le 0$, tomamos $v := -w \ge 0$. Además, $\operatorname{Avg}_{v,[a,b]}(f) = \operatorname{Avg}_{w,[a,b]}(f)$, porque numerador y denominador cambian simultáneamente de signo. El caso ya demostrado para v concluye la prueba. □

La estabilidad uniforme de la segunda parte de Proposición 15.2 se visualiza en Figura 15.7.

EJEMPLO 15.4. **Ejemplo (dos estimaciones sin calcular la integral).**

1. Sea $f \in \mathcal{R}[0,1]$ y supongamos que

$$|f(x) - x| \le \frac{3}{100} \qquad (0 \le x \le 1).$$

En Ejemplo 14.2 se calculó

$$\int_0^1 x\,dx = \frac{1}{2}.$$

La segunda parte de Proposición 15.2 da

$$\left|\int_0^1 f(x)\,dx - \frac{1}{2}\right| \le \frac{3}{100}.$$

Por tanto, sin conocer una fórmula exacta para la integral de f, ya sabemos que

$$\boxed{\frac{47}{100} \le \int_0^1 f(x)\,dx \le \frac{53}{100}.}$$

Como el intervalo tiene longitud 1, la misma banda contiene a $\mathrm{Avg}_{[0,1]}(f)$.

2. Sean $f, w \in \mathcal{R}[a,b]$, con w de signo fijo y $\int_a^b w \neq 0$. Si únicamente sabemos que $2 \le f(x) \le 5$, entonces ya podemos concluir

$$\boxed{2 \le \mathrm{Avg}_{w,[a,b]}(f) \le 5.}$$

No fue necesario calcular ni $\int fw$ ni $\int w$ por separado. Si además f es continua, Teorema 15.2 garantiza que ese promedio ponderado es igual a $f(c)$ para algún $c \in [a,b]$; el teorema añade **existencia de una altura realizadora**, no un procedimiento para localizarla.

Existencia no es cálculo. Para una función continua, Teorema 15.1 afirma que existe $c \in (a,b)$ tal que

$$\int_a^b f(x)\,dx = f(c)(b-a).$$

La identidad puede ser muy informativa, pero no debe leerse como una técnica automática de evaluación. Si el punto c es desconocido, también lo es en general el valor $f(c)$. El teorema transforma una cantidad global en una afirmación de existencia; **no sustituye el cálculo de la integral por el cálculo de un punto que el propio teorema no determina**.

La lectura cuantitativa correcta es otra: cuando conocemos una banda para f, una aproximación uniforme de f por otra función integrable, o un rango compatible con un peso de signo fijo, obtenemos inmediatamente bandas rigurosas para integrales y promedios. En §15.10 usaremos esta distinción para escoger el resultado adecuado y reparar argumentos incorrectos.

15.10. Laboratorio: escoger el teorema y reparar argumentos

Hasta aquí el capítulo ha producido resultados que se parecen entre sí porque todos relacionan una integral con valores de la función. Precisamente por eso es fácil aplicar el teorema correcto por una razón incorrecta, o pedirle a un resultado más de lo que afirma. El objetivo de este laboratorio no es añadir teoría nueva. En cada estación partiremos de un argumento plausible, identificaremos su falla lógica y lo repararemos usando únicamente resultados ya demostrados.

Estación 1. Un promedio numérico no es todavía un punto. Argumento defectuoso. Si $f \in \mathcal{R}[a,b]$, entonces

$$\mathrm{Avg}_{[a,b]}(f) = \frac{1}{b-a}\int_a^b f(x)\,dx,$$

de modo que debe existir $c \in [a,b]$ tal que $f(c) = \mathrm{Avg}_{[a,b]}(f)$.

Diagnóstico. La primera igualdad es solo la definición Definición 15.1. No contiene ninguna afirmación de existencia. De hecho, Ejemplo 15.3 mostró una función Riemann-integrable cuyo promedio no pertenece a su imagen.

Reparación. Si f es continua, entonces sí podemos invocar Teorema 15.1 y concluir que existe $c \in (a,b)$ con $f(c) = \mathrm{Avg}_{[a,b]}(f)$.

La cadena correcta es

$$\text{integrabilidad} \Longrightarrow \text{promedio definido},$$

mientras que

$$\text{continuidad} \Longrightarrow \text{promedio realizado por algún valor de } f.$$

Estación 2. Integrabilidad no sustituye continuidad. Argumento defectuoso. Toda función Riemann-integrable es suficientemente regular para aplicar el teorema del valor medio para integrales.

Diagnóstico. La integrabilidad garantiza que la integral existe y que el promedio de Definición 15.1 está definido, pero no garantiza el valor intermedio necesario para realizar ese promedio. El primer caso de Ejemplo 15.3 separa exactamente ambas propiedades.

Reparación. Antes de citar Teorema 15.1 hay que verificar continuidad en todo el intervalo cerrado. Si solo conocemos integrabilidad, siguen disponibles las cotas de Proposición 15.1 y las estimaciones de Proposición 15.2, pero no la existencia de un punto realizador.

Estación 3. Promedio ordinario y promedio ponderado no son intercambiables. Supongamos que $f, w \in \mathcal{R}[a,b]$ y que $\int_a^b w \neq 0$.

Argumento defectuoso. El promedio ponderado es simplemente el promedio ordinario de fw.

Diagnóstico. Son normalizaciones distintas:

$$\operatorname{Avg}_{w,[a,b]}(f) = \frac{\int_a^b fw}{\int_a^b w},$$

mientras que

$$\operatorname{Avg}_{[a,b]}(fw) = \frac{\int_a^b fw}{b-a}.$$

Si el numerador $\int_a^b fw$ es no nulo, ambos cocientes solo pueden coincidir cuando $\int_a^b w = b-a$. Si, en cambio, $\int_a^b fw = 0$, ambos promedios pueden coincidir trivialmente aunque los denominadores sean distintos. En general son normalizaciones diferentes y no deben identificarse.

Reparación. Primero hay que identificar qué normalización corresponde al problema. Si el peso es constante no nulo, §15.6 mostró que el promedio ponderado recupera el promedio ordinario de f; fuera de ese caso no debemos confundir ambos objetos.

Estación 4. Integral positiva del peso no significa signo fijo. Argumento defectuoso. Si

$$\int_a^b w(x)\,dx > 0,$$

entonces podemos aplicar el teorema del valor medio ponderado.

Diagnóstico. El signo de una integral es información global; no impide que w cambie de signo. Ejemplo 15.2 exhibe precisamente un peso con integral positiva que toma valores negativos y positivos. Además, Ejemplo 15.3 muestra que, con cambio de signo, el promedio ponderado puede incluso salir del rango de f.

Reparación. Para aplicar Teorema 15.2 debemos verificar una condición punto a punto:

$$w \geq 0 \qquad \text{o} \qquad w \leq 0$$

en todo $[a, b]$. El signo de $\int w$ por sí solo no reemplaza esa hipótesis.

Estación 5. El caso $\int w = 0$ se separa antes de dividir. Argumento defectuoso. Del teorema ponderado escribimos inmediatamente

$$f(c) = \frac{\int_a^b fw}{\int_a^b w}.$$

Diagnóstico. Esa división solo es legítima si $\int_a^b w \neq 0$. Cuando w tiene signo fijo y su integral se anula, Teorema 15.2 sigue proporcionando la identidad

$$\int_a^b fw = f(c) \int_a^b w,$$

pero ambos miembros son entonces cero y el cociente de Definición 15.2 no está definido.

Reparación. El orden lógico es obligatorio:

1. establecer primero la identidad factorizada de Teorema 15.2;
2. comprobar si $\int w$ es cero;
3. dividir únicamente en el caso no nulo.

Estación 6. El teorema del valor medio integral no es el teorema fundamental del cálculo. Argumento defectuoso. Como para una función continua existe $c \in (a, b)$ tal que

$$\int_a^b f(x)\,dx = f(c)(b - a),$$

esto explica ya cómo derivar una integral o cómo recuperar una primitiva.

Diagnóstico. Teorema 15.1 es un resultado de **existencia de un valor medio** sobre un intervalo fijo. No define una función de acumulación, no estudia su derivada y no relaciona todavía integración con antiderivación.

Reparación. Debemos mantener separadas dos preguntas:

¿qué valor de f representa el promedio de una integral fija?

y

¿qué ocurre cuando uno de los extremos de integración pasa a ser variable?

La primera pertenece a este capítulo. La segunda abre C17 y será allí donde aparecerá el teorema fundamental del cálculo.

Estación 7. Un puente permitido hacia C17: subintervalos fijos. Sea $f : [a, b] \to \mathbb{R}$ continua y fijemos dos puntos

$$a \leq u < v \leq b.$$

La restricción de f a $[u, v]$ es continua. Por Teorema 15.1 existe algún $c_{u,v} \in (u, v)$ tal que

$$\boxed{\int_u^v f(x)\,dx = f(c_{u,v})(v - u).}$$

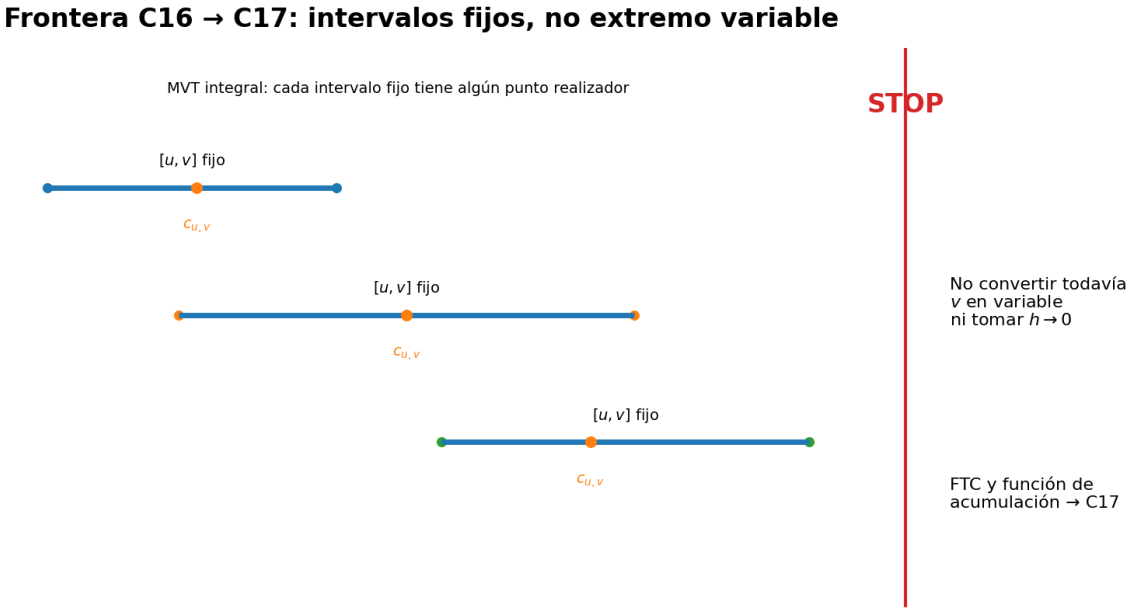


FIGURA 15.8. El teorema del valor medio puede aplicarse a cada subintervalo fijo $[u, v]$, pero no proporciona una elección canónica de $c_{u,v}$ ni autoriza todavía a convertir un extremo de integración en variable.

Para cada subintervalo fijo, la continuidad permite que alguna imagen de la función realice su promedio integral. El punto $c_{u,v}$ puede depender del subintervalo y no está determinado de manera única por el teorema. Si queremos estudiar un límite cuando cambian u y v, necesitaremos justificarlo sin presuponer una elección regular de esos puntos. Esto afecta, en particular, a expresiones como

$$\frac{1}{h}\int_x^{x+h} f(t)\,dt,$$

ni introducimos una función de acumulación del tipo

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

Ambos pasos pertenecen a C17.

La figura Figura 15.8 distingue lo que permite afirmar el valor medio de la cuestión adicional sobre derivar una acumulación con extremo variable.

Mapa de decisión antes de usar un resultado. Ante una integral y una afirmación sobre "promedios", conviene preguntar en este orden:

1. **¿Solo necesito una cota?** Use Proposición 15.1 o Proposición 15.2; la continuidad puede no ser necesaria.
2. **¿Quiero afirmar que el promedio ordinario es un valor de f?** Verifique continuidad y use Teorema 15.1.
3. **¿Hay un peso?** Verifique primero que el producto sea integrable y distinga la identidad factorizada del cociente de Definición 15.2.

4. **¿Quiero realizar un promedio ponderado como $f(c)$?** Verifique continuidad de f, signo fijo de w y, para escribir el cociente, $\int w \neq 0$; entonces use Teorema 15.2.
5. **¿Aparece un extremo variable o un límite de cocientes integrales?** El problema ya apunta a C17 y no debe resolverse adelantando el FTC.

En los ejercicios se distinguirán las cotas del promedio, la existencia de un punto que lo realiza y las condiciones sobre los pesos. El resultado elegido deberá corresponder tanto a la pregunta como a las hipótesis disponibles.

15.11. Ejercicios y soluciones

Los cuarenta ejercicios siguientes recorren la arquitectura del capítulo sin introducir teoría posterior. El banco está organizado en siete niveles y usa únicamente resultados disponibles hasta §15.10. Las soluciones están desarrolladas o razonadas y mantienen separadas las preguntas de **estimación**, **existencia** y **cálculo exacto**.

Nivel A — Promedio integral y normalización. Ejercicio A1. Promedio de una función constante. Sea $f(x) = k$ en $[a, b]$, con $a < b$. Demuestra directamente desde la definición que $\operatorname{Avg}_{[a,b]}(f) = k$.

Explica por qué en este caso todo punto $c \in [a, b]$ realiza el promedio.

Ejercicio A2. Recuperar una altura media a partir de una integral. Sea $f \in \mathcal{R}[2, 6]$ y supón que

$$\int_2^6 f(x)\,dx = 14.$$

1. Calcula $\operatorname{Avg}_{[2,6]}(f)$.
2. Decide si estos datos, por sí solos, garantizan que existe $c \in [2, 6]$ con $f(c) = \operatorname{Avg}_{[2,6]}(f)$.
3. Indica qué hipótesis adicional permite asegurar esa existencia.

Ejercicio A3. Linealidad del promedio. Sean $f, g \in \mathcal{R}[a, b]$ y $\alpha, \beta \in \mathbb{R}$. Demuestra que

$$\operatorname{Avg}_{[a,b]}(\alpha f + \beta g) = \alpha \operatorname{Avg}_{[a,b]}(f) + \beta \operatorname{Avg}_{[a,b]}(g).$$

Ejercicio A4. El promedio total como combinación de promedios parciales. Sean $a < c < b$ y $f \in \mathcal{R}[a, b]$. Demuestra que

$$(b - a) \operatorname{Avg}_{[a,b]}(f) = (c - a) \operatorname{Avg}_{[a,c]}(f) + (b - c) \operatorname{Avg}_{[c,b]}(f).$$

Deduce que el promedio sobre $[a, b]$ es una combinación ponderada por longitudes de los dos promedios parciales.

Ejercicio A5. Dos promedios parciales conocidos. Una función f es Riemann-integrable en $[0, 5]$ y satisface

$$\operatorname{Avg}_{[0,2]}(f) = 1, \qquad \operatorname{Avg}_{[2,5]}(f) = 4.$$

Calcula $\operatorname{Avg}_{[0,5]}(f)$ sin calcular por separado una fórmula para f.

Ejercicio A6. Un promedio que no es un valor de la función. En $[0,3]$ define

$$s(x) = \begin{cases} 1, & 0 \le x < 1, \\ 4, & 1 \le x \le 3. \end{cases}$$

1. Calcula $\int_0^3 s(x)\,dx$.
2. Calcula $\mathrm{Avg}_{[0,3]}(s)$.
3. Decide si existe $c \in [0,3]$ tal que $s(c)$ sea igual al promedio.
4. Explica qué hipótesis del teorema del valor medio integral falta.

Nivel B — Cotas, positividad estricta y comparación. Ejercicio B1. De una banda puntual a una banda integral. Sea $f \in \mathcal{R}[1,4]$ y supón que

$$2 \le f(x) \le 5 \qquad (1 \le x \le 4).$$

Obtén cotas para

$$\int_1^4 f(x)\,dx \qquad \text{y} \qquad \mathrm{Avg}_{[1,4]}(f).$$

Ejercicio B2. Una aproximación uniforme alrededor de una constante. Sea $f \in \mathcal{R}[0,10]$ y supón que

$$|f(x) - 3| \le \frac{1}{5} \qquad (0 \le x \le 10).$$

Encuentra una banda cerrada que contenga a $\mathrm{Avg}_{[0,10]}(f)$ y otra que contenga a $\int_0^{10} f$.

Ejercicio B3. Integral nula de una función continua no negativa. Sea $h : [a,b] \to \mathbb{R}$ continua, $a < b$, con $h \ge 0$. Demuestra que

$$\int_a^b h(x)\,dx = 0 \quad \Longrightarrow \quad h \equiv 0.$$

Tu prueba debe identificar con precisión dónde se usa la continuidad.

Ejercicio B4. Comparación estricta. Sean $f, g : [a,b] \to \mathbb{R}$ continuas y supón

$$f(x) \le g(x) \qquad (x \in [a,b]),$$

con desigualdad estricta en al menos un punto. Demuestra que

$$\int_a^b f(x)\,dx < \int_a^b g(x)\,dx.$$

Ejercicio B5. El promedio de una continua no constante no toca los extremos. Sea $f : [a, b] \to \mathbb{R}$ continua y no constante. Si

$$m = \min_{[a,b]} f, \qquad M = \max_{[a,b]} f,$$

demuestra que $m < \operatorname{Avg}_{[a,b]}(f) < M$.

Ejercicio B6. Por qué la continuidad importa en la positividad estricta. Define en $[-1, 1]$

$$h(x) = \begin{cases} 1, & x = 0, \\ 0, & x \neq 0. \end{cases}$$

Comprueba que $h \geq 0$, $h \not\equiv 0$ y, sin embargo,

$$\int_{-1}^{1} h(x)\, dx = 0.$$

Explica por qué esto no contradice Lema 15.1.

Nivel C — Teorema ordinario e hipótesis. Ejercicio C1. Existencia de una altura concreta. Sea $f : [1, 5] \to \mathbb{R}$ continua y supón

$$\int_{1}^{5} f(x)\, dx = 12.$$

Demuestra que existe $c \in (1, 5)$ tal que $f(c) = 3$.

Ejercicio C2. Localizar el punto medio para la identidad. Para $f(x) = x$ en $[0, 1]$ se sabe que

$$\int_{0}^{1} x\, dx = \frac{1}{2}.$$

Encuentra todos los puntos $c \in (0, 1)$ que satisfacen la identidad del teorema del valor medio integral.

Ejercicio C3. No unicidad extrema. Sea $f(x) = 7$ en $[a, b]$, $a < b$.

1. Verifica la identidad del teorema del valor medio integral.
2. Describe el conjunto de todos los puntos $c \in (a, b)$ que la realizan.
3. Explica por qué el teorema no afirma unicidad.

Ejercicio C4. Integrabilidad sin punto realizador. En $[0, 1]$ define

$$r(x) = \begin{cases} -1, & 0 \leq x < \tfrac{1}{2}, \\ 1, & \tfrac{1}{2} \leq x \leq 1. \end{cases}$$

Calcula su integral y su promedio. Demuestra que no existe $c \in [0, 1]$ tal que $r(c) = \operatorname{Avg}_{[0,1]}(r)$.

Ejercicio C5. Del signo global a un valor puntual. Sea $f : [a, b] \to \mathbb{R}$ continua.

1. Si $\int_a^b f = 0$, demuestra que f se anula en algún punto interior.
2. Si $\int_a^b f > 0$, demuestra que f toma un valor positivo en algún punto interior.
3. Explica por qué la segunda conclusión no implica que $f > 0$ en todo el intervalo.

Ejercicio C6. Cuándo el punto realizador es único. Sea $f : [a, b] \to \mathbb{R}$ continua y estrictamente monótona. Demuestra que existe un único $c \in (a, b)$ tal que $f(c) = \mathrm{Avg}_{[a,b]}(f)$.

Nivel D — Promedio ponderado y teorema ponderado. Ejercicio D1. Peso constante. Sean $f \in \mathcal{R}[a, b]$ y

$$w(x) = \lambda, \qquad \lambda \neq 0.$$

Demuestra que $\mathrm{Avg}_{w,[a,b]}(f) = \mathrm{Avg}_{[a,b]}(f)$.

Ejercicio D2. Invariancia por reescalamiento del peso. Sean $f, w \in \mathcal{R}[a, b]$, con $\int_a^b w \neq 0$, y sea $\mu \neq 0$. Demuestra que $\mathrm{Avg}_{\mu w,[a,b]}(f) = \mathrm{Avg}_{w,[a,b]}(f)$.

Ejercicio D3. Un promedio ponderado de funciones escalonadas. En $[0, 2]$ define

$$f(x) = \begin{cases} 1, & 0 \leq x < 1, \\ 3, & 1 \leq x \leq 2, \end{cases} \qquad w(x) = \begin{cases} 2, & 0 \leq x < 1, \\ 1, & 1 \leq x \leq 2. \end{cases}$$

Calcula

$$\int_0^2 w, \qquad \int_0^2 fw, \qquad \mathrm{Avg}_{w,[0,2]}(f).$$

¿El promedio ponderado obtenido pertenece a la imagen de f?

Ejercicio D4. El teorema ponderado cuando el peso total es cero. En $[-1, 1]$ sea

$$w(x) = \begin{cases} 1, & x = 0, \\ 0, & x \neq 0. \end{cases}$$

y sea f continua.

1. Calcula $\int_{-1}^1 w$.
2. Calcula $\int_{-1}^1 fw$.
3. Verifica que la identidad factorizada de Teorema 15.2 es verdadera.
4. Explica por qué $\mathrm{Avg}_{w,[-1,1]}(f)$ no está definido.

Ejercicio D5. Realización de un promedio ponderado. Sea $f : [a, b] \to \mathbb{R}$ continua y sea $w \in \mathcal{R}[a, b]$ no negativa. Supón

$$\int_a^b w(x)\,dx = 5, \qquad \int_a^b f(x)w(x)\,dx = 15.$$

Demuestra que existe $c \in [a, b]$ tal que $f(c) = 3$.

Ejercicio D6. Un peso con cambio de signo puede romper la conclusión. Retoma las funciones de Ejemplo 15.3 en $[0,3]$:

$$f(x)=\begin{cases}0, & 0\le x\le 1,\\ x-1, & 1\le x\le 2,\\ 1, & 2\le x\le 3,\end{cases}$$

y

$$w(x)=\begin{cases}-1, & 0\le x<1,\\ 0, & 1\le x<2,\\ 2, & 2\le x\le 3.\end{cases}$$

Verifica que $\mathrm{Avg}_{w,[0,3]}(f)=2$ mientras que $f([0,3])=[0,1]$. Identifica exactamente qué hipótesis de Teorema 15.2 falla.

Nivel E — Contraejemplos y fallos lógicos. Ejercicio E1. Reparar una inferencia sobre integrabilidad. Analiza la frase:

> «Si f es Riemann-integrable en $[a,b]$, entonces existe $c\in[a,b]$ tal que $f(c)=\mathrm{Avg}_{[a,b]}(f)$.»

Decide si es verdadera o falsa. Si es falsa, da un contraejemplo y formula una versión correcta.

Ejercicio E2. El signo de una integral no determina el signo del peso. Analiza:

> «Si $\int_a^b w>0$, entonces $w\ge 0$ en todo $[a,b]$.»

Refútala mediante un peso escalonado explícito y explica por qué la condición $\int w>0$ no basta para aplicar Teorema 15.2.

Ejercicio E3. Cancelación. Analiza:

> «Si f es continua y $\int_0^1 f(x)\,dx=0$, entonces $f\equiv 0$.»

Refútala con una función afín y señala qué hipótesis adicional sí permite concluir identidad nula.

Ejercicio E4. Dos normalizaciones distintas. Refuta la identidad general $\mathrm{Avg}_{w,[a,b]}(f)=\mathrm{Avg}_{[a,b]}(fw)$ mediante el caso $[a,b]=[0,1]$, $f\equiv 1$ y $w\equiv 2$. Después escribe la condición algebraica que haría coincidir ambos cocientes cuando el numerador es no nulo.

Ejercicio E5. Existencia no es un método automático de evaluación. Un estudiante escribe:

> «Por el teorema del valor medio integral, $\int_a^b f=f(c)(b-a)$; por tanto, ya sabemos calcular cualquier integral de una función continua.»

Explica por qué el razonamiento es inválido. ¿Qué información no proporciona el teorema?

Ejercicio E6. Una identidad válida con un cociente inválido. Sea f continua y sea $w \geq 0$ Riemann-integrable con

$$\int_a^b w = 0.$$

Un estudiante divide inmediatamente la identidad ponderada por $\int_a^b w$. Explica el error y reconstruye el argumento correcto hasta donde permite Teorema 15.2.

Nivel F — Estimaciones, subintervalos y síntesis local. Ejercicio F1. Propagar un error uniforme. Sean $f, g \in \mathcal{R}[2, 7]$ y supón que

$$|f(x) - g(x)| \leq 0.04 \qquad (2 \leq x \leq 7).$$

Obtén cotas numéricas para

$$\left| \int_2^7 f - \int_2^7 g \right|$$

y para

$$\left| \mathrm{Avg}_{[2,7]}(f) - \mathrm{Avg}_{[2,7]}(g) \right|.$$

Ejercicio F2. Estimar sin integrar. Sea $f \in \mathcal{R}[a, b]$, con

$$b - a = 4, \qquad 3 \leq f(x) \leq 7.$$

Sin calcular la integral exactamente, determina el intervalo más inmediato garantizado por Proposición 15.2 para $\int_a^b f$.

Ejercicio F3. Encierro ponderado. Sean $f, w \in \mathcal{R}[a, b]$, con w de signo fijo y $\int_a^b w \neq 0$. Si $2 \leq f(x) \leq 5$, demuestra que $2 \leq \mathrm{Avg}_{w,[a,b]}(f) \leq 5$.

Explica por qué no hace falta calcular por separado ni $\int fw$ ni $\int w$.

Ejercicio F4. Valor medio en un subintervalo fijo. Sea $f : [a, b] \to \mathbb{R}$ continua y fija

$$a \leq u < v \leq b.$$

Demuestra que existe $c_{u,v} \in (u, v)$ tal que

$$\int_u^v f(x)\, dx = f(c_{u,v})(v - u).$$

Explica por qué esta afirmación no introduce todavía una función de acumulación.

Ejercicio F5. El promedio total queda entre dos promedios parciales. Sean $a < c < b$ y $f \in \mathcal{R}[a,b]$. Define

$$A = \operatorname{Avg}_{[a,c]}(f), \qquad B = \operatorname{Avg}_{[c,b]}(f).$$

Demuestra que $\operatorname{Avg}_{[a,b]}(f) = \frac{c-a}{b-a}A + \frac{b-c}{b-a}B$, y deduce

$$\min\{A,B\} \le \operatorname{Avg}_{[a,b]}(f) \le \max\{A,B\}.$$

Ejercicio F6. Estabilidad del promedio bajo perturbaciones. Sean $f, g \in \mathcal{R}[a,b]$ y $\varepsilon \ge 0$. Demuestra que

$$|f(x) - g(x)| \le \varepsilon \quad \text{para todo } x$$

implica

$$\left|\operatorname{Avg}_{[a,b]}(f) - \operatorname{Avg}_{[a,b]}(g)\right| \le \varepsilon.$$

¿Depende esta conclusión de la longitud del intervalo?

Nivel G — Desafíos integradores y reparación de demostraciones. Ejercicio G1. Caracterizar el caso extremal. Sea $f : [a,b] \to \mathbb{R}$ continua y define

$$m = \min_{[a,b]} f, \qquad M = \max_{[a,b]} f.$$

Demuestra la equivalencia

$$\operatorname{Avg}_{[a,b]}(f) = m \iff f \text{ es constante} \iff \operatorname{Avg}_{[a,b]}(f) = M.$$

Ejercicio G2. Igualdad en una cota integral. Sea $f : [a,b] \to \mathbb{R}$ continua y sea $M = \max_{[a,b]} f$.

Supón que

$$\int_a^b f(x)\,dx = M(b-a).$$

Demuestra que f es constante igual a M.

Ejercicio G3. Auditar una demostración ponderada. Sea f continua y $w \in \mathcal{R}[a,b]$ de signo fijo. Un borrador de prueba dice:

1. de $m \le f \le M$ se obtiene $m \int w \le \int fw \le M \int w$;
2. se divide por $\int w$;
3. por IVT se obtiene c con $\int fw = f(c) \int w$.

Repara la demostración separando todos los casos que el borrador oculta. Debes tratar explícitamente $w \ge 0$, $w \le 0$ y $\int w = 0$.

Ejercicio G4. Detectar el salto hacia el FTC. Sea $f : [a, b] \to \mathbb{R}$ continua. Un estudiante razona:

> «Para cada subintervalo fijo $[u, v]$, el teorema del valor medio produce $c_{u,v}$ con $\int_u^v f = f(c_{u,v})(v - u)$. Por tanto, ya está demostrado que una función de acumulación tiene derivada f.»

Explica exactamente por qué la segunda afirmación **no** se sigue todavía de la primera. Identifica qué objeto nuevo y qué tipo de paso analítico pertenecen a C17.

Soluciones desarrolladas.

Soluciones del nivel A. **Solución A1.** Por la integral de una constante,

$$\int_a^b f(x)\,dx = k(b - a).$$

Usando Definición 15.1, $\operatorname{Avg}_{[a,b]}(f) = \frac{k(b-a)}{b-a} = k$. Como $f(c) = k$ para todo $c \in [a, b]$, todos los puntos realizan el promedio. En particular, cualquier $c \in (a, b)$ satisface la conclusión de Teorema 15.1.

Solución A2. La longitud del intervalo es $6 - 2 = 4$. Por tanto, $\operatorname{Avg}_{[2,6]}(f) = \frac{14}{4} = \frac{7}{2}$.

La integrabilidad solo garantiza que este número está definido; no garantiza que pertenezca a la imagen de f. Para asegurar un punto realizador basta añadir que f sea continua en $[2, 6]$. Entonces Teorema 15.1 produce $c \in (2, 6)$ tal que $f(c) = \frac{7}{2}$.

Solución A3. Por linealidad de la integral,

$$\int_a^b (\alpha f + \beta g) = \alpha \int_a^b f + \beta \int_a^b g.$$

Dividiendo por $b - a > 0$,

$$\begin{aligned}\operatorname{Avg}_{[a,b]}(\alpha f + \beta g) &= \frac{1}{b-a}\left(\alpha \int_a^b f + \beta \int_a^b g\right) \\ &= \alpha \operatorname{Avg}_{[a,b]}(f) + \beta \operatorname{Avg}_{[a,b]}(g).\end{aligned}$$

Solución A4. Por aditividad de la integral,

$$\int_a^b f = \int_a^c f + \int_c^b f.$$

Cada integral puede escribirse como longitud por promedio:

$$\int_a^c f = (c - a) \operatorname{Avg}_{[a,c]}(f),$$

$$\int_c^b f = (b - c) \operatorname{Avg}_{[c,b]}(f),$$

y

$$\int_a^b f = (b - a) \operatorname{Avg}_{[a,b]}(f).$$

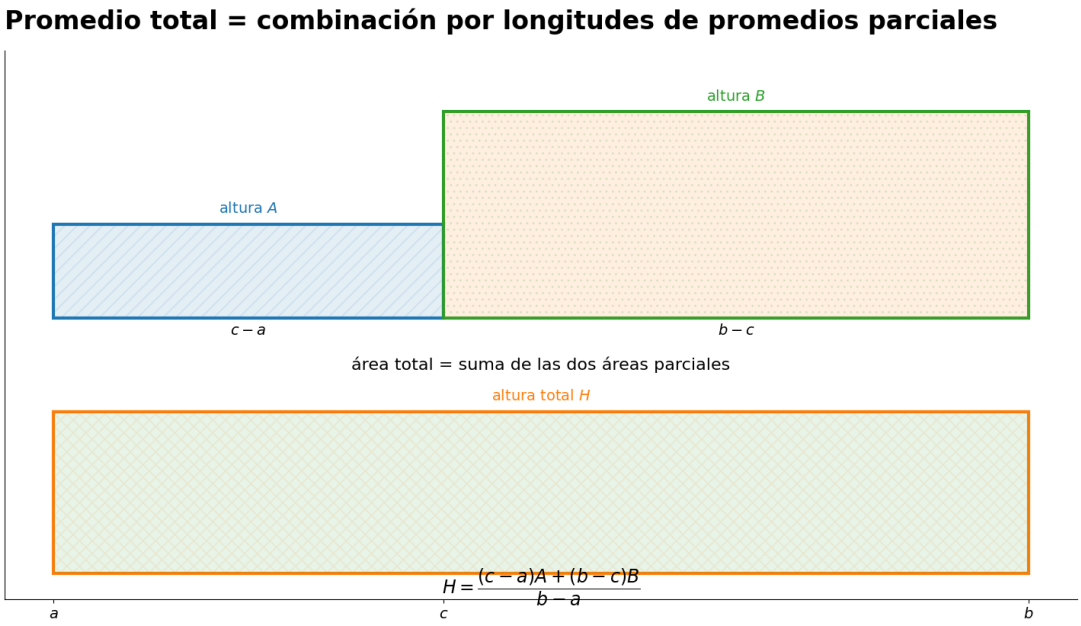


FIGURA 15.9. El promedio sobre $[a,b]$ es la combinación de los promedios parciales sobre $[a,c]$ y $[c,b]$, ponderada por las longitudes respectivas.

Sustituyendo se obtiene la identidad pedida. Al dividir por $b-a$,

$$\operatorname{Avg}_{[a,b]}(f) = \frac{c-a}{b-a}\operatorname{Avg}_{[a,c]}(f) + \frac{b-c}{b-a}\operatorname{Avg}_{[c,b]}(f).$$

Los coeficientes son positivos y suman 1.

La identidad de A4 queda representada en Figura 15.9.

Solución A5. Del dato sobre $[0,2]$,

$$\int_0^2 f = 2\cdot 1 = 2.$$

Del dato sobre $[2,5]$,

$$\int_2^5 f = 3\cdot 4 = 12.$$

Por aditividad,

$$\int_0^5 f = 14.$$

Como la longitud total es 5,

$$\boxed{\operatorname{Avg}_{[0,5]}(f) = \frac{14}{5}.}$$

Solución A6. La función es escalonada. Por tanto,

$$\int_0^3 s(x)\,dx = 1(1-0) + 4(3-1) = 1 + 8 = 9.$$

Así, $\mathrm{Avg}_{[0,3]}(s) = \frac{9}{3} = 3$.

Pero $s([0,3]) = \{1, 4\}$, de modo que no existe c con $s(c) = 3$. Falta continuidad: s tiene un salto en 1, así que Teorema 15.1 no es aplicable.

Soluciones del nivel B. **Solución B1.** El intervalo tiene longitud 3. Por Proposición 15.2,

$$2 \cdot 3 \le \int_1^4 f(x)\,dx \le 5 \cdot 3,$$

es decir,

$$\boxed{6 \le \int_1^4 f \le 15.}$$

Al normalizar,

$$\boxed{2 \le \mathrm{Avg}_{[1,4]}(f) \le 5.}$$

Solución B2. De $|f(x) - 3| \le \frac{1}{5}$ se sigue $\frac{14}{5} \le f(x) \le \frac{16}{5}$.

Por Proposición 15.1,

$$\boxed{\frac{14}{5} \le \mathrm{Avg}_{[0,10]}(f) \le \frac{16}{5}.}$$

Multiplicando por la longitud 10,

$$\boxed{28 \le \int_0^{10} f(x)\,dx \le 32.}$$

Solución B3. Supongamos, por contradicción, que $h \not\equiv 0$. Como $h \ge 0$, existe x_0 con $h(x_0) > 0$. La continuidad permite extender esa positividad desde un punto a un pequeño subintervalo: existe un intervalo no degenerado $[u, v] \subset [a, b]$ y $\eta > 0$ tales que

$$h(x) \ge \eta \qquad (x \in [u, v]).$$

Entonces, por monotonía,

$$\int_u^v h(x)\,dx \ge \eta(v - u) > 0.$$

Como $h \ge 0$ también fuera de $[u, v]$, la aditividad da

$$\int_a^b h(x)\,dx > 0,$$

contradicción. Por tanto, $h \equiv 0$.

La continuidad se usa exactamente para transformar $h(x_0) > 0$ en una cota positiva sobre un intervalo de longitud positiva.

Solución B4. Sea $h := g - f$.

Entonces h es continua, $h \geq 0$ y, por la desigualdad estricta en algún punto, $h \not\equiv 0$. Por Lema 15.1,

$$\int_a^b h(x)\,dx > 0.$$

Usando linealidad,

$$\int_a^b g - \int_a^b f > 0,$$

por lo que

$$\boxed{\int_a^b f < \int_a^b g.}$$

Solución B5. Como f es continua en el compacto $[a, b]$, alcanza m y M. Al no ser constante, existen puntos donde $m < f(x)$ y puntos donde $f(x) < M$.

Aplicando la comparación estricta de Corolario 15.1 a la función constante m y a f,

$$m(b - a) < \int_a^b f.$$

Aplicándola a f y a la función constante M,

$$\int_a^b f < M(b - a).$$

Dividir por $b - a > 0$ da

$$\boxed{m < \operatorname{Avg}_{[a,b]}(f) < M.}$$

Solución B6. La función satisface $h \geq 0$ y $h(0) = 1$, de modo que no es idénticamente nula. Sin embargo, difiere de la función cero en un único punto. Por la invariancia de la integral bajo modificaciones finitas,

$$\int_{-1}^{1} h(x)\,dx = 0.$$

No contradice Lema 15.1 porque h no es continua en 0. Precisamente falta el mecanismo que ensancha la positividad de un punto a un subintervalo.

Soluciones del nivel C. **Solución C1.** La longitud de $[1,5]$ es 4, así que $\operatorname{Avg}_{[1,5]}(f) = \frac{12}{4} = 3$.
Como f es continua, Teorema 15.1 garantiza un punto $c \in (1,5)$ con

$$\boxed{f(c) = 3.}$$

Solución C2. El promedio es $\operatorname{Avg}_{[0,1]}(f) = \frac{1}{2}$.
Como $f(c) = c$, la ecuación $f(c) = \frac{1}{2}$ equivale a $c = \frac{1}{2}$.
Por tanto, el único punto realizador es

$$\boxed{c = \frac{1}{2}.}$$

Solución C3. Para toda constante 7,

$$\int_a^b 7\,dx = 7(b-a).$$

Por tanto, para cualquier $c \in (a,b)$,

$$\int_a^b f(x)\,dx = 7(b-a) = f(c)(b-a).$$

El conjunto de puntos realizadores es todo (a,b). El teorema es existencial: afirma que existe al menos uno, no que sea único.

Solución C4. La función es escalonada y las dos mitades tienen igual longitud. Por tanto,

$$\int_0^1 r(x)\,dx = (-1)\frac{1}{2} + 1\frac{1}{2} = 0.$$

Luego $\operatorname{Avg}_{[0,1]}(r) = 0$.
Sin embargo, $r([0,1]) = \{-1, 1\}$, así que $r(c) \neq 0$ para todo c. La continuidad es esencial para realizar el promedio.

Solución C5. Por Teorema 15.1 existe $c \in (a,b)$ tal que

$$\int_a^b f(x)\,dx = f(c)(b-a).$$

Como $b - a > 0$:

1. si la integral es cero, necesariamente $f(c) = 0$;
2. si la integral es positiva, necesariamente $f(c) > 0$.

La segunda conclusión solo afirma la existencia de **algún** valor positivo. La función puede ser negativa en otras regiones y aun tener acumulación neta positiva.

Solución C6. La existencia viene de Teorema 15.1, pues f es continua. Si hubiera dos puntos distintos $c_1, c_2 \in (a,b)$ que realizaran el promedio, entonces $f(c_1) = \operatorname{Avg}_{[a,b]}(f) = f(c_2)$.
Pero una función estrictamente monótona es inyectiva. Por tanto, $c_1 = c_2$. Así el punto realizador existe y es único.

Soluciones del nivel D. **Solución D1.** Por linealidad,

$$\int_a^b f(x)w(x)\,dx = \lambda \int_a^b f(x)\,dx,$$

mientras

$$\int_a^b w(x)\,dx = \lambda(b-a).$$

Como $\lambda \neq 0$,

$$\begin{aligned}\operatorname{Avg}_{w,[a,b]}(f) &= \frac{\lambda \int_a^b f}{\lambda(b-a)} \\ &= \frac{1}{b-a}\int_a^b f \\ &= \operatorname{Avg}_{[a,b]}(f).\end{aligned}$$

Solución D2. Como $\mu \neq 0$,

$$\int_a^b \mu w = \mu \int_a^b w \neq 0.$$

Entonces

$$\begin{aligned}\operatorname{Avg}_{\mu w,[a,b]}(f) &= \frac{\int_a^b f(\mu w)}{\int_a^b \mu w} \\ &= \frac{\mu \int_a^b fw}{\mu \int_a^b w} \\ &= \operatorname{Avg}_{w,[a,b]}(f).\end{aligned}$$

Solución D3. El peso es escalonado:

$$\int_0^2 w = 2(1) + 1(1) = 3.$$

El producto fw vale 2 en la primera mitad y 3 en la segunda:

$$\int_0^2 fw = 2(1) + 3(1) = 5.$$

Por tanto,

$$\boxed{\operatorname{Avg}_{w,[0,2]}(f) = \frac{5}{3}.}$$

Pero $f([0,2]) = \{1,3\}$, así que $5/3$ no pertenece a la imagen. No hay contradicción con Teorema 15.2 porque allí se exige que f sea continua, y esta f tiene un salto.

El cálculo y la falta de realización puntual se comparan en Figura 15.10.

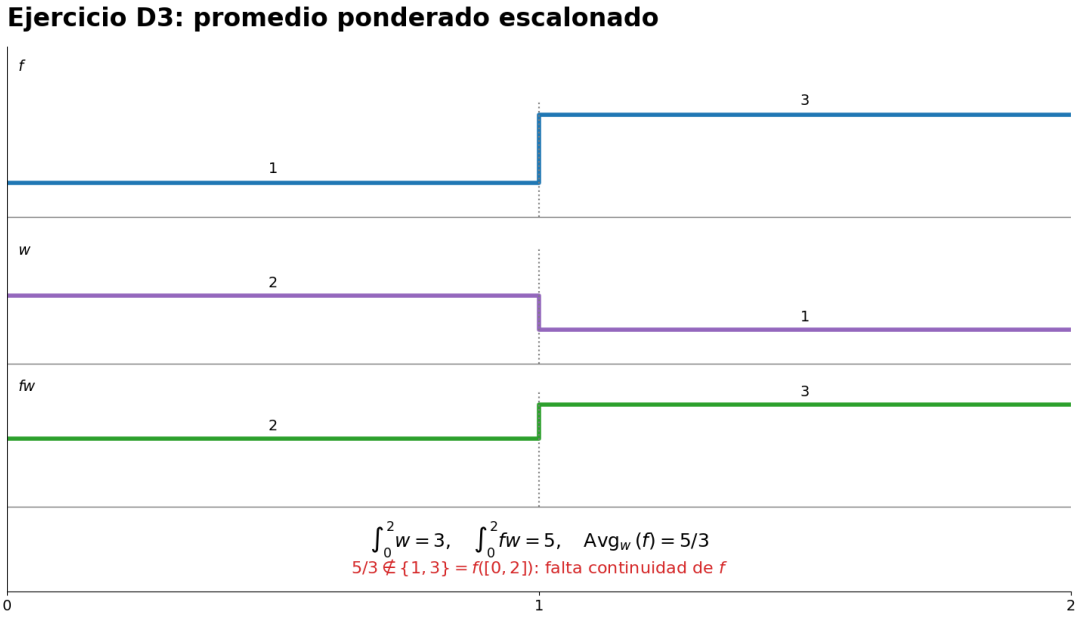


FIGURA 15.10. En D3, el peso escalonado da $\int_0^2 w = 3$ y $\int_0^2 fw = 5$, por lo que el promedio ponderado vale $5/3$; ese número no pertenece a la imagen $\{1, 3\}$ de la función discontinua.

Solución D4. El peso difiere de cero solo en un punto, así que

$$\int_{-1}^{1} w(x)\,dx = 0.$$

El producto fw también difiere de cero a lo sumo en ese punto; por tanto,

$$\int_{-1}^{1} f(x)w(x)\,dx = 0.$$

La identidad ponderada queda $0 = f(c)\,0$, válida para cualquier $c \in [-1, 1]$.
En cambio,

$$\mathrm{Avg}_{w,[-1,1]}(f) = \frac{\int fw}{\int w}$$

no está definido porque el denominador es cero. La identidad factorizada puede seguir teniendo sentido aunque el cociente no exista.

Solución D5. Como

$$\int_a^b w = 5 \neq 0,$$

el promedio ponderado está definido y vale $\mathrm{Avg}_{w,[a,b]}(f) = \frac{15}{5} = 3$. El peso es no negativo y f es continua, de modo que Teorema 15.2 garantiza un $c \in [a, b]$ tal que

$$\boxed{f(c) = 3.}$$

Solución D6. Por cálculo de funciones escalonadas,

$$\int_0^3 w = (-1)(1) + 0(1) + 2(1) = 1.$$

Además, $fw = 0$ en $[0, 2)$ y $fw = 2$ en $[2, 3]$, salvo valores aislados irrelevantes, de modo que

$$\int_0^3 fw = 2.$$

Así, $\mathrm{Avg}_{w,[0,3]}(f) = \frac{2}{1} = 2$.

La función f es continua y su imagen es $[0, 1]$, por lo que ese promedio no puede ser $f(c)$. Falla exactamente la hipótesis de **signo fijo del peso**: w toma valores negativos y positivos.

Soluciones del nivel E. **Solución E1.** La frase es falsa. La función escalón de Ejemplo 15.3,

$$s(x) = \begin{cases} 0, & 0 \le x < 1/2, \\ 1, & 1/2 \le x \le 1, \end{cases}$$

es Riemann-integrable y tiene promedio $1/2$, pero su imagen es $\{0, 1\}$. No existe punto realizador. La versión correcta es: si f es **continua** en $[a, b]$, entonces Teorema 15.1 garantiza $c \in (a, b)$ con $f(c) = \mathrm{Avg}_{[a,b]}(f)$.

Solución E2. En $[-1, 1]$ toma

$$w(x) = \begin{cases} -1, & -1 \le x < 0, \\ 2, & 0 \le x \le 1. \end{cases}$$

Entonces

$$\int_{-1}^1 w = -1 + 2 = 1 > 0,$$

pero w cambia de signo. Por tanto,

$$\int w > 0 \not\Longrightarrow w \ge 0.$$

Teorema 15.2 necesita una condición punto a punto, $w \ge 0$ o $w \le 0$; el signo del peso total no la sustituye.

Solución E3. Toma $f(x) = 2x - 1$.

Es continua y, usando $\int_0^1 x\,dx = 1/2$,

$$\int_0^1 (2x - 1)\,dx = 2 \cdot \frac{1}{2} - 1 = 0.$$

Pero f no es idénticamente nula. La integral cero puede resultar de cancelación. Si además $f \ge 0$ en todo el intervalo, entonces continuidad + no negatividad permiten usar Corolario 15.1 y concluir

$$\int f = 0 \Longrightarrow f \equiv 0.$$

Solución E4. En $[0,1]$, con $f \equiv 1$ y $w \equiv 2$,

$$\operatorname{Avg}_{w,[0,1]}(f) = \frac{\int_0^1 2}{\int_0^1 2} = 1,$$

mientras

$$\operatorname{Avg}_{[0,1]}(fw) = \int_0^1 2\,dx = 2.$$

Por tanto, no son el mismo objeto.

Si $\int fw \neq 0$, los cocientes

$$\frac{\int fw}{\int w} \qquad \text{y} \qquad \frac{\int fw}{b-a}$$

coinciden exactamente cuando

$$\int_a^b w = b - a.$$

Solución E5. El teorema garantiza la existencia de algún c con

$$\int_a^b f = f(c)(b-a),$$

pero no dice cuál es ese c ni proporciona un procedimiento para encontrarlo. En general, si desconocemos la integral, también desconocemos la altura promedio y el punto que la realiza. El teorema es útil para inferencias de existencia, signo y encierro; no reemplaza automáticamente el problema de evaluar una integral por un cálculo explícito de c.

Solución E6. Dividir por

$$\int_a^b w = 0$$

es ilegítimo. El teorema ponderado debe usarse primero en su forma factorizada:

$$\int_a^b fw = f(c)\int_a^b w.$$

Como $w \geq 0$ y su integral es cero, la prueba de Teorema 15.2 muestra además que

$$\int_a^b fw = 0.$$

La identidad queda $0 = f(c)\,0$, que es válida. Pero el cociente

$$\frac{\int fw}{\int w}$$

es $0/0$ y no define ningún promedio ponderado. El argumento correcto termina en la identidad factorizada.

Soluciones del nivel F. **Solución F1.** La longitud del intervalo es $7-2=5$.
Por Proposición 15.2,

$$\left|\int_2^7 f - \int_2^7 g\right| \le 0.04 \cdot 5 = 0.20.$$

Para los promedios, la longitud se cancela:

$$\boxed{\left|\mathrm{Avg}_{[2,7]}(f) - \mathrm{Avg}_{[2,7]}(g)\right| \le 0.04.}$$

Solución F2. Por la cota integral,

$$3(b-a) \le \int_a^b f \le 7(b-a).$$

Como $b-a=4$,

$$\boxed{12 \le \int_a^b f(x)\,dx \le 28.}$$

Solución F3. Si $w \ge 0$, multiplicar $2 \le f \le 5$ por w conserva el orden: $2w \le fw \le 5w$.
Al integrar,

$$2\int w \le \int fw \le 5\int w.$$

Como $\int w \neq 0$ y $w \ge 0$, tenemos $\int w > 0$, así que podemos dividir:

$$2 \le \frac{\int fw}{\int w} \le 5.$$

Si $w \le 0$, sustituimos $v = -w \ge 0$; el promedio ponderado no cambia al multiplicar el peso por -1. Por tanto,

$$\boxed{2 \le \mathrm{Avg}_{w,[a,b]}(f) \le 5.}$$

Solo se usaron desigualdades y normalización; no fue necesario evaluar las dos integrales por separado.

Solución F4. La restricción $f|_{[u,v]}$ es continua. Aplicando Teorema 15.1 al intervalo fijo $[u,v]$, existe $c_{u,v} \in (u,v)$ tal que

$$\boxed{\int_u^v f(x)\,dx = f(c_{u,v})(v-u).}$$

Aquí u y v son números fijados de antemano. No se ha definido una función cuyo extremo superior varíe, ni se ha estudiado la derivada de tal función. Eso pertenece a C17.

Solución F5. Por la identidad de A4,

$$(b-a)\,\mathrm{Avg}_{[a,b]}(f) = (c-a)A + (b-c)B.$$

Dividiendo por $b-a$, $\mathrm{Avg}_{[a,b]}(f) = \lambda A + (1-\lambda)B$, donde $\lambda = \frac{c-a}{b-a} \in (0,1)$.

La geometría de esta combinación por longitudes es la misma de Figura 15.9. Una combinación convexa de dos números queda entre ellos. Así,

$$\boxed{\text{mín}\{A,B\} \le \mathrm{Avg}_{[a,b]}(f) \le \text{máx}\{A,B\}.}$$

Solución F6. De $|f-g| \le \varepsilon$ se sigue $-\varepsilon \le f-g \le \varepsilon$.

Integrando,

$$-\varepsilon(b-a) \le \int_a^b (f-g) \le \varepsilon(b-a).$$

Por tanto,

$$\left|\int_a^b f - \int_a^b g\right| \le \varepsilon(b-a).$$

Dividiendo por $b-a$,

$$\boxed{\left|\mathrm{Avg}_{[a,b]}(f) - \mathrm{Avg}_{[a,b]}(g)\right| \le \varepsilon.}$$

La cota final no depende de la longitud del intervalo: la normalización elimina exactamente ese factor.

Soluciones del nivel G. **Solución G1.** Si f es constante, su único valor coincide con m, con M y con su promedio, de modo que ambas igualdades son inmediatas.

Recíprocamente, si f no fuera constante, Corolario 15.1 daría $m < \mathrm{Avg}_{[a,b]}(f) < M$.

Esto excluye tanto $\mathrm{Avg}_{[a,b]}(f) = m$ como $\mathrm{Avg}_{[a,b]}(f) = M$.

Por tanto, cualquiera de las dos igualdades extremales fuerza que f sea constante, y queda demostrada la equivalencia.

Solución G2. Dividiendo la hipótesis por $b-a > 0$, $\mathrm{Avg}_{[a,b]}(f) = M$.

Si f no fuera constante, por Corolario 15.1 tendríamos $\mathrm{Avg}_{[a,b]}(f) < M$, contradicción. Luego f es constante. Como M es su máximo,

$$\boxed{f(x) = M \quad \text{para todo } x \in [a,b].}$$

Equivalentemente, puede aplicarse comparación estricta entre f y la constante M.

Solución G3. Primero supongamos $w \geq 0$. Si

$$m = \min f, \qquad M = \max f,$$

entonces $mw \leq fw \leq Mw$.
Integrando,

$$m \int w \leq \int fw \leq M \int w.$$

Ahora hay que separar casos.

- Si $\int w = 0$, la cadena obliga a
$$\int fw = 0.$$
Por tanto,
$$\int fw = f(c) \int w$$
para cualquier $c \in [a, b]$. No se divide.
- Si $\int w > 0$, se puede dividir:
$$m \leq \frac{\int fw}{\int w} \leq M.$$
Como f es continua y alcanza todos los valores entre m y M, el IVT proporciona $c \in [a, b]$ con
$$f(c) = \frac{\int fw}{\int w}.$$
Multiplicando se obtiene la identidad factorizada.

Si $w \leq 0$, definimos $v = -w \geq 0$ y aplicamos el caso anterior a v. La identidad resultante, multiplicada por -1, devuelve

$$\int fw = f(c) \int w.$$

El borrador original ocultaba precisamente el caso de denominador cero y el cambio de sentido que surgiría al dividir directamente cuando $\int w < 0$.

Solución G4. La afirmación sobre cada subintervalo fijo dice únicamente que, dados números $u < v$, existe algún punto $c_{u,v}$ con

$$\int_u^v f = f(c_{u,v})(v - u).$$

No se ha definido todavía una función del tipo

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

y tampoco se ha estudiado ningún cociente incremental de F. Para demostrar que una función de acumulación tiene derivada f haría falta introducir ese nuevo objeto con extremo variable y analizar un límite de cocientes que compare integrales sobre intervalos que se hacen pequeños. Ese paso analítico pertenece a C17. Además, Teorema 15.1 no proporciona una elección única ni una regla de continuidad para $c_{u,v}$ cuando cambia el subintervalo. Por tanto, la existencia de alturas medias en intervalos fijos no es todavía una demostración del FTC.

Capítulo 16

Teorema fundamental del cálculo

16.1. De un intervalo fijo a un extremo variable

Hasta ahora la integral de Riemann ha aparecido, sobre todo, como un **número** asociado a una función y a dos extremos fijados. Una vez elegidos $a < b$ y una función integrable $f : [a, b] \to \mathbb{R}$, la expresión

$$\int_a^b f(t)\,dt$$

produce un único valor real. Al fijar el extremo izquierdo y variar el derecho, cada x determina un valor integral. Esa asignación define una función, cuyo incremento podremos estudiar mediante la aditividad:

$$x \longmapsto \int_a^x f(t)\,dt.$$

La función del extremo convierte las estimaciones integrales en información sobre continuidad y derivabilidad. Para aplicarlas será necesario distinguir la variable de entrada de la variable de integración.

Cuatro papeles que conviene separar. En la expresión

$$\int_a^x f(t)\,dt,$$

intervienen objetos que cumplen funciones distintas.

- a es el **punto base fijo**;
- x es el **extremo variable** y será la entrada de la nueva función;
- t es la **variable de integración**: vive dentro del símbolo integral y no es una nueva entrada libre;
- f es la función cuyas contribuciones estamos acumulando entre a y x.

La letra t podría reemplazarse por cualquier otra letra que no cree conflicto:

$$\int_a^x f(t)\,dt = \int_a^x f(u)\,du.$$

Cambiar el nombre de la variable de integración conserva el número obtenido para cada x. El extremo variable sigue siendo el argumento libre que determina la función de acumulación.

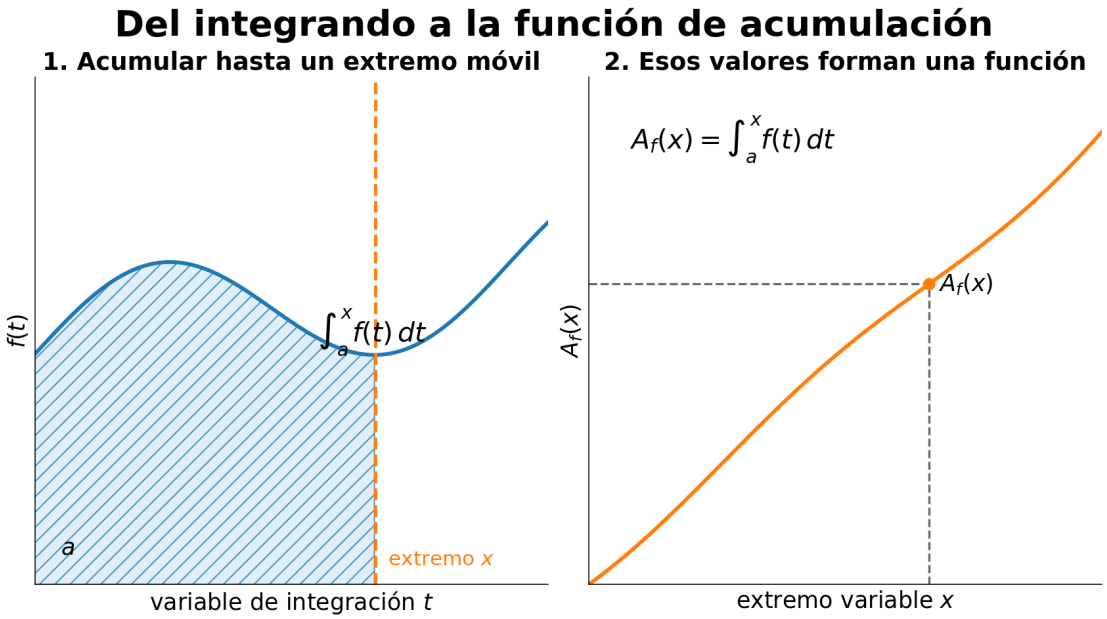


FIGURA 16.1. De un integrando a su función de acumulación: al mover el extremo derecho x, la integral desde el punto base a produce el valor $A_f(x)$; la variable t permanece ligada dentro de la integral.

La misma letra no debe cumplir dos papeles
Cuando estudiamos la dependencia respecto de x, escribiremos

$$\int_a^x f(t)\,dt$$

en lugar de reutilizar x como variable de integración. Así queda visible qué símbolo está ligado por la integral y cuál permanece libre como argumento de la nueva función.

La función de acumulación. La teoría de C15 garantiza que las restricciones de una función Riemann-integrable a subintervalos siguen siendo integrables (Teorema 14.7). Además, la convención de integral orientada establece que una integral con extremos iguales vale cero (Definición 14.3). Por tanto, para cada $x \in [a,b]$ la integral desde a hasta x está bien definida.

DEFINICIÓN 16.1. **Definición (función de acumulación).** Sean $a < b$ y $f \in \mathcal{R}[a,b]$. La **función de acumulación de f con punto base a** es la función

$$A_f : [a,b] \longrightarrow \mathbb{R}$$

definida por

$$\boxed{A_f(x) := \int_a^x f(t)\,dt.}$$

Mientras el punto base a permanezca fijado, omitiremos ese dato del subíndice y escribiremos simplemente A_f.

Figura 16.1 separa visualmente el punto base fijo, el extremo variable, la variable muda de integración y la nueva función A_f.

La definición transforma una construcción integral en una función ordinaria. Para cada entrada x obtenemos el número que resulta de integrar f desde el punto base hasta esa entrada. Dos valores son inmediatos. En el punto base,

$$\boxed{A_f(a) = 0,}$$

porque Definición 14.3 fija $\int_a^a f = 0$. En el extremo derecho,

$$\boxed{A_f(b) = \int_a^b f(t)\,dt.}$$

Así, la integral definida sobre todo $[a, b]$ aparece como **un valor particular** de una función más rica.

Mover el extremo añade exactamente una franja integral. La aditividad orientada de Proposición 14.3 permite comparar dos valores cualesquiera de la acumulación. Si $x, y \in [a, b]$, entonces

$$\int_a^y f(t)\,dt = \int_a^x f(t)\,dt + \int_x^y f(t)\,dt.$$

En la notación recién introducida,

$$\boxed{A_f(y) - A_f(x) = \int_x^y f(t)\,dt.}$$

Esta identidad no afirma todavía que A_f sea continua ni diferenciable. Solo dice, exactamente, cuánto cambia la acumulación cuando el extremo se desplaza de x a y: el incremento de A_f es la integral de f sobre el tramo añadido —con el signo correspondiente si se invierte la orientación—. La identidad expresa el incremento de la acumulación como una integral sobre el tramo añadido o retirado. Una estimación de esa integral dará continuidad; el cociente por $y - x$ permitirá investigar la derivada cuando el tramo se contraiga.

Dos funciones pueden tener la misma integral total y acumular de manera distinta.

Ejemplo 16.1. **Mismo valor final, distinta función de acumulación.** Consideremos en $[0, 1]$ la función escalón de Ejemplo 14.9,

$$s(t) = \begin{cases} 0, & 0 \le t < \frac{1}{2}, \\ 1, & \frac{1}{2} \le t \le 1. \end{cases}$$

Ya sabemos que

$$\int_0^1 s(t)\,dt = \frac{1}{2}.$$

Su función de acumulación con base 0 es

$$A_s(x) = \int_0^x s(t)\,dt.$$

Si $0 \le x \le 1/2$, la restricción solo difiere de la función cero, a lo sumo, en el extremo $1/2$, y la invariancia ante una modificación finita de Corolario 14.1 da $A_s(x) = 0$.

Si $1/2 \le x \le 1$, la aditividad produce

$$\begin{aligned} A_s(x) &= \int_0^{1/2} s(t)\,dt + \int_{1/2}^{x} s(t)\,dt \\ &= 0 + (x - \tfrac{1}{2}). \end{aligned}$$

Por tanto,

$$\boxed{A_s(x) = \begin{cases} 0, & 0 \le x \le \tfrac{1}{2}, \\ x - \tfrac{1}{2}, & \tfrac{1}{2} \le x \le 1. \end{cases}}$$

Comparemos con la función constante $c(t) = \frac{1}{2}$.

También satisface

$$\int_0^1 c(t)\,dt = \frac{1}{2},$$

pero su acumulación es

$$A_c(x) = \int_0^x \frac{1}{2}\,dt = \frac{x}{2}.$$

Las dos funciones tienen el mismo valor integral sobre el intervalo completo, $A_s(1) = A_c(1) = \frac{1}{2}$, pero $A_s \neq A_c$. Conocer únicamente $\int_0^1 f$ no determina **cómo** se va acumulando ese valor a medida que se mueve el extremo.

El ejemplo muestra por qué el nuevo objeto contiene más información que la integral total. La expresión

$$\int_a^b f$$

responde una sola pregunta; la función

$$x \longmapsto \int_a^x f$$

responde simultáneamente una pregunta para cada $x \in [a, b]$.

La primera pregunta analítica del capítulo. Hemos definido A_f usando únicamente resultados ya establecidos sobre la integral de Riemann. Hasta aquí **no** hemos supuesto que f sea continua y tampoco hemos probado ninguna regularidad nueva para A_f.

La identidad

$$A_f(y) - A_f(x) = \int_x^y f(t)\,dt$$

convierte ahora la cuestión en una estimación: si x y y están próximos, ¿podemos obligar a que $A_f(x)$ y $A_f(y)$ también lo estén?

Ésa será la tarea exclusiva de §16.2. Antes de hablar de derivadas, demostraremos que la función de acumulación posee una propiedad de continuidad que proviene directamente de la acotación del integrando.

Antes de seguir. En $A_f(x) = \int_a^x f(t)\,dt$, ¿qué símbolo permanece libre y cuál está ligado por la integral? ¿Por qué $A_f(a) = 0$ no exige ninguna información sobre el valor puntual $f(a)$?

Si $x < y$, ¿qué teorema permite identificar $A_f(y) - A_f(x)$ con la integral sobre $[x, y]$? ¿Cómo demuestra Ejemplo 16.1 que conocer la integral total no basta para conocer la función de acumulación?

16.2. La función de acumulación es continua antes de ser derivable

La sección anterior transformó la integral sobre un intervalo fijo en una función:

$$A_f(x) = \int_a^x f(t)\,dt.$$

La identidad fundamental que obtuvimos fue

$$A_f(y) - A_f(x) = \int_x^y f(t)\,dt.$$

Antes de intentar derivar esta función conviene resolver una pregunta más básica: **¿puede el valor acumulado cambiar mucho cuando movemos poco el extremo?**

La respuesta será más fuerte que la simple continuidad. Y, significativamente, no exigirá que el integrando sea continuo.

La integrabilidad ya contiene una cota global. Sea $f \in \mathcal{R}[a, b]$.

Por la propia definición de integrabilidad de Riemann, f es acotada. Por tanto existe algún número $M \geq 0$ tal que

$$|f(t)| \leq M \qquad (t \in [a, b]).$$

No necesitamos que M sea la mejor cota posible. Basta disponer de **una** cota uniforme válida en todo el intervalo. Esta información parece hablar solo de las alturas de f. Sin embargo, combinada con la identidad de incremento de §16.1, controla inmediatamente cuánto puede variar la función de acumulación.

Proposición 16.1. **Proposición (cota Lipschitz para la función de acumulación).** Sean $a < b$, $f \in \mathcal{R}[a, b]$ y

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

Si $M \geq 0$ satisface

$$|f(t)| \leq M \qquad (t \in [a, b]),$$

entonces, para cualesquiera $x, y \in [a, b]$,

$$\boxed{|A_f(y) - A_f(x)| \leq M|y - x|.}$$

En particular, A_f es uniformemente continua en $[a, b]$ y, por tanto, continua en cada punto del intervalo.

Demostración. Si $x = y$, ambos miembros son cero y la desigualdad es inmediata.

Supongamos primero que $x < y$. Por la identidad de §16.1,

$$A_f(y) - A_f(x) = \int_x^y f(t)\,dt.$$

La restricción de f a $[x, y]$ es Riemann integrable por Teorema 14.7. La desigualdad triangular integral Corolario 14.2 proporciona

$$|A_f(y) - A_f(x)| = \left| \int_x^y f(t)\, dt \right| \leq \int_x^y |f(t)|\, dt.$$

Como

$$|f(t)| \leq M \qquad (t \in [x, y]),$$

la monotonía de la integral en Teorema 14.6 da

$$\int_x^y |f(t)|\, dt \leq \int_x^y M\, dt.$$

La integral de la función constante vale

$$\int_x^y M\, dt = M(y - x).$$

Por consiguiente,

$$|A_f(y) - A_f(x)| \leq M(y - x) = M|y - x|.$$

Si $y < x$, intercambiamos los papeles de x e y. Como

$$|A_f(y) - A_f(x)| = |A_f(x) - A_f(y)|,$$

el caso ya demostrado produce

$$|A_f(y) - A_f(x)| \leq M(x - y) = M|y - x|.$$

Queda establecida la cota para todos los pares $x, y \in [a, b]$. Para deducir continuidad uniforme, fijemos $\varepsilon > 0$. Si $M = 0$, la desigualdad anterior obliga a $|A_f(y) - A_f(x)| = 0$ para todos x, y, de modo que A_f es constante y la conclusión es inmediata.

Si $M > 0$, tomamos $\delta := \frac{\varepsilon}{M}$.

Entonces, para cualesquiera $x, y \in [a, b]$,

$$|x - y| < \delta \quad \Longrightarrow \quad |A_f(y) - A_f(x)| \leq M|y - x| < M\delta = \varepsilon.$$

El mismo δ funciona simultáneamente para todos los puntos del intervalo. Por tanto A_f es uniformemente continua y, en particular, continua. □

La estimación de Proposición 16.1 queda representada en Figura 16.2: la continuidad de la acumulación proviene aquí de la acotación del integrando, no de su continuidad.

La estimación es más informativa que la continuidad. La conclusión

$$A_f(y) \longrightarrow A_f(x) \qquad (y \to x)$$

es solo una parte de la información obtenida. La proposición proporciona una cota cuantitativa:

desplazar el extremo una distancia $\lvert y - x \rvert$ cambia la acumulación a lo sumo en $M\lvert y - x \rvert$.

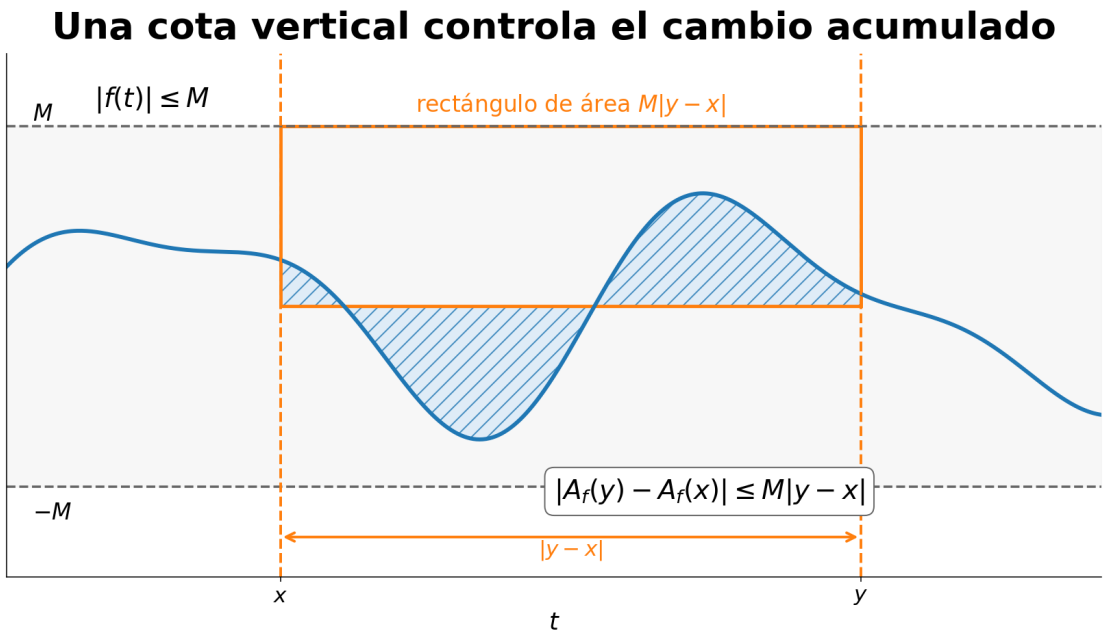


FIGURA 16.2. Si $|f| \le M$, el cambio acumulado entre x e y queda dominado por el rectángulo de base $|y-x|$ y altura M, lo que hace visible la cota $|A_f(y)-A_f(x)| \le M|y-x|$.

No aparece aquí ninguna derivada. Tampoco hemos necesitado que f varíe suavemente entre x e y. La única información puntual empleada ha sido la cota global $|f| \le M$. Esto explica por qué la regularidad de la acumulación puede ser mejor que la del integrando.

Un integrando discontinuo puede tener una acumulación continua. Volvamos a la función escalón de Ejemplo 16.1:

$$s(t) = \begin{cases} 0, & 0 \le t < \frac{1}{2}, \\ 1, & \frac{1}{2} \le t \le 1. \end{cases}$$

La función s es discontinua en $1/2$, pero es Riemann integrable y satisface $|s(t)| \le 1$. Por Proposición 16.1, su función de acumulación verifica $|A_s(y) - A_s(x)| \le |y - x|$ para todos $x, y \in [0,1]$.

En §16.1 calculamos además

$$A_s(x) = \begin{cases} 0, & 0 \le x \le \frac{1}{2}, \\ x - \frac{1}{2}, & \frac{1}{2} \le x \le 1. \end{cases}$$

La fórmula concreta confirma lo que el teorema ya garantiza: A_s es continua incluso en el punto donde s presenta su salto. La lección lógica es importante:

$$\boxed{f \in \mathcal{R}[a,b] \quad \Longrightarrow \quad A_f \text{ es continua},}$$

mientras que $f \in \mathcal{R}[a,b] \not\Longrightarrow f$ es continua.

Por tanto, la continuidad de A_f no puede explicarse diciendo simplemente que «la integral de una función continua cambia continuamente». Esa afirmación sería demasiado débil: la acumulación ya es continua para **todo integrando Riemann integrable**.

Continuidad de la acumulación y continuidad del integrando son preguntas diferentes. Conviene separar tres niveles.

1. **Integrabilidad de f.** Garantiza que f es acotada y que las integrales sobre subintervalos están definidas.
2. **Continuidad de A_f.** Se obtiene de la cota integral de esta sección y no requiere continuidad de f.
3. **Diferenciabilidad de A_f.** No ha sido establecida todavía. Para relacionar la derivada de A_f con el valor puntual $f(c)$ necesitaremos información adicional acerca de f cerca de c.

Esta separación evita una conclusión prematura. De A_f continua no podemos deducir todavía $A_f'(c) = f(c)$. La primera afirmación controla diferencias de valores de A_f; la segunda exige estudiar un cociente.

El cociente que abre la puerta al FTC. Si c y $c+h$ pertenecen a $[a,b]$ y $h \neq 0$, la identidad de §16.1 da

$$A_f(c+h) - A_f(c) = \int_c^{c+h} f(t)\,dt.$$

Por tanto,

$$\boxed{\frac{A_f(c+h) - A_f(c)}{h} = \frac{1}{h}\int_c^{c+h} f(t)\,dt.}$$

El miembro derecho es un **promedio integral local** sobre un intervalo que se contrae hacia c. La cota Lipschitz demuestra que este cociente permanece acotado por M en valor absoluto, pero eso no identifica todavía su límite. La pregunta siguiente será mucho más precisa:

$$\text{si } f \text{ es continua en } c, \quad \text{¿ese promedio local tiende a } f(c)?$$

Responderla será la tarea exclusiva de §16.3.

Antes de seguir. ¿Dónde se usa exactamente que una función Riemann integrable sea acotada? ¿Por qué la prueba de Proposición 16.1 necesita la desigualdad triangular integral antes de aplicar monotonía? ¿Qué demuestra Ejemplo 16.1 acerca de la relación entre continuidad del integrando y continuidad de su acumulación? ¿Por qué la cota $|A_f(y) - A_f(x)| \leq M|y-x|$ no basta, por sí sola, para concluir que $A_f'(c) = f(c)$?

16.3. Derivar una acumulación en un punto de continuidad

La sección anterior dejó preparado el cociente

$$\frac{A_f(c+h) - A_f(c)}{h} = \frac{1}{h}\int_c^{c+h} f(t)\,dt,$$

para incrementos admisibles $h \neq 0$. La forma recuerda a un promedio integral, pero ahora el intervalo de promedio se contrae hacia un único punto c. La pregunta es local. No necesitamos

saber que f sea continua en todo $[a,b]$; necesitamos saber qué ocurre con $f(t)$ cuando t se aproxima al punto concreto c.

El promedio local debe compararse con el valor central. Para descubrir el límite correcto, no basta acotar el cociente. Debemos compararlo con $f(c)$. Restando ese valor obtenemos una identidad exacta:

$$\begin{aligned}\frac{A_f(c+h)-A_f(c)}{h}-f(c) &= \frac{1}{h}\int_c^{c+h} f(t)\,dt - \frac{1}{h}\int_c^{c+h} f(c)\,dt \\ &= \frac{1}{h}\int_c^{c+h} \big(f(t)-f(c)\big)\,dt.\end{aligned}$$

La linealidad de la integral justifica el último paso. La expresión ya muestra la estrategia de la prueba: si $f(t)$ permanece cerca de $f(c)$ en todo el pequeño intervalo entre c y $c+h$, entonces el promedio del error debe ser pequeño. Éste es el punto exacto en el que entra la continuidad de f en c.

TEOREMA 16.1. **Teorema (forma local del teorema fundamental del cálculo).** Sean $a<b$ y $f\in\mathcal{R}[a,b]$. Definamos

$$A_f(x) := \int_a^x f(t)\,dt, \qquad x\in[a,b].$$

Sea $c\in[a,b]$. Si f es continua en c relativamente al dominio $[a,b]$, entonces A_f es diferenciable en c relativamente a $[a,b]$ y

$$\boxed{(A_f)'_{[a,b]}(c) = f(c).}$$

Si $c\in(a,b)$, esta es la derivada bilateral ordinaria. Si $c=a$ o $c=b$, se interpreta como la derivada relativa al dominio, es decir, mediante los incrementos admisibles del intervalo.

Demostración. Consideremos un incremento $h\neq 0$ tal que $c+h\in[a,b]$.

Por la identidad de incremento de §16.1,

$$A_f(c+h)-A_f(c) = \int_c^{c+h} f(t)\,dt.$$

Como la función constante $t\mapsto f(c)$ es integrable y la integral es lineal,

$$\frac{A_f(c+h)-A_f(c)}{h}-f(c) = \frac{1}{h}\int_c^{c+h}\big(f(t)-f(c)\big)\,dt.$$

Queremos demostrar que el miembro derecho tiende a 0 cuando $h\to 0$ a través de los incrementos admisibles. Sea $\varepsilon>0$. Como f es continua en c relativamente a $[a,b]$, existe $\delta>0$ tal que

$$t\in[a,b], \qquad |t-c|<\delta \quad\Longrightarrow\quad |f(t)-f(c)|<\varepsilon.$$

Tomemos ahora un incremento admisible h con $0<|h|<\delta$. Todo punto t situado entre c y $c+h$ pertenece a $[a,b]$ y satisface $|t-c|\le|h|<\delta$. Por tanto, $|f(t)-f(c)|<\varepsilon$ en todo el intervalo comprendido entre ambos extremos.

Escribamos

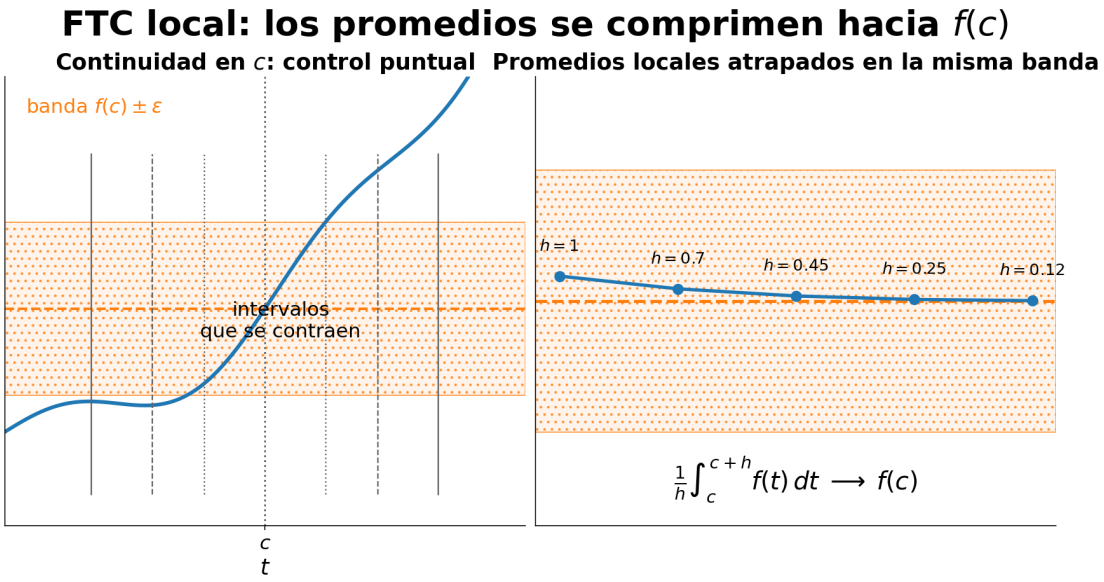


FIGURA 16.3. En un punto de continuidad c, intervalos cada vez más pequeños quedan dentro de una banda donde $f(c)-\varepsilon < f(t) < f(c)+\varepsilon$; el promedio integral local queda atrapado en la misma banda y converge a $f(c)$.

$$u_h := \min\{c, c+h\}, \qquad v_h := \max\{c, c+h\}.$$

La restricción de $f - f(c)$ a $[u_h, v_h]$ es Riemann integrable por Teorema 14.7 y la linealidad de Teorema 14.6. Usando la orientación de la integral, la desigualdad triangular integral Corolario 14.2 y la monotonía,

$$\begin{aligned}\left|\int_c^{c+h} (f(t)-f(c))\,dt\right| &= \left|\int_{u_h}^{v_h} (f(t)-f(c))\,dt\right| \\ &\le \int_{u_h}^{v_h} |f(t)-f(c)|\,dt \\ &\le \int_{u_h}^{v_h} \varepsilon\,dt \\ &= \varepsilon(v_h-u_h) \\ &= \varepsilon|h|.\end{aligned}$$

Dividiendo por $|h| > 0$ obtenemos

$$\left|\frac{A_f(c+h)-A_f(c)}{h} - f(c)\right| \le \varepsilon.$$

Como esto vale para todo incremento admisible suficientemente pequeño,

$$\lim_{\substack{h\to 0 \\ c+h\in[a,b]}} \frac{A_f(c+h)-A_f(c)}{h} = f(c).$$

Ésta es exactamente la derivada relativa de A_f en c. □

Figura 16.3 resume el mecanismo de la prueba de Teorema 16.1: se controla directamente el error integral; no se introduce un punto intermedio mediante el teorema del valor medio para integrales.

La conclusión es verdaderamente local. El teorema no exige continuidad global de f. Si una función Riemann-integrable tiene una discontinuidad lejos de c, esa discontinuidad no interviene en el argumento: al hacer $h \to 0$, el intervalo entre c y $c+h$ termina contenido en cualquier entorno suficientemente pequeño de c. Por eso la hipótesis debe leerse literalmente:

$$\boxed{f \text{ integrable en } [a,b] + f \text{ continua en } c \Longrightarrow (A_f)'(c) = f(c).}$$

La integrabilidad cumple una función global: garantiza que las integrales que definen A_f existen. La continuidad cumple una función local: identifica el límite del promedio integral que se contrae hacia c. La función escalón de Ejemplo 16.1 ilustra esta separación. Es discontinua en $1/2$, pero es continua en todos los demás puntos de $[0,1]$. Por tanto, Teorema 16.1 permite concluir inmediatamente que su acumulación satisface $(A_s)'(c) = s(c)$ en cada punto $c \neq 1/2$, interpretando los extremos relativamente al dominio. En el punto de salto el teorema simplemente **no autoriza ninguna conclusión**; el análisis específico de ese punto quedará para la auditoría de hipótesis posterior.

Una formulación equivalente: los promedios locales recuperan el valor puntual. Para todo punto de continuidad c, el teorema puede releerse como

$$\boxed{\lim_{\substack{h\to 0 \\ c+h\in[a,b] \\ h\neq 0}} \frac{1}{h}\int_c^{c+h} f(t)\,dt = f(c).}$$

Así aparece una relación fundamental entre dos escalas:

promedio integral sobre un intervalo pequeño $\longrightarrow$ valor puntual en el centro.

No hemos usado el teorema del valor medio integral de C16 para seleccionar un punto intermedio. Esa ruta exigiría continuidad en todo el pequeño subintervalo. La prueba anterior necesita únicamente continuidad en el punto c y controla directamente el error integral.

Qué falta para obtener una afirmación global. Teorema 16.1 habla de **un punto** c. Si sabemos que f es continua en cada punto de $[a,b]$, podremos aplicar el resultado simultáneamente a todos ellos y concluir que la función de acumulación es una primitiva de f en todo el intervalo.

Ese paso de una afirmación local, punto por punto, a la primera forma global del teorema fundamental será la tarea exclusiva de §16.4.

Antes de seguir. ¿Qué parte de la prueba usa la integrabilidad global de f y qué parte usa solamente continuidad en c? ¿Por qué se resta $f(c)$ antes de estimar el cociente? ¿Cómo permite la orientación tratar con una única prueba los incrementos positivos y negativos? ¿Por qué Teorema 16.1 sí se aplica a la función escalón de Ejemplo 16.1 fuera de su salto, pero no permite afirmar nada en el propio punto de discontinuidad?

16.4. Primera forma global del teorema fundamental

Teorema 16.1 resolvió el problema punto por punto: si una función Riemann-integrable es continua en un punto c, entonces la derivada de su función de acumulación existe en ese punto y recupera exactamente el valor $f(c)$.

La primera forma global del teorema fundamental aparece cuando esa hipótesis local se cumple **en todos los puntos del intervalo**. No necesitamos una idea nueva de prueba; necesitamos reconocer que el resultado local puede aplicarse sucesivamente a cada punto de $[a, b]$.

De la continuidad global a una identidad de funciones. Supongamos ahora que $f : [a, b] \longrightarrow \mathbb{R}$ es continua. Por Teorema 14.4, toda función continua sobre un intervalo compacto es Riemann integrable. Por tanto la función de acumulación de Definición 16.1,

$$A_f(x) := \int_a^x f(t)\,dt,$$

está bien definida para todo $x \in [a, b]$. Como f es continua en **cada** punto del intervalo, Teorema 16.1 puede aplicarse en cualquier $c \in [a, b]$. Esto produce la primera formulación global del teorema fundamental.

Teorema 16.2. **Teorema (primera forma global del teorema fundamental del cálculo).** Sean $a < b$ y $f : [a, b] \to \mathbb{R}$ continua. Definamos

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

Entonces A_f es diferenciable relativamente a $[a, b]$ en cada punto $c \in [a, b]$ y

$$\boxed{(A_f)'_{[a,b]}(c) = f(c) \qquad (c \in [a, b]).}$$

En particular, para todo $c \in (a, b)$ la derivada ordinaria existe y satisface

$$\boxed{A'_f(c) = f(c).}$$

Demostración. La continuidad de f en $[a, b]$ implica, por Teorema 14.4, que $f \in \mathcal{R}[a, b]$. Fijemos un punto arbitrario $c \in [a, b]$. Como f es continua en c, se cumplen todas las hipótesis de Teorema 16.1; por consiguiente, $(A_f)'_{[a,b]}(c) = f(c)$.

El punto c fue arbitrario, de modo que la igualdad vale en todo $[a, b]$. Si $c \in (a, b)$, la derivada relativa coincide con la derivada bilateral ordinaria. □

Qué significa aquí «global». La palabra **global** no indica que hayamos reemplazado la prueba local por una estimación uniforme en todo el intervalo. La lógica es más sencilla:

$$\boxed{\text{continuidad de } f \text{ en todo } [a, b] \Longrightarrow \text{continuidad de } f \text{ en cada } c \Longrightarrow (A_f)'(c) = f(c) \text{ para cada } c.}$$

La afirmación local de §16.3 sigue siendo el mecanismo esencial. La hipótesis global permite aplicarla sin excepciones. Esta distinción será importante más adelante. Si f es Riemann integrable pero tiene algunos puntos de discontinuidad, la función de acumulación sigue siendo continua por Proposición 16.1, y Teorema 16.1 sigue identificando su derivada en **cada punto donde f sí sea continua**. Lo que falla es la posibilidad de escribir una identidad derivativa válida automáticamente en todo el intervalo.

Interior y extremos: la misma afirmación, límites distintos. En el interior, el teorema afirma que para $c \in (a, b)$,

$$\lim_{h \to 0} \frac{A_f(c+h) - A_f(c)}{h} = f(c).$$

En el extremo izquierdo solo existen incrementos admisibles con $h > 0$ suficientemente pequeño. Por tanto,

$$\boxed{\lim_{h \to 0^+} \frac{A_f(a+h) - A_f(a)}{h} = f(a).}$$

En el extremo derecho solo existen incrementos admisibles con $h < 0$ suficientemente pequeño, y

$$\boxed{\lim_{h \to 0^-} \frac{A_f(b+h) - A_f(b)}{h} = f(b).}$$

Estas dos fórmulas no agregan hipótesis al teorema. Simplemente hacen explícito cómo se lee la derivada relativa cuando el dominio es el intervalo cerrado $[a, b]$.

La convención relativa evita una excepción artificial
Si restringiéramos la palabra «derivable» a la derivada bilateral, tendríamos que formular el resultado solo sobre (a, b) y añadir dos enunciados laterales separados para los extremos. La derivada relativa permite expresar en una sola fórmula la regularidad que realmente posee A_f sobre todo su dominio.

Integrar y después derivar recupera la función original. El contenido conceptual de Teorema 16.2 puede condensarse en el esquema

$$f \quad \xrightarrow{\text{acumular desde } a} \quad A_f(x) = \int_a^x f(t)\,dt \quad \xrightarrow{\text{derivar}} \quad f.$$

Para funciones continuas, la operación de formar la acumulación y luego derivar devuelve el integrando original:

$$\boxed{\frac{d}{dx}\left(\int_a^x f(t)\,dt\right) = f(x), \qquad x \in (a, b),}$$

con la interpretación relativa correspondiente en a y b. Ésta es una de las dos direcciones del teorema fundamental. Todavía no hemos demostrado la dirección inversa que permite evaluar una integral a partir de **otra** función cuya derivada sea f. Esa dirección requerirá formular con precisión qué entendemos por una primitiva y estudiar su falta de unicidad.

El punto base cambia la acumulación solo en una constante. El punto base a determina el valor inicial de la acumulación. Al derivarla en los puntos donde se cumplen las hipótesis del teorema recuperamos f, con independencia de ese punto base.

Sean $p, q \in [a, b]$ y definamos

$$A_p(x) := \int_p^x f(t)\,dt, \qquad A_q(x) := \int_q^x f(t)\,dt.$$

Por la aditividad orientada,

$$\int_p^x f(t)\,dt = \int_p^q f(t)\,dt + \int_q^x f(t)\,dt,$$

de donde

$$\boxed{A_q(x) = A_p(x) - \int_p^q f(t)\,dt.}$$

La cantidad

$$\int_p^q f(t)\,dt$$

no depende de x. Por tanto dos funciones de acumulación construidas a partir del mismo integrando, pero con puntos base distintos, difieren por una constante. Además, si f es continua, Teorema 16.2 da $(A_p)'_{[a,b]}(x) = (A_q)'_{[a,b]}(x) = f(x)$. Estamos viendo ya el fenómeno que la sección siguiente formalizará: una función puede tener muchas funciones asociadas con la misma derivada, y todas ellas están separadas entre sí por constantes.

Lo que el teorema todavía no afirma. Conviene no mezclar las dos direcciones del puente entre derivación e integración.

Teorema 16.2 afirma:

$$\boxed{f \text{ continua} \Longrightarrow \left(x \mapsto \int_a^x f(t)\,dt\right)' = f.}$$

Todavía **no** hemos demostrado una fórmula de la forma

$$\int_a^b f(t)\,dt = F(b) - F(a)$$

a partir de una función F conocida con $F' = f$. Esa será una consecuencia posterior, no una reformulación automática de lo que acabamos de probar. Tampoco hemos definido aún de manera formal el término **primitiva**. La tarea exclusiva de §16.5 será introducir esa noción, demostrar que toda función continua posee al menos una —precisamente gracias a Teorema 16.2— y establecer que dos primitivas de la misma función solo pueden diferir por una constante.

Antes de seguir. ¿Por qué Teorema 16.2 no necesita una nueva prueba analítica distinta de Teorema 16.1? ¿Qué cambia en la interpretación de la derivada cuando $c = a$ o $c = b$? ¿Por qué cambiar el punto base de la acumulación no cambia su derivada? ¿Qué dirección del vínculo entre derivación e integración queda todavía pendiente?

16.5. Primitivas: existencia y unicidad hasta una constante

La sección anterior mostró que, cuando f es continua, la función de acumulación

$$A_f(x) = \int_a^x f(t)\,dt$$

satisface $(A_f)'_{[a,b]}(x) = f(x)$.

La identidad permite investigar el conjunto de todas las funciones cuya derivada es f. Además de construir una primitiva, tendremos que describir en qué pueden diferir dos de ellas.

Una función prescrita como derivada.

Definición 16.2. **Definición (primitiva).** Sean $a < b$ y $f : [a, b] \to \mathbb{R}$. Diremos que una función

$$F : [a, b] \longrightarrow \mathbb{R}$$

es una **primitiva de f en** $[a, b]$ si F es diferenciable relativamente a $[a, b]$ en cada punto del intervalo y

$$\boxed{F'_{[a,b]}(x) = f(x) \qquad (x \in [a, b]).}$$

En los puntos interiores esta condición coincide con la derivada bilateral ordinaria. En los extremos se entiende mediante la derivada relativa al dominio, conforme a la convención usada en este libro.

La definición describe una relación entre dos funciones. La función f está dada y buscamos funciones F cuya tasa de cambio sea exactamente f. Conviene distinguir desde el comienzo dos preguntas diferentes:

1. **Existencia:** ¿hay al menos una primitiva de f?
2. **Descripción:** si existe una, ¿cuántas hay y cómo se relacionan entre sí?

El teorema fundamental ya responde la primera pregunta para una clase importante de funciones.

Corolario 16.1. **Corolario (toda función continua posee una primitiva).** Sean $a < b$ y $f : [a, b] \to \mathbb{R}$ continua. Entonces f posee una primitiva en $[a, b]$.

En particular, la función de acumulación

$$A_f(x) := \int_a^x f(t)\, dt$$

es una primitiva de f.

Demostración. Por Teorema 16.2, la función A_f es diferenciable relativamente a $[a, b]$ en cada punto y satisface

$$(A_f)'_{[a,b]}(x) = f(x) \qquad (x \in [a, b]).$$

Ésta es exactamente la condición de Definición 16.2. □

El corolario es un teorema de **existencia**, no un método general para escribir una fórmula elemental. Afirma que una primitiva existe y construye una mediante una integral con extremo variable. No afirma que esa integral pueda simplificarse usando un repertorio de funciones previamente conocidas.

Existir no significa disponer de una fórmula elemental
Para una función continua f, la expresión

$$x \longmapsto \int_a^x f(t)\, dt$$

ya es una construcción rigurosa de una primitiva. La búsqueda sistemática de fórmulas cerradas, sustituciones, integración por partes y otras técnicas de integración pertenece a capítulos posteriores.

Dos primitivas no pueden diferir de manera variable. La existencia deja abierta la segunda pregunta. Supongamos que F y G son dos primitivas de la misma función f. Entonces, en cada punto interior, $F'(x) = f(x) = G'(x)$.

Al restar, $(F - G)'(x) = 0$.

Una derivada nula en todo un intervalo impide que la diferencia cambie de valor. El teorema del valor medio permite convertir esta intuición en una afirmación global.

Proposición 16.2. **Proposición (unicidad de primitivas hasta una constante).** Sean $a < b$ y $f : [a, b] \to \mathbb{R}$. Si F y G son primitivas de f en $[a, b]$, entonces existe una constante $C \in \mathbb{R}$ tal que

$$\boxed{F(x) = G(x) + C \qquad (x \in [a, b]).}$$

Recíprocamente, si G es una primitiva de f y $C \in \mathbb{R}$, entonces $G + C$ también es una primitiva de f.

Demostración. Definamos $H := F - G$.

Como F y G son diferenciables relativamente a $[a, b]$ en cada punto, Teorema 6.1 implica que ambas son continuas en $[a, b]$. Por las reglas de continuidad, H es continua en $[a, b]$. Además, para cada $x \in (a, b)$, la linealidad de la derivada da

$$H'(x) = F'(x) - G'(x) = f(x) - f(x) = 0.$$

Tomemos dos puntos cualesquiera $x, y \in [a, b]$ con $x < y$. La restricción de H a $[x, y]$ es continua en ese intervalo y diferenciable en (x, y). Por el teorema del valor medio Teorema 9.3, existe $c \in (x, y)$ tal que $\frac{H(y)-H(x)}{y-x} = H'(c) = 0$.

Como $y - x \neq 0$, $H(y) = H(x)$.

Los puntos x e y fueron arbitrarios; por tanto H es constante en todo $[a, b]$. Existe, pues, $C \in \mathbb{R}$ tal que $H(x) = C$, y, en consecuencia, $F(x) = G(x) + C$. Para la recíproca, si G es una primitiva y C es constante, la linealidad de la derivada y la derivada nula de una constante proporcionan $(G + C)'_{[a,b]}(x) = G'_{[a,b]}(x) = f(x)$. Así, $G + C$ también es una primitiva. □

La libertad aditiva descrita por Proposición 16.2 se hace visible en Figura 16.4: cambiar C desplaza la gráfica verticalmente sin alterar ninguna derivada.

La constante es toda la libertad disponible. La proposición anterior clasifica por completo las primitivas una vez conocida una de ellas. Si F_0 es una primitiva de f, entonces el conjunto de todas las primitivas es

$$\boxed{\{F_0 + C : C \in \mathbb{R}\}.}$$

No pueden aparecer dos primitivas cuya diferencia dependa de x. La derivada elimina constantes, pero no elimina ninguna otra variación sobre un intervalo. Esto permite fijar una única primitiva mediante una condición de normalización. Si elegimos $x_0 \in [a, b]$ y un número $y_0 \in \mathbb{R}$, entonces **como máximo una** primitiva de f puede satisfacer $F(x_0) = y_0$. Y si ya disponemos de alguna primitiva F_0, existe exactamente una con ese valor: basta elegir $C := y_0 - F_0(x_0)$. La condición en un solo punto elimina precisamente la libertad aditiva.

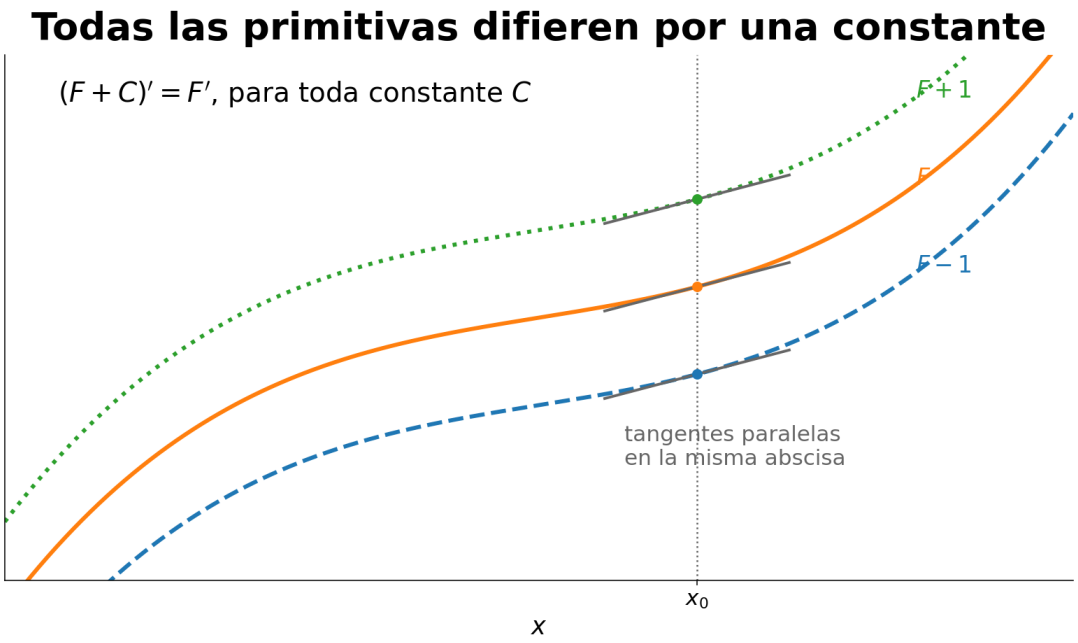


FIGURA 16.4. Las primitivas de una misma función forman la familia $F + C$: son traslaciones verticales de una curva base y conservan la misma pendiente en cada abscisa.

EJEMPLO 16.2. **Una familia completa de primitivas.** Consideremos

$$f(x) = 3x^2 - 2x + 1 \qquad (-1 \le x \le 2).$$

La función polinómica $F_0(x) = x^3 - x^2 + x$ satisface $F_0'(x) = 3x^2 - 2x + 1 = f(x)$. Por Proposición 16.2, todas las primitivas de f en $[-1, 2]$ son exactamente

$$\boxed{F_C(x) = x^3 - x^2 + x + C, \qquad C \in \mathbb{R}.}$$

Por ejemplo, la condición $F(0) = 5$ selecciona únicamente $F(x) = x^3 - x^2 + x + 5$. La familia no surge de una convención de notación: la proposición demuestra que **no existen otras primitivas** fuera de esta familia.

Existencia, conocimiento y evaluación son problemas distintos. Ya podemos separar tres niveles que a menudo se mezclan.

- Decir que f **posee una primitiva** es una afirmación de existencia.
- Exhibir una función F y verificar $F' = f$ proporciona una primitiva concreta.
- Relacionar una integral definida con los valores de una primitiva en los extremos es una afirmación adicional que todavía debemos demostrar.

La acumulación proporciona una primitiva para funciones continuas (Corolario 16.1), y Proposición 16.2 determina la libertad de elección. Para evaluar una integral a partir de una primitiva conocida necesitaremos justificar además la fórmula con las hipótesis de integrabilidad indicadas a continuación. Ésa será la tarea exclusiva de §16.6. Allí estudiaremos la dirección complementaria del teorema fundamental —**derivar y luego integrar**— y demostraremos

la forma fuerte de Newton–Leibniz para una derivada Riemann-integrable, sin reducirla de antemano al caso continuo.

Antes de seguir. ¿Por qué Corolario 16.1 es un resultado de existencia aunque no simplifique la integral que define A_f? ¿Dónde se usa el teorema del valor medio en la prueba de Proposición 16.2? ¿Por qué una condición $F(x_0) = y_0$ elimina exactamente una constante? ¿Qué afirmación sobre integrales definidas permanece todavía sin demostrar?

16.6. Newton–Leibniz: de una derivada a una integral

Hasta ahora recorrimos una dirección del puente entre integración y derivación. Para una función continua f, construimos su acumulación

$$A_f(x) = \int_a^x f(t)\,dt$$

y demostramos que $A_f'(x) = f(x)$.

Ahora partiremos del extremo opuesto. Supondremos conocida una función G cuya derivada es f y preguntaremos si la integral de f sobre $[a, b]$ puede recuperarse únicamente a partir del cambio neto de G entre los extremos. La respuesta será afirmativa bajo una hipótesis que conviene formular con cuidado: **no necesitaremos que f sea continua**. Bastará que f sea Riemann integrable y que coincida en el interior con la derivada de una función continua en el intervalo cerrado.

La forma fuerte de Newton–Leibniz.

Teorema 16.3. **Teorema (Newton–Leibniz para una derivada Riemann-integrable).** Sean $a < b$ y

$$f \in \mathcal{R}[a, b].$$

Sea

$$G : [a, b] \longrightarrow \mathbb{R}$$

continua en $[a, b]$ y diferenciable en (a, b). Supongamos que

$$G'(x) = f(x) \qquad (x \in (a, b)).$$

Entonces

$$\boxed{\int_a^b f(x)\,dx = G(b) - G(a).}$$

La formulación merece dos observaciones antes de la prueba. Primero, no hemos supuesto que f sea continua. La hipótesis analítica sobre el integrando es exactamente $f \in \mathcal{R}[a, b]$. Segundo, no exigimos que G tenga derivadas relativas en a y b. Para recuperar el incremento total $G(b) - G(a)$ bastará aplicar el teorema del valor medio en el **interior** de cada pieza de una partición.

El mapa lógico de Figura 16.5 mantiene separadas las dos direcciones del FTC: Teorema 16.2 parte de continuidad del integrando; Teorema 16.3 parte de integrabilidad de f y de una función G cuya derivada coincide con f en el interior.

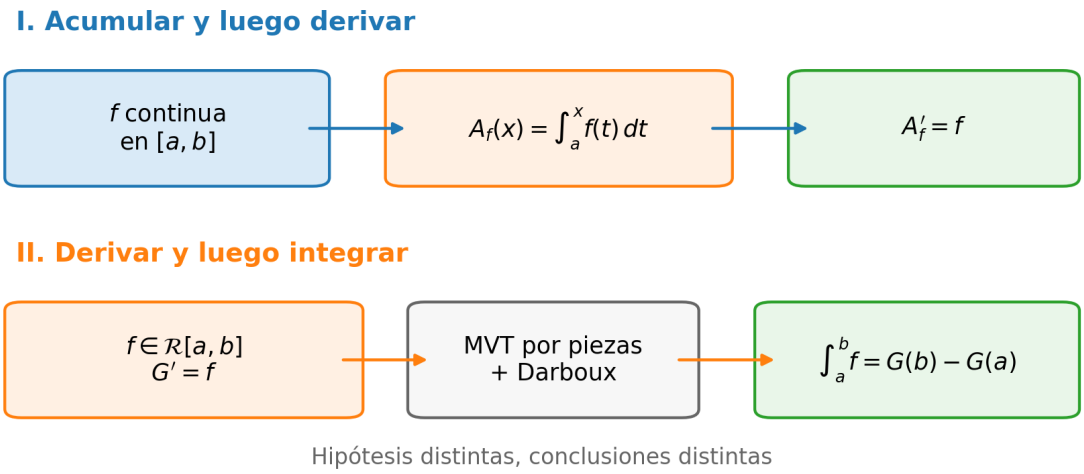


FIGURA 16.5. Las dos direcciones del teorema fundamental del cálculo tienen hipótesis distintas: una función continua genera una acumulación cuya derivada recupera al integrando, mientras que una primitiva de una función Riemann-integrable permite evaluar su integral por el incremento de la primitiva.

La prueba: de incrementos locales a un telescopado global. Demostración. Como f es Riemann integrable, en particular es acotada. Tomemos una partición arbitraria

$$P = \{a = x_0 < x_1 < \cdots < x_n = b\}$$

de $[a, b]$.

Para cada $k = 1, \ldots, n$, la restricción de G al subintervalo $[x_{k-1}, x_k]$ es continua, y G es diferenciable en (x_{k-1}, x_k). Por el teorema del valor medio Teorema 9.3, existe un punto $c_k \in (x_{k-1}, x_k)$ tal que

$$G(x_k) - G(x_{k-1}) = G'(c_k)(x_k - x_{k-1}).$$

Como $c_k \in (a, b)$ y $G'(c_k) = f(c_k)$,

$$\boxed{G(x_k) - G(x_{k-1}) = f(c_k)\,\Delta x_k,}$$

donde $\Delta x_k := x_k - x_{k-1} > 0$.

Figura 16.6 muestra el puente de la prueba entre incrementos de G y sumas etiquetadas de f. La figura no identifica la integral con una sola partición: el paso decisivo posterior es el encierro por Darboux para particiones con brecha arbitrariamente pequeña.

Definamos, como en las sumas de Darboux,

$$m_k := \inf_{x \in [x_{k-1}, x_k]} f(x), \qquad M_k := \sup_{x \in [x_{k-1}, x_k]} f(x).$$

Como c_k pertenece a ese subintervalo, $m_k \le f(c_k) \le M_k$.

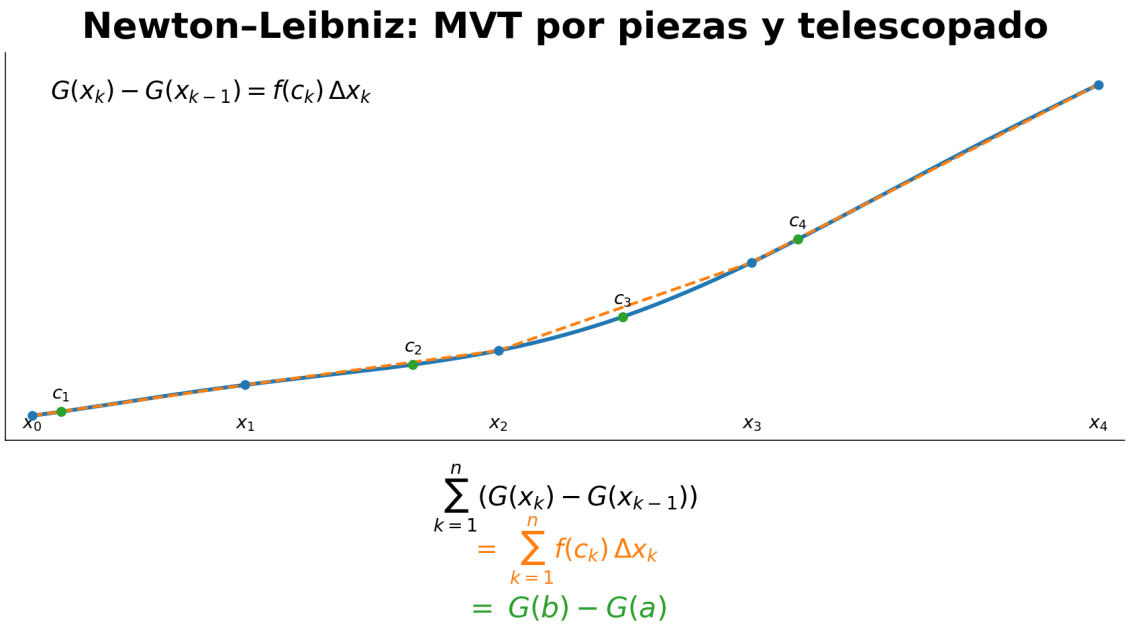


FIGURA 16.6. En cada subintervalo de una partición, el teorema del valor medio produce un punto c_k con $G(x_k) - G(x_{k-1}) = f(c_k)\Delta x_k$; al sumar, los incrementos de G telescopan y forman una suma etiquetada de f.

Multiplicando por $\Delta x_k > 0$,

$$m_k \Delta x_k \le G(x_k) - G(x_{k-1}) \le M_k \Delta x_k.$$

Sumamos estas desigualdades para $k = 1, \dots, n$. Los miembros exteriores son precisamente las sumas inferior y superior de Darboux:

$$L(f, P) \le \sum_{k=1}^{n} (G(x_k) - G(x_{k-1})) \le U(f, P).$$

La suma central es telescópica:

$$\begin{aligned} \sum_{k=1}^{n} (G(x_k) - G(x_{k-1})) &= G(x_1) - G(x_0) + G(x_2) - G(x_1) + \cdots \\ &\qquad + G(x_n) - G(x_{n-1}) \\ &= G(x_n) - G(x_0) \\ &= G(b) - G(a). \end{aligned}$$

Por tanto, para **toda** partición P,

$$\boxed{L(f, P) \le G(b) - G(a) \le U(f, P).}$$

Como la desigualdad izquierda vale para todas las sumas inferiores, la definición de integral inferior Definición 14.1 da

$$\underline{\int_a^b} f(x)\,dx \le G(b) - G(a).$$

Análogamente, como la desigualdad derecha vale para todas las sumas superiores,

$$G(b) - G(a) \le \overline{\int_a^b} f(x)\,dx.$$

Así,

$$\underline{\int_a^b} f(x)\,dx \le G(b) - G(a) \le \overline{\int_a^b} f(x)\,dx.$$

Pero $f \in \mathcal{R}[a,b]$. Por Definición 14.2, las integrales inferior y superior coinciden con la integral de Riemann:

$$\underline{\int_a^b} f(x)\,dx = \int_a^b f(x)\,dx = \overline{\int_a^b} f(x)\,dx.$$

El encierro anterior obliga entonces a

$$\int_a^b f(x)\,dx = G(b) - G(a).$$

□

Qué hizo realmente la prueba. La demostración puede leerse como una cadena de cuatro mecanismos ya disponibles:

$$\boxed{\begin{array}{c} \text{MVT en cada pieza} \\ \Downarrow \\ G(x_k) - G(x_{k-1}) = f(c_k)\Delta x_k \\ \Downarrow \\ m_k\Delta x_k \le G(x_k) - G(x_{k-1}) \le M_k\Delta x_k \\ \Downarrow \\ L(f,P) \le G(b) - G(a) \le U(f,P) \\ \Downarrow \\ \underline{\int_a^b} f \le G(b) - G(a) \le \overline{\int_a^b} f. \end{array}}$$

La integrabilidad de f no interviene para producir los puntos c_k. Esos puntos provienen del teorema del valor medio aplicado a G. La integrabilidad entra al final, cuando hace colapsar el encierro global de Darboux a un único número. Ésta es la razón por la cual la continuidad de f no aparece entre las hipótesis.

La hipótesis fuerte no es continuidad, sino integrabilidad
El teorema no dice

$$f \text{ continua} \Longrightarrow \int_a^b f = G(b) - G(a).$$

Dice algo más general:

$$\boxed{f \in \mathcal{R}[a,b], \qquad G' = f \text{ en } (a,b) \Longrightarrow \int_a^b f = G(b) - G(a),}$$

siempre que G sea continua en $[a,b]$ y diferenciable en su interior.

Los valores del integrando en los extremos no participan. La igualdad $G'(x) = f(x)$ solo se exige para $x \in (a,b)$. Esto no es una omisión. El argumento del valor medio selecciona puntos interiores c_k, nunca los extremos de las piezas. Además, cambiar únicamente $f(a)$ o $f(b)$ constituye una modificación finita y, por Corolario 14.1, no cambia la integral de Riemann. Por eso el valor de

$$\int_a^b f$$

puede estar determinado por la derivada interior de G sin imponer una ecuación derivativa adicional en a o b.

Las dos direcciones no tienen las mismas hipótesis. Ya disponemos de dos resultados que deben mantenerse conceptualmente separados. La primera forma global, Teorema 16.2, afirma:

$$\boxed{f \text{ continua} \Longrightarrow \left(x \mapsto \int_a^x f(t)\,dt\right)' = f.}$$

Newton–Leibniz, Teorema 16.3, afirma:

$$\boxed{f \in \mathcal{R}[a,b], \quad G' = f \Longrightarrow \int_a^b f = G(b) - G(a),}$$

con las condiciones de continuidad y diferenciabilidad sobre G explicitadas en el teorema. En la primera dirección, la integral **construye** una función cuya derivada recupera el integrando. En la segunda, una función cuya derivada ya conocemos permite **recuperar su cambio neto mediante una integral**. No son simplemente la misma frase escrita al revés: las hipótesis son distintas y las pruebas usan mecanismos distintos.

Lo que Newton–Leibniz no proporciona. Teorema 16.3 no afirma que toda función Riemann integrable posea una primitiva. Tampoco afirma que toda derivada sea automáticamente Riemann integrable. La integrabilidad de f es una hipótesis independiente del teorema.

El resultado tampoco explica cómo **encontrar** una función G tal que $G' = f$. Si ya conocemos una, Newton–Leibniz certifica la relación entre su incremento y la integral. La búsqueda sistemática de primitivas pertenece a C19. La sección siguiente se ocupará de una tarea más concreta: cuando una primitiva ya es conocida, aprenderemos a leer la fórmula

$$\int_a^b f(x)\,dx = G(b) - G(a)$$

como un procedimiento de evaluación, sin convertir todavía esa lectura en una teoría de técnicas de integración.

Antes de seguir. ¿Por qué la prueba aplica el teorema del valor medio a G y no a f? ¿En qué paso exacto se usa que f sea Riemann integrable? ¿Por qué no hace falta exigir $G'(a) = f(a)$

ni $G'(b) = f(b)$? ¿En qué sentido Teorema 16.3 es más general que la versión de Newton–Leibniz restringida a integrandos continuos?

16.7. Evaluar integrales cuando una primitiva es conocida

Teorema 16.3 resolvió la cuestión conceptual: si una función Riemann-integrable coincide en el interior con la derivada de una función adecuada, entonces su integral es exactamente el cambio neto de esa función.

En esta sección no buscaremos nuevas primitivas. Partiremos de una función G **ya conocida o ya propuesta** que satisfaga las hipótesis de Newton–Leibniz; toda primitiva en el sentido de Definición 16.2 es un caso particular. Aprenderemos a usar G como un certificado de evaluación.

La distinción es importante:

encontrar una primitiva $\neq$ evaluar una integral a partir de una primitiva conocida.

La primera tarea pertenece sistemáticamente a C19. La segunda ya está completamente justificada por el teorema fundamental.

La fórmula sobre cualquier subintervalo orientado. El enunciado de Teorema 16.3 se formuló sobre un intervalo $[a, b]$ con $a < b$. La convención de integral orientada permite convertirlo inmediatamente en una fórmula válida para cualesquiera dos puntos del intervalo original.

Corolario 16.2. **Corolario (fórmula de evaluación orientada).** Sean $a < b$, $f \in \mathcal{R}[a, b]$ y

$$G : [a, b] \longrightarrow \mathbb{R}$$

continua en $[a, b]$ y diferenciable en (a, b), con

$$G'(x) = f(x) \qquad (x \in (a, b)).$$

Entonces, para cualesquiera $u, v \in [a, b]$,

$$\boxed{\int_u^v f(x)\,dx = G(v) - G(u).}$$

Demostración. Si $u < v$, la restricción de f a $[u, v]$ es Riemann integrable por Teorema 14.7. La restricción de G es continua en $[u, v]$, diferenciable en (u, v) y satisface $G' = f$ allí. Por Teorema 16.3,

$$\int_u^v f(x)\,dx = G(v) - G(u).$$

Si $u = v$, ambos miembros son cero: la integral orientada vale cero por Definición 14.3 y $G(v) - G(u) = 0$.

Finalmente, si $u > v$, aplicamos el caso ya demostrado al intervalo $[v, u]$ y usamos la orientación:

$$\begin{aligned}\int_u^v f(x)\,dx &= -\int_v^u f(x)\,dx \\ &= -\big(G(u) - G(v)\big) \\ &= G(v) - G(u).\end{aligned}$$

Así, la misma fórmula vale para cualquier orden de los extremos. □

Notación de evaluación en los extremos. Cuando una primitiva G ya está identificada, es frecuente abreviar la diferencia $G(v) - G(u)$ mediante la notación

$$\boxed{\big[G(x)\big]_u^v := G(v) - G(u).}$$

Con esta convención, Corolario 16.2 puede escribirse como

$$\boxed{\int_u^v f(x)\,dx = \big[G(x)\big]_u^v, \qquad G' = f.}$$

Los corchetes no constituyen una nueva operación analítica. Son solamente una abreviatura para **evaluar en el extremo final y restar el valor en el extremo inicial**.

El orden de los extremos forma parte de la fórmula
La regla es $\big[G(x)\big]_u^v = G(v) - G(u)$, no $G(u) - G(v)$. Si se invierten los extremos de la integral, la diferencia cambia automáticamente de signo.

Un protocolo de evaluación, no un método para hallar primitivas. Cuando la función G está suministrada, la evaluación puede organizarse en cuatro pasos lógicos:

1. verificar que la integral de Riemann está definida en el intervalo considerado;
2. verificar que $G'(x) = f(x)$ en el interior pertinente;
3. evaluar G en los dos extremos;
4. formar **valor final menos valor inicial**.

El paso 2 es esencial. Una función parecida a una primitiva, o una fórmula recordada sin comprobación, no basta. Newton–Leibniz certifica la evaluación solo después de que la relación derivativa haya sido establecida.

Ejemplo 16.3. **Evaluación polinómica con una primitiva suministrada.** Consideremos

$$f(x) = 4x^3 - 4x + 1 \qquad (-1 \le x \le 2)$$

y la función $G(x) = x^4 - 2x^2 + x$.

Por las reglas de derivación de polinomios, $G'(x) = 4x^3 - 4x + 1 = f(x)$. Además, f es continua y, por Teorema 14.4, Riemann integrable en $[-1, 2]$. Por Corolario 16.2,

$$\begin{aligned}\int_{-1}^{2} (4x^3 - 4x + 1)\,dx &= \big[x^4 - 2x^2 + x\big]_{-1}^{2} \\ &= (16 - 8 + 2) - (1 - 2 - 1) \\ &= 10 - (-2) \\ &= \boxed{12}.\end{aligned}$$

La parte conceptual de la evaluación no consiste en manipular el símbolo integral: consiste en verificar primero que la función propuesta es efectivamente una primitiva.

La orientación se lee directamente en los valores de la primitiva.

Ejemplo 16.4. **Una integral con extremos invertidos.** En $[0,2]$ consideremos $f(x) = -\frac{1}{(x+1)^2}$ y $G(x) = \frac{1}{x+1}$.

Como $x+1 \neq 0$ en todo $[0,2]$, ambas funciones están bien definidas allí. La regla del recíproco da $G'(x) = -\frac{1}{(x+1)^2} = f(x)$. La función f es continua en $[0,2]$ y por tanto Riemann integrable. Si pedimos la integral con extremos invertidos,

$$\begin{aligned}\int_2^0 -\frac{1}{(x+1)^2}\,dx &= \Big[\frac{1}{x+1}\Big]_2^0 \\ &= 1 - \frac{1}{3} \\ &= \boxed{\frac{2}{3}}.\end{aligned}$$

En cambio,

$$\int_0^2 -\frac{1}{(x+1)^2}\,dx = -\frac{2}{3}.$$

No hace falta añadir una regla de signos después de evaluar: la diferencia $G(v) - G(u)$ ya incorpora la orientación correcta.

La constante de integración desaparece en una integral definida. En §16.5 demostramos que todas las primitivas de una misma función difieren por una constante. Esa libertad no produce ambigüedad al evaluar una integral definida.

Si $H(x) = G(x) + C$, entonces

$$\begin{aligned}H(v) - H(u) &= \big(G(v) + C\big) - \big(G(u) + C\big) \\ &= G(v) - G(u).\end{aligned}$$

Por tanto,

$$\boxed{\big[G(x) + C\big]_u^v = \big[G(x)\big]_u^v.}$$

Esto es exactamente lo que debe ocurrir: Proposición 16.2 afirma que la constante es toda la libertad disponible entre primitivas, y Newton–Leibniz muestra que esa libertad se cancela al formar el incremento entre dos extremos.

Ejemplo 16.5. **Dos primitivas, una sola integral.** Retomemos la función de Ejemplo 16.2,

$$f(x) = 3x^2 - 2x + 1 \qquad (-1 \le x \le 2).$$

Dos primitivas son $F_0(x) = x^3 - x^2 + x$ y $F_5(x) = x^3 - x^2 + x + 5$.

Usando F_0,

$$\begin{aligned}\int_{-1}^2 f(x)\,dx &= F_0(2) - F_0(-1) \\ &= 6 - (-3) \\ &= 9.\end{aligned}$$

Usando F_5 obtenemos el mismo resultado:

$$F_5(2) - F_5(-1) = 11 - 2$$
$$= 9.$$

La integral definida depende del incremento de la primitiva, no de la constante elegida para representarla.

Evaluar no es todavía una teoría de integración simbólica. Los ejemplos anteriores pueden dar la impresión de que el problema general consiste simplemente en «mirar» el integrando y escribir una primitiva. Esa inferencia sería prematura. En esta sección la primitiva fue **suministrada** o ya era conocida. Hemos aprendido a verificarla y a convertirla en el valor de una integral definida. No hemos desarrollado reglas para descubrir primitivas de productos, cocientes o composiciones, ni sustitución, integración por partes, fracciones parciales o repertorios de fórmulas. La frontera puede resumirse así:

$$\boxed{\text{primitiva conocida} \Longrightarrow \text{evaluación certificada por Newton–Leibniz,}}$$

mientras que

$$\boxed{\text{integrando dado} \Longrightarrow \text{¿cómo encontrar una primitiva?}}$$

es una pregunta distinta que se desarrollará sistemáticamente en C19. La sección siguiente cambiará nuevamente el tipo de problema. Ya no fijaremos ambos extremos: permitiremos que uno o los dos dependan de una variable y combinaremos el teorema fundamental con la regla de la cadena.

Antes de seguir. ¿Qué debe verificarse antes de reemplazar una integral por $[G]_u^v$? ¿Por qué una constante añadida a una primitiva no cambia la integral definida? ¿Cómo incorpora la fórmula $G(v) - G(u)$ la orientación de los extremos? ¿Qué parte del problema se ha resuelto aquí y qué parte se reserva todavía a C19?

16.8. Extremos variables y regla de la cadena

Hasta ahora hemos considerado dos situaciones distintas:

- una integral con extremos fijos, evaluada mediante una primitiva conocida;
- una acumulación con extremo derecho igual a la propia variable, cuya derivada recupera el integrando.

Ahora combinaremos ambas ideas con la regla de la cadena. El extremo ya no será necesariamente x: podrá ser una función $\phi(x)$. La estructura que conviene reconocer es

$$x \longmapsto \phi(x) \longmapsto \int_p^{\phi(x)} f(t)\,dt.$$

La integral exterior es una función de acumulación; por tanto, una vez demostrada su derivabilidad, el resto del argumento es exactamente una composición de funciones.

Un extremo variable.

Teorema 16.4. **Teorema (FTC con un extremo variable).** Sean $a < b$, $f : [a, b] \to \mathbb{R}$ continua y $p \in [a, b]$. Sean además $A \subseteq \mathbb{R}$, $c \in A$ un punto de acumulación de A y

$$\phi : A \longrightarrow [a, b]$$

diferenciable en c relativamente a A. Definamos

$$H(x) := \int_p^{\phi(x)} f(t)\,dt.$$

Entonces H es diferenciable en c relativamente a A y

$$\boxed{H'_A(c) = f(\phi(c))\,\phi'_A(c).}$$

Demostración. Consideremos la función de acumulación con base a,

$$A_f(y) := \int_a^y f(t)\,dt, \qquad y \in [a,b].$$

Como f es continua, Teorema 16.2 da

$$(A_f)'_{[a,b]}(y) = f(y) \qquad (y \in [a,b]).$$

Por aditividad orientada,

$$\int_p^{\phi(x)} f(t)\,dt = \int_a^{\phi(x)} f(t)\,dt - \int_a^p f(t)\,dt,$$

de modo que $H(x) = A_f(\phi(x)) - A_f(p)$.

El segundo término es constante respecto de x. Para el primero aplicamos la regla de la cadena Teorema 7.4 a $\phi : A \to [a,b]$ y $A_f : [a,b] \to \mathbb{R}$. Como todo punto de $[a,b]$ es punto de acumulación relativo del intervalo y A_f es diferenciable relativamente a $[a,b]$ en $\phi(c)$,

$$\begin{aligned} H'_A(c) &= (A_f \circ \phi)'_A(c) \\ &= (A_f)'_{[a,b]}(\phi(c))\,\phi'_A(c) \\ &= f(\phi(c))\,\phi'_A(c). \end{aligned}$$

□

El factor $f(\phi(c))$ proviene del teorema fundamental: es la derivada de la acumulación evaluada en el extremo alcanzado. El factor $\phi'_A(c)$ proviene de la regla de la cadena: mide la velocidad con la que ese extremo se mueve. Por eso la fórmula no es simplemente

$$\frac{d}{dx}\int_p^{\phi(x)} f(t)\,dt = f(x).$$

El integrando debe evaluarse en **el extremo real de acumulación**, $\phi(x)$, y después multiplicarse por la derivada de ese extremo.

Ejemplo 16.6. **Un extremo cuadrático.** En $[0,1]$ tomemos $f(t) = 3t^2 + 1$, y en $[-1,1]$ definamos $\phi(x) = x^2$.

Como $\phi([-1,1]) \subseteq [0,1]$, la función

$$H(x) = \int_0^{x^2} (3t^2 + 1)\,dt$$

está bien definida. Por Teorema 16.4,

$$\begin{aligned} H'(x) &= f(x^2)\,(x^2)' \\ &= (3x^4+1)(2x) \\ &= \boxed{6x^5+2x} \end{aligned}$$

para $x \in (-1,1)$, con la correspondiente interpretación relativa en $x = \pm 1$. Podemos verificar el resultado independientemente usando la evaluación de §16.7. Una primitiva conocida de f es $G(t) = t^3 + t$, de modo que $H(x) = G(x^2) - G(0) = x^6 + x^2$. Derivando esta expresión obtenemos nuevamente $H'(x) = 6x^5 + 2x$. Las dos rutas coinciden: una usa FTC más cadena; la otra evalúa primero la integral y deriva después la fórmula obtenida.

Dos extremos variables. Si se mueven ambos extremos, la orientación de la integral introduce automáticamente una diferencia de dos acumulaciones.

Proposición 16.3. **Proposición (FTC con dos extremos variables).** Sean $a < b$ y $f : [a,b] \to \mathbb{R}$ continua. Sean $A \subseteq \mathbb{R}$, $c \in A$ un punto de acumulación y

$$\alpha, \beta : A \longrightarrow [a,b]$$

diferenciables en c relativamente a A. Definamos

$$K(x) := \int_{\alpha(x)}^{\beta(x)} f(t)\,dt.$$

Entonces K es diferenciable en c relativamente a A y

$$\boxed{K'_A(c) = f(\beta(c))\,\beta'_A(c) - f(\alpha(c))\,\alpha'_A(c).}$$

Demostración. Fijemos cualquier $p \in [a,b]$. Por aditividad orientada,

$$\int_{\alpha(x)}^{\beta(x)} f(t)\,dt = \int_{p}^{\beta(x)} f(t)\,dt - \int_{p}^{\alpha(x)} f(t)\,dt.$$

Cada término del miembro derecho está cubierto por Teorema 16.4. Usando además la linealidad de la derivada,

$$K'_A(c) = f(\beta(c))\,\beta'_A(c) - f(\alpha(c))\,\alpha'_A(c).$$

□

Figura 16.7 hace visible el origen de los signos en Proposición 16.3: la contribución del extremo inferior se resta porque la integral orientada se descompone como acumulación superior menos acumulación inferior.

El signo menos del extremo inferior no es una regla memorizada añadida desde fuera. Proviene de la identidad orientada

$$\int_{\alpha(x)}^{\beta(x)} f = \int_{p}^{\beta(x)} f - \int_{p}^{\alpha(x)} f.$$

Tampoco hace falta suponer que $\alpha(x) \le \beta(x)$ para todos los valores de x. La integral orientada ya registra correctamente qué ocurre cuando los extremos se cruzan.

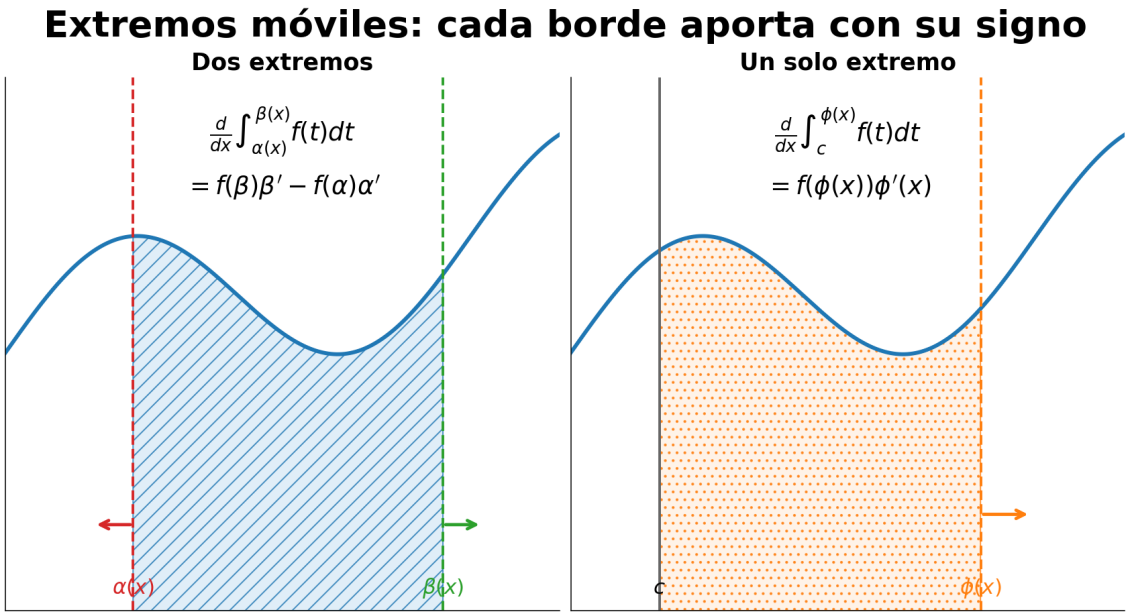


FIGURA 16.7. Con dos extremos móviles, el extremo superior aporta $f(\beta(x))\beta'(x)$ y el inferior aporta $-f(\alpha(x))\alpha'(x)$; el caso de un solo extremo aparece como especialización de la misma regla orientada.

EJEMPLO 16.7. **Dos extremos que se cruzan.** En $[0,1]$ tomemos $f(t) = 2t$, y para $x \in [0,1]$ definamos

$$\alpha(x) = x, \qquad \beta(x) = 1 - x.$$

Los extremos coinciden en $x = 1/2$ y cambian de orden al atravesar ese punto. Sin embargo,

$$K(x) = \int_x^{1-x} 2t\,dt$$

está definido para todo $x \in [0,1]$ por la convención orientada.

Como

$$\alpha'(x) = 1, \qquad \beta'(x) = -1,$$

Proposición 16.3 da

$$\begin{aligned} K'(x) &= 2(1-x)(-1) - 2x(1) \\ &= -2 + 2x - 2x \\ &= \boxed{-2} \end{aligned}$$

para $x \in (0,1)$, con derivadas relativas en los extremos. La evaluación directa confirma el resultado:

$$\begin{aligned} K(x) &= [t^2]_x^{1-x} \\ &= (1-x)^2 - x^2 \\ &= 1 - 2x, \end{aligned}$$

por lo que $K'(x) = -2$ incluso en el punto $x = 1/2$, donde los extremos se igualan. El cruce no produce singularidad alguna.

Dominio primero, fórmula después. Las fórmulas anteriores solo tienen sentido después de comprobar que los extremos móviles permanecen dentro del dominio del integrando. Si $f : [a, b] \to \mathbb{R}$, entonces debemos verificar $\phi(A) \subseteq [a, b]$ o, en el caso de dos extremos, $\alpha(A) \cup \beta(A) \subseteq [a, b]$. Los extremos móviles deben permanecer en el intervalo donde el integrando está definido e integrable. De lo contrario la expresión integral puede carecer de valor y no habrá una función a la que aplicar la regla de la cadena. La cadena lógica correcta es

$$\boxed{\begin{array}{c}\text{dominio de los extremos}\\ \Longrightarrow \text{FTC para la acumulación}\\ \Longrightarrow \text{regla de la cadena}\\ \Longrightarrow \text{fórmula derivativa.}\end{array}}$$

Esto no es diferenciación paramétrica bajo el signo integral. En Teorema 16.4 y Proposición 16.3 el integrando $f(t)$ es una función **fija**. La variable exterior x entra solamente a través de uno o dos extremos. Es conceptualmente distinto estudiar una expresión como

$$\int_{\alpha(x)}^{\beta(x)} F(x, t)\, dt,$$

donde el propio integrando cambia con x. En ese problema aparecen nuevas preguntas: cuándo puede derivarse respecto del parámetro, qué regularidad debe tener F, cómo se controla la derivada respecto de x y cuándo puede intercambiarse derivación e integración. Nada de eso ha sido demostrado aquí. La teoría de integrales dependientes de parámetros permanece reservada al Tomo II. En §16.8 hemos usado solamente dos herramientas ya establecidas:

$$\boxed{\text{FTC para una acumulación} + \text{regla de la cadena.}}$$

La sección siguiente someterá precisamente estas conclusiones a una auditoría de hipótesis: veremos qué puede fallar cuando el integrando no es continuo en el punto pertinente y qué parte del teorema fundamental sobrevive pese a ello.

Antes de seguir. ¿Por qué aparece $f(\phi(x))$ y no $f(x)$ cuando el extremo superior es $\phi(x)$? ¿De dónde proviene el signo menos asociado al extremo inferior? ¿Por qué el cruce de $\alpha(x)$ y $\beta(x)$ no invalida la fórmula? ¿Qué diferencia lógica hay entre un extremo dependiente de x y un integrando $F(x, t)$ dependiente de parámetro?

16.9. Por qué las hipótesis importan

Los resultados anteriores forman una cadena potente, pero cada eslabón tiene hipótesis distintas. En particular, conviene evitar dos errores opuestos:

1. **borrar una hipótesis** porque en algunos ejemplos la conclusión sigue siendo cierta sin ella;
2. **creer que la conclusión debe fallar** cada vez que una hipótesis no se cumple.

Una hipótesis de un teorema es una condición suficiente dentro de ese enunciado. Si la retiramos, debemos volver a analizar el problema. La sección estará dedicada a esa auditoría.

Una modificación puntual puede destruir la identidad derivativa.

Ejemplo 16.8. **Una función integrable cuya acumulación no recupera el valor modificado.** Definamos

$$f : [-1, 1] \longrightarrow \mathbb{R}, \qquad f(x) = \begin{cases} 1, & x = 0, \\ 0, & x \neq 0. \end{cases}$$

La función f difiere de la función cero solamente en un punto. Por la invariancia ante modificaciones finitas de Corolario 14.1, $f \in \mathcal{R}[-1, 1]$ y, para todo subintervalo $[u, v] \subseteq [-1, 1]$,

$$\int_u^v f(t)\, dt = 0.$$

Tomemos como punto base -1. Su función de acumulación es

$$A_f(x) = \int_{-1}^{x} f(t)\, dt.$$

Para cada $x \in [-1, 1]$ la integral anterior vale cero, de modo que

$$\boxed{A_f(x) = 0 \qquad (-1 \leq x \leq 1).}$$

Por tanto A_f es diferenciable en todo el intervalo y $A_f'(0) = 0$.

Pero $f(0) = 1$.

Así,

$$\boxed{A_f'(0) \neq f(0).}$$

La función es Riemann integrable, pero no es continua en 0. La conclusión local de Teorema 16.1 falla precisamente en el punto donde hemos alterado el valor.

El contraste de Figura 16.8 muestra por qué Teorema 16.1 necesita una hipótesis local sobre el integrando: la integral ignora modificaciones puntuales, mientras que la identidad $A_f'(c) = f(c)$ contiene el valor puntual $f(c)$. El ejemplo muestra que la continuidad en el punto no puede simplemente eliminarse de Teorema 16.1 y reemplazarse por la sola integrabilidad. Sin embargo, hay que formular la conclusión con cuidado. No hemos demostrado la conversa $A_f'(c) = f(c) \Longrightarrow f$ continua en c.

Teorema 16.1 proporciona una **condición suficiente**; Ejemplo 16.8 demuestra que esa condición no puede omitirse de un teorema universal que pretenda concluir la identidad para toda función Riemann integrable.

Un salto puede producir una esquina en la acumulación. Volvamos a la función escalón utilizada desde C15:

$$s(t) = \begin{cases} 0, & 0 \leq t < \frac{1}{2}, \\ 1, & \frac{1}{2} \leq t \leq 1. \end{cases}$$

Su integrabilidad ya está establecida, y en Ejemplo 16.1 calculamos exactamente su acumulación con base 0:

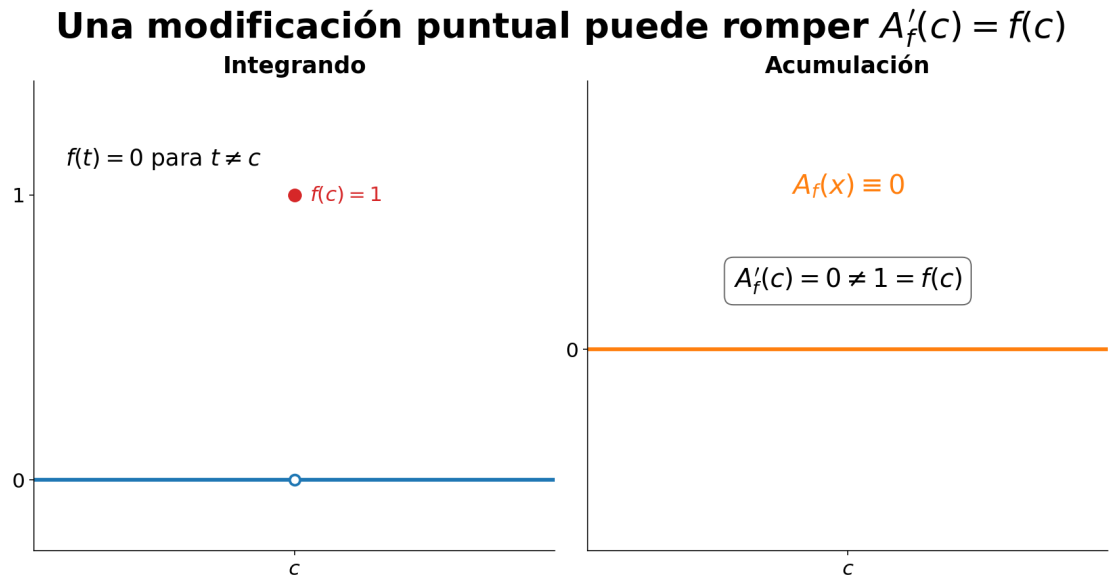


FIGURA 16.8. Una modificación en un único punto puede cambiar el valor puntual de f sin cambiar ninguna integral: la acumulación permanece idénticamente nula, de modo que $A_f'(0)$ no tiene por qué coincidir con el valor modificado $f(0)$.

$$A_s(x) = \begin{cases} 0, & 0 \le x \le \frac{1}{2}, \\ x - \frac{1}{2}, & \frac{1}{2} \le x \le 1. \end{cases}$$

EJEMPLO 16.9. **La acumulación es continua, pero no derivable en el salto.** En $c = 1/2$, el cociente incremental por la izquierda satisface, para $h < 0$ suficientemente pequeño,

$$\frac{A_s(\frac{1}{2} + h) - A_s(\frac{1}{2})}{h} = \frac{0 - 0}{h} = 0.$$

Por tanto,

$$\lim_{h \to 0^-} \frac{A_s(\frac{1}{2} + h) - A_s(\frac{1}{2})}{h} = 0.$$

Por la derecha, si $h > 0$,

$$\frac{A_s(\frac{1}{2} + h) - A_s(\frac{1}{2})}{h} = \frac{h}{h} = 1,$$

y entonces

$$\lim_{h \to 0^+} \frac{A_s(\frac{1}{2} + h) - A_s(\frac{1}{2})}{h} = 1.$$

Las derivadas laterales son distintas. En consecuencia,

$$\boxed{A_s'\left(\frac{1}{2}\right) \text{ no existe.}}$$

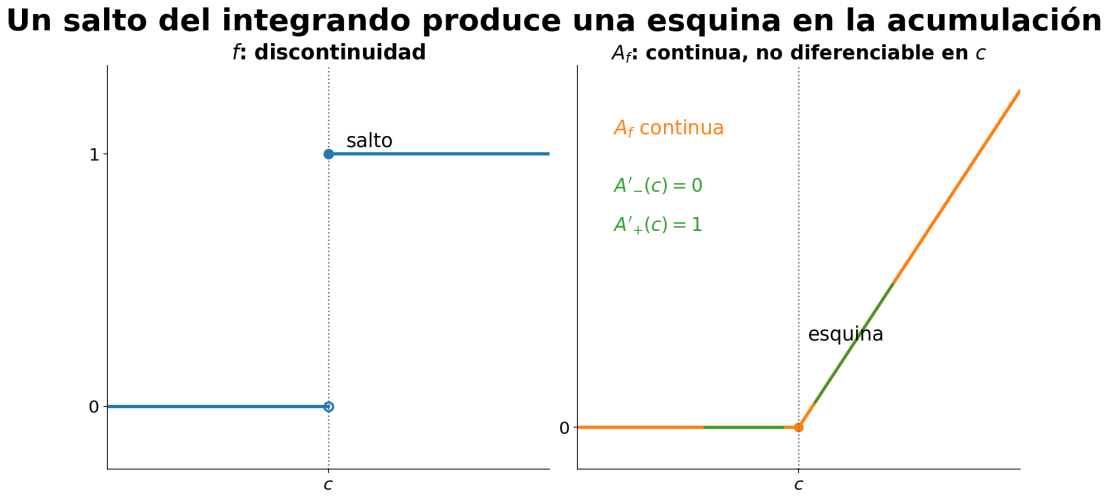


FIGURA 16.9. Un salto del integrando puede producir una acumulación continua con una esquina: en el punto del salto las pendientes laterales son distintas, aunque la acumulación siga siendo Lipschitz.

Esto ocurre aunque A_s sea continua —de hecho Lipschitz— en todo $[0,1]$ por Proposición 16.1.

Figura 16.9 representa el fenómeno de Ejemplo 16.9: integrabilidad basta para la continuidad de la acumulación, pero no para su diferenciabilidad en cada punto.

El ejemplo separa con nitidez dos niveles de regularidad:

$$\boxed{\text{integrabilidad de } s \Longrightarrow \text{continuidad de } A_s,}$$

pero no

$$\boxed{\text{integrabilidad de } s \Longrightarrow \text{diferenciabilidad de } A_s \text{ en todo punto.}}$$

Fuera de $1/2$, la función s sí es continua. Por Teorema 16.1,

$$A_s'(c) = s(c) \qquad (c \neq \tfrac{1}{2}),$$

con interpretación relativa en los extremos. En el salto, el teorema no se aplica y el cálculo directo muestra que la derivada falla.

La integrabilidad tampoco garantiza la existencia de una primitiva. El mismo escalón permite distinguir otra afirmación que a menudo se mezcla con el FTC. Supongamos, por contradicción, que existiera una primitiva $F : [0,1] \longrightarrow \mathbb{R}$ de s en el sentido de Definición 16.2. Como F es diferenciable relativamente a $[0,1]$ en cada punto, Teorema 6.1 garantiza que F es continua en todo $[0,1]$. Además,

$$F'(x) = 0 \qquad (0 < x < \tfrac{1}{2}).$$

La restricción de F a $[0,1/2]$ es, por tanto, continua en el cerrado y diferenciable en el interior. Por el teorema del valor medio Teorema 9.3, F sería constante en $[0,1/2]$.

En cambio,

$$F'(x) = 1 \qquad (\tfrac{1}{2} < x < 1).$$

La función $H(x) := F(x) - x$ es continua en $[1/2, 1]$ porque F y la identidad lo son, y es diferenciable en $(1/2, 1)$. Allí, $H'(x) = F'(x) - 1 = 0$. Otro uso del teorema del valor medio muestra entonces que H es constante en $[1/2, 1]$. Así, F tendría la forma

$$F(x) = \begin{cases} C, & 0 \le x \le \frac{1}{2}, \\ x + D, & \frac{1}{2} \le x \le 1, \end{cases}$$

con las constantes ajustadas por continuidad en $1/2$. Pero entonces la derivada lateral izquierda de F en $1/2$ sería 0 y la derecha sería 1. La derivada bilateral no existiría, en contradicción con que F fuera una primitiva.

Por tanto,

$$\boxed{s \in \mathcal{R}[0,1] \quad \text{pero} \quad s \text{ no posee primitiva en } [0,1].}$$

Esto prueba que la implicación

$$f \in \mathcal{R}[a,b] \Longrightarrow f \text{ posee una primitiva}$$

es falsa.

Lo que sí demostramos en Corolario 16.1 fue una afirmación más específica:

$$\boxed{f \text{ continua} \Longrightarrow f \text{ posee una primitiva},}$$

y la primitiva se construye mediante la acumulación.

Newton–Leibniz puede sobrevivir sin continuidad global del integrando. La situación cambia cuando **ya conocemos** una función cuya derivada coincide con el integrando en el interior. La forma fuerte Teorema 16.3 no exige continuidad global de f.

Ejemplo 16.10. **Una discontinuidad en el extremo no impide Newton–Leibniz.** Definamos en $[0, 1]$

$$f(x) = \begin{cases} 2x, & 0 \le x < 1, \\ 100, & x = 1. \end{cases}$$

La función f difiere de $g(x) = 2x$ únicamente en el punto 1. Como g es continua, es Riemann integrable por Teorema 14.4; luego Corolario 14.1 da $f \in \mathcal{R}[0, 1]$ y

$$\int_0^1 f(x)\,dx = \int_0^1 2x\,dx.$$

Tomemos $G(x) = x^2$.

Entonces G es continua en $[0, 1]$, diferenciable en $(0, 1)$ y

$$G'(x) = 2x = f(x) \qquad (0 < x < 1).$$

Todas las hipótesis de Teorema 16.3 están satisfechas, aunque f es discontinua en 1. Por tanto,

$$\boxed{\int_0^1 f(x)\,dx = G(1) - G(0) = 1.}$$

El valor aislado $f(1) = 100$ no interviene en la identidad derivativa interior y tampoco cambia la integral.

Este ejemplo no demuestra que toda discontinuidad sea irrelevante. Demuestra algo más preciso: **la continuidad global de f no es una hipótesis de Newton–Leibniz en su forma fuerte**. La diferencia lógica entre las dos direcciones del FTC queda ahora especialmente visible:

- para **construir** una primitiva mediante
$$x \mapsto \int_a^x f(t)\,dt,$$
la continuidad del integrando ofrece una hipótesis suficiente que garantiza la identidad derivativa en todo el intervalo;
- para **evaluar** una integral mediante una función G ya disponible, Teorema 16.3 solo exige integrabilidad de f y la relación $G' = f$ en el interior, además de las condiciones sobre G.

Mapa de supervivencia de las conclusiones. Las tres pruebas anteriores pueden resumirse así:

Hipótesis disponible	Qué podemos afirmar en C17
$f \in \mathcal{R}[a,b]$	A_f está bien definida y es Lipschitz, luego continua (Proposición 16.1).
$f \in \mathcal{R}[a,b]$ y f continua en c	$(A_f)'(c) = f(c)$ (Teorema 16.1).
f continua en todo $[a,b]$	A_f es una primitiva de f en todo el intervalo (Teorema 16.2).
$f \in \mathcal{R}[a,b]$ y existe G continua en $[a,b]$, diferenciable en (a,b), con $G' = f$ allí	$\int_a^b f = G(b) - G(a)$ (Teorema 16.3), sin exigir continuidad global de f.
Solo $f \in \mathcal{R}[a,b]$	No podemos concluir en general ni que A_f sea derivable en todo punto ni que f posea una primitiva.

La tabla no añade teoremas nuevos. Organiza el alcance lógico de los ya demostrados.

Una regla de lectura para las hipótesis. Cuando una hipótesis falla, la pregunta correcta no es

> «¿Entonces la conclusión es falsa?»

sino

> «¿Qué resultado anterior sigue siendo aplicable y qué parte debe analizarse de nuevo?»

En Ejemplo 16.8, la acumulación sigue siendo continua y derivable, pero no recupera el valor puntual alterado. En Ejemplo 16.9, la acumulación sigue siendo continua, pero pierde derivabilidad en el salto. En Ejemplo 16.10, la discontinuidad global de f no impide aplicar Newton–Leibniz porque sus hipótesis concretas siguen satisfechas. Cada hipótesis permite un paso de la prueba. Comprobarla identifica el resultado que podemos aplicar y evita atribuir a

la fórmula una conclusión más amplia que la demostrada. La sección siguiente convertirá esta auditoría en un laboratorio de decisión: habrá que identificar qué dirección del FTC corresponde, detectar argumentos inválidos y reparar cada uno con el resultado correcto.

Antes de seguir. ¿Qué demuestra Ejemplo 16.8 y qué no demuestra acerca de la continuidad como hipótesis? ¿Por qué la acumulación del escalón de Ejemplo 16.9 puede ser continua y no derivable? ¿Cómo muestra el mismo escalón que la integrabilidad no garantiza existencia de primitiva? ¿Por qué Ejemplo 16.10 sí permite usar Newton–Leibniz aunque el integrando no sea continuo en todo el intervalo?

16.10. Laboratorio: escoger la dirección correcta y reparar argumentos

En el laboratorio tendremos que distinguir qué objeto representa cada integral y cómo interviene la variable exterior. Esa identificación determinará cuál de los resultados anteriores puede utilizarse. En cada estación aparecerá una inferencia tentadora. La tarea será separar tres preguntas:

1. ¿qué dato está realmente disponible?;
2. ¿qué conclusión se quiere obtener?;
3. ¿qué resultado de C17 conecta legítimamente ambas cosas?

La estrategia general puede resumirse así:

identificar el tipo de objeto
$\longrightarrow$ verificar hipótesis
$\longrightarrow$ escoger la dirección correcta del FTC
$\longrightarrow$ aplicar la herramienta.

Estación 1 — Un número integral no es una función de acumulación. Supongamos que f es continua en $[a, b]$. Consideremos el argumento:

«Por el teorema fundamental,

$$\frac{d}{dx}\int_a^b f(t)\,dt = f(x).$$

»

El problema aparece antes de derivar. Los extremos a y b están fijados, de modo que

$$\int_a^b f(t)\,dt$$

es un **número real**, no una función de x. Respecto de una variable exterior x, ese número es constante. Por tanto,

$$\frac{d}{dx}\left(\int_a^b f(t)\,dt\right) = 0.$$

La reparación consiste en cambiar el objeto:

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

Ahora sí hay una función de x, y Teorema 16.2 permite afirmar

$$\boxed{A'_f(x) = f(x)}$$

en los puntos interiores, con la interpretación relativa correspondiente en los extremos. La diferencia no es notacional:

$$\boxed{\int_a^b f \quad \text{es un valor,} \qquad x \longmapsto \int_a^x f \quad \text{es una función.}}$$

Diagnóstico. Antes de aplicar el FTC, hay que localizar qué símbolos son constantes, cuál es la variable libre y cuál es la variable muda de integración.

Estación 2 — La acumulación puede ser continua aunque el integrando no lo sea. Consideremos ahora la inferencia:

> «Si
> $$A_f(x) = \int_a^x f(t)\,dt$$
> es continua, entonces f debe ser continua.»

Proposición 16.1 muestra que toda función Riemann-integrable tiene una acumulación Lipschitz. Por tanto, $f \in \mathcal{R}[a,b] \Longrightarrow A_f$ continua, sin exigir continuidad de f.

La función escalón de Ejemplo 16.9 lo hace visible: s tiene un salto en $1/2$, mientras que su acumulación

$$A_s(x) = \begin{cases} 0, & 0 \le x \le \frac{1}{2}, \\ x - \frac{1}{2}, & \frac{1}{2} \le x \le 1 \end{cases}$$

es continua en todo el intervalo. La reparación lógica es, por tanto,

$$\boxed{\text{continuidad de } A_f \not\Longrightarrow \text{continuidad de } f.}$$

Lo que sí puede recuperar información puntual sobre f es la **derivada** de A_f en circunstancias adecuadas, no su mera continuidad.

Diagnóstico. No invertir una implicación. La regularidad de una acumulación puede ser mayor que la del integrando.

Estación 3 — Una hipótesis local produce una conclusión local. Sea $f \in \mathcal{R}[a,b]$ y supongamos que sabemos únicamente que f es continua en un punto concreto $c \in [a,b]$.

Un argumento incorrecto sería:

> «Como f es continua en c, el teorema fundamental da $A'_f(x) = f(x)$ para todo $x \in [a,b]$.»

La continuidad disponible es puntual. Por tanto el resultado pertinente es la forma local Teorema 16.1, que autoriza solamente

$$\boxed{(A_f)'_{[a,b]}(c) = f(c).}$$

Para convertir esta afirmación en

$$(A_f)'_{[a,b]}(x) = f(x) \qquad \text{para todo } x \in [a,b],$$

necesitamos poder repetir el argumento en **cada** punto. Ésa es precisamente la hipótesis de continuidad global usada en Teorema 16.2. La estructura lógica correcta es

$$\boxed{\begin{array}{c} f \text{ continua en } c \\ \Downarrow \\ A_f'(c) = f(c) \end{array}} \qquad \text{frente a} \qquad \boxed{\begin{array}{c} f \text{ continua en todo } [a,b] \\ \Downarrow \\ A_f' = f \text{ en todo } [a,b]. \end{array}}$$

Diagnóstico. El alcance de la conclusión debe conservar el alcance de la hipótesis: continuidad en un punto no se convierte por sí sola en una identidad global.

Estación 4 — Existencia de primitiva no significa fórmula explícita. Sea $f : [a,b] \to \mathbb{R}$ continua. Alguien podría razonar:

> «Como f es continua, debe existir una fórmula conocida F construida con las funciones que ya manejamos tal que $F' = f$.»

Corolario 16.1 afirma algo distinto. Garantiza la **existencia** de una primitiva y proporciona una construcción:

$$\boxed{F(x) := \int_a^x f(t)\,dt.}$$

Esta expresión ya define una función rigurosa y Teorema 16.2 demuestra que $F' = f$. Nada de ello implica que podamos transformar F en una fórmula cerrada usando el repertorio algebraico disponible en este punto del libro. Por ejemplo, si se nos entrega una función continua

$$f(t) = \frac{1}{1+t^4} \qquad (0 \le t \le 1),$$

C17 permite afirmar que

$$F(x) = \int_0^x \frac{dt}{1+t^4}$$

es una primitiva. **No necesitamos** disponer aquí de una expresión alternativa para F.

La distinción correcta es

$$\boxed{\text{existencia de primitiva} \neq \text{fórmula explícita conocida} \neq \text{método sistemático para encontrarla.}}$$

La tercera tarea se reserva a C19.

Diagnóstico. Una definición mediante integral puede constituir por sí misma una construcción exacta; «no haber simplificado la fórmula» no significa «no haber construido la función».

Estación 5 — FTC I y Newton–Leibniz no se prueban intercambiando sus papeles. Supongamos que $f \in \mathcal{R}[a,b]$ y que ya conocemos una función $G : [a,b] \to \mathbb{R}$ continua en $[a,b]$, diferenciable en (a,b) y tal que

$$G' = f \qquad \text{en } (a,b).$$

Queremos demostrar

$$\int_a^b f = G(b) - G(a).$$

Una ruta tentadora sería:

1. definir $A_f(x) = \int_a^x f$;
2. afirmar $A_f' = f$;
3. concluir que A_f y G difieren por una constante;
4. evaluar en los extremos.

El paso 2 no está justificado con las hipótesis disponibles: Teorema 16.2 exige continuidad de f en todo el intervalo, mientras que Newton–Leibniz Teorema 16.3 solo supone que f es Riemann integrable. Por tanto, usar FTC I para demostrar la forma fuerte de Newton–Leibniz perdería precisamente la generalidad que §16.6 estableció. La reparación es escoger la dirección correcta:

$$\boxed{\begin{array}{c} f \text{ continua} \\ \text{y queremos derivar } x \mapsto \int_a^x f \end{array} \implies \text{derivar la acumulación,}}$$

mientras que

$$\boxed{\begin{array}{c} f \in \mathcal{R}[a,b],\ G' = f \\ \text{y queremos evaluar } \int_a^b f \end{array} \implies \text{evaluar por la primitiva.}}$$

La prueba fuerte de Teorema 16.3 usa MVT sobre una partición y encierro de Darboux; no necesita convertir antes la acumulación en una primitiva.

Diagnóstico. Las dos direcciones del FTC están conectadas, pero sus hipótesis y mecanismos de prueba no son idénticos.

Estación 6 — Extremo variable no es integrando paramétrico. Comparemos

$$H(x) := \int_0^{x^2} (t^2 + 1)\, dt$$

con

$$J(x) := \int_0^1 (x + t)\, dt.$$

En H, el integrando $t \longmapsto t^2 + 1$ es fijo; la variable x entra únicamente por el extremo x^2. Estamos exactamente en la situación de Teorema 16.4. Por tanto,

$$H'(x) = ((x^2)^2 + 1)(2x) = 2x(x^4 + 1),$$

en todo dominio donde el extremo permanezca dentro del intervalo sobre el que se ha fijado el integrando. En J, en cambio, los extremos son constantes y el propio integrando depende de x: $F(x,t) = x + t$.

Teorema 16.4 y Proposición 16.3 **no** autorizan a escribir una fórmula general para

$$\frac{d}{dx} \int_0^1 F(x,t)\, dt.$$

En este ejemplo particular podemos resolver la integral usando únicamente linealidad y la integral de funciones elementales ya conocidas:

$$J(x) = \int_0^1 x\,dt + \int_0^1 t\,dt$$
$$= x + \frac{1}{2},$$

y entonces $J'(x) = 1$.

Pero esta comprobación particular no constituye un teorema de diferenciación bajo el signo integral.

La frontera es

$$\boxed{x \text{ mueve extremos} \neq x \text{ modifica el integrando.}}$$

El segundo problema pertenece a la teoría de integrales dependientes de parámetros del Tomo II.

Diagnóstico. Antes de derivar una integral dependiente de x, localizar **dónde** entra x: en los extremos, en el integrando o en ambos.

Estación 7 — El patrón que abre C18 sin construir todavía el logaritmo. Consideremos, para $x > 0$,

$$L(x) := \int_1^x \frac{dt}{t}.$$

Esta expresión ya tiene sentido con la teoría disponible. Para cualquier $x > 0$, la función $t \longmapsto \frac{1}{t}$ es continua en el intervalo compacto cuyos extremos son 1 y x, de modo que la integral orientada está definida. Además, si fijamos un punto $c > 0$, podemos escoger un intervalo cerrado $[u, v] \subset (0, \infty)$ tal que $1 \in [u, v]$ y $c \in (u, v)$. En un entorno de c contenido en $[u, v]$, podemos trabajar con la primera forma global del FTC. Localmente,

$$\boxed{L'(c) = \frac{1}{c}.}$$

Lo que **no** haremos en C17 es convertir esta observación en una teoría nueva. En particular, aquí no introduciremos todavía:

- el logaritmo como función distinguida;
- sus propiedades funcionales;
- su inversa exponencial;
- potencias reales construidas a partir de ambas.

Ese desarrollo pertenece a C18.

El papel de esta estación es reconocer el puente:

$$\boxed{\text{FTC} \longrightarrow x \mapsto \int_1^x \frac{dt}{t} \longrightarrow \text{construcción futura de una función elemental.}}$$

El capítulo siguiente podrá usar este objeto como punto de partida; C17 solo certifica la maquinaria analítica que lo hace posible.

Mapa de decisión. Las siete estaciones pueden condensarse en una tabla operativa.

Datos disponibles	Pregunta	Herramienta correcta
$f \in \mathcal{R}[a,b]$	¿Es continua la acumulación?	Proposición 16.1
$f \in \mathcal{R}[a,b]$ y f continua en c	¿Cuál es $A'_f(c)$?	Teorema 16.1
f continua en todo $[a,b]$	¿Es A_f una primitiva?	Teorema 16.2
f continua	¿Existe alguna primitiva?	Corolario 16.1
$f \in \mathcal{R}[a,b]$ y ya existe G con $G' = f$ en (a,b)	¿Cuánto vale $\int_a^b f$?	Teorema 16.3
Integrando fijo y extremo $\phi(x)$	¿Cómo derivar la integral?	Teorema 16.4 + cadena
Integrando fijo y extremos $\alpha(x), \beta(x)$	¿Cómo derivar la integral?	Proposición 16.3
Integrando $F(x,t)$ dependiente de parámetro	¿Cómo intercambiar derivación e integración?	**No resuelto en C17**

Para usar la tabla, identificamos primero los datos y la pregunta. Solo entonces podemos comprobar si las hipótesis del teorema correspondiente están satisfechas.

El nombre «teorema fundamental» no autoriza una regla única
C17 no ha producido una fórmula universal que pueda aplicarse mecánicamente a cualquier expresión con un signo integral. Ha construido varios resultados con hipótesis diferentes. La habilidad matemática consiste en distinguir cuál de ellos corresponde al problema presente.

Reparación final: cuatro preguntas antes de usar el FTC. Ante una expresión que combine derivadas e integrales, conviene preguntar, en este orden:

1. **¿La integral representa un número o una función?**
2. **¿La variable exterior entra por un extremo, por el integrando o por ambos?**
3. **¿Quiero derivar una acumulación o evaluar una integral usando una derivada ya conocida?**
4. **¿Las hipótesis del resultado elegido están realmente verificadas?**

Solo después de responder esas cuatro preguntas tiene sentido manipular fórmulas. Una integral con extremos fijos representa un número; con extremo variable puede definir una acumulación. Si la variable exterior entra en el integrando, los teoremas de derivación por extremos no justifican por sí solos intercambiar derivada e integral. Mantener esas distinciones permite separar los resultados demostrados de las cuestiones que requieren hipótesis nuevas. El teorema Teorema 16.2 construye una primitiva para una función continua; Teorema 16.3 evalúa la integral bajo sus hipótesis de integrabilidad y primitiva existente. La derivación por extremos de Teorema 16.4 no es un teorema de derivación de integrandos que dependan de un parámetro.

16.11. Ejercicios y soluciones

Los cuarenta ejercicios siguientes recorren la arquitectura completa del capítulo sin introducir teoría posterior. El banco está organizado en siete niveles y utiliza únicamente resultados disponibles hasta §16.10. Las soluciones mantienen separadas las tareas de **definir una acumulación**, **demostrar regularidad**, **recuperar un integrando mediante derivación**, **evaluar una integral mediante una primitiva** y **derivar extremos variables**.

Nivel A — Función de acumulación, orientación y lectura de variables. Ejercicio A1. Una acumulación lineal. Sea $f(t) = 3$ en $[0, 2]$ y define

$$A(x) := \int_0^x f(t)\,dt.$$

1. Calcula $A(x)$ para $x \in [0, 2]$.
2. Verifica $A(0)$ y $A(2)$.
3. Para $x, y \in [0, 2]$, comprueba directamente que
$$A(y) - A(x) = \int_x^y f(t)\,dt$$
también cuando $y < x$.

Ejercicio A2. Cambiar el punto base dentro de una zona nula. Sea

$$s(t) = \begin{cases} 0, & 0 \le t < \frac{1}{2}, \\ 1, & \frac{1}{2} \le t \le 1, \end{cases}$$

y define

$$B(x) := \int_{1/4}^x s(t)\,dt.$$

Calcula $B(x)$ para todo $x \in [0, 1]$ y compárala con la acumulación A_s de Ejemplo 16.1.

Ejercicio A3. Dos puntos base producen una diferencia constante. Sea $f \in \mathcal{R}[a, b]$ y sean $p, q \in [a, b]$. Define

$$A_p(x) := \int_p^x f(t)\,dt, \qquad A_q(x) := \int_q^x f(t)\,dt.$$

Demuestra que

$$\boxed{A_p(x) - A_q(x) = \int_p^q f(t)\,dt}$$

para todo $x \in [a, b]$. Explica por qué la diferencia no depende de x.

Ejercicio A4. Variable muda y variable libre. Considera

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

1. Explica por qué
$$\int_a^x f(t)\,dt = \int_a^x f(u)\,du.$$
2. Identifica cuál de a, x, t es el punto base, cuál es el extremo variable y cuál es la variable muda.
3. Reescribe sin ambigüedad la expresión informal
$$\int_a^x f(x)\,dx$$
cuando se pretende que x sea el argumento libre de la función de acumulación.

Ejercicio A5. Un integrando no negativo produce una acumulación no decreciente. Sea $f \in \mathcal{R}[a,b]$ y supón que $f(t) \geq 0$ para todo $t \in [a,b]$. Demuestra que su función de acumulación

$$A_f(x) = \int_a^x f(t)\,dt$$

es no decreciente en $[a,b]$. No utilices derivadas.

Ejercicio A6. La integral total no determina la acumulación. En $[0,1]$ considera

$$f(t) = 1, \qquad g(t) = 2t.$$

1. Verifica que
$$\int_0^1 f(t)\,dt = \int_0^1 g(t)\,dt = 1.$$
2. Calcula las funciones de acumulación A_f y A_g con base 0.
3. Determina en qué puntos de $[0,1]$ coinciden ambas acumulaciones.

Nivel B — Continuidad, Lipschitz y cotas de incrementos. Ejercicio B1. Una cota numérica inmediata. Sea $f \in \mathcal{R}[-2,3]$ y supón que $|f(t)| \leq 4$ en todo el intervalo. Si

$$A(x) = \int_{-2}^x f(t)\,dt,$$

1. acota $|A(1.2) - A(1.1)|$;
2. encuentra un $\delta > 0$ que garantice
$$|x-y| < \delta \Longrightarrow |A(x) - A(y)| < 0.01$$
para todos $x, y \in [-2,3]$.

Ejercicio B2. La constante Lipschitz no depende del punto base. Sea $f \in \mathcal{R}[a,b]$ con $|f| \leq M$ y fija $p \in [a,b]$. Define

$$A_p(x) := \int_p^x f(t)\,dt.$$

Demuestra directamente que $|A_p(y) - A_p(x)| \leq M|y-x|$ para todos $x, y \in [a,b]$.

Ejercicio B3. Verificar Lipschitz sin invocar el teorema general. Para la función escalón s de Ejemplo 16.1 se tiene

$$A_s(x) = \begin{cases} 0, & 0 \leq x \leq \frac{1}{2}, \\ x - \frac{1}{2}, & \frac{1}{2} \leq x \leq 1. \end{cases}$$

Demuestra directamente, separando los casos necesarios, que $|A_s(y) - A_s(x)| \leq |y-x|$ para todos $x, y \in [0,1]$.

Ejercicio B4. Dos integrandos uniformemente próximos. Sean $f, g \in \mathcal{R}[a,b]$ y supón que

$$|f(t) - g(t)| \le \varepsilon \qquad (t \in [a,b]).$$

Con base común a, define A_f y A_g. Demuestra que, para todos $x, y \in [a,b]$,

$$|(A_f - A_g)(y) - (A_f - A_g)(x)| \le \varepsilon |y - x|.$$

Deduce en particular $|A_f(x) - A_g(x)| \le \varepsilon |x - a|$.

Ejercicio B5. Una cota inferior fuerza crecimiento estricto. Sea $f \in \mathcal{R}[a,b]$ y supón que existe $m > 0$ tal que

$$f(t) \ge m \qquad (t \in [a,b]).$$

Demuestra que, si $x < y$, $A_f(y) - A_f(x) \ge m(y - x) > 0$.

Concluye que A_f es estrictamente creciente y, por tanto, inyectiva.

Ejercicio B6. Cuánto puede oscilar una acumulación. Sea $f \in \mathcal{R}[a,b]$ con $|f| \le M$. Demuestra:

1. para todo $x \in [a,b]$, $|A_f(x)| \le M|x - a|$;
2. para cualesquiera $x, y \in [a,b]$, $|A_f(x) - A_f(y)| \le M(b - a)$.

Interpreta la segunda desigualdad como una cota para el diámetro de la imagen $A_f([a,b])$.

Nivel C — FTC local y global: cocientes y continuidad puntual. Ejercicio C1. Derivar sin calcular primero la integral. Sea

$$A(x) := \int_0^x t^2 \, dt, \qquad 0 \le x \le 2.$$

Usa Teorema 16.1 o Teorema 16.2 para determinar $A'(c)$ en cada $c \in [0,2]$, interpretando correctamente los extremos. No calcules una fórmula cerrada para A antes de derivar.

Ejercicio C2. El FTC local alrededor de un salto. Para el escalón s de Ejemplo 16.1:

1. determina en qué puntos de $[0,1]$ puede aplicarse Teorema 16.1;
2. calcula $A_s'(c)$ en todos esos puntos;
3. usando la fórmula explícita de A_s, calcula las derivadas laterales en $c = 1/2$ y decide si existe allí la derivada bilateral.

Ejercicio C3. Una discontinuidad lejana no estorba al FTC local. Sea

$$f(t) = \begin{cases} 0, & 0 \le t < \frac{1}{2}, \\ 1, & \frac{1}{2} \le t \le 1, \end{cases}$$

y $A_f(x) = \int_0^x f(t)\,dt$. Justifica, sin exigir continuidad global, que

$$A_f'\left(\frac{1}{4}\right) = 0, \qquad A_f'\left(\frac{3}{4}\right) = 1.$$

Explica qué hipótesis de Teorema 16.1 se verifica localmente en cada punto.

Ejercicio C4. Una versión cuantitativa del promedio local. Supón que $f \in \mathcal{R}[a,b]$ y que para cierto $c \in (a,b)$ existen $K, \rho > 0$ tales que $|f(t) - f(c)| \leq K|t-c|$ si $t \in [a,b]$ y $|t-c| < \rho$. Demuestra que, para todo $h \neq 0$ admisible con $|h| < \rho$,

$$\left| \frac{A_f(c+h) - A_f(c)}{h} - f(c) \right| \leq K|h|.$$

Deduce de nuevo $A_f'(c) = f(c)$.

Ejercicio C5. Del resultado local a una identidad global. Sea $f : [a,b] \to \mathbb{R}$ continua y

$$A_f(x) = \int_a^x f(t)\,dt.$$

Explica por qué Teorema 16.1 puede aplicarse en **cada** punto de $[a,b]$ y concluye que

$$(A_f)'_{[a,b]}(x) = f(x) \qquad (x \in [a,b]).$$

Aclara qué significa esta igualdad en a y en b.

Ejercicio C6. La derivabilidad de la acumulación no recupera valores puntuales arbitrarios. En $[-1,1]$ define

$$f(t) = \begin{cases} 7, & t = 0, \\ 0, & t \neq 0. \end{cases}$$

1. Demuestra que f es Riemann integrable y que su función de acumulación con base -1 es idénticamente cero.
2. Calcula $A_f'(0)$.
3. Compara ese valor con $f(0)$ y explica por qué esto no contradice Teorema 16.1.

Nivel D — Primitivas, unicidad y Newton–Leibniz. Ejercicio D1. Verificar una primitiva y evaluar. Sea $f(x) = 3x^2 - 2$ en $[-1,2]$ y define $F(x) = x^3 - 2x + 5$.

1. Verifica que F es una primitiva de f.
2. Usa Newton–Leibniz para calcular

$$\int_{-1}^{2} (3x^2 - 2)\,dx.$$

Ejercicio D2. Describir todas las primitivas. Demuestra que todas las primitivas de $f(x) = 4x^3$ en cualquier intervalo cerrado $[a,b]$ son exactamente

$$F_C(x) = x^4 + C, \qquad C \in \mathbb{R}.$$

Indica dónde utilizas Proposición 16.2.

Ejercicio D3. Una condición inicial elimina la constante. Encuentra la única primitiva F de $f(x) = 2x + 1$ que satisface $F(0) = 7$. Después calcula

$$\int_0^2 (2x + 1)\,dx.$$

Ejercicio D4. Una discontinuidad en el extremo no cambia la evaluación. Define en $[0, 1]$

$$f(x) = \begin{cases} 6x^2, & 0 \le x < 1, \\ -100, & x = 1. \end{cases}$$

1. Justifica que $f \in \mathcal{R}[0, 1]$.
2. Verifica que $G(x) = 2x^3$ satisface las hipótesis de Teorema 16.3.
3. Calcula $\int_0^1 f(x)\,dx$.
4. Explica por qué el valor $f(1)$ no entra en la identidad derivativa exigida por el teorema.

Ejercicio D5. Reconstruir el mecanismo de Newton–Leibniz. Sean $f \in \mathcal{R}[a, b]$ y G como en Teorema 16.3. Para una partición

$$P = \{a = x_0 < x_1 < \cdots < x_n = b\},$$

1. usa el teorema del valor medio diferencial en cada $[x_{k-1}, x_k]$ para obtener puntos ξ_k tales que $G(x_k) - G(x_{k-1}) = f(\xi_k)\Delta x_k$;
2. suma las identidades y demuestra que $G(b) - G(a) = R(f; P, \xi)$;
3. explica por qué el encierro de Darboux y la integrabilidad de f obligan a

$$G(b) - G(a) = \int_a^b f.$$

Ejercicio D6. Existencia de primitiva no equivale a fórmula cerrada. Sea

$$f(t) = \frac{1}{1 + t^2}, \qquad 0 \le t \le 1.$$

Sin introducir ninguna función trascendente nueva:

1. justifica que f posee una primitiva en $[0, 1]$;
2. proporciona una primitiva usando la construcción de acumulación;
3. explica por qué este resultado de existencia, por sí solo, no constituye todavía un procedimiento para obtener un valor cerrado de

$$\int_0^1 \frac{dt}{1 + t^2}.$$

Nivel E — Extremos variables y regla de la cadena. Ejercicio E1. Extremo cuadrático. Para $x \in [-1, 1]$ define

$$H(x) = \int_0^{x^2} (t^2 + 1)\, dt.$$

Calcula $H'(x)$ usando Teorema 16.4, sin evaluar primero la integral.

Ejercicio E2. Mover solo el extremo inferior. En $[0, 1]$ define

$$H(x) = \int_x^1 (1 + t^2)\, dt.$$

Demuestra que $H'(x) = -(1 + x^2)$, con interpretación relativa en los extremos.

Ejercicio E3. Dos extremos que se mueven a velocidades distintas. Para $x \in [0, 1]$ define

$$K(x) = \int_{x^2}^{2x} (1 + t^3)\, dt.$$

Calcula $K'(x)$ mediante Proposición 16.3 y simplifica el resultado.

Ejercicio E4. El cruce de los extremos no crea una excepción. Para $x \in [0, 1]$ sea

$$H(x) = \int_{1-x}^{x} t\, dt.$$

1. Usa Proposición 16.3 para calcular $H'(x)$.
2. Explica por qué la fórmula sigue siendo válida en $x = 1/2$, donde los dos extremos coinciden.
3. Verifica el resultado evaluando la integral mediante una primitiva conocida.

Ejercicio E5. El dominio se determina antes de derivar. Sea $f : [-2, 2] \to \mathbb{R}$ continua y define $\phi(x) = \frac{1}{x-1}$.

Queremos considerar

$$H(x) = \int_0^{\phi(x)} f(t)\, dt.$$

1. Determina el conjunto de números reales x para los que $\phi(x) \in [-2, 2]$.
2. En ese dominio, calcula la derivada relativa de H en cualquier punto de acumulación.
3. Explica por qué no es lícito escribir una fórmula derivativa antes de comprobar que el extremo pertenece al dominio de f.

Ejercicio E6. Una variable dentro del integrando no es un extremo variable. Define

$$J(x) = \int_0^1 (x + t)\, dt.$$

1. Explica por qué Teorema 16.4 no se aplica directamente a esta expresión.
2. Evalúa primero la integral como función de x y calcula $J'(x)$.

3. Indica qué teoría más general sería necesaria para justificar, en abstracto, intercambiar derivación e integración cuando el integrando depende de un parámetro.

Nivel F — Contraejemplos, hipótesis y reparación de pruebas. Ejercicio F1. Integrabilidad no basta para recuperar el valor puntual. En $[-1,1]$ define

$$f(t) = \begin{cases} 5, & t = 0, \\ 0, & t \neq 0. \end{cases}$$

Refuta con este ejemplo la afirmación:

«Si f es Riemann integrable, entonces $A_f'(c) = f(c)$ para todo c.»

Identifica exactamente la hipótesis faltante en $c = 0$.

Ejercicio F2. La continuidad de la acumulación no implica continuidad del integrando. Usa la función escalón de Ejemplo 16.1 para refutar A_f continua $\Longrightarrow$ f continua.

Explica qué resultado sí garantiza la continuidad de A_f.

Ejercicio F3. Una función integrable sin primitiva. Sea

$$s(x) = \begin{cases} 0, & 0 \leq x < \frac{1}{2}, \\ 1, & \frac{1}{2} \leq x \leq 1. \end{cases}$$

Demuestra por contradicción que s no posee una primitiva en $[0,1]$. Puedes usar el teorema del valor medio para demostrar que una función con derivada cero en un intervalo es constante.

Ejercicio F4. Newton–Leibniz no exige continuidad global del integrando. En $[0,1]$ define

$$f(x) = \begin{cases} 2x, & 0 \leq x < 1, \\ 37, & x = 1. \end{cases}$$

Demuestra que

$$\int_0^1 f(x)\,dx = 1$$

aplicando Teorema 16.3 con una función G adecuada. Explica por qué este ejemplo refuta la afirmación «Newton–Leibniz solo puede usarse cuando f es continua».

Ejercicio F5. Invariancia de la integral no significa invariancia del FTC puntual. Sea $g(t) = t$ en $[-1,1]$ y define f modificando solo el valor en 0:

$$f(0) = 10, \qquad f(t) = t \quad (t \neq 0).$$

1. Demuestra que f y g tienen la misma función de acumulación con base -1.
2. Calcula la derivada de esa acumulación en 0.
3. Explica por qué la invariancia de la integral bajo cambios finitos no autoriza a sustituir $f(0)$ por ese valor derivativo.

Ejercicio F6. Reparar una regla demasiado amplia. Un estudiante escribe:

«Si ϕ es diferenciable, entonces siempre

$$\frac{d}{dx}\int_a^{\phi(x)} f(t)\,dt = f(\phi(x))\phi'(x).$$

»

1. Señala al menos dos datos que faltan para que la expresión y la conclusión estén justificadas por Teorema 16.4.
2. Escribe una versión correcta del enunciado usando las hipótesis del teorema.
3. Explica por qué el resultado tampoco cubre automáticamente integrandos de la forma $F(x,t)$.

Nivel G — Desafíos integradores y síntesis C15–C17. Ejercicio G1. De integral a acumulación, derivada y forma global. En $[0,1]$ considera

$$f(t) = 2t - 1, \qquad A(x) = \int_0^x f(t)\,dt.$$

1. Usando una primitiva conocida y Corolario 16.2, demuestra que $A(x) = x^2 - x$.
2. Verifica mediante Teorema 16.2 que $A'(x) = 2x - 1$.
3. Calcula $\int_0^1 f(t)\,dt$.
4. Usa el signo de A' para localizar el mínimo absoluto de A en $[0,1]$ y calcular su valor.

Ejercicio G2. Diagnóstico completo de un escalón. Para

$$s(t) = \begin{cases} 0, & 0 \le t < \frac{1}{2}, \\ 1, & \frac{1}{2} \le t \le 1, \end{cases}$$

realiza un diagnóstico completo:

1. prueba que s es Riemann integrable y calcula su integral total;
2. escribe su acumulación con base 0;
3. demuestra que la acumulación es continua en todo $[0,1]$;
4. determina dónde es diferenciable y compara su derivada con s;
5. decide si s posee una primitiva en todo $[0,1]$.

Ejercicio G3. Dos rutas para una integral con extremos variables. Para $x \in [0,1]$ define

$$H(x) = \int_{x^2}^{x} (3t^2 + 1)\,dt.$$

1. Calcula $H'(x)$ directamente mediante Proposición 16.3.
2. Evalúa la integral usando la primitiva $G(t) = t^3 + t$ y deriva la fórmula obtenida.
3. Comprueba que las dos rutas coinciden.
4. Calcula $H'(1/2)$.

Ejercicio G4. Reconstruir el puente completo entre derivación e integración. Supón que $f \in \mathcal{R}[a,b]$, que G es continua en $[a,b]$, diferenciable en (a,b) y que $G' = f$ en el interior. Sin citar Teorema 16.3 como una caja negra, demuestra

$$\int_a^b f(x)\,dx = G(b) - G(a)$$

siguiendo esta ruta:

1. partición arbitraria;
2. teorema del valor medio diferencial en cada pieza;
3. suma etiquetada igual al incremento total de G;
4. encierro entre $L(f,P)$ y $U(f,P)$;
5. criterio de integrabilidad de Darboux.

Señala en qué paso se usa cada hipótesis.

Soluciones desarrolladas.

Soluciones del nivel A. **Solución A1.** La integral de una constante ya está disponible en C15. Para $x \in [0,2]$,

$$A(x) = \int_0^x 3\,dt = 3x.$$

Por tanto,

$$A(0) = 0, \qquad A(2) = 6.$$

Si $x, y \in [0,2]$,

$$A(y) - A(x) = 3y - 3x = 3(y - x).$$

Cuando $x < y$,

$$\int_x^y 3\,dt = 3(y - x).$$

Si $y < x$, la orientación de Definición 14.3 da

$$\int_x^y 3\,dt = -\int_y^x 3\,dt = -3(x - y) = 3(y - x).$$

Así, en ambos órdenes,

$$A(y) - A(x) = \int_x^y 3\,dt.$$

Solución A2. Si $0 \le x \le 1/2$, el intervalo orientado entre $1/4$ y x queda contenido en una región donde s es cero, salvo posiblemente el valor aislado en $1/2$. Por Corolario 14.1, $B(x) = 0$.

Si $1/2 \le x \le 1$, la aditividad orientada da

$$\begin{aligned} B(x) &= \int_{1/4}^{1/2} s(t)\,dt + \int_{1/2}^{x} s(t)\,dt \\ &= 0 + (x - \tfrac{1}{2}). \end{aligned}$$

Luego

$$\boxed{B(x) = \begin{cases} 0, & 0 \le x \le \frac{1}{2}, \\ x - \frac{1}{2}, & \frac{1}{2} \le x \le 1. \end{cases}}$$

Ésta es exactamente la misma fórmula obtenida para A_s en Ejemplo 16.1. El cambio de base de 0 a 1/4 no altera la acumulación porque

$$\int_0^{1/4} s = 0.$$

Solución A3. Por la aditividad orientada Proposición 14.3,

$$\int_p^x f = \int_p^q f + \int_q^x f.$$

Por tanto,

$$A_p(x) = \int_p^q f + A_q(x),$$

y al restar,

$$\boxed{A_p(x) - A_q(x) = \int_p^q f.}$$

El miembro derecho depende de p, q y de f, pero no de x. Por eso cambiar el punto base traslada verticalmente la función de acumulación por una constante.

Solución A4. La variable de integración está ligada por el símbolo integral. Por eso cambiar su nombre no cambia el valor:

$$\int_a^x f(t)\,dt = \int_a^x f(u)\,du.$$

Aquí a es el punto base fijo, x es el extremo variable y argumento libre de F, y t o u es una variable muda.

La escritura

$$\int_a^x f(x)\,dx$$

reutiliza la misma letra como extremo libre y variable ligada, ocultando los papeles. La forma clara es

$$\boxed{F(x) = \int_a^x f(t)\,dt.}$$

Solución A5. Sean $x < y$. Por §16.1,

$$A_f(y) - A_f(x) = \int_x^y f(t)\,dt.$$

Como $f \geq 0$, la monotonía de la integral Teorema 14.6 produce

$$\int_x^y f(t)\,dt \geq 0.$$

Luego $A_f(x) \leq A_f(y)$.

La desigualdad vale para todo par $x < y$, de modo que A_f es no decreciente. No se ha usado ninguna derivada.

Solución A6. Para $f(t) = 1$,

$$\int_0^1 f(t)\,dt = 1.$$

Por linealidad y Ejemplo 14.2,

$$\int_0^1 2t\,dt = 2\int_0^1 t\,dt = 2 \cdot \frac{1}{2} = 1.$$

Sin embargo,

$$A_f(x) = \int_0^x 1\,dt = x,$$

mientras que una primitiva de $2t$ es t^2, así que Corolario 16.2 da

$$A_g(x) = \int_0^x 2t\,dt = x^2.$$

Las acumulaciones coinciden cuando $x = x^2$, es decir,

$$\boxed{x = 0 \quad \text{o} \quad x = 1.}$$

Por tanto, compartir el valor integral total no determina la historia de acumulación dentro del intervalo.

Soluciones del nivel B. **Solución B1.** Por Proposición 16.1,

$$|A(1.2) - A(1.1)| \leq 4|1.2 - 1.1| = 0.4.$$

Para garantizar un error menor que 0.01 basta imponer $4|x - y| < 0.01$.

Una elección es

$$\boxed{\delta = \frac{0.01}{4} = 0.0025.}$$

Entonces $|x - y| < \delta$ implica

$$|A(x) - A(y)| \leq 4|x - y| < 4\delta = 0.01.$$

Solución B2. Para cualesquiera $x, y \in [a, b]$, la aditividad orientada da

$$A_p(y) - A_p(x) = \int_x^y f(t)\,dt.$$

La prueba de Proposición 16.1 depende solo de esta integral sobre el tramo entre x e y, no del punto base. Si I es el intervalo cerrado cuyos extremos son x e y, entonces

$$\left|\int_x^y f(t)\,dt\right| \leq \int_I |f(t)|\,dt \leq M|y - x|.$$

Por tanto,

$$\boxed{|A_p(y) - A_p(x)| \leq M|y - x|.}$$

Solución B3. Podemos suponer $x \leq y$, pues la desigualdad es simétrica.
Si $y \leq 1/2$, entonces $A_s(x) = A_s(y) = 0$, y la diferencia es cero.
Si $1/2 \leq x \leq y$, entonces

$$|A_s(y) - A_s(x)| = |(y - \tfrac{1}{2}) - (x - \tfrac{1}{2})| = y - x.$$

Queda el caso cruzado $x \leq 1/2 \leq y$. Allí $|A_s(y) - A_s(x)| = y - \frac{1}{2}$.
Como $x \leq 1/2$, $y - \frac{1}{2} \leq y - x$.
En todos los casos,

$$\boxed{|A_s(y) - A_s(x)| \leq |y - x|.}$$

Solución B4. Definamos $h = f - g$. Como $f, g \in \mathcal{R}[a, b]$, la linealidad de la integral en Teorema 14.6 garantiza que $h = f - g \in \mathcal{R}[a, b]$.

Además, por la misma linealidad,

$$(A_f - A_g)(y) - (A_f - A_g)(x) = \int_x^y h(t)\,dt.$$

Como $|h(t)| \leq \varepsilon$, Proposición 16.1 aplicado a h da

$$\boxed{|(A_f - A_g)(y) - (A_f - A_g)(x)| \leq \varepsilon|y - x|.}$$

Aplicamos ahora esta desigualdad a los puntos a y x. Como $(A_f - A_g)(a) = 0$, obtenemos

$$|A_f(x) - A_g(x)| = |(A_f - A_g)(x) - (A_f - A_g)(a)| \leq \varepsilon|x - a|.$$

Solución B5. Si $x < y$, entonces

$$A_f(y) - A_f(x) = \int_x^y f(t)\,dt.$$

La hipótesis $f \geq m$ y la monotonía integral implican

$$\int_x^y f(t)\,dt \geq \int_x^y m\,dt = m(y - x).$$

Como $m > 0$ y $y - x > 0$,

$$\boxed{A_f(y) - A_f(x) \geq m(y-x) > 0.}$$

Así $x < y$ implica $A_f(x) < A_f(y)$: la acumulación es estrictamente creciente. Toda función estrictamente creciente es inyectiva.

Solución B6. Como $A_f(a) = 0$, Proposición 16.1 con los puntos a y x da

$$|A_f(x)| = |A_f(x) - A_f(a)| \leq M|x-a|.$$

Para dos puntos arbitrarios $x, y \in [a,b]$, $|x-y| \leq b-a$.

Otra aplicación de Proposición 16.1 produce

$$\boxed{|A_f(x) - A_f(y)| \leq M|x-y| \leq M(b-a).}$$

El diámetro de la imagen es el supremo de las distancias entre dos valores de A_f; todas esas distancias están acotadas por $M(b-a)$.

Soluciones del nivel C. **Solución C1.** La función $t \mapsto t^2$ es continua en todo $[0,2]$. Por Teorema 16.2,

$$\boxed{A'_{[0,2]}(c) = c^2 \qquad (0 \leq c \leq 2).}$$

Si $0 < c < 2$, se trata de la derivada bilateral ordinaria. En $c = 0$ la derivada es relativa al dominio y usa incrementos positivos; en $c = 2$, incrementos negativos. No fue necesario evaluar primero $\int_0^x t^2\,dt$.

Solución C2. El escalón s es continuo en cada punto de

$$[0,1] \setminus \left\{\frac{1}{2}\right\}.$$

Por Teorema 16.1, $A'_s(c) = s(c)$ en todos esos puntos, con derivadas relativas en 0 y 1. Por tanto,

$$A'_s(c) = 0 \quad (0 \leq c < \tfrac{1}{2}),$$

y

$$A'_s(c) = 1 \quad (\tfrac{1}{2} < c \leq 1).$$

En el salto, la fórmula explícita de Ejemplo 16.1 da

$$\frac{A_s(\frac{1}{2}+h) - A_s(\frac{1}{2})}{h} = \begin{cases} 0, & h < 0, \\ 1, & h > 0, \end{cases}$$

para h suficientemente pequeño. Las derivadas laterales son 0 y 1, de modo que

$$\boxed{A'_s(1/2) \text{ no existe.}}$$

Solución C3. La función f es Riemann integrable, y es continua en $c = 1/4$ porque en un entorno relativo de ese punto es constantemente cero. Por Teorema 16.1,

$$A'_f\left(\frac{1}{4}\right) = f\left(\frac{1}{4}\right) = 0.$$

Análogamente, alrededor de $c = 3/4$ la función es constantemente uno, de modo que es continua en ese punto y

$$A'_f\left(\frac{3}{4}\right) = f\left(\frac{3}{4}\right) = 1.$$

La discontinuidad situada en $1/2$ no interviene: el teorema local exige continuidad únicamente en el punto donde se deriva la acumulación.

Solución C4. Para $h \neq 0$ admisible,

$$\frac{A_f(c+h) - A_f(c)}{h} - f(c) = \frac{1}{h}\int_c^{c+h} (f(t) - f(c))\,dt.$$

Sea I_h el intervalo cerrado cuyos extremos son c y $c+h$. Si $|h| < \rho$, entonces todo $t \in I_h$ satisface $|t - c| \leq |h| < \rho$, y por hipótesis $|f(t) - f(c)| \leq K|t - c| \leq K|h|$. Por la desigualdad triangular integral,

$$\begin{aligned}\left|\frac{A_f(c+h) - A_f(c)}{h} - f(c)\right| &\leq \frac{1}{|h|}\int_{I_h} |f(t) - f(c)|\,dt \\ &\leq \frac{1}{|h|}(K|h|)|h| \\ &= K|h|.\end{aligned}$$

El último miembro tiende a cero cuando $h \to 0$. Luego

$$\boxed{A'_f(c) = f(c).}$$

Solución C5. La continuidad global de f significa que f es continua en cada punto $c \in [a, b]$ relativamente al dominio. Además, Teorema 14.4 garantiza $f \in \mathcal{R}[a, b]$. Por tanto, Teorema 16.1 se aplica para cada c y produce

$$\boxed{(A_f)'_{[a,b]}(c) = f(c) \qquad (c \in [a, b]).}$$

En $c \in (a, b)$ la derivada es bilateral. En a es la derivada relativa desde la derecha y en b la derivada relativa desde la izquierda. Ésta es precisamente la formulación global de Teorema 16.2.

Solución C6. La función f difiere de la función cero únicamente en un punto. Por Corolario 14.1 es Riemann integrable y todas sus integrales coinciden con las de la función cero. Así, para todo $x \in [-1, 1]$,

$$A_f(x) = \int_{-1}^{x} f(t)\,dt = 0.$$

Por tanto A_f es constante y

$$\boxed{A'_f(0) = 0.}$$

Sin embargo, $f(0) = 7$.

No hay contradicción con Teorema 16.1 porque f no es continua en 0. El teorema no afirma la identidad derivativa en puntos de discontinuidad.

Soluciones del nivel D. **Solución D1.** Las reglas de derivación dan $F'(x) = 3x^2 - 2 = f(x)$, así que F es una primitiva. Por Teorema 16.3 o Corolario 16.2,

$$\int_{-1}^{2} (3x^2 - 2)\,dx = F(2) - F(-1).$$

Calculamos $F(2) = 8 - 4 + 5 = 9$, y $F(-1) = -1 + 2 + 5 = 6$.

Luego

$$\boxed{\int_{-1}^{2} (3x^2 - 2)\,dx = 3.}$$

Solución D2. Para cualquier $C \in \mathbb{R}$, $F'_C(x) = 4x^3$, de modo que cada $F_C(x) = x^4 + C$ es una primitiva de f.

Recíprocamente, sea H cualquier primitiva de $4x^3$. Como x^4 también lo es, Proposición 16.2 afirma que dos primitivas de la misma función difieren por una constante. Por tanto existe $C \in \mathbb{R}$ tal que $H(x) - x^4 = C$, es decir,

$$\boxed{H(x) = x^4 + C.}$$

Solución D3. Una primitiva conocida de $2x + 1$ es $x^2 + x$.

Por Proposición 16.2, todas las primitivas son $F(x) = x^2 + x + C$. La condición $F(0) = 7$ da $C = 7$. Por tanto,

$$\boxed{F(x) = x^2 + x + 7.}$$

Luego

$$\int_{0}^{2} (2x + 1)\,dx = F(2) - F(0) = (4 + 2 + 7) - 7 = \boxed{6}.$$

Solución D4. La función f difiere de la función continua $6x^2$ solo en el punto 1. Por Corolario 14.1, $f \in \mathcal{R}[0, 1]$.

La función $G(x) = 2x^3$ es continua en $[0, 1]$, diferenciable en $(0, 1)$ y, para $0 < x < 1$, $G'(x) = 6x^2 = f(x)$. Así se cumplen las hipótesis de Teorema 16.3 y

$$\int_{0}^{1} f(x)\,dx = G(1) - G(0) = 2.$$

El teorema exige la identidad $G' = f$ únicamente en el intervalo abierto $(0, 1)$. El valor aislado $f(1) = -100$ no participa en esa identidad y, además, una modificación finita no altera la integral.

Solución D5. En cada subintervalo $[x_{k-1}, x_k]$, la función G es continua y diferenciable en el interior. Por Teorema 9.3 existe $\xi_k \in (x_{k-1}, x_k)$ tal que

$$\frac{G(x_k) - G(x_{k-1})}{x_k - x_{k-1}} = G'(\xi_k) = f(\xi_k).$$

Multiplicando por Δx_k, $G(x_k) - G(x_{k-1}) = f(\xi_k)\Delta x_k$.

Al sumar desde $k = 1$ hasta n, el miembro izquierdo telescopa:

$$G(b) - G(a) = \sum_{k=1}^{n} f(\xi_k)\Delta x_k = R(f; P, \xi).$$

Toda suma etiquetada queda encerrada por Darboux: $L(f, P) \le G(b) - G(a) \le U(f, P)$. Como f es Riemann integrable, Teorema 14.1 permite escoger particiones con $U(f, P) - L(f, P)$ arbitrariamente pequeño, mientras la integral queda encerrada en el mismo intervalo. El único valor compatible con todos esos encierros es

$$\boxed{G(b) - G(a) = \int_a^b f.}$$

Solución D6. La función $f(t) = \frac{1}{1+t^2}$ es continua en $[0, 1]$. Por Corolario 16.1 posee una primitiva. Una construcción canónica es

$$F(x) := \int_0^x \frac{dt}{1 + t^2}.$$

Por Teorema 16.2, $F'(x) = \frac{1}{1+x^2}$ relativamente a $[0, 1]$. Esto demuestra existencia sin introducir ninguna función elemental nueva. Pero escribir

$$\int_0^1 \frac{dt}{1 + t^2} = F(1) - F(0) = F(1)$$

no produce un valor cerrado independiente de la integral, porque F ha sido definida precisamente mediante esa acumulación. El resultado es estructural, no todavía una técnica de integración simbólica.

Soluciones del nivel E. **Solución E1.** El integrando $f(t) = t^2 + 1$ es continuo en $[0, 1]$, y $\phi(x) = x^2$ envía $[-1, 1]$ en $[0, 1]$. Por Teorema 16.4, $H'(x) = f(x^2) \cdot 2x$.

Por tanto,

$$\boxed{H'(x) = 2x(x^4 + 1).}$$

Solución E2. Por orientación,

$$H(x) = -\int_1^x (1 + t^2)\, dt.$$

Aplicando Teorema 16.4 a la acumulación con extremo x,

$$\frac{d}{dx} \int_1^x (1 + t^2)\, dt = 1 + x^2.$$

Así,

$$\boxed{H'(x) = -(1 + x^2).}$$

En 0 y 1 la derivada se entiende relativamente al dominio $[0, 1]$.

Solución E3. Tomamos

$$f(t) = 1 + t^3, \qquad \alpha(x) = x^2, \qquad \beta(x) = 2x.$$

Para $x \in [0, 1]$, ambos extremos pertenecen a $[0, 2]$. Por Proposición 16.3,

$$\begin{aligned} K'(x) &= f(2x) \cdot 2 - f(x^2) \cdot 2x \\ &= 2(1 + 8x^3) - 2x(1 + x^6). \end{aligned}$$

Por tanto,

$$\boxed{K'(x) = 2 + 16x^3 - 2x - 2x^7.}$$

Solución E4. Aquí

$$\alpha(x) = 1 - x, \qquad \beta(x) = x, \qquad f(t) = t.$$

Entonces

$$\alpha'(x) = -1, \qquad \beta'(x) = 1.$$

Por Proposición 16.3,

$$H'(x) = f(x) \cdot 1 - f(1 - x)(-1) = x + (1 - x) = \boxed{1}.$$

En $x = 1/2$ los extremos coinciden, pero la integral orientada vale cero y la fórmula sigue siendo válida: el teorema no exige un orden fijo entre $\alpha(x)$ y $\beta(x)$. Como una primitiva de t es $t^2/2$, $H(x) = \frac{x^2}{2} - \frac{(1-x)^2}{2} = x - \frac{1}{2}$, y derivar esta expresión vuelve a dar $H'(x) = 1$.

Solución E5. Necesitamos

$$\left| \frac{1}{x - 1} \right| \leq 2, \qquad x \neq 1.$$

Esto equivale a $|x - 1| \geq \frac{1}{2}$.

Por tanto el dominio natural es

$$\boxed{D = (-\infty, \tfrac{1}{2}] \cup [\tfrac{3}{2}, \infty).}$$

En todo punto de acumulación $c \in D$, la función ϕ es diferenciable relativamente a D y $\phi'(c) = -\frac{1}{(c-1)^2}$. Como f es continua en $[-2, 2]$, Teorema 16.4 da

$$\boxed{H'_D(c) = -\frac{f(1/(c-1))}{(c-1)^2}.}$$

La comprobación del dominio precede a la fórmula porque el símbolo integral requiere que el extremo $\phi(x)$ pertenezca al intervalo donde f está definida.

Solución E6. Teorema 16.4 estudia expresiones del tipo

$$\int_p^{\phi(x)} f(t)\,dt$$

con **integrando fijo** $f(t)$. En $J(x)$ la variable x aparece dentro del integrando y los extremos son constantes, de modo que ese teorema no se aplica directamente.

Aquí podemos evaluar primero:

$$\begin{aligned} J(x) &= \int_0^1 x\,dt + \int_0^1 t\,dt \\ &= x + \frac{1}{2}. \end{aligned}$$

Por tanto,

$$\boxed{J'(x) = 1.}$$

La justificación abstracta del intercambio entre derivación e integración para integrandos $F(x,t)$ dependientes de parámetro pertenece a la teoría de integrales con parámetros del Tomo II.

Soluciones del nivel F. **Solución F1.** La función difiere de la función cero solo en el punto 0. Por Corolario 14.1 es Riemann integrable y todas sus integrales coinciden con las de la función cero. Su acumulación con cualquier punto base es, por tanto, constante; en particular, $A_f'(0) = 0$.

Pero $f(0) = 5$.

Así, $A_f'(0) \neq f(0)$.

La hipótesis faltante de Teorema 16.1 es la **continuidad de f en el punto** $c = 0$. La integrabilidad por sí sola no controla el valor puntual.

Solución F2. El escalón s de Ejemplo 16.1 es discontinuo en $1/2$. Sin embargo, por §16.1,

$$A_s(x) = \begin{cases} 0, & 0 \le x \le \frac{1}{2}, \\ x - \frac{1}{2}, & \frac{1}{2} \le x \le 1, \end{cases}$$

que es continua en todo $[0,1]$. Por tanto, la implicación propuesta es falsa. Lo que sí garantiza la continuidad de A_f es Proposición 16.1: para todo integrando Riemann-integrable, la acumulación es incluso Lipschitz.

Solución F3. Supongamos que existiera una primitiva F de s en $[0,1]$. Por Definición 16.2, F es diferenciable relativamente a $[0,1]$ en cada punto; por Teorema 6.1, F es continua en todo $[0,1]$.

En $(0, 1/2)$, $F'(x) = 0$.

La restricción de F a $[0,1/2]$ es continua en el intervalo cerrado y diferenciable en su interior. Por Teorema 9.3, para cualesquiera $u < v$ en $[0,1/2]$ existe $c \in (u,v)$ tal que $\frac{F(v)-F(u)}{v-u} = F'(c) = 0$. Por tanto $F(v) = F(u)$ y F es constante en $[0,1/2]$.

En $(1/2, 1)$ se tiene $F'(x) = 1$. Definamos $H(x) = F(x) - x$.

Como F y $x \mapsto x$ son continuas, H es continua en $[1/2, 1]$; además es diferenciable en $(1/2, 1)$ y $H'(x) = F'(x) - 1 = 0$. Aplicando nuevamente Teorema 9.3, H es constante en $[1/2, 1]$. Así, para constantes C, D,

$$F(x) = \begin{cases} C, & 0 \le x \le \frac{1}{2}, \\ x + D, & \frac{1}{2} \le x \le 1. \end{cases}$$

La continuidad de F fuerza el empalme de valores, pero en $1/2$ la derivada lateral izquierda es 0 y la derecha es 1. La derivada bilateral no existe. Esto contradice la definición de primitiva Definición 16.2. Por tanto,

$$\boxed{s \text{ no posee primitiva en } [0, 1].}$$

Solución F4. La función f difiere de $g(x) = 2x$ únicamente en el extremo 1. Como g es continua e integrable, Corolario 14.1 implica $f \in \mathcal{R}[0, 1]$.

Tomemos $G(x) = x^2$.

G es continua en $[0, 1]$, diferenciable en $(0, 1)$ y

$$G'(x) = 2x = f(x) \qquad (0 < x < 1).$$

Así Teorema 16.3 da

$$\boxed{\int_0^1 f(x)\,dx = G(1) - G(0) = 1.}$$

El integrando no es continuo en 1, pero la forma fuerte de Newton–Leibniz no exige continuidad global de f.

Solución F5. Por Corolario 14.1, modificar el valor de $g(t) = t$ en el único punto 0 no cambia ninguna integral sobre subintervalos. Por tanto, con base -1, $A_f(x) = A_g(x)$ para todo $x \in [-1, 1]$.

Como g es continua, Teorema 16.2 da $A_g'(0) = g(0) = 0$.

Así también

$$\boxed{A_f'(0) = 0,}$$

mientras $f(0) = 10$.

La invariancia bajo cambios finitos es una propiedad de **valores integrales**. No convierte al valor modificado $f(0)$ en la derivada de la acumulación; para esa identidad puntual se necesita la continuidad exigida por Teorema 16.1.

Solución F6. La fórmula escrita no es automática. Para aplicar Teorema 16.4 debemos especificar, entre otras cosas:

- un intervalo $[a, b]$ en el que f sea continua;
- un dominio A para ϕ y un punto de acumulación $c \in A$;
- que $\phi : A \to [a, b]$, para que el extremo variable sea admisible;
- que ϕ sea diferenciable en c relativamente a A.

Una versión correcta es: si $f : [a, b] \to \mathbb{R}$ es continua, $p \in [a, b]$, $\phi : A \to [a, b]$ es diferenciable en $c \in A$ y c es punto de acumulación de A, entonces

$$H(x) = \int_p^{\phi(x)} f(t)\,dt$$

es diferenciable en c y

$$\boxed{H'_A(c) = f(\phi(c))\phi'_A(c).}$$

El teorema no cubre un integrando $F(x, t)$ porque allí x modifica la función integrada, no solo el extremo. Ése es un problema paramétrico diferente.

Soluciones del nivel G. **Solución G1.** Una primitiva de $f(t) = 2t - 1$ es $G(t) = t^2 - t$.

Por Corolario 16.2, $A(x) = G(x) - G(0) = x^2 - x$. Como f es continua, Teorema 16.2 proporciona independientemente

$$\boxed{A'(x) = 2x - 1.}$$

La integral total es

$$\int_0^1 f(t)\,dt = A(1) = 1 - 1 = \boxed{0}.$$

Para estudiar el mínimo, observamos

$$A'(x) < 0 \quad (0 \le x < \tfrac{1}{2}),$$

$$A'(1/2) = 0,$$

y

$$A'(x) > 0 \quad (\tfrac{1}{2} < x \le 1).$$

Por los criterios de monotonía de C11, A decrece hasta $1/2$ y crece después. El mínimo absoluto se alcanza en $x = \frac{1}{2}$, con valor

$$A(1/2) = \frac{1}{4} - \frac{1}{2} = \boxed{-\frac{1}{4}}.$$

Solución G2. La función s es escalonada y, por tanto, Riemann integrable. Su integral total es

$$\int_0^1 s(t)\,dt = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \boxed{\frac{1}{2}}.$$

Su acumulación, ya calculada en Ejemplo 16.1, es

$$A_s(x) = \begin{cases} 0, & 0 \le x \le \frac{1}{2}, \\ x - \frac{1}{2}, & \frac{1}{2} \le x \le 1. \end{cases}$$

Las dos piezas coinciden en $x = 1/2$, así que A_s es continua; también lo garantiza Proposición 16.1. Para $c < 1/2$, $A_s'(c) = 0 = s(c)$; para $c > 1/2$, $A_s'(c) = 1 = s(c)$. En los extremos se interpreta la derivada relativamente a $[0,1]$. En $1/2$, las derivadas laterales son 0 y 1, por lo que la derivada no existe.

Finalmente, el argumento de §16.9 —reconstruido en F3— muestra que s no posee una primitiva en todo $[0,1]$. Así se separan los hechos correctos:

$$\boxed{\begin{array}{c} s \in \mathcal{R}[0,1], \qquad A_s \text{ es continua}, \\ s \text{ no es continua en } \tfrac{1}{2}, \qquad A_s \text{ no es derivable en } \tfrac{1}{2}, \\ s \text{ no posee primitiva en } [0,1]. \end{array}}$$

En particular, la integrabilidad de un integrando no implica su continuidad, ni la diferenciabilidad global de su acumulación, ni la existencia de una primitiva; sí implica la continuidad de la acumulación por Proposición 16.1.

Solución G3. Con

$$f(t) = 3t^2 + 1, \qquad \alpha(x) = x^2, \qquad \beta(x) = x,$$

Proposición 16.3 da

$$\begin{aligned} H'(x) &= f(x) \cdot 1 - f(x^2) \cdot 2x \\ &= (3x^2 + 1) - 2x(3x^4 + 1). \end{aligned}$$

Por tanto,

$$\boxed{H'(x) = 3x^2 + 1 - 6x^5 - 2x.}$$

Como $G(t) = t^3 + t$ es una primitiva de f, Corolario 16.2 produce

$$\begin{aligned} H(x) &= G(x) - G(x^2) \\ &= (x^3 + x) - (x^6 + x^2). \end{aligned}$$

Derivando, $H'(x) = 3x^2 + 1 - 6x^5 - 2x$, que coincide con la primera ruta.
En $x = 1/2$, $H'(1/2) = \frac{3}{4} + 1 - \frac{6}{32} - 1 = \frac{9}{16}$.
Así,

$$\boxed{H'(1/2) = \frac{9}{16}.}$$

Solución G4. Tomemos una partición arbitraria

$$P = \{a = x_0 < x_1 < \cdots < x_n = b\}.$$

La continuidad de G en cada intervalo cerrado y su diferenciabilidad en cada interior permiten aplicar Teorema 9.3. Para cada k existe $\xi_k \in (x_{k-1}, x_k)$ tal que

$$G(x_k) - G(x_{k-1}) = G'(\xi_k)\Delta x_k = f(\xi_k)\Delta x_k.$$

Sumando,

$$\begin{aligned} G(b) - G(a) &= \sum_{k=1}^{n} \big(G(x_k) - G(x_{k-1})\big) \\ &= \sum_{k=1}^{n} f(\xi_k)\Delta x_k \\ &= R(f; P, \xi). \end{aligned}$$

Como f es Riemann integrable, está acotada y sus sumas de Darboux están definidas. El encierro de sumas etiquetadas proporciona $L(f, P) \leq G(b) - G(a) \leq U(f, P)$.

Por otra parte,

$$L(f, P) \leq \int_a^b f \leq U(f, P).$$

Dado $\varepsilon > 0$, Teorema 14.1 permite escoger P con $U(f, P) - L(f, P) < \varepsilon$. Los dos números $G(b) - G(a)$ y $\int_a^b f$ quedan entonces dentro de un mismo intervalo de longitud menor que ε, por lo que

$$\left| G(b) - G(a) - \int_a^b f \right| < \varepsilon.$$

Como esto vale para toda $\varepsilon > 0$,

$$\boxed{\int_a^b f(x)\, dx = G(b) - G(a).}$$

Las hipótesis se usan así:

- continuidad de G: MVT en cada pieza cerrada;
- diferenciabilidad de G en el interior: MVT y relación $G' = f$;
- igualdad $G' = f$: transforma cada incremento de G en una contribución de f;
- integrabilidad de f: proporciona Darboux y brechas arbitrariamente pequeñas.

No se necesitó continuidad global de f.

Capítulo 17

Logaritmo, exponencial y funciones relacionadas desde el cálculo

17.1. Por qué construir una función nueva

El teorema fundamental del cálculo permite construir una primitiva de una función continua mediante acumulación. Para encontrarla ya no dependemos de reconocer la derivada de una fórmula conocida: fijado un punto base a, tenemos

$$A_f(x) = \int_a^x f(t)\,dt$$

es una primitiva de f en el sentido preciso de Corolario 16.1, y Teorema 16.2 explica por qué su derivada recupera al integrando. La construcción permite introducir una función nueva a partir de la tasa que queremos recuperar. Podremos estudiar sus propiedades desde la definición integral, aunque no dispongamos de una fórmula previa para ella.

El hueco de la regla de la potencia. Consideremos primero el patrón más familiar. Para un entero $n \neq -1$, en cualquier intervalo donde la expresión tenga sentido,

$$\frac{d}{dx}\left(\frac{x^{n+1}}{n+1}\right) = x^n.$$

Por tanto la familia de potencias proporciona, de manera uniforme, primitivas para $\ldots, x^{-3}, x^{-2}, 1, x, x^2, x^3, \ldots$ con una única excepción aparente: cuando $n = -1$ la receta exigiría escribir $\frac{x^0}{0}$, que no es una expresión definida. El integrando que queda fuera del patrón es precisamente

$$\boxed{\frac{1}{x}.}$$

Conviene interpretar correctamente este fallo.

Que falle una fórmula no significa que falle la existencia
La imposibilidad de sustituir $n = -1$ en $x^{n+1}/(n+1)$ **no demuestra** que $1/x$ carezca de primitiva. Solo demuestra que esa familia concreta de fórmulas deja de proporcionarla. De hecho, C17 ya garantiza existencia en todo intervalo compacto contenido en $(0, \infty)$, porque $x \mapsto 1/x$ es continua allí.

Ésta es una diferencia conceptual fundamental. Una cosa es preguntar

> «¿Cuál de las funciones que ya conozco tiene derivada $1/x$?»

y otra muy distinta es preguntar

> «¿Puedo construir rigurosamente una función cuya derivada sea $1/x$, aunque todavía no pertenezca a mi repertorio de funciones nombradas?»

El teorema fundamental responde afirmativamente a la segunda pregunta.

La integral como máquina de construcción. La función $t \longmapsto \frac{1}{t}$ es continua en $(0,\infty)$. En particular, si $0 < a < b$, su restricción a $[a,b]$ es continua y, por tanto, Riemann-integrable. Podemos entonces usar el mecanismo de acumulación de C17. El punto base más natural para nuestro propósito es 1. Esto sugiere considerar, para $x > 0$, la expresión

$$\boxed{L(x) := \int_1^x \frac{dt}{t}.}$$

Usaremos provisionalmente la letra L mientras comprobamos que la integral define una función en todo el dominio positivo. Esa verificación permitirá introducir la definición de logaritmo natural y la notación $\ln x$. La expresión anterior está bien definida para cada $x > 0$. Si $x \geq 1$, integramos sobre el intervalo compacto $[1,x]$. Si $0 < x < 1$, la convención de integral orientada de Definición 14.3 da

$$\int_1^x \frac{dt}{t} = -\int_x^1 \frac{dt}{t},$$

y nuevamente el integrando es continuo en el intervalo compacto $[x,1]$. Así, sin haber supuesto todavía ninguna propiedad de los logaritmos, hemos obtenido una función real definida para todos los reales positivos: $L : (0,\infty) \longrightarrow \mathbb{R}$.

Lo que C17 ya permite anticipar. La elección anterior no es arbitraria. Si fijamos un punto $c > 0$, podemos encerrar c y 1 en algún intervalo compacto contenido en $(0,\infty)$. Sobre ese intervalo, $t \mapsto 1/t$ es continua. Por tanto, el teorema fundamental del cálculo será exactamente la herramienta que permitirá demostrar en §17.2 que $L'(c) = \frac{1}{c}$. No necesitamos adivinar una fórmula cerrada antes de diferenciar. La integral **es** la construcción de la función buscada. Este cambio de perspectiva es uno de los puntos estructurales del libro:

integrando conocido $\longrightarrow$ función de acumulación $\longrightarrow$ nueva función trascendente.

La función se construye a partir de la integral. Sus propiedades deberán deducirse de esa definición y de los resultados ya disponibles, sin presuponer las leyes logarítmicas que queremos demostrar.

Por qué el punto base 1 es especial. Podríamos haber comenzado en cualquier punto base positivo a y definido

$$L_a(x) := \int_a^x \frac{dt}{t}.$$

Por Proposición 16.2, una vez que se verifique que estas funciones tienen la misma derivada $1/x$, dos elecciones de punto base solo podrán diferir por una constante. Por tanto, escoger 1 no es necesario para la existencia de una primitiva. Sin embargo, 1 tiene dos ventajas que serán decisivas:

- es el elemento neutro de la multiplicación en los reales positivos;
- normaliza automáticamente la construcción, porque

$$L(1) = \int_1^1 \frac{dt}{t} = 0.$$

Todavía **no** hemos demostrado que L convierta productos en sumas. Esa propiedad será un resultado posterior, no una razón válida para introducirla clandestinamente como si ya conociéramos el logaritmo escolar.

No podemos usar todavía las propiedades que queremos demostrar
En este punto no están disponibles como hechos de la teoría de C18:

- $\ln(xy) = \ln x + \ln y$;
- la existencia de una función exponencial inversa;
- el número e como base de esa exponencial;
- la notación e^x para exponentes reales arbitrarios;
- la definición general a^x para $a > 0$ y $x \in \mathbb{R}$.

Todos esos objetos y propiedades deben construirse en el orden lógico fijado para el capítulo.

De una primitiva a una teoría completa. La integral

$$\int_1^x \frac{dt}{t}$$

resuelve únicamente el primer problema: proporciona el candidato natural a primitiva de $1/x$ en los reales positivos. Para convertir esta construcción en la teoría del logaritmo y la exponencial tendremos que responder, en orden, preguntas adicionales.

1. ¿Qué regularidad posee esta función y cuál es exactamente su derivada?
2. ¿Cómo se comporta frente al producto y al cociente de números positivos?
3. ¿Es estrictamente monótona y cuál es su imagen?
4. ¿Su imagen es todo $\mathbb{R}$, de modo que pueda invertirse globalmente?
5. ¿Qué propiedades hereda su inversa y cómo nace de allí la exponencial?
6. ¿Cómo se extienden después, sin circularidad, las potencias a exponentes reales?

El resto de C18 responderá estas preguntas sin usar sustitución integral, series de potencias ni ecuaciones diferenciales como fundamento. La cadena será deliberadamente autorreferencial: cada paso utilizará únicamente resultados ya construidos en el libro. La definición de **logaritmo natural** recogerá esta construcción. Su dominio, su normalización y su derivada se justificarán mediante las propiedades de la integral y el teorema fundamental.

Antes de seguir. ¿Por qué el fallo de la fórmula $x^{n+1}/(n+1)$ en $n = -1$ no demuestra inexistencia de una primitiva de $1/x$? ¿Qué resultado de C17 garantiza que podemos construir una? ¿Por qué la orientación de la integral permite usar el mismo punto base 1 tanto para $x > 1$ como para $0 < x < 1$? ¿Qué propiedades familiares del logaritmo hemos evitado suponer y por qué sería circular usarlas ya?

17.2. El logaritmo natural como función de acumulación

La sección anterior aisló la función

$$L(x) = \int_1^x \frac{dt}{t}, \qquad x > 0,$$

como una primitiva de $1/x$ en los reales positivos. Al comprobar que la integral está definida para cada entrada positiva podremos adoptar la notación del logaritmo y deducir sus propiedades desde esa construcción.

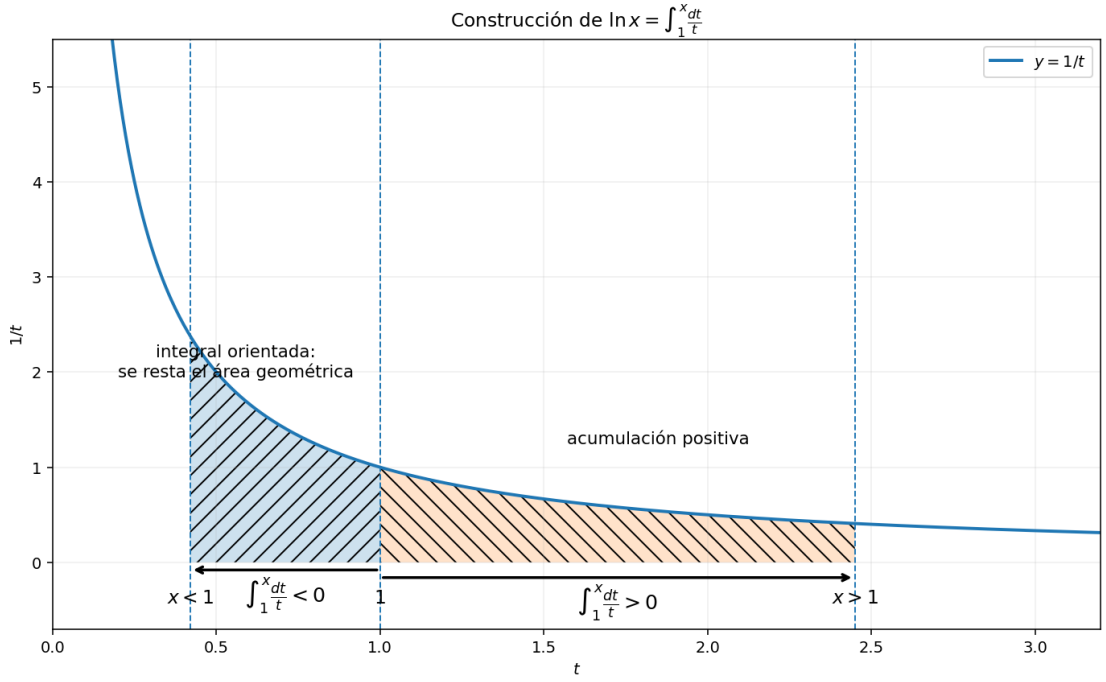


FIGURA 17.1. El logaritmo natural nace como una función de acumulación: para $x > 1$ la integral orientada desde 1 añade área positiva bajo $1/t$, mientras que para $0 < x < 1$ la orientación invierte el signo del área geométrica recorrida.

Definición y dominio.

DEFINICIÓN 17.1. **Definición (logaritmo natural).** Para todo $x > 0$, definimos el **logaritmo natural de** x por

$$\boxed{\ln x := \int_1^x \frac{dt}{t}.}$$

Así obtenemos una función

$$\ln : (0, \infty) \longrightarrow \mathbb{R}.$$

Figura 17.1 hace visible que el mismo punto base 1 construye $\ln x$ a ambos lados de 1 sin abandonar el dominio positivo ni confundir integral orientada con área geométrica.

La definición exige $x > 0$ por una razón estructural, no por convención. Si $x > 0$, el intervalo cerrado cuyos extremos son 1 y x está contenido en $(0, \infty)$, donde $t \mapsto 1/t$ es continua. Por tanto la integral de la definición existe. En cambio, si intentáramos atravesar 0, el integrando $1/t$ dejaría de estar definido en el intervalo de integración; la teoría de la integral de Riemann desarrollada hasta aquí no autoriza esa expresión.

Dominio no significa todavía imagen
Escribir $\ln : (0, \infty) \to \mathbb{R}$ significa que cada entrada positiva recibe un valor real. **No** hemos demostrado todavía que todo número real sea alcanzado por ln. La determinación de la

imagen, y en particular la sobreyectividad sobre $\mathbb{R}$, pertenece a §17.4 y será necesaria antes de definir una inversa global.

La elección del punto base 1 fija inmediatamente la normalización:

$$\boxed{\ln 1 = \int_1^1 \frac{dt}{t} = 0.}$$

Esta igualdad procede únicamente de la convención de integral orientada de Definición 14.3. No hemos usado ninguna ley del producto ni ninguna propiedad previa de un logaritmo.

La derivada se obtiene del teorema fundamental. La construcción integral fue elegida precisamente para que el teorema fundamental del cálculo determine su derivada. Conviene escribir el argumento sin ocultar el cambio de punto base.

Teorema 17.1. **Teorema (derivada del logaritmo natural).** La función

$$\ln : (0,\infty) \longrightarrow \mathbb{R}$$

es diferenciable en todo punto de su dominio y satisface

$$\boxed{(\ln x)' = \frac{1}{x} \qquad (x > 0).}$$

Demostración. Fijemos arbitrariamente $c > 0$. Elegimos números a, b tales que

$$0 < a < \min\{1, c\} \qquad \text{y} \qquad b > \max\{1, c\}.$$

Entonces 1 y c pertenecen al interior de $[a, b]$, y la función $f(t) = \frac{1}{t}$ es continua en todo $[a, b]$. Definamos la función de acumulación con punto base a:

$$A(x) := \int_a^x \frac{dt}{t}, \qquad x \in [a, b].$$

Por la primera forma global del teorema fundamental, Teorema 16.2,

$$A'(x) = \frac{1}{x} \qquad (x \in (a, b)).$$

Ahora usamos únicamente la aditividad orientada de la integral, Proposición 14.3. Para todo $x \in [a, b]$,

$$\int_a^x \frac{dt}{t} = \int_a^1 \frac{dt}{t} + \int_1^x \frac{dt}{t}.$$

Por la definición de A y Definición 17.1, $A(x) = A(1) + \ln x$, de modo que $\ln x = A(x) - A(1)$. El término $A(1)$ es constante respecto de x. Como $c \in (a, b)$, $(\ln)'(c) = A'(c) = \frac{1}{c}$. El punto $c > 0$ fue arbitrario, por lo que la igualdad vale en todo $(0, \infty)$. □

La prueba merece ser leída en su orden lógico. Primero se trabaja dentro de un intervalo compacto donde la teoría de C17 se aplica literalmente; después se cambia el punto base mediante una identidad de integrales; solo al final se concluye una afirmación para todo el dominio positivo. No hemos extendido informalmente el FTC desde un intervalo compacto a una semirrecta.

Diferenciabilidad y continuidad. El teorema anterior produce más que una fórmula de cálculo. Como ln es diferenciable en cada $x > 0$, la implicación diferenciabilidad $\Rightarrow$ continuidad de Teorema 6.1 da

$$\boxed{\ln \text{ es continua en } (0, \infty).}$$

Ésta es la primera propiedad de regularidad global de la función recién construida. Obsérvese nuevamente el sentido de la dependencia:

$$\text{integral} \Longrightarrow \text{derivada por FTC} \Longrightarrow \text{continuidad.}$$

No estamos usando continuidad del logaritmo para justificar la integral que lo define; usamos continuidad de $1/t$ para construir la integral y después deducimos continuidad de la nueva función.

Los incrementos del logaritmo son integrales sobre el intervalo recorrido. La aditividad orientada permite registrar una identidad que será útil repetidamente. Si $x, y > 0$, elegimos un intervalo compacto positivo que contenga a $1, x, y$. Entonces

$$\int_1^y \frac{dt}{t} = \int_1^x \frac{dt}{t} + \int_x^y \frac{dt}{t},$$

y, por Definición 17.1,

$$\boxed{\ln y - \ln x = \int_x^y \frac{dt}{t}.}$$

Esta fórmula no es todavía una ley logarítmica de productos. Solo expresa que la diferencia entre dos valores de una función de acumulación es la integral sobre el tramo añadido o retirado. En §17.4 permitirá estudiar su geometría; por ahora la conservamos como consecuencia directa de la definición y de la aditividad de la integral.

Qué hemos demostrado y qué sigue prohibido usar. En este punto la teoría disponible de ln consta exactamente de los siguientes hechos:

- su dominio es $(0, \infty)$;
- está definida por $\ln x = \int_1^x dt/t$;
- $\ln 1 = 0$;
- es diferenciable y $(\ln x)' = 1/x$;
- en consecuencia es continua;
- sus incrementos satisfacen

$$\ln y - \ln x = \int_x^y \frac{dt}{t}.$$

Todavía no hemos demostrado $\ln(xy) = \ln x + \ln y$, ni las fórmulas para cocientes o potencias, ni la monotonía global, ni la concavidad, ni el comportamiento en 0^+ o en $+\infty$, ni que la imagen sea todo $\mathbb{R}$.

La fórmula de la derivada no reemplaza las demostraciones posteriores
De $(\ln x)' = 1/x > 0$ podremos deducir crecimiento estricto usando la teoría ya establecida de monotonía. Pero esa deducción y sus consecuencias globales se organizarán en §17.4. Antes debemos resolver en §17.3 una propiedad algebraica fundamental: cómo responde ln a la multiplicación de números positivos.

Para demostrar la ley del producto compararemos dos funciones de una variable, calcularemos sus derivadas por la regla de la cadena y utilizaremos la unicidad de primitivas hasta una constante. Así podremos obtener la ecuación funcional con los resultados disponibles, antes de introducir sustituciones integrales.

Antes de seguir. ¿Por qué la definición de $\ln x$ obliga a trabajar con $x > 0$? ¿Por qué en la prueba de Teorema 17.1 encerramos primero 1 y c en un intervalo compacto? ¿Qué papel cumple Proposición 14.3 al cambiar el punto base? ¿Por qué escribir $\ln : (0,\infty) \to \mathbb{R}$ no demuestra que ln sea sobreyectiva? ¿Qué propiedades del logaritmo continúan todavía pendientes de demostración?

17.3. La ecuación funcional del logaritmo

Hasta ahora el logaritmo natural ha sido construido únicamente como una función de acumulación y sabemos que

$$(\ln x)' = \frac{1}{x} \qquad (x > 0).$$

La siguiente pregunta es algebraica: ¿qué ocurre cuando el argumento es un producto? La identidad familiar $\ln(xy) = \ln x + \ln y$ no puede suponerse, porque precisamente forma parte de la teoría que estamos construyendo. Tampoco usaremos un cambio de variable en la integral que define a ln, pues la sustitución sistemática pertenece a C19. La idea será distinta: fijaremos uno de los factores, compararemos dos funciones de la otra variable y mostraremos que tienen la misma derivada. La unicidad de primitivas de Proposición 16.2 reducirá entonces la diferencia entre ambas a una constante, que podremos determinar en $x = 1$.

Un parámetro fijo convierte el producto en un problema de primitivas. Fijemos $y > 0$. Para $x > 0$ consideremos

$$F_y(x) := \ln(xy) \qquad \text{y} \qquad G(x) := \ln x.$$

Las dos expresiones están bien definidas porque $x > 0$ y $y > 0$ implican $xy > 0$. La función interior $x \longmapsto xy$ es diferenciable y tiene derivada constante y. Por la regla de la cadena Teorema 7.4 y la fórmula de Teorema 17.1,

$$\begin{aligned} F_y'(x) &= (\ln)'(xy)\, y \\ &= \frac{1}{xy}\, y \\ &= \frac{1}{x}. \end{aligned}$$

Por otra parte, $G'(x) = \frac{1}{x}$.

Así, para $y > 0$ fijo, las funciones F_y y G tienen exactamente la misma derivada en todo $(0,\infty)$. El punto delicado es convertir esta igualdad local de derivadas en una identidad global de funciones. No diremos simplemente «por tanto difieren en una constante» sobre una semirrecta

sin justificar el paso. Tomaremos un intervalo compacto positivo que contenga los puntos que queremos comparar y aplicaremos allí el resultado ya demostrado para primitivas.

Proposición 17.1. **Proposición (ecuación funcional del logaritmo natural).** Para todos $x, y > 0$,

$$\boxed{\ln(xy) = \ln x + \ln y.}$$

Demostración. Fijemos $x, y > 0$. Elegimos a, b tales que

$$0 < a < \text{mín}\{1, x\} \qquad \text{y} \qquad b > \text{máx}\{1, x\}.$$

En el intervalo $[a, b]$ consideramos las restricciones de

$$F_y(t) = \ln(ty) \qquad \text{y} \qquad G(t) = \ln t.$$

Como ambas funciones son diferenciables en un intervalo abierto que contiene a $[a, b]$, sus restricciones son diferenciables relativamente a $[a, b]$: en los extremos, los cocientes relativos son sublímites del límite bilateral. El cálculo anterior da, para todo $t \in [a, b]$, $(F_y)'_{[a,b]}(t) = G'_{[a,b]}(t) = \frac{1}{t}$. Por tanto, ambas son primitivas de la misma función $t \mapsto 1/t$ en $[a, b]$. Por Proposición 16.2 existe una constante C tal que

$$F_y(t) = G(t) + C \qquad (t \in [a, b]).$$

Evaluamos en $t = 1$. Como $\ln 1 = 0$,

$$C = F_y(1) - G(1) = \ln y - \ln 1 = \ln y.$$

En consecuencia,

$$F_y(t) = G(t) + \ln y \qquad (t \in [a, b]).$$

Como $x \in [a, b]$, $\ln(xy) = \ln x + \ln y$.

Los números $x, y > 0$ fueron arbitrarios. □

Figura 17.2 representa el mecanismo de la prueba de Proposición 17.1 mediante el caso ilustrativo $y = 2$; la igualdad de derivadas y la determinación de la constante son generales.

La prueba utiliza únicamente tres ingredientes ya disponibles: la derivada de ln, la regla de la cadena y la unicidad de primitivas hasta una constante. La multiplicación aparece dentro del argumento de ln, pero no hemos transformado ninguna integral mediante sustitución.

La ley del producto no se ha importado: se ha demostrado
La definición

$$\ln x = \int_1^x \frac{dt}{t}$$

no contiene por sí sola la identidad $\ln(xy) = \ln x + \ln y$. Esa identidad emerge porque, al fijar y, las funciones $x \mapsto \ln(xy)$ y $x \mapsto \ln x$ poseen la misma derivada y la normalización en $x = 1$ determina la constante restante.

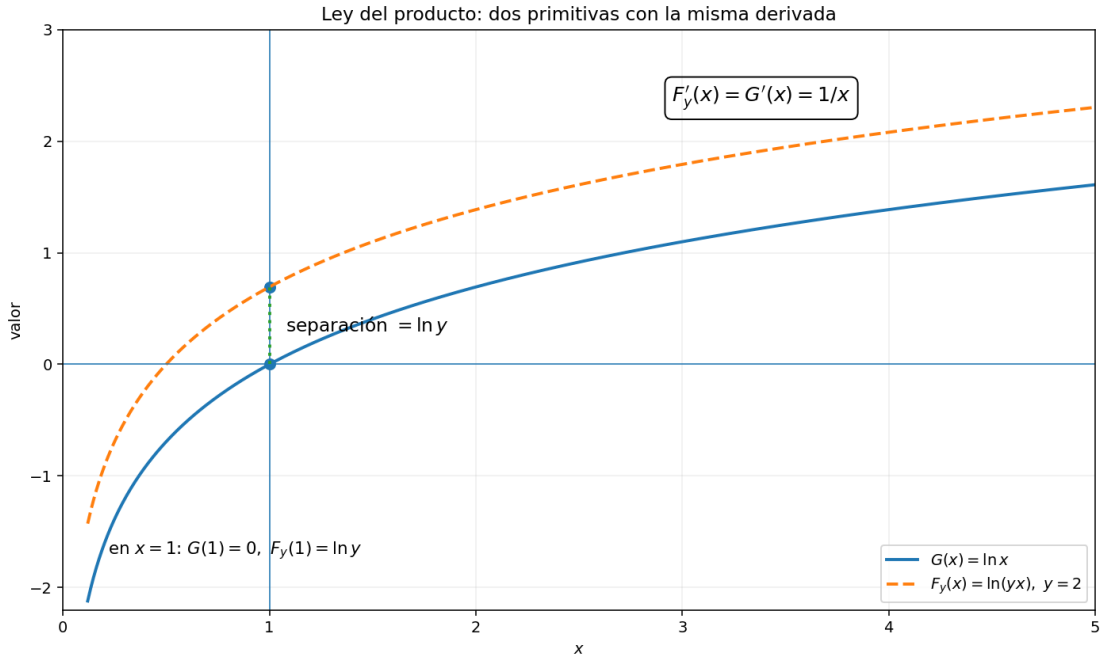


FIGURA 17.2. Para un factor positivo y fijo, las funciones $F_y(x) = \ln(xy)$ y $G(x) = \ln x$ tienen la misma derivada $1/x$; por eso sus gráficas difieren por una traslación vertical constante, determinada en $x = 1$ como $\ln y$.

Recíprocos, cocientes y potencias. Una vez demostrada la ley del producto, varias fórmulas familiares dejan de ser reglas memorizadas y pasan a ser consecuencias formales.

COROLARIO 17.1. **Corolario (recíprocos, cocientes y potencias racionales).** Sean $x, y > 0$.

1. Para el recíproco,
$$\boxed{\ln\left(\frac{1}{x}\right) = -\ln x.}$$
2. Para el cociente,
$$\boxed{\ln\left(\frac{x}{y}\right) = \ln x - \ln y.}$$
3. Para todo entero $n \in \mathbb{Z}$,
$$\boxed{\ln(x^n) = n \ln x.}$$
4. Para todo racional $q \in \mathbb{Q}$, interpretando x^q mediante las potencias racionales positivas construidas en C09,
$$\boxed{\ln(x^q) = q \ln x.}$$

Demostración. Como $x(1/x) = 1$, Proposición 17.1 da
$$0 = \ln 1 = \ln x + \ln\left(\frac{1}{x}\right),$$
de donde

$$\ln\left(\frac{1}{x}\right) = -\ln x.$$

Aplicando de nuevo Proposición 17.1,

$$\begin{aligned}\ln\left(\frac{x}{y}\right) &= \ln\left(x\frac{1}{y}\right)\\ &= \ln x + \ln\left(\frac{1}{y}\right)\\ &= \ln x - \ln y.\end{aligned}$$

Para $n \geq 1$, la identidad

$$x^n = \underbrace{x \cdots x}_{n \text{ factores}}$$

y una inducción usando Proposición 17.1 producen $\ln(x^n) = n \ln x$.

Para $n = 0$,

$$\ln(x^0) = \ln 1 = 0 = 0 \ln x.$$

Si $n < 0$, escribimos $n = -m$ con $m > 0$ y usamos el caso del recíproco:

$$\ln(x^n) = \ln\left(\frac{1}{x^m}\right) = -\ln(x^m) = -m \ln x = n \ln x.$$

Queda el caso racional. Sea

$$q = \frac{m}{n}, \qquad m \in \mathbb{Z}, \quad n \in \mathbb{N}, \quad n > 0,$$

en forma reducida, y sea $r := x^{1/n} > 0$, la raíz positiva construida en C09. Entonces $r^n = x$, y por el caso entero, $\ln x = \ln(r^n) = n \ln r$.

Por tanto, $\ln r = \frac{1}{n} \ln x$.

La definición de potencia racional de Sección 8.5 da $x^{m/n} = r^m$. Aplicando nuevamente el caso entero,

$$\begin{aligned}\ln(x^{m/n}) &= \ln(r^m)\\ &= m \ln r\\ &= \frac{m}{n} \ln x.\end{aligned}$$

Así, $\ln(x^q) = q \ln x$ para todo $q \in \mathbb{Q}$. □

La compatibilidad con C09 es una consecuencia, no una redefinición. Es importante notar qué ocurrió en el último paso. Las potencias racionales ya existían antes de C18: C09 las construyó algebraicamente a partir de raíces y potencias enteras. Aquí **no** las hemos redefinido usando logaritmos. Hemos demostrado que el logaritmo recién construido es compatible con aquella teoría anterior:

$$\boxed{\ln(x^q) = q \ln x \qquad (x > 0,\ q \in \mathbb{Q}).}$$

Más adelante, cuando definamos potencias con exponentes reales arbitrarios, esta compatibilidad será una condición indispensable: la teoría nueva deberá prolongar la antigua sin modificarla.

Ejemplo 17.1. **Tres lecturas de una misma ley.** Sin calcular numéricamente ningún logaritmo, las identidades anteriores permiten escribir

$$\ln\left(\frac{27}{8}\right) = \ln\left(\left(\frac{3}{2}\right)^3\right) = 3\ln\left(\frac{3}{2}\right) = 3(\ln 3 - \ln 2),$$

y también

$$\ln \sqrt[3]{4} = \ln(4^{1/3}) = \frac{1}{3}\ln 4 = \frac{2}{3}\ln 2.$$

En ambos casos solo se usan la ecuación funcional y la compatibilidad con potencias racionales previamente construidas.

Qué no se ha usado. Podría parecer natural intentar demostrar $\ln(xy) = \ln x + \ln y$ partiendo de

$$\int_1^{xy} \frac{dt}{t}$$

y efectuando un cambio de variable. Esa ruta es válida una vez desarrollada la teoría de sustitución, pero aquí sería un adelantamiento curricular: la sustitución pertenece a C19. La prueba adoptada mantiene la dependencia estrictamente regresiva:

$$\boxed{\begin{array}{c}\text{FTC}\\ \Longrightarrow \ln' = \dfrac{1}{x}\\ \Longrightarrow \text{cadena}\\ \Longrightarrow \text{misma derivada}\\ \Longrightarrow \text{unicidad de primitivas}\\ \Longrightarrow \ln(xy) = \ln x + \ln y.\end{array}}$$

Tampoco se ha usado la exponencial, el número e, la monotonía de ln, su rango completo ni ninguna teoría de exponentes irracionales.

Lo disponible al terminar §17.3. La teoría del logaritmo natural contiene ahora:

- la definición integral de Definición 17.1;
- la derivada $(\ln x)' = 1/x$ de Teorema 17.1;
- continuidad en $(0, \infty)$;
- la ecuación funcional del producto Proposición 17.1;
- las reglas de recíproco y cociente;
- la compatibilidad $\ln(x^q) = q \ln x$ para todo $q \in \mathbb{Q}$ y $x > 0$.

El siguiente paso será geométrico y global. En §17.4 usaremos la derivada positiva para demostrar crecimiento estricto, estudiaremos la concavidad y, apoyándonos en la ecuación funcional, probaremos el comportamiento no acotado necesario para determinar exactamente la imagen de ln antes de invertirla.

Antes de seguir. ¿Por qué fijar $y > 0$ convierte la ley del producto en una comparación entre primitivas? ¿Dónde entra exactamente la regla de la cadena? ¿Por qué la constante de diferencia se determina evaluando en $x = 1$? ¿Cómo se deduce la fórmula del cociente a partir del recíproco? ¿Por qué la ley racional no redefine las potencias de C09? ¿Qué técnica de integración hemos evitado deliberadamente y por qué pertenece al capítulo siguiente?

17.4. Geometría y rango del logaritmo

La ecuación funcional de §17.3 describió cómo responde el logaritmo a la multiplicación. Ahora debemos resolver una cuestión distinta y global: **qué forma tiene la función y qué valores alcanza**. Este paso no puede omitirse. Para definir después una inversa global no basta con saber que ln es continua o que satisface una ley algebraica. Debemos probar dos hechos separados:

1. que dos entradas positivas distintas no pueden tener el mismo logaritmo;
2. que cada número real aparece efectivamente como logaritmo de algún número positivo.

El primer problema será consecuencia del signo de la derivada. El segundo exigirá combinar continuidad, la ecuación funcional, la propiedad arquimediana y el teorema del valor intermedio.

Una derivada positiva en todo el dominio. Por Teorema 17.1,

$$(\ln x)' = \frac{1}{x} \qquad (x > 0).$$

Como $x > 0$, también $\frac{1}{x} > 0$.

La positividad no ocurre en un punto aislado: vale en **todo** el intervalo $(0, \infty)$. Podemos, por tanto, aplicar el criterio global de monotonía de Teorema 10.1.

Proposición 17.2. **Proposición (geometría básica del logaritmo natural).** La función

$$\ln : (0, \infty) \longrightarrow \mathbb{R}$$

es estrictamente creciente y estrictamente cóncava. En particular,

$$0 < x < 1 \quad \Longrightarrow \quad \ln x < 0,$$

$$\ln 1 = 0,$$

y

$$x > 1 \quad \Longrightarrow \quad \ln x > 0.$$

Demostración. La función ln es continua y diferenciable en $(0, \infty)$ por Teorema 17.1 y Teorema 6.1. Además, $(\ln x)' = \frac{1}{x} > 0$ para todo $x > 0$. Aplicando Teorema 10.1 al intervalo $I = (0, \infty)$ concluimos que ln es estrictamente creciente.

Como $\ln 1 = 0$, si $0 < x < 1$, el crecimiento estricto da $\ln x < \ln 1 = 0$, mientras que, si $x > 1$, $\ln x > \ln 1 = 0$. Para estudiar la curvatura derivamos una vez más. La regla para potencias enteras negativas de Corolario 7.2 aplicada a x^{-1} proporciona

$$(\ln x)'' = \frac{d}{dx}\left(x^{-1}\right) = -x^{-2} = -\frac{1}{x^2}.$$

Como

$$-\frac{1}{x^2} < 0 \qquad (x > 0),$$

el criterio de la segunda derivada de Teorema 10.4 muestra que ln es estrictamente cóncava en $(0, \infty)$. □

Las dos conclusiones geométricas son compatibles pero expresan información diferente:

- $(\ln x)' > 0$ dice que la gráfica siempre asciende;
- $(\ln x)'' < 0$ dice que sus pendientes disminuyen al avanzar hacia la derecha.

Así, el logaritmo crece en todo su dominio, pero lo hace cada vez con menor pendiente.

Creciente no significa todavía sobreyectiva
El crecimiento estricto prueba inmediatamente la **inyectividad** de ln: si $x_1 \neq x_2$, entonces $\ln x_1 \neq \ln x_2$. Pero una función puede ser continua y estrictamente creciente sin alcanzar todos los números reales. Para construir una inversa $\mathbb{R} \longrightarrow (0, \infty)$ todavía debemos demostrar que el rango de ln es exactamente $\mathbb{R}$.

Un valor positivo que podemos amplificar. La proposición anterior aplicada a $2 > 1$ da

$$\boxed{\ln 2 > 0.}$$

La ecuación funcional ya permite amplificar este único valor. Por Corolario 17.1, para cada entero $n \geq 1$,

$$\boxed{\ln(2^n) = n \ln 2.}$$

Como $\ln 2$ es un número real estrictamente positivo, los múltiplos $n \ln 2$ superan cualquier barrera real si n es suficientemente grande. Este paso es exactamente una aplicación de la propiedad arquimediana Teorema 1.4. En efecto, dada una barrera $M \in \mathbb{R}$, elegimos $n \in \mathbb{N}$ tal que $n > \frac{|M|}{\ln 2}$. La división es legítima porque $\ln 2 > 0$. Al multiplicar por ese número positivo obtenemos $n \ln 2 > |M| \geq M$.

Por tanto, $\ln(2^n) = n \ln 2 > M$.

Esto ya demuestra que ln **no está acotada superiormente**. La identidad del recíproco de Corolario 17.1 proporciona al mismo tiempo $\ln(2^{-n}) = -n \ln 2$. Como $n \ln 2$ puede ser arbitrariamente grande, los valores $-n \ln 2$ pueden ser arbitrariamente negativos. Así, ln tampoco está acotada inferiormente. La no acotación garantiza valores arbitrariamente grandes en magnitud. Para describir el límite completo debemos asegurar que toda una región suficientemente próxima al extremo del dominio satisface la barrera exigida.

Cuando x se hace arbitrariamente grande.

Teorema 17.2. **Teorema (comportamiento extremo y rango completo del logaritmo).** El logaritmo natural satisface

$$\boxed{\lim_{x \to +\infty} \ln x = +\infty}$$

y

$$\boxed{\lim_{x \to 0^+} \ln x = -\infty.}$$

Además,

$$\boxed{\ln : (0,\infty) \longrightarrow \mathbb{R}}$$

es una biyección.

Demostración.

1. Límite cuando $x \to +\infty$. Sea $M \in \mathbb{R}$ arbitrario. Por Teorema 1.4 y la positividad de $\ln 2$, podemos elegir $n \in \mathbb{N}$ tal que $n \ln 2 > |M| \geq M$.

Definimos $R := 2^n > 0$.

Si $x > R$, entonces, por el crecimiento estricto de Proposición 17.2, $\ln x > \ln R$. Por Corolario 17.1, $\ln R = \ln(2^n) = n \ln 2$.

En consecuencia,

$$x > R \quad \Longrightarrow \quad \ln x > n \ln 2 > M.$$

Como la barrera M fue arbitraria, ésta es exactamente la condición de Definición 4.5 para concluir

$$\lim_{x \to +\infty} \ln x = +\infty.$$

2. Límite cuando $x \to 0^+$. Sea nuevamente $M \in \mathbb{R}$ arbitrario. Elegimos $n \in \mathbb{N}$ tal que $n \ln 2 > |M|$.

Definimos $\delta := 2^{-n} > 0$.

Si $0 < x < \delta$, el crecimiento estricto da $\ln x < \ln \delta$. Por Corolario 17.1, $\ln \delta = \ln(2^{-n}) = -n \ln 2$.

Y, como $n \ln 2 > |M|$, $-n \ln 2 < -|M| \leq M$.

Por tanto,

$$0 < x < \delta \quad \Longrightarrow \quad \ln x < M.$$

Aplicando Definición 4.4 al dominio $A = (0,\infty)$ y al punto de acumulación 0, obtenemos

$$\lim_{x \to 0^+} \ln x = -\infty.$$

La notación 0^+ solo hace explícito algo que el dominio ya impone: todos los puntos admisibles próximos a 0 son positivos.

3. Inyectividad. Ya sabemos por Proposición 17.2 que ln es estrictamente creciente. Si $0 < x_1 < x_2$, entonces $\ln x_1 < \ln x_2$.

Por consiguiente, entradas distintas tienen imágenes distintas y ln es inyectiva.

4. Sobreyectividad. Sea $y \in \mathbb{R}$ arbitrario. Como $\ln 2 > 0$, la propiedad arquimediana Teorema 1.4 permite elegir $n \in \mathbb{N}$ tal que $n > \frac{|y|}{\ln 2}$.

Entonces $n \ln 2 > |y|$.

Consideremos los números positivos

$$a := 2^{-n}, \qquad b := 2^n.$$

Por Corolario 17.1,

$$\ln a = -n \ln 2, \qquad \ln b = n \ln 2.$$

La elección de n da $-n\ln 2 < -|y| \le y$ y $y \le |y| < n\ln 2$.

Luego $\ln a < y < \ln b$.

La función ln es continua en el intervalo compacto $[a, b]$, pues es continua en todo $(0, \infty)$. El teorema del valor intermedio Teorema 5.2 garantiza entonces la existencia de algún $c \in (a, b)$ tal que $\ln c = y$. Como $y \in \mathbb{R}$ fue arbitrario, todo número real pertenece a la imagen de ln. Por tanto ln es sobreyectiva sobre $\mathbb{R}$. Uniendo inyectividad y sobreyectividad concluimos que

$$\boxed{\ln : (0, \infty) \longrightarrow \mathbb{R} \text{ es biyectiva.}}$$

□

Por qué no bastaba con escribir «no acotada». Conviene distinguir tres afirmaciones que ahora sabemos verdaderas, pero que no son idénticas:

$$\ln \text{ no está acotada superiormente,}$$

$$\lim_{x \to +\infty} \ln x = +\infty,$$

y $\mathrm{Im}(\ln) = \mathbb{R}$.

La primera solo dice que aparecen valores arbitrariamente grandes en algún lugar del dominio. La segunda es más fuerte: una vez que x supera cierto umbral, **todos** los valores posteriores quedan por encima de la barrera prescrita. El crecimiento estricto es lo que permite pasar de la primera información a esta forma eventual. La tercera afirmación tampoco se sigue de la no acotación por sí sola. Para asegurar que no quedan huecos entre valores muy negativos y muy positivos usamos la **continuidad** y Teorema 5.2. La cadena lógica completa es, por tanto,

$$\boxed{\begin{array}{c} (\ln x)' > 0 \Longrightarrow \text{crecimiento estricto} \Longrightarrow \text{inyectividad}, \\ \ln(2^n) = n\ln 2 + \text{ propiedad arquimediana} \Longrightarrow \text{escape hacia } \pm\infty, \\ \text{continuidad} + \text{ valores a ambos lados de } y \Longrightarrow \text{valor intermedio} \Longrightarrow \text{sobreyectividad}. \end{array}}$$

Un intervalo que encierra cualquier nivel prescrito.

Ejemplo 17.2. **Encierro multiplicativo de niveles.** Sea $n \in \mathbb{N}_{>0}$. En el intervalo $[2^{-n}, 2^n]$ el logaritmo toma los valores extremos

$$\ln(2^{-n}) = -n\ln 2, \qquad \ln(2^n) = n\ln 2.$$

Como ln es continua y estrictamente creciente, su restricción $\ln : [2^{-n}, 2^n] \longrightarrow [-n\ln 2, n\ln 2]$ es biyectiva. En particular, para todo y que satisfaga $|y| \le n\ln 2$ existe un único $x \in [2^{-n}, 2^n]$ tal que $\ln x = y$. El resultado no introduce todavía un nombre para ese x como función de y; solo certifica existencia y unicidad dentro de un intervalo explícito.

Figura 17.3 condensa la cadena de Proposición 17.2, Teorema 17.2 y Ejemplo 17.2: monotonía para la inyectividad, crecimiento arquimediano de $n\ln 2$ y continuidad para llenar todos los niveles intermedios.

Este ejemplo hace visible la función de la propiedad arquimediana en la prueba global. Al aumentar n, los intervalos de salida $[-n\ln 2, n\ln 2]$ terminan conteniendo cualquier real prescrito, mientras los intervalos de entrada $[2^{-n}, 2^n]$ se extienden simultáneamente hacia 0 y hacia valores arbitrariamente grandes.

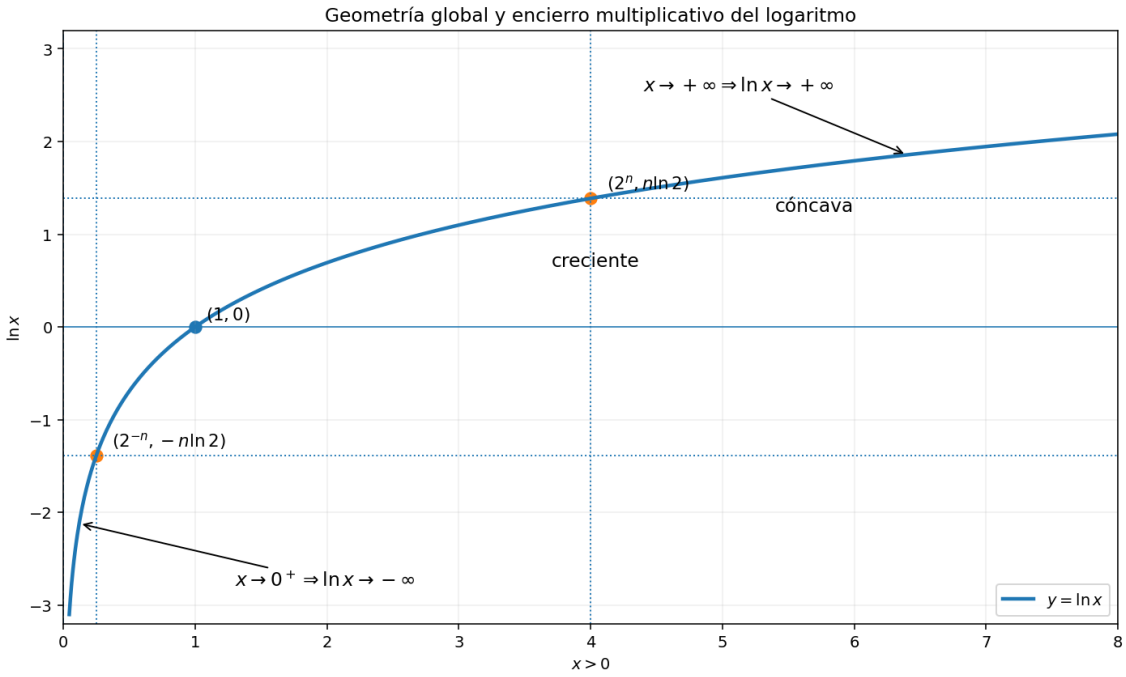


FIGURA 17.3. La gráfica de $\ln x$ reúne su geometría global: dominio $x > 0$, crecimiento estricto, concavidad, escape hacia $-\infty$ al aproximarse a 0^+ y hacia $+\infty$ cuando $x \to +\infty$; los puntos 2^{-n} y 2^n producen el encierro de niveles usado en la prueba del rango completo.

La puerta para la función inversa está abierta, pero aún no la cruzamos. Al comienzo de §17.2 escribíamos $\ln : (0, \infty) \to \mathbb{R}$ sin saber si $\mathbb{R}$ era realmente la imagen. Teorema 17.2 ha cerrado esa laguna. Ahora sí sabemos que la correspondencia $x \longmapsto \ln x$ asocia a cada positivo un real distinto y alcanza todos los reales exactamente una vez. Eso **justifica** que pueda construirse una función inversa global. Pero todavía no la hemos definido, no le hemos asignado la notación exp, no hemos introducido el número e y no hemos demostrado ninguna ley exponencial. Podemos invertir la biyección establecida. La teoría de funciones inversas de C09 proporcionará las primeras propiedades de la nueva función a partir de las del logaritmo.

Antes de seguir. ¿Por qué $(\ln x)' > 0$ basta para obtener inyectividad pero no sobreyectividad? ¿Dónde se usa que $\ln 2 > 0$? ¿Qué aporta exactamente la propiedad arquimediana a la prueba de los límites? ¿Por qué la no acotación superior no es, por sí sola, la definición de límite $+\infty$? ¿En qué paso interviene la continuidad para probar que todo $y \in \mathbb{R}$ es alcanzado? ¿Por qué todavía no hemos definido la exponencial, aunque la existencia de una inversa ya esté garantizada?

17.5. La exponencial como función inversa

La sección anterior cerró la condición que todavía faltaba para invertir globalmente al logaritmo. Ya no sabemos solo que $\ln : (0, \infty) \longrightarrow \mathbb{R}$ es continua y estrictamente creciente: Teorema 17.2 demuestra que es una **biyección**. Por tanto, para cada número real x existe un único número positivo y tal que $\ln y = x$. Éste es el momento lógico exacto en que podemos convertir esa existencia y unicidad en una nueva función.

Definir la inversa, no adivinar una fórmula.

Definición 17.2. **Definición (función exponencial natural y número e).** Definimos la **función exponencial natural** como la inversa global del logaritmo natural:

$$\boxed{\exp := \ln^{-1} : \mathbb{R} \longrightarrow (0,\infty).}$$

Equivalentemente, para cada $x \in \mathbb{R}$, $\exp(x)$ es el único número positivo que satisface

$$\boxed{\ln(\exp x) = x.}$$

Definimos además

$$\boxed{e := \exp(1).}$$

La definición contiene dos decisiones y conviene no confundirlas. Primero, exp no se introduce mediante una serie, una ecuación diferencial ni una expresión e^x previamente conocida. Se introduce porque Teorema 17.2 ya ha demostrado que ln posee una inversa global. Segundo, el símbolo e acaba de nombrar un número real concreto: el único positivo cuyo logaritmo natural vale 1. Todavía **no** hemos definido e^x para un exponente real arbitrario. Esa notación se justificará en §17.6, después de demostrar la ley funcional de la exponencial.

La notación no puede adelantarse a la construcción
En este punto están disponibles $\exp(x)$ y el número $e = \exp(1)$, pero la igualdad $\exp(x) = e^x$ todavía no es un hecho de la teoría. Usarla ahora convertiría en premisa precisamente una identificación que debe demostrarse después.

Las dos identidades inversas. Como exp y ln son funciones inversas, se obtienen inmediatamente las dos composiciones fundamentales:

$$\boxed{\ln(\exp x) = x \qquad (x \in \mathbb{R}),}$$

y

$$\boxed{\exp(\ln y) = y \qquad (y > 0).}$$

La primera identidad comienza con una entrada real, la envía mediante exp a un número positivo y regresa mediante ln. La segunda comienza con una entrada positiva, la envía mediante ln a un real y regresa mediante exp. La diferencia de dominios es esencial:

$$\exp : \mathbb{R} \to (0,\infty), \qquad \ln : (0,\infty) \to \mathbb{R}.$$

No escribiremos ninguna de las dos composiciones fuera del dominio en que está definida.

Continuidad y monotonía heredadas de la inversión.

Proposición 17.3. **Proposición (primeras propiedades de la exponencial).** La función

$$\exp : \mathbb{R} \longrightarrow (0,\infty)$$

es continua y estrictamente creciente. Además,

$$\boxed{\exp(0) = 1}, \qquad \boxed{\exp(1) = e}, \qquad \boxed{\ln e = 1},$$

y

$$\boxed{e > 1}.$$

En particular,

$$\exp(x) > 0 \qquad \text{para todo } x \in \mathbb{R}.$$

Demostración. La continuidad de exp no se supone. La función $\ln : (0,\infty) \to \mathbb{R}$ es continua e inyectiva sobre el intervalo $(0,\infty)$, y su imagen es todo $\mathbb{R}$ por Teorema 17.2. El teorema de continuidad de la inversa Teorema 5.7 se aplica, por tanto, con

$$I = (0,\infty), \qquad J = \mathbb{R},$$

y muestra que $\exp = \ln^{-1}$ es continua en todo $\mathbb{R}$. Probemos ahora el crecimiento estricto sin apelar a ninguna ley exponencial todavía inexistente. Sean $x < y$.

Pongamos

$$u := \exp x, \qquad v := \exp y.$$

Por definición, $u, v > 0$ y las identidades inversas dan

$$\ln u = x, \qquad \ln v = y.$$

Si ocurriera $u \geq v$, el crecimiento estricto de ln probado en Proposición 17.2 implicaría $\ln u \geq \ln v$, es decir, $x \geq y$, contradicción. Luego

$$x < y \quad \Longrightarrow \quad \exp x < \exp y.$$

Así, exp es estrictamente creciente.

Para evaluar en 0, observamos que $\ln 1 = 0$. Como exp es la inversa de ln, el único positivo cuyo logaritmo vale 0 es 1. Por tanto, $\exp(0) = 1$. La igualdad $\exp(1) = e$ es la propia definición de e, y aplicando ln a ella obtenemos $\ln e = \ln(\exp 1) = 1$.

Finalmente, $0 < 1$ y el crecimiento estricto de exp dan $\exp 0 < \exp 1$.

Por las identidades anteriores, $1 < e$.

La positividad de todos los valores de exp forma parte de su codominio $(0,\infty)$. □

Figura 17.4 representa la inversión global ya justificada por Teorema 17.2; no usa la ley exponencial como premisa.

La prueba deja visible el orden de dependencias:

$$\boxed{\ln \text{ biyectiva} \Longrightarrow \exp = \ln^{-1} \Longrightarrow \text{continuidad por la teoría de inversas} \Longrightarrow \text{monotonía por orden.}}$$

No hemos usado todavía $\exp(x+y) = \exp x \, \exp y$.

La derivada de la exponencial nace de la derivada de la inversa. La nueva función es una inversa y, por tanto, podemos aplicar el resultado específico construido en C09. Debemos hacerlo con los puntos de evaluación correctamente tipados.

Teorema 17.3. **Teorema (derivada de la función exponencial natural).** La función

$$\exp : \mathbb{R} \longrightarrow (0,\infty)$$

es diferenciable en todo $\mathbb{R}$ y satisface

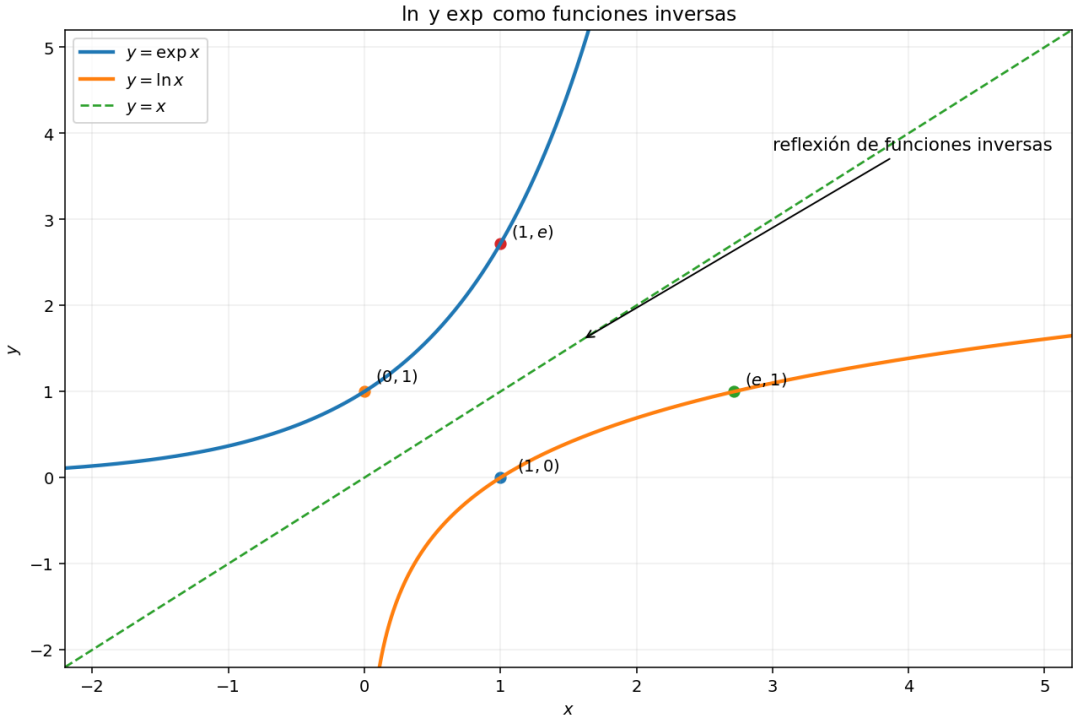


FIGURA 17.4. Las gráficas de $y = \ln x$ y $y = \exp x$ son reflejos respecto de $y = x$: sus dominios y rangos se intercambian, y los puntos $(1, 0)$, $(0, 1)$, $(e, 1)$ y $(1, e)$ aparecen por pares bajo la inversión.

$$\boxed{(\exp x)' = \exp x \qquad (x \in \mathbb{R}).}$$

Demostración. Fijemos arbitrariamente $x \in \mathbb{R}$ y definamos $a := \exp x$.

Por la definición de exp, $a > 0$ y $\ln a = x$.

Consideremos ahora la función original $f = \ln : (0, \infty) \longrightarrow \mathbb{R}$. El punto a pertenece al intervalo abierto $(0, \infty)$ y, por tanto, es un punto de acumulación de su dominio. Por Teorema 17.1, $f'(a) = (\ln)'(a) = \frac{1}{a}$.

Como $a > 0$, $\frac{1}{a} \neq 0$.

Todas las hipótesis de Teorema 8.1 están verificadas para la función $f = \ln$, su inversa $g = f^{-1} = \exp$, el punto a del dominio de f y el punto correspondiente

$$b = f(a) = \ln a = x.$$

Por ese teorema, $g'(b) = \frac{1}{f'(a)}$.

Sustituyendo los objetos concretos,

$$\begin{aligned}(\exp)'(x) &= \frac{1}{(\ln)'(a)}\\ &= \frac{1}{1/a}\\ &= a\\ &= \exp x.\end{aligned}$$

Como $x \in \mathbb{R}$ fue arbitrario, $(\exp x)' = \exp x$ en toda la recta real. □

La demostración evita una circularidad frecuente. Sería tentador comenzar por $\ln(\exp x) = x$ y diferenciar ambos lados con la regla de la cadena para obtener

$$\frac{\exp'(x)}{\exp x} = 1.$$

Pero ese cálculo **ya presupone la existencia de** $\exp'(x)$, que es precisamente lo que debemos demostrar. Teorema 8.1 fue construido en C09 para resolver esta dificultad: primero acredita la diferenciabilidad de la inversa y solo entonces entrega la fórmula de su derivada.

La ecuación diferencial aparece como teorema, no como definición
Ahora hemos probado $y' = y$ para la función $y = \exp x$. Eso no significa que hayamos definido la exponencial como solución de una ecuación diferencial. La dirección lógica de C18 es la contraria:

$$\ln \text{ construido por integral} \Longrightarrow \exp = \ln^{-1} \Longrightarrow \exp' = \exp.$$

La teoría sistemática de ecuaciones diferenciales y los problemas de valor inicial pertenecen a C21.

Consecuencias inmediatas que no requieren la ley exponencial. Del teorema anterior y de Proposición 17.3 obtenemos, por ejemplo, $\exp'(0) = \exp 0 = 1$ y $\exp'(1) = \exp 1 = e$. Estas igualdades describen pendientes de la nueva función sin haber utilizado aún ninguna regla de suma para sus argumentos.

Ejemplo 17.3. **Cuatro valores obtenidos solo por inversión.** Las identidades inversas permiten calcular exactamente algunos valores sin usar todavía la ley exponencial.

Como $\ln 1 = 0$, tenemos $\exp 0 = 1$.

Por definición, $\exp 1 = e$.

Como $\ln e = 1$, la regla del recíproco de Corolario 17.1 da

$$\ln\left(\frac{1}{e}\right) = -1.$$

La unicidad del antecedente bajo ln implica entonces

$$\boxed{\exp(-1) = \frac{1}{e}.}$$

Finalmente,

$$\boxed{\exp(\ln 2) = 2}$$

es simplemente una de las identidades inversas. Ninguna de estas igualdades usa la fórmula $\exp(x+y) = \exp x \exp y$. En particular, $\exp(-1) = 1/e$ se ha deducido desde la teoría ya demostrada del logaritmo y la unicidad de la inversa, no desde una ley exponencial adelantada.

Qué está disponible y qué sigue pendiente. Al terminar §17.5 disponemos exactamente de la cadena

$$\boxed{\begin{array}{c}\ln : (0,\infty) \to \mathbb{R} \text{ biyectiva} \\ \Downarrow \\ \exp := \ln^{-1} : \mathbb{R} \to (0,\infty) \\ \Downarrow \\ \exp \text{ continua y estrictamente creciente} \\ \Downarrow \\ (\exp x)' = \exp x.\end{array}}$$

También hemos definido $e = \exp 1 > 1$ y sabemos que $\ln e = 1$. Todavía no hemos demostrado la ley de suma de la exponencial, ni hemos identificado $\exp(x)$ con una potencia e^x para exponentes reales, ni hemos definido a^x para una base positiva arbitraria y exponente real. Para obtener la ley exponencial utilizaremos la ecuación funcional de ln y su inyectividad: $\exp(x+y) = \exp x \exp y$ sin series y sin ecuaciones diferenciales como fundamento. Solo después de esa ley podremos justificar rigurosamente la notación e^x.

Antes de seguir. ¿Por qué la definición de exp solo es legítima después de Teorema 17.2? ¿Por qué la continuidad de exp no necesita una nueva prueba ε–δ desde cero? ¿Cómo se demuestra su crecimiento estricto usando el orden y las identidades inversas? ¿En qué punto exacto de la prueba de Teorema 17.3 se verifica que la derivada de ln no es cero? ¿Por qué derivar directamente $\ln(\exp x) = x$ habría sido circular? ¿Qué significa exactamente definir $e = \exp 1$ antes de haber definido e^x para exponentes reales?

17.6. Ley exponencial y notación e^x

La sección anterior construyó la función $\exp : \mathbb{R} \longrightarrow (0,\infty)$ como inversa global del logaritmo natural y demostró que $(\exp x)' = \exp x$.

También definió el número $e := \exp 1$.

Todavía falta, sin embargo, la propiedad que convierte a esta función en una verdadera exponencial en sentido algebraico. Antes de escribir potencias con exponente real debemos demostrar que sumar argumentos corresponde a multiplicar valores.

La suma en la entrada debe convertirse en producto en la salida. La ecuación funcional de Proposición 17.1 afirma que, para positivos u, v, $\ln(uv) = \ln u + \ln v$. Como todos los valores de exp son positivos, podemos aplicarla a

$$u = \exp x, \qquad v = \exp y.$$

Esto permite transportar la estructura aditiva de $\mathbb{R}$ a la estructura multiplicativa de $(0,\infty)$ sin introducir ninguna potencia real de antemano.

Teorema 17.4. **Teorema (ley exponencial).** Para todos $x, y \in \mathbb{R}$,

$$\boxed{\exp(x+y) = \exp x \, \exp y.}$$

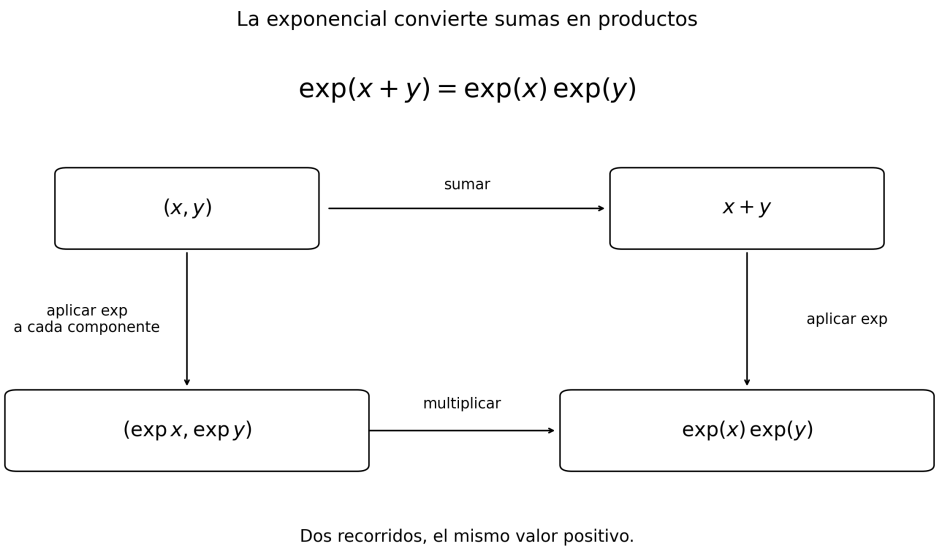


FIGURA 17.5. La exponencial transporta la suma de argumentos en $\mathbb{R}$ al producto de valores positivos: la ley $\exp(x+y) = \exp x \, \exp y$ es la contraparte multiplicativa de la ecuación funcional del logaritmo.

Demostración. Sean $x, y \in \mathbb{R}$. Como

$$\exp x > 0, \qquad \exp y > 0,$$

el producto $\exp x \, \exp y$ pertenece al dominio de ln. Por la ecuación funcional Proposición 17.1,

$$\begin{aligned} \ln(\exp x \, \exp y) &= \ln(\exp x) + \ln(\exp y) \\ &= x + y, \end{aligned}$$

donde en la última igualdad usamos la identidad inversa $\ln(\exp t) = t$.

Por otra parte, $\ln(\exp(x + y)) = x + y$.

Hemos obtenido $\ln(\exp x \, \exp y) = \ln(\exp(x + y))$.

La función ln es inyectiva por Proposición 17.2 —equivalentemente, por la biyección de Teorema 17.2—. Por tanto sus argumentos son iguales: $\exp x \, \exp y = \exp(x + y)$. Esto demuestra la ley para todos $x, y \in \mathbb{R}$. □

Figura 17.5 separa las dos operaciones: arriba se suman argumentos; abajo se multiplican imágenes bajo exp.

La prueba merece atención porque no utiliza ni la derivada de exp ni una teoría de ecuaciones diferenciales. La ley aparece por inversión directa de la ecuación funcional del logaritmo:

$$\boxed{\ln(uv) = \ln u + \ln v \quad \Longrightarrow \quad \exp(x + y) = \exp x \, \exp y.}$$

La ley exponencial se deduce de la construcción de ln, de su ley del producto y de la existencia de su inversa global. Esas dependencias permiten justificar la nueva identidad sin presuponerla.

Consecuencias antes de introducir la nueva notación. La ley anterior permite deducir inmediatamente las fórmulas de signo y diferencia. Tomando $y = 0$, $\exp(x + 0) = \exp x \exp 0$, lo cual es compatible con $\exp 0 = 1$ ya demostrado.

Tomando $y = -x$, $\exp 0 = \exp x \exp(-x)$.

Como $\exp 0 = 1$ y $\exp x > 0$, podemos dividir y obtener

$$\boxed{\exp(-x) = \frac{1}{\exp x}.}$$

En consecuencia, para $x, y \in \mathbb{R}$,

$$\begin{aligned} \exp(x - y) &= \exp(x + (-y)) \\ &= \exp x \, \exp(-y) \\ &= \frac{\exp x}{\exp y}. \end{aligned}$$

Así,

$$\boxed{\exp(x - y) = \frac{\exp x}{\exp y}.}$$

Estas fórmulas son teoremas sobre la función exp. Todavía no hemos necesitado escribir ninguna potencia e^x.

Ahora sí: exponentes reales para la base e. Hasta este punto, e era simplemente el número positivo definido por $e = \exp 1$. Para exponentes enteros o racionales, la teoría anterior del libro ya sabía interpretar potencias de una base positiva como e. Lo que aún no existía era un significado canónico para e^x cuando x es un real arbitrario. La función exp proporciona precisamente esa extensión.

Definición 17.3. **Definición (potencia real de base e).** Para todo $x \in \mathbb{R}$, definimos

$$\boxed{e^x := \exp x.}$$

Esta definición amplía la notación exponencial a **todos** los exponentes reales para la base específica e.

La igualdad $\exp x = e^x$ es, a partir de ahora, una igualdad por definición. Pero todavía debemos comprobar que esta nueva convención no entra en conflicto con las potencias racionales que ya existían antes de C18.

Extender una notación exige comprobar compatibilidad
Para $q \in \mathbb{Q}$, el símbolo e^q ya tenía un significado algebraico en C09, porque $e > 0$ y allí se construyeron las potencias racionales de bases positivas. La nueva definición sería inaceptable si asignara a ese mismo símbolo un valor distinto. Debemos demostrar que ambas construcciones coinciden.

Compatibilidad con las potencias racionales anteriores.

COROLARIO 17.2. **Corolario (compatibilidad racional y leyes básicas de e^x).** Con la definición de Definición 17.3 se cumplen las siguientes afirmaciones.

1. Para todos $x, y \in \mathbb{R}$,
$$\boxed{e^{x+y} = e^x e^y.}$$
2. Para todo $x \in \mathbb{R}$,
$$\boxed{e^{-x} = \frac{1}{e^x}}$$
y, para todos $x, y \in \mathbb{R}$,
$$\boxed{e^{x-y} = \frac{e^x}{e^y}.}$$
3. Si $q \in \mathbb{Q}$, el valor de e^q dado por Definición 17.3 coincide con la potencia racional de la base positiva e construida en C09.
4. Más generalmente, si $x \in \mathbb{R}$ y $q \in \mathbb{Q}$, entonces
$$\boxed{(e^x)^q = e^{qx},}$$
donde el miembro izquierdo usa la potencia racional de la base positiva e^x.

Demostración. Los puntos 1 y 2 son traducciones directas de Teorema 17.4 y de las fórmulas precedentes mediante la definición $e^t = \exp t$.

Probemos el punto 3 con cuidado. Sea $q \in \mathbb{Q}$. Para evitar ambigüedad durante la demostración, llamemos p_q al valor de la **potencia racional anterior** de e construida en C09. Como $e > 0$, tenemos $p_q > 0$. La compatibilidad logarítmica con potencias racionales, Corolario 17.1, da $\ln p_q = q \ln e$. Por Proposición 17.3, $\ln e = 1$, y por tanto $\ln p_q = q$.

Por otra parte, usando la nueva definición y la identidad inversa, $\ln(e^q) = \ln(\exp q) = q$.

Así, $\ln p_q = \ln(e^q)$.

La inyectividad de ln implica $p_q = e^q$.

Por consiguiente, la nueva definición prolonga exactamente la potencia racional previa; no la reemplaza por una convención incompatible. Para el punto 4, fijemos $x \in \mathbb{R}$ y $q \in \mathbb{Q}$. Como $e^x > 0$, la potencia racional $(e^x)^q$ está definida por la teoría de C09. Aplicando nuevamente Corolario 17.1, $\ln\big((e^x)^q\big) = q \ln(e^x)$.

Pero $e^x = \exp x$, de modo que $\ln(e^x) = x$.

Por tanto, $\ln\big((e^x)^q\big) = qx$.

A la vez, $\ln(e^{qx}) = \ln(\exp(qx)) = qx$.

La inyectividad de ln vuelve a dar $(e^x)^q = e^{qx}$. Quedan demostradas las cuatro afirmaciones. □

El tercer punto es el puente que necesitamos para evitar una ruptura de significado. Para exponentes racionales, la potencia de C09 y la nueva notación de C18 son **el mismo número**. Para exponentes irracionales, en cambio, C09 no había asignado ningún valor: allí la expresión e^x no estaba definida. Definición 17.3 llena precisamente ese hueco usando la función exp.

Qué significa realmente un exponente irracional en este punto. Sea, por ejemplo, $\alpha \in \mathbb{R} \setminus \mathbb{Q}$.

Ahora la expresión e^α tiene un significado exacto: $e^\alpha := \exp \alpha$. No la hemos definido como un límite de potencias racionales aproximantes, ni mediante una serie infinita, ni mediante una ecuación diferencial. Es el único número positivo cuyo logaritmo vale α:

$$\boxed{\ln(e^\alpha) = \alpha.}$$

La continuidad de exp demostrada en §17.5 garantizará más adelante que esta definición también es compatible con aproximaciones por exponentes racionales, si decidimos estudiar ese punto de vista. Pero esa caracterización por límites no es necesaria para la existencia ni para la definición canónica actual.

Las identidades inversas adquieren una nueva forma. La definición de Definición 17.3 permite reescribir las dos identidades inversas como

$$\boxed{\ln(e^x) = x \qquad (x \in \mathbb{R}),}$$

y

$$\boxed{e^{\ln y} = y \qquad (y > 0).}$$

La segunda igualdad no define todavía potencias de una base y arbitraria. La base sigue siendo e y el exponente es el real $\ln y$. De la derivada demostrada en §17.5 obtenemos además, ahora en la nueva notación,

$$\boxed{\frac{d}{dx} e^x = e^x.}$$

Aquí tampoco aparece una nueva demostración: es exactamente Teorema 17.3 traducido mediante Definición 17.3.

Ejemplo 17.4. **Compatibilidad visible entre las dos teorías.** La nueva notación permite reunir resultados que antes procedían de construcciones distintas.

Como $1/2 \in \mathbb{Q}$, el punto 3 de Corolario 17.2 asegura que $e^{1/2}$ es exactamente la raíz cuadrada positiva de e construida mediante C09. Por tanto,

$$\boxed{e^{1/2} = \sqrt{e}.}$$

Análogamente, $e^{-1} = \frac{1}{e}$, en acuerdo con el cálculo de Ejemplo 17.3.

Por otra parte, para cualquier $y > 0$, $e^{\ln y} = \exp(\ln y) = y$.

En particular,

$$e^{\ln 2} = 2, \qquad e^{\ln(27/8)} = \frac{27}{8}.$$

El último valor no requiere convertir $\ln(27/8)$ en un número decimal ni decidir si es racional: la identidad inversa vale para todo exponente real.

Lo que esta sección no ha definido. La notación e^x ya tiene ahora un significado para todo $x \in \mathbb{R}$. Pero de aquí no se sigue automáticamente que, para una base arbitraria $a > 0$, la expresión a^x esté definida para todo real x. La extensión de las potencias reales desde la base especial e hacia una base positiva cualquiera requiere una construcción adicional. La próxima sección hará exactamente eso. Usará $\ln a$ como factor de escala y definirá $a^x := \exp(x \ln a) = e^{x \ln a}$, para después demostrar sus leyes, distinguir los casos $a > 1$, $a = 1$ y $0 < a < 1$, y construir los logaritmos de base arbitraria sin circularidad.

Antes de seguir. ¿Por qué la prueba de Teorema 17.4 aplica primero ln a un producto de valores de exp? ¿Dónde se usa la inyectividad de ln? ¿Por qué era importante demostrar la ley en notación exp antes de definir e^x? ¿Cómo comprobamos que la nueva definición de e^q coincide con la potencia racional anterior? ¿Qué cambia conceptualmente cuando el exponente es irracional? ¿Por qué todavía no estamos autorizados a escribir a^x para una base positiva arbitraria?

17.7. Potencias reales y logaritmos de base arbitraria

La sección anterior extendió la notación exponencial a todos los exponentes reales cuando la base es e: $e^x := \exp x$. El siguiente paso consiste en abandonar esa base privilegiada sin perder la estructura ya construida. Si $a > 0$, entonces $\ln a \in \mathbb{R}$, y podemos usar ese número para reescalar el exponente antes de aplicar exp.

La idea central será

$$\boxed{a^x = \exp(x \ln a).}$$

Esta fórmula no se toma como una identidad conocida de antemano. Será la **definición** de la potencia real para bases positivas, y después tendremos que verificar que prolonga exactamente las potencias enteras y racionales construidas en C09.

Definición de la potencia real de base positiva.

Definición 17.4. **Definición (potencia real de base positiva).** Sea

$$a > 0.$$

Para todo exponente real $x \in \mathbb{R}$, definimos

$$\boxed{a^x := \exp(x \ln a) = e^{x \ln a}.}$$

Así, para cada base positiva fija a, obtenemos una función

$$x \longmapsto a^x$$

definida en toda $\mathbb{R}$ y con valores en $(0, \infty)$.

La restricción $a > 0$ es estructural. La definición utiliza $\ln a$, y el logaritmo natural está definido exactamente en $(0, \infty)$. No estamos afirmando que las potencias racionales de bases negativas hayan dejado de existir en los casos permitidos por C09. Simplemente no las prolongaremos aquí a exponentes reales arbitrarios: una teoría general de potencias reales con base negativa no surge de la fórmula anterior dentro de los reales.

Las leyes de exponentes deben volver a demostrarse. Ahora que a^x está definido para todo exponente real, las leyes familiares ya no pueden darse por memoria escolar. Deben deducirse de la definición y de la ley exponencial de Teorema 17.4.

PROPOSICIÓN 17.4. **Proposición (estructura de las potencias reales de base positiva).** Sean

$$a, b > 0, \qquad x, y \in \mathbb{R}.$$

Entonces:

1. normalización:
$$\boxed{a^0 = 1}, \qquad \boxed{a^1 = a};$$
2. suma de exponentes:
$$\boxed{a^{x+y} = a^x a^y};$$
3. signo y diferencia:
$$\boxed{a^{-x} = \frac{1}{a^x}, \qquad a^{x-y} = \frac{a^x}{a^y}};$$
4. producto de bases:
$$\boxed{(ab)^x = a^x b^x};$$
5. potencia de una potencia:
$$\boxed{(a^x)^y = a^{xy}};$$
6. logaritmo de una potencia real:
$$\boxed{\ln(a^x) = x \ln a.}$$

Además, la nueva definición es compatible con la teoría anterior: si $q \in \mathbb{Q}$, el valor de a^q dado por Definición 17.4 coincide exactamente con la potencia racional de la base positiva a construida en C09. Para cada base fija $a > 0$, la función $x \mapsto a^x$ es continua. Su monotonía queda completamente determinada por la base:

- si $a > 1$, es estrictamente creciente;
- si $a = 1$, es constante e igual a 1;
- si $0 < a < 1$, es estrictamente decreciente.

Si $a \neq 1$, la función

$$x \longmapsto a^x$$

es una biyección de $\mathbb{R}$ sobre $(0, \infty)$.

Demostración.
Por Definición 17.4,

$$a^0 = \exp(0 \ln a) = \exp 0 = 1.$$

Asimismo, $a^1 = \exp(\ln a) = a$ por la identidad inversa de exp y ln. Para la suma de exponentes,

$$\begin{aligned} a^{x+y} &= \exp((x+y)\ln a) \\ &= \exp(x\ln a + y\ln a) \\ &= \exp(x\ln a)\exp(y\ln a) \\ &= a^x a^y, \end{aligned}$$

donde usamos Teorema 17.4.

Tomando $y = -x$ y usando $a^0 = 1$ obtenemos $1 = a^x a^{-x}$, y como $a^x > 0$, $a^{-x} = \frac{1}{a^x}$.

Luego

$$a^{x-y} = a^{x+(-y)} = a^x a^{-y} = \frac{a^x}{a^y}.$$

Para el producto de bases usamos primero la ecuación funcional del logaritmo Proposición 17.1: $\ln(ab) = \ln a + \ln b$.

Por tanto,

$$\begin{aligned} (ab)^x &= \exp(x\ln(ab)) \\ &= \exp(x\ln a + x\ln b) \\ &= \exp(x\ln a)\exp(x\ln b) \\ &= a^x b^x. \end{aligned}$$

La identidad $\ln(a^x) = x\ln a$ sale directamente de la definición, pues $a^x = \exp(x\ln a)$ y $\ln(\exp t) = t$ para todo $t \in \mathbb{R}$. Con esa identidad podemos demostrar la ley de potencia de una potencia:

$$\begin{aligned} (a^x)^y &= \exp(y\ln(a^x)) \\ &= \exp(yx\ln a) \\ &= a^{xy}. \end{aligned}$$

Queda comprobar la compatibilidad racional. Sea $q \in \mathbb{Q}$ y llamemos p_q al valor de la potencia racional antigua de a construida en C09. Como $a > 0$, también $p_q > 0$. Por Corolario 17.1, $\ln p_q = q\ln a$.

Por otra parte, la nueva definición da $\ln(a^q) = q\ln a$.

Así, $\ln p_q = \ln(a^q)$.

La inyectividad de ln implica $p_q = a^q$.

Por tanto la nueva teoría prolonga exactamente la anterior.

Consideremos ahora la regularidad. Para $a > 0$ fijo, la función $x \longmapsto x\ln a$ es lineal y continua. La función exp es continua por Proposición 17.3. Por las reglas de continuidad y composición ya establecidas en Proposición 5.2, $x \longmapsto \exp(x\ln a) = a^x$ es continua en toda $\mathbb{R}$. Para la monotonía distinguimos los tres casos. Si $a > 1$, Proposición 17.2 da $\ln a > 0$.

Entonces

$$x < y \quad \Longrightarrow \quad x\ln a < y\ln a.$$

Como exp es estrictamente creciente,

$$a^x = \exp(x\ln a) < \exp(y\ln a) = a^y.$$

Luego $x \mapsto a^x$ es estrictamente creciente.

Si $a = 1$, $\ln 1 = 0$, y para todo $x \in \mathbb{R}$,

$$1^x = \exp(x \cdot 0) = \exp 0 = 1.$$

Si $0 < a < 1$, Proposición 17.2 da $\ln a < 0$.

Ahora

$$x < y \quad \Longrightarrow \quad x \ln a > y \ln a,$$

y el crecimiento de exp produce $a^x > a^y$.

Así, la función es estrictamente decreciente.

Finalmente supongamos $a \neq 1$.

Entonces $\ln a \neq 0$.

La monotonía estricta ya demuestra la inyectividad. Para probar la sobreyectividad sobre $(0, \infty)$, sea $u > 0$. Definimos $x := \frac{\ln u}{\ln a}$. Este número es real porque $\ln a \neq 0$. Entonces

$$\begin{aligned} a^x &= \exp\left(\frac{\ln u}{\ln a} \ln a\right) \\ &= \exp(\ln u) \\ &= u. \end{aligned}$$

Todo positivo es, por tanto, imagen de algún real. Luego, si $a \neq 1$, $x \longmapsto a^x$ es biyectiva de $\mathbb{R}$ sobre $(0, \infty)$. □

La proposición ha resuelto dos problemas a la vez. Primero, ha extendido las potencias a exponentes reales sin romper la teoría racional previa. Segundo, ha clasificado completamente la geometría de las funciones exponenciales reales de base positiva:

$a > 1$	a^x estrictamente creciente
$a = 1$	$a^x \equiv 1$
$0 < a < 1$	a^x estrictamente decreciente.

La exclusión de $a = 1$ será precisamente la condición necesaria para construir un logaritmo de base a como función inversa.

Definición de los logaritmos de base arbitraria. Sea ahora

$$a > 0, \qquad a \neq 1.$$

Como acabamos de demostrar que $x \longmapsto a^x$ es una biyección de $\mathbb{R}$ sobre $(0, \infty)$, posee una inversa global. La fórmula de esa inversa puede leerse directamente desde la construcción de a^x.

Definición 17.5. **Definición (logaritmo de base a).** Sea

$$a > 0, \qquad a \neq 1.$$

Para $x > 0$, definimos

$$\boxed{\log_a x := \frac{\ln x}{\ln a}.}$$

Así obtenemos una función

$$\log_a : (0,\infty) \longrightarrow \mathbb{R}.$$

La condición $a \neq 1$ es indispensable porque $\ln 1 = 0$, y la fórmula exigiría dividir por cero. Esta exclusión coincide exactamente con el fallo estructural de la función $x \longmapsto 1^x$, que es constante y por tanto no puede tener una inversa global.

TEOREMA 17.5. **Teorema (potencia real y logaritmo de base a como funciones inversas).** Sea

$$a > 0, \qquad a \neq 1.$$

Entonces las funciones

$$E_a : \mathbb{R} \longrightarrow (0,\infty), \qquad E_a(x) = a^x,$$

y

$$\log_a : (0,\infty) \longrightarrow \mathbb{R}$$

son inversas entre sí. En particular,

$$\boxed{\log_a(a^x) = x \qquad (x \in \mathbb{R}),}$$

y

$$\boxed{a^{\log_a y} = y \qquad (y > 0).}$$

Además, $\log_a$ es continua y:

- si $a > 1$, es estrictamente creciente;
- si $0 < a < 1$, es estrictamente decreciente.

Demostración. Sea $x \in \mathbb{R}$. Por Proposición 17.4, $\ln(a^x) = x \ln a$.
Como $\ln a \neq 0$,

$$\log_a(a^x) = \frac{\ln(a^x)}{\ln a} = \frac{x \ln a}{\ln a} = x.$$

Ahora sea $y > 0$. Entonces

$$\begin{aligned} a^{\log_a y} &= \exp\left(\log_a y \, \ln a\right) \\ &= \exp\left(\frac{\ln y}{\ln a} \ln a\right) \\ &= \exp(\ln y) \\ &= y. \end{aligned}$$

Las dos composiciones son identidades en sus dominios respectivos. Por tanto E_a y $\log_a$ son funciones inversas. La continuidad de $\log_a$ sigue de la continuidad de ln y de la multiplicación por la constante no nula $\frac{1}{\ln a}$. Si $a > 1$, entonces $\ln a > 0$. Para $0 < u < v$ el crecimiento estricto de ln da $\ln u < \ln v$, y dividir por el número positivo $\ln a$ conserva el orden: $\log_a u < \log_a v$. Si $0 < a < 1$, entonces $\ln a < 0$, y la misma división invierte el sentido: $\log_a u > \log_a v$.

Queda demostrada la clasificación. □

La inversión preserva, por tanto, la orientación cuando la base supera a 1 y la orientación decreciente cuando la base está entre 0 y 1:

	$x \mapsto a^x$	$x \mapsto \log_a x$
$a > 1$	creciente	creciente
$0 < a < 1$	decreciente	decreciente.

Las leyes logarítmicas en cualquier base. Las reglas de $\log_a$ ya no requieren pruebas independientes desde cero. Se transportan desde las leyes de ln mediante el factor constante $1/\ln a$.

Corolario 17.3. **Corolario (leyes de los logaritmos de base arbitraria y cambio de base).** Sean

$$a > 0, \qquad a \neq 1,$$

y $x, y > 0$. Entonces:

$$\boxed{\log_a(xy) = \log_a x + \log_a y,}$$

$$\boxed{\log_a\left(\frac{x}{y}\right) = \log_a x - \log_a y,}$$

y, para todo $r \in \mathbb{R}$,

$$\boxed{\log_a(x^r) = r \log_a x.}$$

Además,

$$\boxed{\log_a 1 = 0, \qquad \log_a a = 1.}$$

Si también

$$b > 0, \qquad b \neq 1,$$

entonces

$$\boxed{\log_a x = \frac{\log_b x}{\log_b a}.}$$

Demostración. Por Definición 17.5 y Proposición 17.1,

$$\begin{aligned} \log_a(xy) &= \frac{\ln(xy)}{\ln a} \\ &= \frac{\ln x + \ln y}{\ln a} \\ &= \log_a x + \log_a y. \end{aligned}$$

La fórmula del cociente se obtiene del mismo modo usando Corolario 17.1. Para $r \in \mathbb{R}$, Proposición 17.4 da $\ln(x^r) = r \ln x$.

Por tanto,

$$\log_a(x^r) = \frac{\ln(x^r)}{\ln a} = r\frac{\ln x}{\ln a} = r\log_a x.$$

Las normalizaciones son inmediatas: $\log_a 1 = \frac{\ln 1}{\ln a} = 0$, y $\log_a a = \frac{\ln a}{\ln a} = 1$.

Finalmente, si $b > 0$ y $b \neq 1$, $\log_b x = \frac{\ln x}{\ln b}$ y $\log_b a = \frac{\ln a}{\ln b} \neq 0$.

Entonces

$$\frac{\log_b x}{\log_b a} = \frac{\ln x/\ln b}{\ln a/\ln b} = \frac{\ln x}{\ln a} = \log_a x.$$

□

Un ejemplo que separa base, exponente e inversa.

Ejemplo 17.5. **Dos bases con orientaciones opuestas.** Consideremos primero $a = 4 > 1$.

La compatibilidad racional de Proposición 17.4 garantiza que $4^{1/2} = 2$, exactamente con el significado previo de raíz positiva de C09. Como $4^{3/2} = 4 \cdot 4^{1/2} = 8$, el teorema de inversión da

$$\boxed{\log_4 8 = \frac{3}{2}.}$$

Ahora consideremos $b = \frac{1}{2}$, que satisface $0 < b < 1$.

La función

$$x \longmapsto \left(\frac{1}{2}\right)^x$$

es estrictamente decreciente. Como

$$\left(\frac{1}{2}\right)^{-3} = 8,$$

su inversa satisface

$$\boxed{\log_{1/2} 8 = -3.}$$

Los dos logaritmos toman el mismo argumento 8, pero el signo y el tamaño del resultado dependen de la base. No existe una regla de «logaritmo sin base» que permita ignorar esa información: ln es simplemente el caso particular de base e.

Figura 17.6 hace visible por qué $a = 1$ queda excluido de la construcción de $\log_a$: la función 1^x no es inyectiva y, por tanto, no admite inversa global.

Qué queda pendiente para el cálculo diferencial. La construcción algebraica y funcional ya está cerrada:

$$\boxed{a^x = \exp(x\ln a), \qquad \log_a x = \frac{\ln x}{\ln a}.}$$

Tenemos dominios, codominios, leyes algebraicas, compatibilidad con C09, continuidad, monotonía e inversión global para $a \neq 1$. Pero todavía no hemos convertido estas fórmulas en un repertorio sistemático de reglas de derivación. En particular, esta sección **no** demostrará todavía como resultado general $\frac{d}{dx}a^x = (\ln a)a^x$, ni $\frac{d}{dx}\log_a x = \frac{1}{x\ln a}$, ni la extensión de la regla

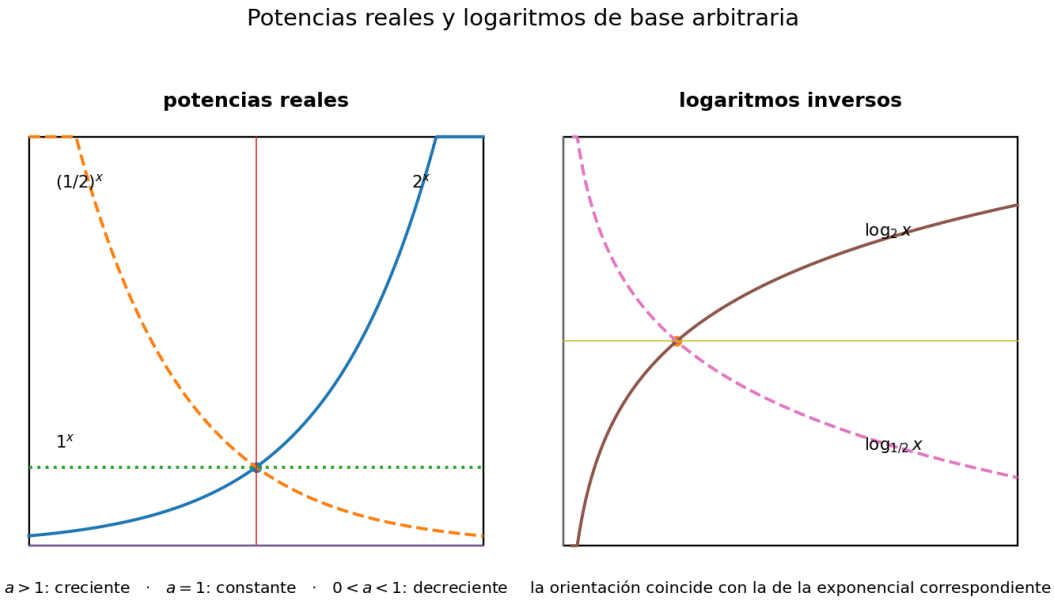


FIGURA 17.6. Para bases positivas, la orientación de $x \mapsto a^x$ y de su logaritmo inverso depende del signo de $\ln a$: bases mayores que 1 producen funciones crecientes, $a = 1$ produce la constante 1, y bases entre 0 y 1 producen funciones decrecientes.

$\frac{d}{dx}x^\alpha = \alpha x^{\alpha-1}$ a exponentes reales arbitrarios. Tampoco hemos tratado aún funciones con exponente variable como $u(x)^{v(x)}$. Todo ese bloque pertenece a §17.8, donde las funciones ya construidas podrán diferenciarse mediante las reglas de C08 sin ninguna circularidad.

Antes de seguir. ¿Por qué la definición $a^x = \exp(x \ln a)$ obliga a exigir $a > 0$? ¿Cómo se demuestra que no contradice la potencia racional de C09? ¿Por qué el signo de $\ln a$ determina la monotonía de a^x? ¿Qué falla exactamente cuando $a = 1$ y se intenta construir $\log_a$? ¿Por qué $\log_a(x^r) = r \log_a x$ ya puede formularse para todo $r \in \mathbb{R}$? ¿Qué fórmulas diferenciales hemos dejado deliberadamente para la sección siguiente?

17.8. Derivadas de logaritmos, exponenciales y potencias

Las secciones anteriores hicieron primero el trabajo de **construcción**. Ahora están disponibles, con dominios y leyes ya demostrados,

$$\ln x, \qquad \exp x = e^x, \qquad a^x \ (a > 0), \qquad \log_a x \ (a > 0,\ a \neq 1).$$

Solo después de haber definido esos objetos tiene sentido organizar sus reglas de derivación. Este orden evita una circularidad frecuente: no utilizaremos fórmulas diferenciales para justificar la existencia de las funciones que aparecen en ellas. La herramienta principal será la regla de la cadena Teorema 7.4, junto con la linealidad Teorema 7.1, la regla del producto Teorema 7.2 y las derivadas ya demostradas de ln y exp.

Las fórmulas básicas.

TEOREMA 17.6. **Teorema (derivadas de exponenciales, logaritmos y potencias reales).** Se cumplen las siguientes fórmulas.

1. Si $a > 0$, entonces
$$\boxed{\frac{d}{dx}a^x = (\ln a)a^x \qquad (x \in \mathbb{R}).}$$
2. Si $a > 0$ y $a \neq 1$, entonces
$$\boxed{\frac{d}{dx}\log_a x = \frac{1}{x \ln a} \qquad (x > 0).}$$
3. Si $\alpha \in \mathbb{R}$, entonces
$$\boxed{\frac{d}{dx}x^\alpha = \alpha x^{\alpha-1} \qquad (x > 0).}$$

Demostración.

1. Derivada de a^x. Por Definición 17.4, $a^x = \exp(x \ln a)$, donde $\ln a$ es una constante real. La función interior $u(x) = x \ln a$ es diferenciable y, por Teorema 7.1, $u'(x) = \ln a$.

La derivada de la función exterior exp es Teorema 17.3. Aplicando la regla de la cadena Teorema 7.4,

$$\begin{aligned}\frac{d}{dx}a^x &= \exp(x \ln a)\ \ln a \\ &= (\ln a)a^x.\end{aligned}$$

La fórmula incluye correctamente el caso $a = 1$: como $\ln 1 = 0$ y $1^x = 1$, $\frac{d}{dx}1^x = 0$.

2. Derivada de $\log_a x$. Por Definición 17.5, $\log_a x = \frac{1}{\ln a}\ln x$.

Como $a \neq 1$, tenemos $\ln a \neq 0$ y el factor $1/\ln a$ es una constante real bien definida. Por la linealidad Teorema 7.1 y la derivada del logaritmo Teorema 17.1, $\frac{d}{dx}\log_a x = \frac{1}{\ln a}\frac{1}{x} = \frac{1}{x \ln a}$, para todo $x > 0$.

3. Derivada de x^α. Fijemos $\alpha \in \mathbb{R}$. Sobre el dominio positivo, Definición 17.4 da $x^\alpha = \exp(\alpha \ln x)$.

La función interior $h(x) = \alpha \ln x$ es diferenciable para $x > 0$ y $h'(x) = \frac{\alpha}{x}$. Por la regla de la cadena,

$$\begin{aligned}\frac{d}{dx}x^\alpha &= \exp(\alpha \ln x)\frac{\alpha}{x} \\ &= x^\alpha \frac{\alpha}{x}.\end{aligned}$$

Como $x > 0$, Proposición 17.4 permite usar la ley de exponentes reales: $\frac{x^\alpha}{x} = \frac{x^\alpha}{x^1} = x^{\alpha-1}$.

Por tanto, $\frac{d}{dx}x^\alpha = \alpha x^{\alpha-1}$. □

La tercera fórmula merece una lectura cuidadosa. Para exponentes enteros o racionales ya existían reglas anteriores en dominios más amplios, incluidos ciertos valores negativos de la base. El nuevo enunciado no elimina esos resultados. Lo que añade es una regla uniforme para **todo** exponente real α, y esa extensión tiene como dominio canónico $x > 0$. No es legítimo transportar automáticamente esta fórmula a bases negativas cuando α es un real arbitrario, porque allí la potencia real general ni siquiera ha sido definida en C18.

Composición con funciones positivas. Las fórmulas anteriores se vuelven realmente útiles cuando el argumento deja de ser la variable identidad.

COROLARIO 17.4. **Corolario (reglas de composición).** Sea $I \subseteq \mathbb{R}$ un intervalo abierto y sea

$$u : I \longrightarrow (0, \infty)$$

diferenciable. Sea además $a > 0$, y en el caso logarítmico de base a supongamos $a \neq 1$. Entonces, para todo $x \in I$,

$$\boxed{\frac{d}{dx} \ln(u(x)) = \frac{u'(x)}{u(x)},}$$

$$\boxed{\frac{d}{dx} \log_a(u(x)) = \frac{u'(x)}{u(x) \ln a},}$$

$$\boxed{\frac{d}{dx} a^{u(x)} = (\ln a) a^{u(x)} u'(x),}$$

y, para todo $\alpha \in \mathbb{R}$,

$$\boxed{\frac{d}{dx} (u(x))^{\alpha} = \alpha (u(x))^{\alpha - 1} u'(x).}$$

En particular,

$$\boxed{\frac{d}{dx} \ln |x| = \frac{1}{x} \qquad (x \neq 0).}$$

Demostración. Las cuatro primeras fórmulas son aplicaciones directas de la regla de la cadena Teorema 7.4 al teorema Teorema 17.6 y a Teorema 17.1.

La última afirmación requiere separar el dominio en sus dos componentes. Si $x > 0$, entonces $\ln |x| = \ln x$, y por Teorema 17.1, $\frac{d}{dx} \ln |x| = \frac{1}{x}$. Si $x < 0$, entonces $|x| = -x > 0$, de modo que $\ln |x| = \ln(-x)$. Aplicando la regla de la cadena, $\frac{d}{dx} \ln(-x) = \frac{1}{-x}(-1) = \frac{1}{x}$. La función no está definida en $x = 0$, por lo que no existe ninguna afirmación diferencial allí. □

El valor absoluto no es un adorno
La fórmula

$$\frac{d}{dx} \ln |x| = \frac{1}{x} \qquad (x \neq 0)$$

es la forma correcta de reunir en una sola expresión las dos primitivas locales de $1/x$ sobre

$$(-\infty, 0) \qquad \text{y} \qquad (0, \infty).$$

Esto será útil en C19. No estamos integrando todavía: solo hemos identificado rigurosamente una derivada en el dominio desconectado $\mathbb{R} \setminus \{0\}$.

Cuando base y exponente varían a la vez. La expresión $u(x)^{v(x)}$ exige más cuidado que las anteriores. La regla para $a^{u(x)}$ supone una base constante, y la regla para $(u(x))^{\alpha}$ supone un exponente constante. Si ambas partes varían, ninguna de esas dos fórmulas puede aplicarse por sí sola. La construcción de §17.7 proporciona una definición punto por punto siempre que la base sea positiva.

DEFINICIÓN 17.6. **Definición (potencia de funciones con base positiva).** Sea $I \subseteq \mathbb{R}$ y sean

$$u : I \longrightarrow (0, \infty), \qquad v : I \longrightarrow \mathbb{R}.$$

Definimos la función

$$u^v : I \longrightarrow (0, \infty)$$

por

$$\boxed{(u^v)(x) := u(x)^{v(x)} := \exp(v(x) \ln u(x)).}$$

La hipótesis $u(x) > 0$ para todo $x \in I$ no es una comodidad técnica. Es exactamente lo que hace que $\ln u(x)$ esté definido y, por tanto, que la potencia real anterior tenga significado uniforme.

Diferenciación logarítmica sin pasos formales ocultos.

PROPOSICIÓN 17.5. **Proposición (exponente variable y diferenciación logarítmica).** Sea $I \subseteq \mathbb{R}$ un intervalo abierto. Supongamos que

$$u : I \longrightarrow (0, \infty), \qquad v : I \longrightarrow \mathbb{R}$$

son diferenciables. Definamos

$$w(x) := u(x)^{v(x)}.$$

Entonces w es diferenciable en I y

$$\boxed{w'(x) = u(x)^{v(x)} \left[v'(x) \ln u(x) + v(x) \frac{u'(x)}{u(x)} \right].}$$

Equivalentemente, como $w(x) > 0$,

$$\boxed{\frac{w'(x)}{w(x)} = v'(x) \ln u(x) + v(x) \frac{u'(x)}{u(x)}.}$$

Demostración. Por Definición 17.6, $w(x) = \exp(h(x))$, donde $h(x) = v(x) \ln u(x)$.
Como u es positiva y diferenciable, Corolario 17.4 muestra que

$$\frac{d}{dx} \ln u(x) = \frac{u'(x)}{u(x)}.$$

Las funciones v y $\ln u$ son diferenciables. Por la regla del producto Teorema 7.2,

$$h'(x) = v'(x) \ln u(x) + v(x) \frac{u'(x)}{u(x)}.$$

Ahora aplicamos la regla de la cadena Teorema 7.4 y la derivada de exp de Teorema 17.3:

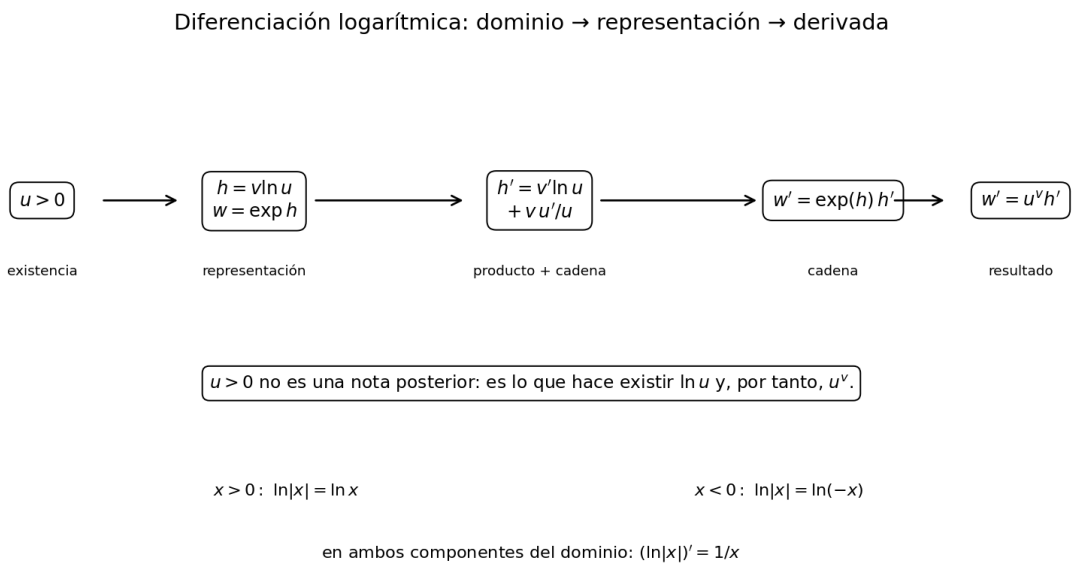


FIGURA 17.7. La diferenciación logarítmica comienza por una condición de existencia: $u > 0$ permite escribir $u^v = \exp(v \ln u)$; solo después se diferencian $v \ln u$ y la exponencial exterior. El mismo control de dominio distingue las dos componentes de $\ln|x|$.

$$\begin{aligned} w'(x) &= \exp(h(x))\, h'(x) \\ &= u(x)^{v(x)} \left[v'(x) \ln u(x) + v(x) \frac{u'(x)}{u(x)} \right]. \end{aligned}$$

Como $u(x)^{v(x)} > 0$, podemos dividir por $w(x)$ y obtener la segunda forma. □

Figura 17.7 resume la lógica de Proposición 17.5: dominio antes que representación y representación antes que regla de derivación.

Esta prueba muestra qué significa realmente **diferenciar logarítmicamente**. No se trata de una regla formal que permita «tomar logaritmos y derivar» sin verificar nada. La secuencia correcta es:

$$\boxed{u > 0 \Longrightarrow u^v = \exp(v \ln u) \Longrightarrow u^v \text{ es diferenciable} \Longrightarrow \frac{(u^v)'}{u^v} = (v \ln u)'.}$$

Solo después de acreditar la diferenciabilidad de la función podemos interpretar la identidad logarítmica como una herramienta de cálculo.

Casos particulares que deben coincidir. La fórmula general contiene como subcasos las reglas anteriores.

Si $v(x) \equiv \alpha$ es constante, entonces $v'(x) = 0$ y

$$w'(x) = u(x)^\alpha \left[\alpha \frac{u'(x)}{u(x)} \right] = \alpha u(x)^{\alpha - 1} u'(x).$$

Si $u(x) \equiv a > 0$ es constante, entonces $u'(x) = 0$ y $w'(x) = a^{v(x)} v'(x) \ln a$.

Es exactamente la regla para una exponencial de argumento compuesto. Así, la fórmula de exponente variable no reemplaza las reglas anteriores: las contiene y las hace compatibles.

Ejemplo 17.6. **Dos exponentes variables.**

1. La función x^x. En el dominio $x > 0$ definimos $f(x) = x^x$.

Tomamos

$$u(x) = x, \qquad v(x) = x.$$

Entonces

$$u'(x) = 1, \qquad v'(x) = 1.$$

Por Proposición 17.5,

$$\boxed{f'(x) = x^x(\ln x + 1).}$$

El dominio positivo no puede borrarse después del cálculo: la función real x^x que estamos diferenciando es la construida por Definición 17.6, cuya base debe ser positiva.

2. La función $(x^2+1)^x$. Ahora $g(x) = (x^2+1)^x$ está definida para todo $x \in \mathbb{R}$, porque $x^2 + 1 > 0$.

Tomamos

$$u(x) = x^2 + 1, \qquad v(x) = x.$$

Entonces

$$u'(x) = 2x, \qquad v'(x) = 1,$$

y la fórmula general da

$$\boxed{g'(x) = (x^2+1)^x \left[\ln(x^2+1) + \frac{2x^2}{x^2+1}\right].}$$

En este segundo ejemplo la positividad de la base no exige restringir el dominio, porque está garantizada algebraicamente.

Un límite de alcance que importa. Las nuevas reglas no autorizan a ignorar dominios. Por ejemplo, la expresión $(-x)^\alpha$ puede tener significado real para ciertos exponentes racionales y ciertos valores de x, según la teoría de C09. Pero para un parámetro real arbitrario $\alpha \in \mathbb{R}$ no existe en C18 una potencia real general de base negativa. Del mismo modo, escribir $u(x)^{v(x)} = e^{v(x)\ln u(x)}$ requiere $u(x) > 0$. Si la base se anula o cambia de signo, la fórmula no puede utilizarse simplemente «donde parezca algebraicamente conveniente». Hay que volver al dominio de la función original y analizar cada caso con la teoría que realmente esté disponible.

El repertorio diferencial queda cerrado. Al terminar §17.8 sabemos derivar, dentro de sus dominios correctos,

$$\ln u, \qquad \log_a u, \qquad a^u, \qquad u^\alpha, \qquad u^v.$$

La cadena conceptual completa es ahora

$$\boxed{\begin{array}{c}\text{construcción de } \ln \\ \Longrightarrow \text{construcción de } \exp \\ \Longrightarrow \text{potencias reales} \\ \Longrightarrow \text{logaritmos de base } a \\ \Longrightarrow \text{reglas diferenciales.}\end{array}}$$

No se ha necesitado recurrir a series de potencias, ecuaciones diferenciales como fundamento ni técnicas de integración. Las funciones exponenciales ya construidas permiten definir las hiperbólicas. Derivaremos sus identidades y, en las ramas invertibles correspondientes, expresaremos sus inversas mediante ln.

Antes de seguir. ¿Por qué la fórmula para x^α con $\alpha \in \mathbb{R}$ se formula canónicamente en $x > 0$? ¿Qué resultado permite derivar $\ln u$? ¿Por qué $\ln|x|$ tiene derivada $1/x$ en ambos componentes de su dominio? ¿Qué diferencia lógica hay entre $a^{u(x)}$, $(u(x))^\alpha$ y $u(x)^{v(x)}$? ¿Dónde entra la regla del producto en la diferenciación logarítmica? ¿Por qué la positividad de la base debe verificarse antes de escribir $u^v = \exp(v \ln u)$?

17.9. Funciones hiperbólicas y sus inversas elementales

La exponencial ya construida permite formar nuevas funciones sin introducir ningún objeto fundacional adicional. Entre las combinaciones más importantes aparecen las funciones hiperbólicas. Su teoría elemental nace directamente de las dos expresiones

$$e^x \qquad \text{y} \qquad e^{-x},$$

que están definidas y son positivas para todo $x \in \mathbb{R}$. El adjetivo «hiperbólico» no modificará aquí el punto de partida lógico. Definiremos las funciones mediante e^x, deduciremos sus identidades algebraicamente y solo después estudiaremos sus inversas.

Definiciones desde la exponencial.

Definición 17.7. **Definición (funciones hiperbólicas elementales).** Para todo $x \in \mathbb{R}$, definimos

$$\boxed{\sinh x := \frac{e^x - e^{-x}}{2},}$$

$$\boxed{\cosh x := \frac{e^x + e^{-x}}{2},}$$

y

$$\boxed{\tanh x := \frac{\sinh x}{\cosh x}.}$$

Las dos primeras funciones están definidas en toda $\mathbb{R}$. Además,

$$e^x > 0, \qquad e^{-x} > 0,$$

de modo que $\cosh x = \frac{e^x + e^{-x}}{2} > 0$.

Por tanto el denominador de $\tanh x$ nunca se anula y $\tanh : \mathbb{R} \longrightarrow \mathbb{R}$ también está bien definida en toda la recta.

Identidades que proceden de la ley exponencial.

PROPOSICIÓN 17.6. **Proposición (identidades hiperbólicas básicas).** Para todos $x, y \in \mathbb{R}$ se cumplen:

1. paridad:
$$\boxed{\sinh(-x) = -\sinh x, \qquad \cosh(-x) = \cosh x, \qquad \tanh(-x) = -\tanh x;}$$
2. identidad fundamental:
$$\boxed{\cosh^2 x - \sinh^2 x = 1;}$$
3. fórmulas de adición:
$$\boxed{\sinh(x+y) = \sinh x\, \cosh y + \cosh x\, \sinh y,}$$
$$\boxed{\cosh(x+y) = \cosh x\, \cosh y + \sinh x\, \sinh y;}$$
4. representación racional de tanh:
$$\boxed{\tanh x = \frac{e^{2x}-1}{e^{2x}+1};}$$
5. cota estricta:
$$\boxed{|\tanh x| < 1.}$$

Demostración. Las identidades de paridad salen inmediatamente de Definición 17.7 al sustituir $-x$ y observar que los dos términos e^x y e^{-x} intercambian sus papeles.

Para la identidad fundamental,

$$\begin{aligned}\cosh^2 x - \sinh^2 x &= \frac{(e^x + e^{-x})^2 - (e^x - e^{-x})^2}{4} \\ &= \frac{4e^x e^{-x}}{4} \\ &= 1,\end{aligned}$$

porque Corolario 17.2 da $e^x e^{-x} = e^0 = 1$. Para la fórmula de adición del seno hiperbólico usamos la ley exponencial:

$$e^{x+y} = e^x e^y, \qquad e^{-(x+y)} = e^{-x} e^{-y}.$$

Entonces

$$\begin{aligned}\sinh(x+y) &= \frac{e^x e^y - e^{-x} e^{-y}}{2} \\ &= \frac{(e^x - e^{-x})(e^y + e^{-y}) + (e^x + e^{-x})(e^y - e^{-y})}{4} \\ &= \sinh x\, \cosh y + \cosh x\, \sinh y.\end{aligned}$$

El cálculo para $\cosh(x+y)$ es análogo:

$$\begin{aligned}\cosh(x+y) &= \frac{e^x e^y + e^{-x}e^{-y}}{2}\\ &= \frac{(e^x+e^{-x})(e^y+e^{-y}) + (e^x-e^{-x})(e^y-e^{-y})}{4}\\ &= \cosh x\,\cosh y + \sinh x\,\sinh y.\end{aligned}$$

Para tanh, multiplicamos numerador y denominador por el número positivo e^x: $\tanh x = \frac{e^x-e^{-x}}{e^x+e^{-x}} = \frac{e^{2x}-1}{e^{2x}+1}$.

Finalmente, la identidad fundamental y la positividad de $\cosh x$ dan $\cosh^2 x = 1+\sinh^2 x > \sinh^2 x$. Por tanto $|\sinh x| < \cosh x$, y al dividir por $\cosh x > 0$ obtenemos $|\tanh x| < 1$. □

La identidad $\cosh^2 x - \sinh^2 x = 1$ explica la terminología: el punto $(\cosh x, \sinh x)$ pertenece a la rama derecha de la hipérbola $X^2 - Y^2 = 1$. No necesitamos convertir esta observación geométrica en una nueva definición; la parametrización ha surgido de la exponencial.

Figura 17.8 da contenido geométrico exacto a la terminología «hiperbólica» sin convertir la parametrización en una definición alternativa.

Derivadas, monotonía y rangos.

TEOREMA 17.7. **Teorema (cálculo y geometría global de las funciones hiperbólicas).** Las funciones de Definición 17.7 son diferenciables en toda $\mathbb{R}$ y satisfacen

$$\boxed{(\sinh x)' = \cosh x,}$$

$$\boxed{(\cosh x)' = \sinh x,}$$

y

$$\boxed{(\tanh x)' = \frac{1}{\cosh^2 x} = 1 - \tanh^2 x.}$$

Además:

1. $\sinh : \mathbb{R} \to \mathbb{R}$ es una biyección estrictamente creciente;
2. cosh es par, satisface $\cosh x \geq 1$, decrece estrictamente en $(-\infty, 0]$ y crece estrictamente en $[0, \infty)$; su restricción
$$\cosh : [0,\infty) \longrightarrow [1,\infty)$$
es una biyección;
3. $\tanh : \mathbb{R} \to (-1, 1)$ es una biyección estrictamente creciente.

Demostración.

Por Teorema 17.3 y la regla de la cadena Teorema 7.4 aplicada a $x \mapsto -x$,

$$\frac{d}{dx}e^x = e^x, \qquad \frac{d}{dx}e^{-x} = -e^{-x}.$$

La linealidad Teorema 7.1 da $(\sinh x)' = \frac{e^x+e^{-x}}{2} = \cosh x$, y $(\cosh x)' = \frac{e^x-e^{-x}}{2} = \sinh x$. Como $\cosh x > 0$, la regla del cociente Teorema 7.3 puede aplicarse a $\tanh x = \frac{\sinh x}{\cosh x}$.

Así,

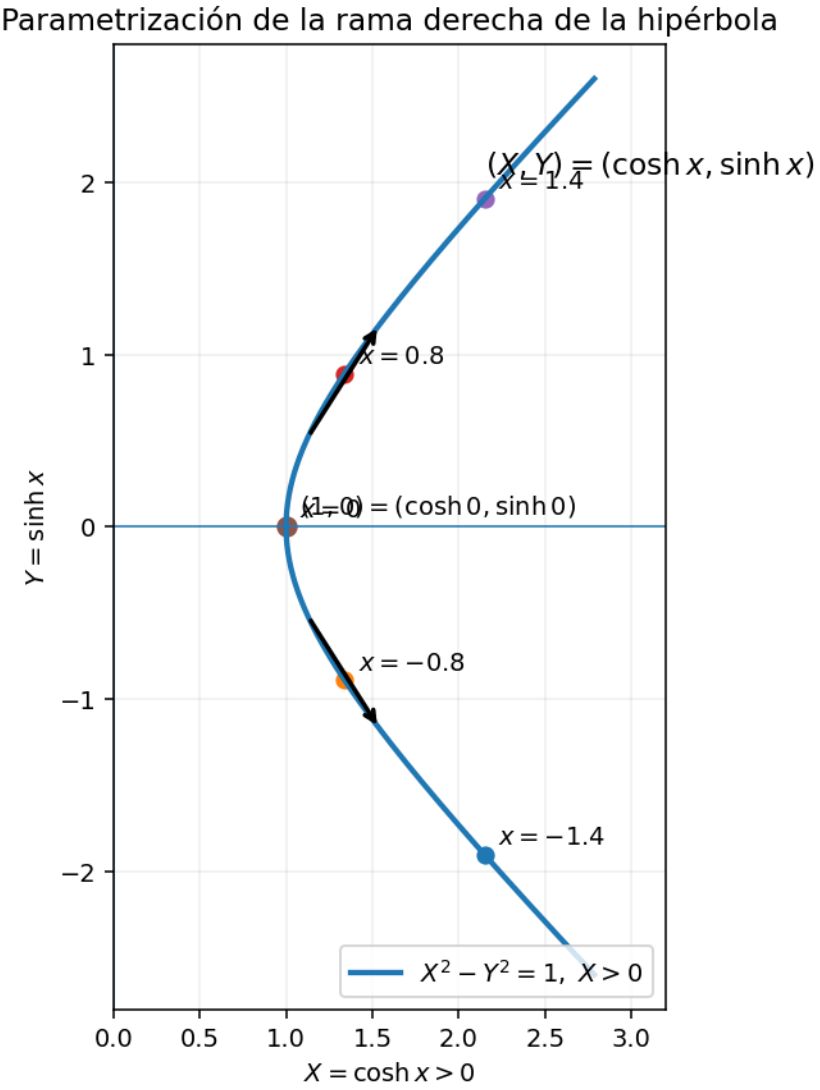


FIGURA 17.8. La identidad $\cosh^2 x - \sinh^2 x = 1$ parametriza la rama derecha $X > 0$ de la hipérbola $X^2 - Y^2 = 1$ mediante $(X, Y) = (\cosh x, \sinh x)$; $x = 0$ corresponde al vértice $(1, 0)$ y los signos de x recorren las dos mitades de esa rama.

$$\begin{aligned}(\tanh x)' &= \frac{\cosh^2 x - \sinh^2 x}{\cosh^2 x} \\ &= \frac{1}{\cosh^2 x}.\end{aligned}$$

Por la identidad fundamental,

$$1 - \tanh^2 x = \frac{\cosh^2 x - \sinh^2 x}{\cosh^2 x} = \frac{1}{\cosh^2 x}.$$

Esto prueba las tres fórmulas diferenciales. Como $(\sinh x)' = \cosh x > 0$ en toda $\mathbb{R}$, Teorema 10.1 muestra que sinh es estrictamente creciente. Para probar que alcanza todo real no necesitamos importar límites nuevos. Sea $y \in \mathbb{R}$ y definamos $z := y + \sqrt{y^2+1}$. Como $\sqrt{y^2+1} > |y|$, tenemos $z > 0$. Además, $(y + \sqrt{y^2+1})(\sqrt{y^2+1} - y) = 1$, de modo que $\frac{1}{z} = \sqrt{y^2+1} - y$.

Tomemos $x := \ln z$.

Entonces

$$e^x = z, \qquad e^{-x} = \frac{1}{z},$$

y

$$\begin{aligned}\sinh x &= \frac{1}{2}\left(z - \frac{1}{z}\right)\\ &= \frac{1}{2}\left(y + \sqrt{y^2+1} - \sqrt{y^2+1} + y\right)\\ &= y.\end{aligned}$$

Por tanto sinh es sobreyectiva sobre $\mathbb{R}$; junto con la monotonía estricta, es biyectiva. Para cosh, la identidad fundamental implica $\cosh^2 x = 1 + \sinh^2 x \geq 1$.

Como $\cosh x > 0$, $\cosh x \geq 1$.

Además, $\sinh 0 = 0$, y como sinh es estrictamente creciente,

$$x < 0 \Longrightarrow \sinh x < 0, \qquad x > 0 \Longrightarrow \sinh x > 0.$$

Puesto que $(\cosh x)' = \sinh x$,

Teorema 10.1 muestra que cosh decrece estrictamente en $(-\infty, 0]$ y crece estrictamente en $[0, \infty)$. En particular, su mínimo es $\cosh 0 = 1$.

Queda comprobar que la restricción a $[0, \infty)$ alcanza todo $y \geq 1$. Sea, pues, $y \geq 1$ y definamos $z := y + \sqrt{y^2-1}$.

Tenemos $z \geq 1$, por lo que $x := \ln z \geq 0$.

Además, $(y + \sqrt{y^2-1})(y - \sqrt{y^2-1}) = 1$, de modo que $\frac{1}{z} = y - \sqrt{y^2-1}$.

Entonces

$$\begin{aligned}\cosh x &= \frac{1}{2}\left(z + \frac{1}{z}\right)\\ &= y.\end{aligned}$$

La restricción $\cosh : [0, \infty) \to [1, \infty)$ es, por tanto, sobreyectiva; su crecimiento estricto da inyectividad.

Finalmente, $(\tanh x)' = \frac{1}{\cosh^2 x} > 0$, así que tanh es estrictamente creciente. Por Proposición 17.6 su imagen está contenida en $(-1, 1)$. Sea ahora $y \in (-1, 1)$.

Entonces

$$1 + y > 0, \qquad 1 - y > 0,$$

y podemos definir

$$x := \frac{1}{2}\ln\left(\frac{1+y}{1-y}\right).$$

La definición de potencia real da

$$e^{2x} = \exp\left(\ln\frac{1+y}{1-y}\right) = \frac{1+y}{1-y}.$$

Usando la representación racional de Proposición 17.6,

$$\begin{aligned}\tanh x &= \frac{e^{2x}-1}{e^{2x}+1}\\ &= \frac{\frac{1+y}{1-y}-1}{\frac{1+y}{1-y}+1}\\ &= y.\end{aligned}$$

Por tanto todo $y \in (-1,1)$ es alcanzado. La función $\tanh : \mathbb{R} \to (-1,1)$ es biyectiva. □

El teorema muestra una diferencia que será importante al invertir:

- sinh y tanh son inyectivas en toda $\mathbb{R}$;
- cosh no lo es, porque es par.

Por eso la inversa del coseno hiperbólico requiere escoger una rama, exactamente como ocurrió con otras funciones no inyectivas en C09.

Las inversas hiperbólicas.

Definición 17.8. **Definición (funciones hiperbólicas inversas reales).** A partir de las biyecciones de Teorema 17.7 definimos

$$\boxed{\operatorname{arsinh} := \sinh^{-1} : \mathbb{R} \longrightarrow \mathbb{R},}$$

$$\boxed{\operatorname{arcosh} := \left(\cosh|_{[0,\infty)}\right)^{-1} : [1,\infty) \longrightarrow [0,\infty),}$$

y

$$\boxed{\operatorname{artanh} := \tanh^{-1} : (-1,1) \longrightarrow \mathbb{R}.}$$

Figura 17.9 muestra simultáneamente por qué arsinh y artanh se invierten globalmente, mientras que arcosh exige seleccionar una rama de cosh.

La elección de la rama no es decorativa. La función cosh satisface $\cosh(-x) = \cosh x$, por lo que ningún intento de invertirla sobre toda $\mathbb{R}$ podría ser unívoco. La restricción a $[0,\infty)$ selecciona la rama creciente y produce una inversa bien definida.

Fórmulas logarítmicas y derivadas de las inversas.

Corolario 17.5. **Corolario (expresiones elementales de las inversas hiperbólicas).** Se cumplen:

$$\boxed{\operatorname{arsinh} y = \ln\left(y + \sqrt{y^2+1}\right) \qquad (y \in \mathbb{R}),}$$

$$\boxed{\operatorname{arcosh} y = \ln\left(y + \sqrt{y^2-1}\right) \qquad (y \geq 1),}$$

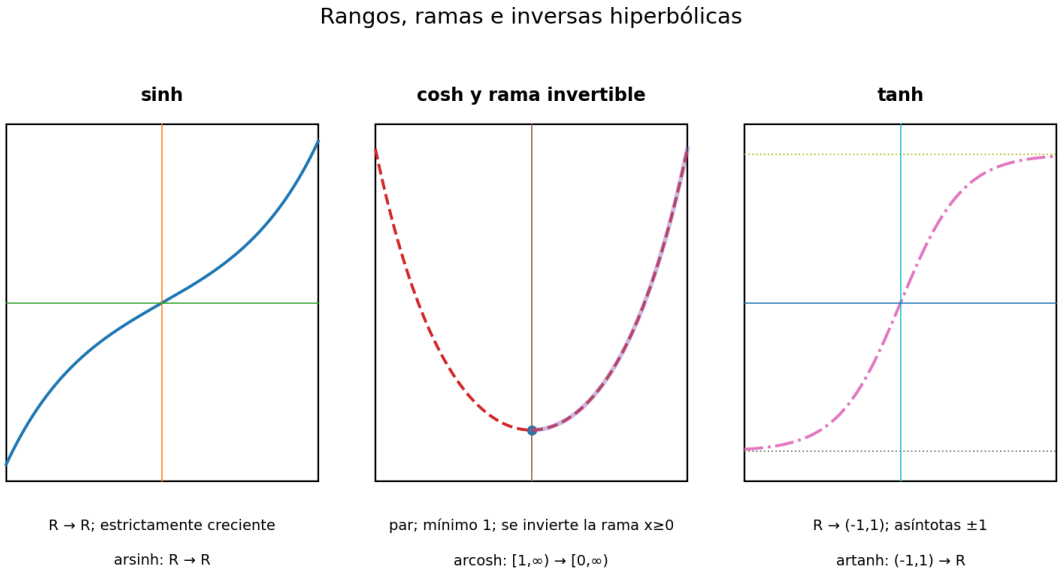


FIGURA 17.9. Los rangos determinan las inversas hiperbólicas: sinh es biyectiva $\mathbb{R} \to \mathbb{R}$, cosh necesita restringirse a la rama $[0,\infty) \to [1,\infty)$, y tanh es biyectiva $\mathbb{R} \to (-1,1)$ con asíntotas horizontales ± 1.

y

$$\boxed{\operatorname{artanh} y = \frac{1}{2}\ln\left(\frac{1+y}{1-y}\right) \qquad (|y|<1).}$$

Además,

$$\boxed{\frac{d}{dy}\operatorname{arsinh} y = \frac{1}{\sqrt{1+y^2}} \qquad (y\in\mathbb{R}),}$$

$$\boxed{\frac{d}{dy}\operatorname{arcosh} y = \frac{1}{\sqrt{y^2-1}} \qquad (y>1),}$$

y

$$\boxed{\frac{d}{dy}\operatorname{artanh} y = \frac{1}{1-y^2} \qquad (|y|<1).}$$

Demostración. Las tres fórmulas logarítmicas ya aparecieron, respectivamente, como los valores de x construidos en la demostración de sobreyectividad de Teorema 17.7. La unicidad proporcionada por las biyecciones permite identificarlos con las inversas recién definidas.

Conviene verificar también que los logaritmos están siempre bien tipados.

Para $y\in\mathbb{R}$, $\sqrt{y^2+1} > |y|$, de modo que $y+\sqrt{y^2+1} > 0$.

Para $y \geq 1$, $y+\sqrt{y^2-1} \geq 1 > 0$.

Si $|y|<1$, entonces

$$1+y>0, \qquad 1-y>0,$$

y por tanto $\frac{1+y}{1-y}>0$.

Las derivadas pueden deducirse sin volver a diferenciar las fórmulas logarítmicas. Aplicamos el teorema de la derivada de la inversa Teorema 8.1.

Sea $x=\operatorname{arsinh} y$.

Entonces $y=\sinh x$, y $\sinh'(x)=\cosh x>0$.

Por Teorema 8.1, $(\operatorname{arsinh})'(y)=\frac{1}{\cosh x}$.

La identidad fundamental y $\cosh x>0$ dan $\cosh x=\sqrt{1+\sinh^2 x}=\sqrt{1+y^2}$.

Luego

$$(\operatorname{arsinh})'(y)=\frac{1}{\sqrt{1+y^2}}.$$

Para arcosh tomamos $y>1$ y $x=\operatorname{arcosh} y$.

Entonces $x>0$, porque $\cosh 0=1$ y la rama es estrictamente creciente. Así, $\sinh x>0$.

Como $\cosh'(x)=\sinh x\neq 0$,

Teorema 8.1 se aplica a la restricción creciente de cosh. Obtenemos $(\operatorname{arcosh})'(y)=\frac{1}{\sinh x}$.

La identidad fundamental y la positividad de $\sinh x$ proporcionan $\sinh x=\sqrt{\cosh^2 x-1}=\sqrt{y^2-1}$, de donde

$$(\operatorname{arcosh})'(y)=\frac{1}{\sqrt{y^2-1}}.$$

No afirmamos esta fórmula en $y=1$: allí el antecedente es $x=0$ y $\cosh'(0)=\sinh 0=0$, de modo que la hipótesis de derivada no nula de Teorema 8.1 falla exactamente en el extremo.

Finalmente, sea

$$x=\operatorname{artanh} y, \qquad |y|<1.$$

Entonces $y=\tanh x$ y

$$\tanh'(x)=1-\tanh^2 x=1-y^2>0.$$

Por Teorema 8.1, $(\operatorname{artanh})'(y)=\frac{1}{1-y^2}$. □

Un mismo punto visto por las tres funciones.

Ejemplo 17.7. **El punto generado por $x=\ln 2$.** Como $e^{\ln 2}=2$ y $e^{-\ln 2}=\frac{1}{2}$, obtenemos

$$\sinh(\ln 2)=\frac{2-\frac{1}{2}}{2}=\frac{3}{4},$$

$$\cosh(\ln 2)=\frac{2+\frac{1}{2}}{2}=\frac{5}{4},$$

y

$$\tanh(\ln 2)=\frac{3/4}{5/4}=\frac{3}{5}.$$

Por las inversas correspondientes,

$$\boxed{\operatorname{arsinh}\frac{3}{4} = \ln 2,}$$

$$\boxed{\operatorname{arcosh}\frac{5}{4} = \ln 2,}$$

y

$$\boxed{\operatorname{artanh}\frac{3}{5} = \ln 2.}$$

Las tres igualdades también pueden comprobarse directamente con las fórmulas logarítmicas de Corolario 17.5. Por ejemplo,

$$\operatorname{arsinh}\frac{3}{4} = \ln\left(\frac{3}{4} + \sqrt{\frac{9}{16}+1}\right) = \ln\left(\frac{3}{4}+\frac{5}{4}\right) = \ln 2.$$

La familia se cierra sin añadir un nuevo fundamento. Las funciones hiperbólicas no han requerido una construcción independiente. Toda la sección se apoya en objetos ya obtenidos:

$$\boxed{\ln \Longrightarrow \exp \Longrightarrow e^x \Longrightarrow \sinh, \cosh, \tanh \Longrightarrow \text{inversas hiperbólicas.}}$$

Las expresiones logarítmicas de las inversas tampoco son definiciones arbitrarias: aparecen al resolver explícitamente las ecuaciones

$$\sinh x = y, \qquad \cosh x = y, \qquad \tanh x = y$$

en los dominios donde cada función es biyectiva. La próxima sección no añadirá una nueva familia elemental. Funcionará como laboratorio integrador: auditará dominios, representaciones equivalentes, inferencias falsas y compatibilidades entre todas las construcciones de C18 antes de abrir el banco final de ejercicios.

Antes de seguir. ¿Por qué $\cosh x$ nunca se anula? ¿Cómo se obtiene $\cosh^2 x - \sinh^2 x = 1$ únicamente desde la ley exponencial? ¿Por qué sinh puede invertirse en toda $\mathbb{R}$ pero cosh exige restringir el dominio? ¿Cómo prueba la fórmula explícita de artanh que cada $y \in (-1,1)$ tiene antecedente? ¿Por qué la fórmula de la derivada de arcosh se restringe a $y > 1$? ¿Qué papel cumple otra vez Teorema 8.1 al derivar las inversas?

17.10. Laboratorio integrador y auditoría de dominios

Las nueve secciones anteriores han construido una cadena de funciones que en cursos elementales suele presentarse como un repertorio ya disponible. Aquí ocurrió lo contrario: cada objeto nació solamente cuando las dependencias anteriores lo hicieron legítimo. Antes de cerrar el capítulo con ejercicios conviene detenerse y auditar esa cadena. El objetivo de esta sección no es añadir una función nueva ni demostrar otro teorema central. Es aprender a decidir **qué representación usar, qué dominio está realmente disponible y qué inferencias no están autorizadas**.

Estación 1 — Reconstruir el orden de construcción. La secuencia canónica de C18 puede resumirse así:

$$\ln x := \int_1^x \frac{dt}{t}$$
$$\Downarrow$$
$$(\ln x)' = \frac{1}{x}$$
$$\Downarrow$$
$$\ln(xy) = \ln x + \ln y$$
$$\Downarrow$$
$$\ln : (0,\infty) \to \mathbb{R} \text{ biyectiva}$$
$$\Downarrow$$
$$\exp := \ln^{-1}$$
$$\Downarrow$$
$$\exp' = \exp$$
$$\Downarrow$$
$$\exp(x+y) = \exp x \, \exp y$$
$$\Downarrow$$
$$e^x := \exp x$$
$$\Downarrow$$
$$a^x := \exp(x \ln a)$$
$$\Downarrow$$
$$\log_a x := \frac{\ln x}{\ln a}$$
$$\Downarrow$$
potencias de funciones e hiperbólicas.

Cada flecha contiene una justificación que ya fue demostrada:

- Definición 17.1 construye ln por acumulación;
- Teorema 17.1 obtiene su derivada;
- Proposición 17.1 obtiene la ley del producto sin sustitución integral;
- Teorema 17.2 acredita la biyección global;
- Definición 17.2 introduce exp solo entonces;
- Teorema 17.3 prueba $\exp' = \exp$ mediante el teorema de la inversa;
- Teorema 17.4 transporta suma a producto;
- Definición 17.3 legitima la notación e^x;
- Definición 17.4 y Definición 17.5 construyen a^x y $\log_a$;
- Definición 17.6 formaliza u^v para base positiva;
- Definición 17.7 y Definición 17.8 construyen las funciones hiperbólicas y sus inversas.

Esta reconstrucción permite diagnosticar un error frecuente: **una fórmula que aparece tarde en la cadena no puede usarse para justificar un objeto que aparece antes**, salvo que se proporcione una construcción independiente.

Estación 2 — Auditar antes el dominio que la fórmula. Consideremos las siguientes expresiones:

$$\ln u(x), \qquad \log_a u(x), \qquad a^{v(x)}, \qquad u(x)^{\alpha}, \qquad u(x)^{v(x)}.$$

No basta reconocer su forma. Debemos preguntar primero dónde existen dentro de la teoría construida.

Para $\ln u(x)$ necesitamos $u(x) > 0$.

Para $\log_a u(x)$ necesitamos simultáneamente

$$a > 0, \qquad a \neq 1, \qquad u(x) > 0.$$

Para $a^{v(x)}$ con exponente real arbitrario necesitamos $a > 0$. Para $(u(x))^{\alpha}$ con $\alpha \in \mathbb{R}$ arbitrario, la construcción uniforme de C18 exige $u(x) > 0$. Y la expresión general $u(x)^{v(x)} = \exp(v(x) \ln u(x))$ requiere nuevamente $u(x) > 0$.

La conclusión operativa es:

$$\boxed{\text{dominio de la expresión} \longrightarrow \text{representación permitida} \longrightarrow \text{cálculo.}}$$

El orden inverso es peligroso. Simplificar o derivar una fórmula y preguntar por el dominio solo al final puede eliminar restricciones esenciales.

Estación 3 — C09 no fue reemplazado por C18. C09 ya había construido potencias racionales con un control fino del signo de la base, la paridad de los denominadores y los puntos donde una potencia negativa deja de existir. En particular, ciertas expresiones racionales pueden tener sentido para bases negativas. C18 hace algo diferente. Definición 17.4 define $a^x = \exp(x \ln a)$ para **todo** $x \in \mathbb{R}$, pero exige $a > 0$. No hay contradicción. La teoría nueva extiende la antigua en el dominio donde ambas conviven. Proposición 17.4 demuestra precisamente que, para

$$a > 0, \qquad q \in \mathbb{Q},$$

la potencia real nueva a^q coincide con la potencia racional definida en Sección 8.5. Así debemos leer, por ejemplo, $(-8)^{1/3} = -2$. Esta igualdad pertenece legítimamente a la teoría racional de C09. Pero de ella no se puede concluir que $(-8)^{\alpha}$ esté definido como número real para un $\alpha \in \mathbb{R}$ arbitrario. C18 no ha construido tal objeto.

Una extensión no borra las excepciones del dominio anterior
La fórmula $x^{\alpha} = \exp(\alpha \ln x)$ es una construcción uniforme para $x > 0$ y $\alpha \in \mathbb{R}$. No convierte automáticamente en inválidas las potencias racionales de base negativa de C09, ni las convierte en una exponenciación real general sobre bases negativas.

Estación 4 — Elegir la representación que revela la estructura. Una misma función puede admitir varias representaciones correctas. La elección adecuada depende de la pregunta. Si queremos explotar productos, $\ln(xy) = \ln x + \ln y$ es preferible a volver a la integral de Definición 17.1. Si queremos explotar sumas en el exponente, $e^{x+y} = e^x e^y$ es la forma estructural adecuada. Si la base es positiva pero distinta de e, $a^x = \exp(x \ln a)$ convierte el problema en uno sobre la exponencial natural. Si queremos comparar bases logarítmicas, $\log_a x = \frac{\ln x}{\ln a}$ reduce toda la teoría a ln. Si base y exponente varían, $u^v = \exp(v \ln u)$ es la representación que permite aplicar producto y cadena, siempre bajo $u > 0$. Y para las inversas hiperbólicas, Corolario 17.5 convierte el problema de inversión en fórmulas construidas con ln y raíces positivas.

No existe, por tanto, una «forma más verdadera» de cada función. Existen representaciones equivalentes **en un dominio común**, y cada representación expone operaciones diferentes.

Estación 5 — Ocho inferencias que deben rechazarse. Examinemos ocho razonamientos plausibles pero incorrectos.

1. «Como $\ln : (0,\infty) \to \mathbb{R}$**, su rango es** $\mathbb{R}$**.»**

Incorrecto. El codominio escrito en una declaración de función no prueba sobreyectividad. La imagen completa se demostró recién en Teorema 17.2 mediante crecimiento, propiedad arquimediana, continuidad y valor intermedio.

2. «Como $(\ln x)' = 1/x$**, entonces necesariamente** $\ln(xy) = \ln x + \ln y$**.»**

Incompleto. La derivada sugiere la comparación correcta, pero Proposición 17.1 requiere fijar un parámetro, aplicar la cadena, demostrar igualdad de derivadas y determinar la constante mediante normalización.

3. «Como $\exp = \ln^{-1}$**, entonces automáticamente** $\exp(x+y) = \exp x \exp y$**.»**

Incorrecto. Ser inversa no basta. Teorema 17.4 usa además la ecuación funcional del logaritmo y su inyectividad.

4. «En cuanto definimos $e = \exp 1$**, ya podemos escribir** e^x **para todo real** x**.»**

Incorrecto dentro del orden canónico. El número e existe antes que la notación de potencia real e^x. Definición 17.3 aparece solo después de que la ley exponencial haya sido demostrada.

5. «La fórmula $(x^\alpha)' = \alpha x^{\alpha-1}$ **vale para todo** $x \neq 0$ **y todo** $\alpha \in \mathbb{R}$**.»**

Incorrecto. Teorema 17.6 formula la extensión a exponentes reales arbitrarios en $x > 0$. Otros dominios pueden existir para exponentes racionales o enteros particulares, pero deben analizarse desde la teoría correspondiente.

6. «Como $(\ln|x|)' = 1/x$ **para** $x \neq 0$**, entonces** $\ln x$ **está definido también para** $x < 0$**.»**

Incorrecto. Corolario 17.4 usa $|x| > 0$ y aplica ln al número positivo $|x|$. El dominio de ln sigue siendo $(0,\infty)$.

7. «Como cosh **tiene fórmula explícita, posee una inversa sobre toda** $\mathbb{R}$**.»**

Incorrecto. Proposición 17.6 muestra que cosh es par, luego $\cosh(-x) = \cosh x$. La inversa real de Definición 17.8 requiere restringir a $[0,\infty)$.

8. «Podemos diferenciar $u(x)^{v(x)}$ **tomando logaritmos, aunque** u **cambie de signo.»**

Incorrecto en la teoría de C18. Definición 17.6 y Proposición 17.5 requieren $u(x) > 0$. Sin esa condición, ni $\ln u(x)$ ni la potencia real general utilizada en la demostración están disponibles.

Estación 6 — Una cancelación algebraica no repara un dominio perdido. Consideremos $f(x) := \frac{x^2-1}{x-1}$.

Como expresión algebraica, $x^2 - 1 = (x-1)(x+1)$, y para $x \neq 1$, $f(x) = x + 1$.

Pero la función original no está definida en $x = 1$. La simplificación produce una fórmula coincidente **en el dominio original**, no una extensión automática. El mismo principio reaparece con logaritmos y potencias. Por ejemplo, $\ln(x^2) = 2\ln x$ no es una identidad válida para todo $x \neq 0$, porque el miembro derecho exige $x > 0$. Lo que sí podemos afirmar para $x \neq 0$ es $x^2 = |x|^2$ y, como $|x| > 0$,

$$\boxed{\ln(x^2) = 2\ln|x|.}$$

La diferencia no es cosmética: la segunda fórmula respeta los dominios de todos sus términos. Análogamente, de $\sqrt{x^2} = |x|$ no podemos sustituir $|x|$ por x sin la hipótesis adicional $x \geq 0$.

Estación 7 — Diferenciar no es construir. En C18 aparecen dos tipos de argumentos que conviene separar. Por una parte están las **construcciones**:

$$\ln x := \int_1^x \frac{dt}{t},$$

$$\exp := \ln^{-1},$$

$$a^x := \exp(x \ln a).$$

Por otra están las **consecuencias diferenciales**: $(\ln x)' = \frac{1}{x}$,

$$\exp' = \exp,$$

$$(a^x)' = (\ln a)a^x.$$

El hecho de que una función satisfaga una fórmula diferencial no significa que esa fórmula haya sido usada para definirla. En particular, $y' = y$ es una propiedad ya probada de $y = e^x$, pero C18 no ha demostrado todavía que una condición inicial como $y(0) = 1$ caracterice de manera única a esa función entre todas las soluciones posibles de una ecuación diferencial. Ese problema pertenece a C21. Esta separación impide razonar circularmente:

> «La exponencial es la función cuya derivada es ella misma; por eso su derivada es ella misma.»

Ese enunciado confunde construcción y caracterización. En CPM la dirección es:

$$\boxed{\text{integral} \to \ln \to \text{inversa} \to \exp \to \exp' = \exp\,.}$$

Funciones circulares: construir antes de usar sus derivadas. Las funciones hiperbólicas no sustituyen a seno y coseno. Para las técnicas del capítulo siguiente necesitamos construir también las funciones circulares, con su normalización en radianes. La idea será invertir una integral positiva y, después, prolongar por simetrías. Solo emplearemos el teorema fundamental, la derivada de una inversa y el teorema del valor medio, ya disponibles.

Definamos

$$A(x) = \int_0^x \frac{dt}{1+t^2} \qquad (x \in \mathbb{R}).$$

Por el teorema fundamental, $A'(x) = 1/(1+x^2) > 0$; por tanto A es estrictamente creciente. El integrando es par, así que el cambio $t = -s$ prueba $A(-x) = -A(x)$. Para $x \geq 1$,

$$A(x) \leq A(1) + \int_1^x \frac{dt}{t^2} < 2.$$

El conjunto de sus valores positivos está acotado. Sea $L = \sup\{A(x) : x \geq 0\}$. Es positivo y, por monotonía y la propiedad del supremo, $A(x) \to L$ cuando $x \to +\infty$. Ningún argumento finito alcanza L, porque A sigue creciendo. Continuidad, valor intermedio y paridad muestran que A es una biyección de $\mathbb{R}$ sobre $(-L, L)$. Su inversa T es continua y diferenciable, con

$$T'(y) = 1 + T(y)^2, \qquad T(0) = 0.$$

En ese intervalo pongamos

$$C(y)=\frac{1}{\sqrt{1+T(y)^2}},\qquad S(y)=\frac{T(y)}{\sqrt{1+T(y)^2}}.$$

La raíz positiva está construida en C09. La regla de la cadena da $S'=C$, $C'=-S$ y $S^2+C^2=1$. Además, $S(0)=0$, $C(0)=1$, S es impar y C par. Cuando $y\to L^-$, necesariamente $T(y)\to+\infty$: si permaneciera acotada, la continuidad de A daría un argumento finito con imagen L. Así $S(y)\to 1$ y $C(y)\to 0$.

Definimos $S(L)=1$, $C(L)=0$ y, para $L\le y\le 2L$, prolongamos mediante

$$S(y)=S(2L-y),\qquad C(y)=-C(2L-y).$$

Para $2L\le y\le 4L$ ponemos $S(y)=-S(y-2L)$ y $C(y)=-C(y-2L)$; finalmente prolongamos ambas funciones a toda la recta con período $4L$. Los valores en los extremos coinciden: en $0,L,2L,3L,4L$ son, respectivamente, $(S,C)=(0,1),(1,0),(0,-1),(-1,0),(0,1)$. En el interior de cada tramo siguen cumpliéndose $S'=C$ y $C'=-S$ por la regla de la cadena.

También se conservan las derivadas en los empalmes. Para verlo sin suponerlo, aplicamos el valor medio al cociente incremental de cada función en un intervalo lateral: su derivada interior tiende al valor continuo de C o de $-S$, respectivamente. Los dos cocientes laterales tienen el mismo límite. Esto prueba que la prolongación es diferenciable en toda la recta y satisface las mismas ecuaciones, incluso en los extremos de los tramos.

Llamamos **seno** a S, **coseno** a C y definimos $\pi=2L$. Esta normalización analítica es la de los radianes: la curva $(C(y),S(y))$ recorre la circunferencia unidad en sentido positivo con velocidad de módulo $\sqrt{(-S)^2+C^2}=1$; la longitud del arco se justificará al estudiar longitudes. No necesitamos usar ahora una fórmula de longitud para definir las funciones. En particular, hemos demostrado, y no incorporado desde una tabla,

$$(\sin y)'=\cos y,\qquad (\cos y)'=-\sin y,\qquad \sin^2 y+\cos^2 y=1.$$

Adición, períodos e inversas circulares. Para obtener las identidades de adición basta una unicidad elemental. Si dos parejas diferenciables satisfacen $s'=c$, $c'=-s$ y tienen los mismos valores en un punto, sus diferencias p,q cumplen $p'=q$, $q'=-p$. Entonces $(p^2+q^2)'=0$. Por el valor medio, p^2+q^2 es constante y su valor inicial es cero; ambas diferencias son idénticamente nulas.

Fijado b, las parejas $(\sin(y+b),\cos(y+b))$ y

$$(\sin y\cos b+\cos y\sin b,\cos y\cos b-\sin y\sin b)$$

satisfacen esas ecuaciones y coinciden en $y=0$. Por unicidad,

$$\sin(y+b)=\sin y\cos b+\cos y\sin b,$$
$$\cos(y+b)=\cos y\cos b-\sin y\sin b.$$

Estas fórmulas justifican los ángulos dobles y las identidades producto-a-suma que utilizaremos en C19. La construcción muestra además que sin es impar, cos par, ambas tienen período 2π, $\sin(k\pi)=0$ y $\cos(k\pi)=(-1)^k$ para todo entero k. En cada primer cuarto de vuelta S crece de cero a uno y C decrece de uno a cero; las simetrías describen el resto. Por ello los ceros de seno son exactamente $k\pi$ y los de coseno exactamente $\pi/2+k\pi$. Donde sus denominadores no se anulan definimos $\tan=\sin/\cos$, $\cot=\cos/\sin$, $\sec=1/\cos$ y $\csc=1/\sin$. La regla del cociente y la identidad pitagórica dan

$$\tan' = \sec^2 = 1 + \tan^2, \quad \cot' = -\csc^2, \quad \sec' = \sec\tan, \quad \csc' = -\csc\cot.$$

En $(-\pi/2, \pi/2)$ tenemos $\tan = T$, de modo que su inversa principal es precisamente $\arctan = A$ y

$$(\arctan x)' = \frac{1}{1+x^2} \qquad (x \in \mathbb{R}).$$

El seno restringido a $[-\pi/2, \pi/2]$ es una biyección creciente sobre $[-1, 1]$; el coseno restringido a $[0, \pi]$ es una biyección decreciente sobre el mismo intervalo. Sus inversas principales son arcsin y $\arccos$. En el interior $(-1, 1)$, la derivada de la inversa y los signos del coseno y seno en las ramas elegidas dan

$$(\arcsin x)' = \frac{1}{\sqrt{1-x^2}}, \qquad (\arccos x)' = -\frac{1}{\sqrt{1-x^2}}.$$

Estas fórmulas no afirman una derivada finita en ± 1. En particular, el paso $x = a\sin y$, con $a > 0$ y $y \in [-\pi/2, \pi/2]$, permite escribir $\sqrt{a^2 - x^2} = a\cos y$ sin perder el signo. Esa elección de rama será esencial en las sustituciones trigonométricas. Como comprobación de los límites elementales, los cocientes incrementales en cero dan $\sin h/h \to 1$ y $(\cos h - 1)/h \to 0$. Los ejemplos tempranos de oscilación utilizaron el modelo triangular y no requirieron estas derivadas; desde aquí, el repertorio diferencial circular y sus ramas inversas queda demostrado dentro del libro.

Comprobación activa. Reconstruye por qué L es finito y no se alcanza, verifica la derivada de C, y señala en qué paso la rama de arcsin obliga a tomar la raíz positiva. Si se cambia la rama, no se puede conservar automáticamente el signo de la fórmula inversa.

Estación 8 — El puente exacto hacia C19. La sección §17.8 dejó probado

$$\boxed{\frac{d}{dx}\ln|x| = \frac{1}{x} \qquad (x \neq 0).}$$

Esto identifica una primitiva de $1/x$ en cada componente del dominio $\mathbb{R}\setminus\{0\} = (-\infty, 0)\cup(0, \infty)$. Pero todavía no hemos convertido esa observación en una técnica sistemática de integración. C19 será dueño de cuestiones como:

- reconocer primitivas mediante sustitución;
- integración por partes;
- transformación sistemática de integrandos;
- fracciones parciales;
- uso operativo de
$$\int \frac{dx}{x} = \ln|x| + C$$
dentro del marco apropiado.

La frontera es importante: C18 construye las funciones y sus derivadas; C19 estudia cómo utilizarlas dentro de un repertorio sistemático de técnicas de integración.

Estación 9 — Lo que se difiere a C20 y C21. Tampoco debemos transformar el capítulo en una colección de aplicaciones antes de tiempo. C20 será dueño de las aplicaciones sistemáticas de la integral y de las funciones construidas aquí a magnitudes geométricas o físicas. C21 será dueño de las ecuaciones diferenciales elementales. Allí podrá estudiarse de manera sistemática por qué la exponencial aparece en crecimiento y decrecimiento, cómo se resuelven

problemas de valor inicial y qué papel desempeña la unicidad. Por tanto, aunque ya sabemos que $(e^x)' = e^x$, no abriremos aquí el desarrollo general de $y' = ky$, ni modelos de población, decaimiento, enfriamiento u otras aplicaciones.

Estación 10 — Fronteras hacia el Tomo II. Hay además propiedades importantes de ln y exp que **no** pertenecen todavía a este capítulo:

- desarrollos en series de potencias;
- series de Taylor infinitas;
- analiticidad;
- integrales impropias;
- diferenciación bajo el signo integral con parámetros;
- extensiones complejas del logaritmo y la exponencial.

La ausencia de estos temas no es una laguna en las demostraciones actuales. Ninguno fue usado para construir los objetos de C18. Esto permite registrar una distinción metodológica:

una función puede estar rigurosamente construida antes de conocer todas sus representaciones futuras.

Mapa de decisión. Ante una expresión de la familia de C18 podemos seguir el siguiente orden.

Pregunta	Control que debe hacerse primero
¿Aparece $\ln u$?	Verificar $u > 0$.
¿Aparece $\log_a u$?	Verificar $a > 0$, $a \neq 1$ y $u > 0$.
¿Aparece a^x con x real arbitrario?	Verificar $a > 0$.
¿Aparece u^α con α real arbitrario?	Verificar $u > 0$ en la construcción uniforme de C18.
¿Aparece u^v?	Verificar $u > 0$ antes de usar $\exp(v \ln u)$.
¿Se quiere invertir cosh?	Elegir una rama inyectiva; en C18, $[0, \infty)$.
¿Se usa una fórmula racional de C09 con base negativa?	Conservar las restricciones racionales originales; no sustituirlas por la teoría real positiva.
¿Una simplificación elimina un denominador o un logaritmo?	Conservar el dominio de la expresión original salvo que se defina explícitamente una extensión.
¿Se quiere integrar usando ln o exp?	Verificar si el paso requiere una técnica reservada a C19.
¿Se quiere caracterizar exp por una EDO?	Diferir la teoría sistemática a C21.

El laboratorio puede condensarse en una sola regla de lectura:

objeto → dominio → representación → teorema aplicable → conclusión.

Figura 17.10 resume la disciplina de dominio que gobierna todo el capítulo y mantiene explícita la frontera entre las potencias racionales anteriores y la teoría real de base positiva de C18.

Una fórmula puede conservar su apariencia tras una simplificación y perder las condiciones de dominio que la justificaban. Esas condiciones deben acompañar cada operación, incluidas las que introducen exponentes o inversas.

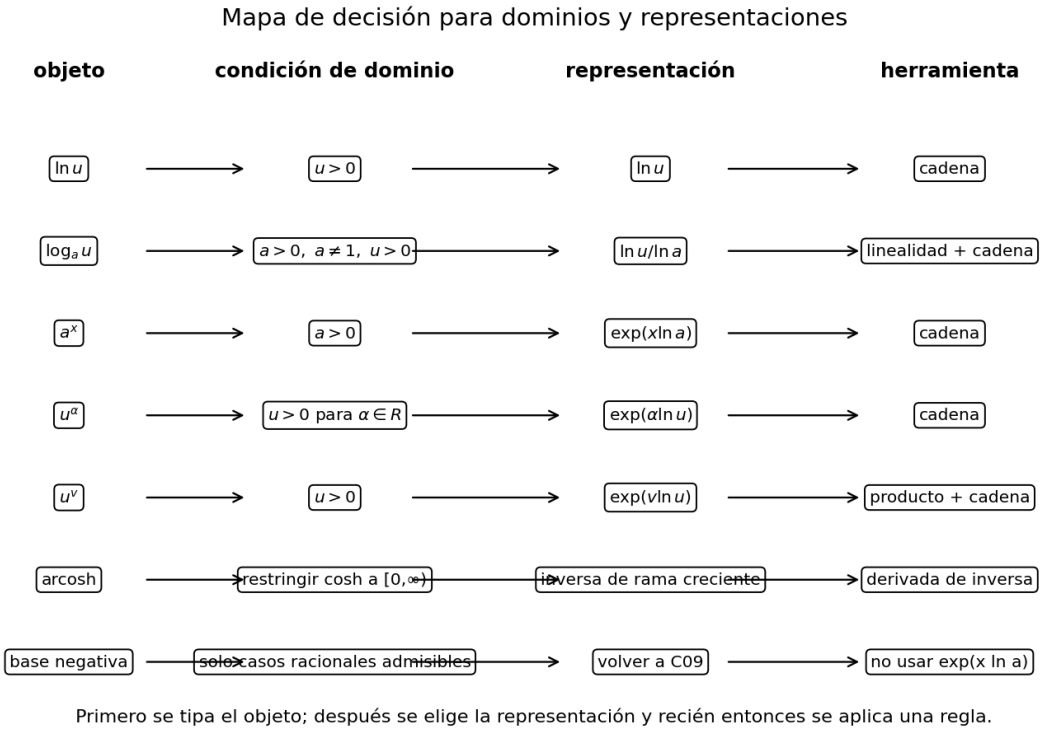


FIGURA 17.10. Mapa operativo de C18: identificar el objeto, verificar primero su dominio, escoger una representación ya construida y solo entonces aplicar el teorema correspondiente. La potencia racional con base negativa permanece en la teoría de C09 y no se sustituye por $\exp(x \ln a)$.

Cierre de la exposición. C18 comenzó con un integrando que parecía quedar fuera de la regla ordinaria de potencias: $\frac{1}{x}$. La integral permitió construir ln; la geometría de ln permitió invertirlo; la inversa produjo exp; la ley funcional permitió construir exponentes reales; y esas funciones generaron logaritmos de base arbitraria, diferenciación logarítmica y funciones hiperbólicas. La arquitectura completa puede verse ahora como

$$\boxed{\frac{1}{x} \xrightarrow{\text{FTC}} \ln \xrightarrow{\text{inversa}} \exp \longrightarrow \text{potencias reales} \longrightarrow \text{logaritmos} \longrightarrow \text{familias derivadas.}}$$

En los ejercicios tendremos que distinguir codominio e imagen al invertir ln, conservar las bases positivas para exponentes reales arbitrarios y verificar el signo antes de aplicar diferenciación logarítmica. Las potencias racionales previamente construidas conservan sus dominios propios; una simplificación no incorpora por sí sola entradas nuevas.

17.11. Ejercicios y soluciones

Los cuarenta ejercicios siguientes recorren la arquitectura completa del capítulo sin introducir teoría posterior. El banco está organizado en siete niveles y utiliza únicamente resultados disponibles hasta §17.10 y capítulos anteriores. Las soluciones distinguen sistemáticamente **existencia**, **dominio**, **representación**, **identidad algebraica**, **inversión** y **diferenciación**.

Nivel A — Construcción integral y propiedades básicas del logaritmo. Ejercicio A1. Normalización y signo. A partir de Definición 17.1 y Proposición 17.2:

1. justifica $\ln 1 = 0$;
2. demuestra que $0 < x < 1$ implica $\ln x < 0$;
3. demuestra que $x > 1$ implica $\ln x > 0$.

Indica qué parte procede de la definición integral y qué parte procede de la monotonía.

Ejercicio A2. Diferencias de logaritmos como integrales. Sean $a, b > 0$. Demuestra, incluyendo el caso $b < a$, que

$$\boxed{\int_a^b \frac{dt}{t} = \ln b - \ln a.}$$

No utilices sustitución de variables.

Ejercicio A3. Reconstruir la derivada desde el FTC. Fija $c > 0$. Elige explícitamente un intervalo compacto positivo que contenga a 1 y a c, y reconstruye la prueba de $(\ln)'(c) = \frac{1}{c}$ a partir de Teorema 16.2 y la aditividad orientada de Proposición 14.3. Explica por qué no basta citar informalmente el FTC «sobre $(0, \infty)$».

Ejercicio A4. Cálculo exacto sin decimales. Simplifica hasta una combinación de $\ln 2$ y $\ln 3$:

$$\ln\left(\frac{81}{16}\right), \qquad \ln\left(\sqrt[3]{\frac{9}{8}}\right).$$

Justifica cada paso mediante Corolario 17.1.

Ejercicio A5. Una desigualdad fundamental. Demuestra, usando el teorema del valor medio diferencial, que para todo $x > 0$,

$$\boxed{\ln x \leq x - 1,}$$

y que la igualdad ocurre únicamente cuando $x = 1$.

Ejercicio A6. Una cota complementaria. Deduce del ejercicio anterior que, para todo $x > 0$,

$$\boxed{1 - \frac{1}{x} \leq \ln x \leq x - 1.}$$

Determina cuándo hay igualdad en cada lado.

Nivel B — Ecuación funcional, rango e inversión. Ejercicio B1. Tres factores y un cociente. Sean $x, y, z > 0$. Demuestra $\ln(xyz) = \ln x + \ln y + \ln z$ y $\ln\left(\frac{xy}{z}\right) = \ln x + \ln y - \ln z$.

No vuelvas a la definición integral: usa la estructura ya demostrada.

Ejercicio B2. Resolver una ecuación logarítmica exactamente. Resuelve en $x > 0$ $\ln\left(\frac{x}{2}\right) = \ln 3 - \ln 4$.

La solución debe justificarse mediante la inyectividad de ln, no mediante una regla informal de «cancelar logaritmos».

Ejercicio B3. Encerrar un nivel real. Sea $y \in \mathbb{R}$. Demuestra que existe $n \in \mathbb{N}_{>0}$ tal que $-n \ln 2 < y < n \ln 2$.

Concluye que existe un único $x \in (2^{-n}, 2^n)$ tal que $\ln x = y$. Señala dónde intervienen la propiedad arquimediana, la continuidad y la monotonía.

Ejercicio B4. Creciente no significa sobreyectiva. Considera

$$f : (0, \infty) \longrightarrow \mathbb{R}, \qquad f(x) = \frac{x}{1 + x}.$$

1. Demuestra que f es estrictamente creciente.
2. Prueba que $0 < f(x) < 1$ para todo $x > 0$.
3. Explica por qué este ejemplo muestra que el crecimiento estricto de ln no bastaba para concluir $\mathrm{Im}(\ln) = \mathbb{R}$.

Ejercicio B5. Leer correctamente una función inversa. Simplifica, indicando el dominio que legitima cada identidad:

$$\ln(\exp(-5)), \qquad \exp(\ln 7), \qquad \exp\left(\ln \frac{2}{5}\right).$$

Explica por qué no tendría sentido real escribir $\exp(\ln(-7))$ dentro de la teoría de C18.

Ejercicio B6. Comportamiento extremo de la inversa. Usando la monotonía de exp y las identidades inversas, demuestra

$$\lim_{x \to +\infty} \exp x = +\infty$$

y

$$\lim_{x \to -\infty} \exp x = 0.$$

No uses series ni ecuaciones diferenciales.

Nivel C — Exponencial, número e y leyes funcionales. Ejercicio C1. Iterar la ley exponencial. Sea $n \in \mathbb{Z}$. Demuestra que para todo $x \in \mathbb{R}$,

$$\boxed{e^{nx} = (e^x)^n,}$$

interpretando el miembro derecho como potencia entera de un número positivo.

Ejercicio C2. Una expresión mixta. Simplifica exactamente $e^{\ln 5 - \ln 2}$ y justifica cada transformación sin convertir ningún logaritmo en decimal.

Ejercicio C3. Resolver una ecuación exponencial. Resuelve en $\mathbb{R}$ $e^{2x-1} = 7$.
Justifica la unicidad de la solución.

Ejercicio C4. Compatibilidad racional visible. Demuestra $e^{3/2} = e\sqrt{e}$ y

$$\left(e^{2/3}\right)^3 = e^2.$$

Indica en qué punto se usa que la definición de C18 coincide con las potencias racionales de C09.

Ejercicio C5. La exponencial queda sobre su tangente en cero. Usa la desigualdad de A5 y la identidad $\ln(e^x) = x$ para demostrar

$$\boxed{e^x \geq 1 + x \qquad (x \in \mathbb{R}),}$$

con igualdad únicamente en $x = 0$.

Ejercicio C6. La ecuación $y' = y$ no basta como definición. Sea $h(x) = 3e^x$.

1. Demuestra que $h' = h$.
2. Compara $h(0)$ con $\exp(0)$.
3. Explica por qué el hecho $\exp' = \exp$ no caracteriza por sí solo a la función exponencial construida en C18.

Nivel D — Potencias reales y logaritmos de base arbitraria. Ejercicio D1. La extensión conserva el valor racional. Calcula $27^{2/3}$ usando primero la teoría racional de C09 y después la definición $a^x = \exp(x \ln a)$.
Explica por qué ambos caminos deben producir el mismo número.

Ejercicio D2. Reconstruir dos leyes desde la definición. Sean $a, b > 0$ y $x, y \in \mathbb{R}$. Demuestra directamente desde Definición 17.4 que $(ab)^x = a^x b^x$ y $(a^x)^y = a^{xy}$.
Especifica dónde usas la ecuación funcional de ln.

Ejercicio D3. Dos bases, dos orientaciones. Compara las funciones

$$f(x) = 2^x, \qquad g(x) = \left(\frac{1}{2}\right)^x.$$

1. Determina cuál es creciente y cuál decreciente.
2. Demuestra que $g(x) = f(-x)$.
3. Resuelve las desigualdades $2^x > 1$ y $(1/2)^x > 1$.

Ejercicio D4. Una ecuación con base arbitraria. Resuelve $9^x = 27$ usando logaritmos. Justifica la unicidad de la solución sin recurrir a tanteo.

Ejercicio D5. Dos logaritmos exactos. Calcula

$$\log_4 32 \qquad \text{y} \qquad \log_{1/3} 27.$$

Usa el cambio de base y explica cómo el signo de $\ln(1/3)$ concuerda con el segundo resultado.

Ejercicio D6. Base negativa: distinguir teorías. Dentro de los reales:

1. calcula $(-8)^{1/3}$ y $(-8)^{2/3}$ usando C09;
2. explica por qué esos valores no contradicen Definición 17.4;
3. decide si C18 ha definido $(-8)^{\sqrt{2}}$ y justifica la respuesta.

Nivel E — Derivación, dominios y exponentes variables. Ejercicio E1. Exponencial de argumento cuadrático. Calcula la derivada de $f(x) = 3^{x^2}$ para todo $x \in \mathbb{R}$.

Ejercicio E2. Logaritmo de una expresión siempre positiva. Calcula la derivada de $g(x) = \log_5(x^2 + 1)$ y justifica por qué su dominio es toda $\mathbb{R}$.

Ejercicio E3. Una potencia de exponente irracional. Para $h(x) = x^{\sqrt{2}}$,

1. fija el dominio canónico proporcionado por C18;
2. calcula $h'(x)$;
3. explica por qué no has demostrado una fórmula sobre $x < 0$.

Ejercicio E4. Base y exponente variables sin restricción adicional. Sea $F(x) = (x^2 + 1)^x$.

Demuestra que F está definida para todo real x y calcula $F'(x)$ mediante Proposición 17.5.

Ejercicio E5. Una potencia más sensible al dominio. Para $x > 0$ define $G(x) = x^{x^2}$.

Calcula $G'(x)$ y factoriza el resultado de la forma

$$G'(x) = x^{x^2} \cdot (\text{expresión elemental}).$$

Ejercicio E6. El valor absoluto conserva tres componentes del dominio. Define $H(x) = \ln|x^2 - 1|$.

1. determina el dominio real de H;
2. calcula $H'(x)$ en ese dominio;
3. explica por qué no se puede simplificar globalmente H como $\ln(x^2 - 1)$.

Nivel F — Funciones hiperbólicas, inversas y contraejemplos. Ejercicio F1. Identidad y no anulación. A partir de Definición 17.7, demuestra directamente $\cosh^2 x - \sinh^2 x = 1$ y deduce que $\cosh x \neq 0$ para todo $x \in \mathbb{R}$.

Ejercicio F2. Un punto hiperbólico generado por $\ln 3$**.** Calcula exactamente

$$\sinh(\ln 3), \qquad \cosh(\ln 3), \qquad \tanh(\ln 3).$$

Verifica la identidad fundamental con los tres valores obtenidos.

Ejercicio F3. Recuperar un argumento con arsinh**.** Demuestra

$$\operatorname{arsinh}\left(\frac{4}{3}\right) = \ln 3$$

de dos maneras:

1. usando que arsinh es la inversa de sinh;
2. usando la fórmula logarítmica de Corolario 17.5.

Ejercicio F4. Un valor exacto de la tangente hiperbólica inversa. Demuestra

$$\boxed{\operatorname{artanh}\left(\frac{1}{2}\right) = \frac{1}{2}\ln 3.}$$

Verifica primero que $1/2$ pertenece al dominio de artanh.

Ejercicio F5. La rama positiva del coseno hiperbólico. Demuestra

$$\operatorname{arcosh}\left(\frac{17}{8}\right) = \ln 4.$$

Justifica por qué $\ln 4$ pertenece a la rama escogida para invertir cosh.

Ejercicio F6. Dos obstrucciones distintas.

1. Demuestra que cosh no posee inversa global $\mathbb{R} \to \mathbb{R}$ por falta de inyectividad.
2. Explica por qué la rama $[0, \infty)$ sí puede invertirse.
3. Justifica por qué la fórmula

$$(\operatorname{arcosh} y)' = \frac{1}{\sqrt{y^2 - 1}}$$

se afirma para $y > 1$ y no en $y = 1$.

Nivel G — Síntesis y demostraciones completas. Ejercicio G1. Reconstruir la ley del producto sin sustitución. Sean $x, y > 0$. Sin usar cambio de variable en integrales, demuestra $\ln(xy) = \ln x + \ln y$ siguiendo esta ruta:

1. fija y;
2. compara $F_y(t) = \ln(ty)$ con $G(t) = \ln t$;
3. calcula ambas derivadas;
4. trabaja en un intervalo compacto positivo que contenga a 1 y a x;
5. usa unicidad de primitivas hasta una constante;
6. determina la constante evaluando en $t = 1$.

Explica por qué cada paso evita circularidad.

Ejercicio G2. De la definición de a^x a la inversión global. Sea $a > 0$, $a \neq 1$.

1. Demuestra desde $a^x = \exp(x \ln a)$ que $\ln(a^x) = x \ln a$.
2. Demuestra que $x \mapsto a^x$ es inyectiva.
3. Para cada $u > 0$, construye explícitamente el único x tal que $a^x = u$.
4. Deduce $\log_a u = \frac{\ln u}{\ln a}$.

No cites Teorema 17.5 como caja negra.

Ejercicio G3. Demostrar desde cero la regla de exponente variable. Sea I un intervalo abierto y sean

$$u : I \to (0, \infty), \qquad v : I \to \mathbb{R}$$

diferenciables. Define $w(x) = u(x)^{v(x)}$.
Sin citar Proposición 17.5 como caja negra, demuestra que

$$w'(x) = u(x)^{v(x)} \left[v'(x) \ln u(x) + v(x) \frac{u'(x)}{u(x)} \right].$$

Señala exactamente dónde se usan la positividad de u, la regla del producto y la regla de la cadena.

Ejercicio G4. Reconstruir artanh **y su derivada.** Parte de $\tanh x = \frac{e^{2x}-1}{e^{2x}+1}$.
Para $y \in (-1, 1)$:

1. resuelve algebraicamente $\tanh x = y$ y obtiene
$$x = \frac{1}{2} \ln\left(\frac{1+y}{1-y} \right);$$
2. demuestra que la expresión logarítmica está bien definida;
3. concluye que $\tanh : \mathbb{R} \to (-1, 1)$ es sobreyectiva y, usando su monotonía, biyectiva;
4. deriva la inversa mediante Teorema 8.1 y demuestra $(\operatorname{artanh} y)' = \frac{1}{1-y^2}$.

No uses la fórmula de Corolario 17.5 como punto de partida.

Soluciones desarrolladas.

Soluciones del nivel A. **Solución A1.** Por Definición 17.1,

$$\ln 1 = \int_1^1 \frac{dt}{t} = 0.$$

Ésta es una consecuencia directa de la convención para una integral con extremos iguales. La función ln es estrictamente creciente por Proposición 17.2. Por tanto, si $0 < x < 1$, $\ln x < \ln 1 = 0$, mientras que si $x > 1$, $\ln x > \ln 1 = 0$. Así, la normalización procede de la **definición integral** y el signo procede de la **monotonía** ya demostrada.

Solución A2. La aditividad orientada Proposición 14.3 da, para $a, b > 0$,

$$\int_1^b \frac{dt}{t} = \int_1^a \frac{dt}{t} + \int_a^b \frac{dt}{t}.$$

Por Definición 17.1,

$$\ln b = \ln a + \int_a^b \frac{dt}{t},$$

y por tanto

$$\boxed{\int_a^b \frac{dt}{t} = \ln b - \ln a.}$$

La identidad ya incorpora el caso $b < a$, porque Proposición 14.3 está formulada para integrales orientadas. En efecto,

$$\int_a^b \frac{dt}{t} = -\int_b^a \frac{dt}{t} = -(\ln a - \ln b).$$

No se ha efectuado ningún cambio de variable.

Solución A3. Fijado $c > 0$, por ejemplo podemos elegir

$$a = \frac{1}{2}\,\mín\{1, c\} > 0, \qquad b = 1 + \máx\{1, c\}.$$

Entonces $1, c \in (a, b)$ y $t \mapsto 1/t$ es continua en $[a, b]$. Definimos

$$A(x) = \int_a^x \frac{dt}{t}.$$

Por Teorema 16.2, $A'(c) = \frac{1}{c}$.
La aditividad orientada produce

$$A(x) = \int_a^1 \frac{dt}{t} + \int_1^x \frac{dt}{t} = A(1) + \ln x.$$

Por tanto, $\ln x = A(x) - A(1)$ en ese intervalo, y $A(1)$ es constante. Luego

$$\boxed{(\ln)'(c) = A'(c) = \frac{1}{c}.}$$

El FTC de C17 está formulado sobre intervalos donde el integrando es Riemann integrable y continuo en el punto pertinente. Encerrar primero 1 y c en un intervalo compacto positivo verifica literalmente esas hipótesis; decir simplemente «aplicamos el FTC en $(0, \infty)$» ocultaría ese paso.

Solución A4. Como $\frac{81}{16} = \frac{3^4}{2^4}$,
Corolario 17.1 da

$$\ln\left(\frac{81}{16}\right) = 4\ln 3 - 4\ln 2 = \boxed{4(\ln 3 - \ln 2)}.$$

Además,

$$\sqrt[3]{\frac{9}{8}} = \left(\frac{9}{8}\right)^{1/3}.$$

Por la compatibilidad racional,

$$\begin{aligned}\ln\left(\sqrt[3]{\frac{9}{8}}\right) &= \frac{1}{3}\ln\left(\frac{9}{8}\right)\\ &= \frac{1}{3}(2\ln 3 - 3\ln 2)\\ &= \boxed{\frac{2}{3}\ln 3 - \ln 2}.\end{aligned}$$

Todas las potencias involucradas tienen base positiva, de modo que las identidades están bien tipadas.

Solución A5. Para $x = 1$, $\ln 1 = 0 = 1 - 1$.

Supongamos primero $x > 1$. El teorema del valor medio aplicado a ln en $[1, x]$ proporciona $c \in (1, x)$ tal que $\ln x - \ln 1 = \frac{1}{c}(x - 1)$.

Como $c > 1$, $0 < \frac{1}{c} < 1$, y como $x - 1 > 0$, $\ln x = \frac{x-1}{c} < x - 1$.

Si $0 < x < 1$, aplicamos el valor medio en $[x, 1]$. Existe $c \in (x, 1)$ tal que, escrito nuevamente entre 1 y x, $\ln x - \ln 1 = \frac{1}{c}(x-1)$. Ahora $c < 1$, por lo que $1/c > 1$, mientras $x - 1 < 0$. Multiplicar el número negativo $x - 1$ por un factor mayor que 1 lo hace más negativo: $\ln x = \frac{x-1}{c} < x - 1$.

Así,

$$\boxed{\ln x \le x - 1}$$

para todo $x > 0$, y la igualdad ocurre solo en $x = 1$.

Solución A6. La cota superior es A5. Para obtener la inferior aplicamos A5 al número positivo $1/x$:

$$\ln\left(\frac{1}{x}\right) \le \frac{1}{x} - 1.$$

Por Corolario 17.1, $-\ln x \le \frac{1}{x} - 1$. Multiplicando por -1 se invierte la desigualdad:

$$\boxed{\ln x \ge 1 - \frac{1}{x}.}$$

Junto con A5,

$$\boxed{1 - \frac{1}{x} \le \ln x \le x - 1.}$$

En A5 la igualdad ocurre únicamente cuando el argumento es 1. Para la cota inferior, el argumento aplicado fue $1/x$, así que también hay igualdad únicamente cuando $1/x = 1$, es decir, cuando

$$\boxed{x = 1}.$$

Soluciones del nivel B. **Solución B1.** Por Proposición 17.1,

$$\ln(xyz) = \ln((xy)z) = \ln(xy) + \ln z.$$

Aplicando una vez más la misma proposición,

$$\boxed{\ln(xyz) = \ln x + \ln y + \ln z.}$$

Para el cociente usamos Corolario 17.1:

$$\begin{aligned}\ln\left(\frac{xy}{z}\right) &= \ln(xy) - \ln z \\ &= \boxed{\ln x + \ln y - \ln z}.\end{aligned}$$

Todos los argumentos son positivos, como exigen las identidades.

Solución B2. Por la ley del cociente,

$$\ln 3 - \ln 4 = \ln\left(\frac{3}{4}\right).$$

La ecuación queda

$$\ln\left(\frac{x}{2}\right) = \ln\left(\frac{3}{4}\right).$$

Ambos argumentos son positivos. Como ln es inyectiva por Proposición 17.2, $\frac{x}{2} = \frac{3}{4}$.

Luego

$$\boxed{x = \frac{3}{2}}.$$

No hemos «cancelado» símbolos: hemos aplicado explícitamente la propiedad de inyectividad de una función.

Solución B3. Sabemos que $\ln 2 > 0$. Por la propiedad arquimediana Teorema 1.4 existe $n \in \mathbb{N}_{>0}$ tal que $n > \frac{|y|}{\ln 2}$.

Multiplicando por $\ln 2 > 0$, $n \ln 2 > |y|$.

Por tanto,

$$-n \ln 2 < -|y| \leq y \leq |y| < n \ln 2.$$

Usando Corolario 17.1,

$$\ln(2^{-n}) = -n \ln 2, \qquad \ln(2^n) = n \ln 2.$$

Así, $\ln(2^{-n}) < y < \ln(2^n)$.

La continuidad de ln y el teorema del valor intermedio Teorema 5.2 garantizan algún $x \in (2^{-n}, 2^n)$ con $\ln x = y$. La estricta monotonía de Proposición 17.2 garantiza que ese x es único. La propiedad arquimediana produce el intervalo suficientemente grande; la continuidad evita huecos en la imagen; la monotonía aporta unicidad.

Solución B4. La función racional está definida en $(0, \infty)$ y

$$f'(x) = \frac{(1+x) - x}{(1+x)^2} = \frac{1}{(1+x)^2} > 0.$$

Por Teorema 10.1, f es estrictamente creciente.

Además, para $x > 0$,

$$x > 0 \quad \Longrightarrow \quad f(x) > 0,$$

y como $x < 1 + x$ con denominador positivo, $\frac{x}{1+x} < 1$.

Luego $0 < f(x) < 1$.

Así, f es continua y estrictamente creciente, pero su imagen está contenida en $(0, 1)$ y no puede ser $\mathbb{R}$. Esto muestra por qué para ln fue necesario probar separadamente el rango completo en Teorema 17.2.

Solución B5. Como $-5 \in \mathbb{R}$, $\exp(-5) > 0$, de modo que puede aplicarse ln:

$$\boxed{\ln(\exp(-5)) = -5}.$$

Como $7 > 0$, $\ln 7$ está definido y la composición inversa opuesta da

$$\boxed{\exp(\ln 7) = 7}.$$

También $2/5 > 0$, así que

$$\boxed{\exp\left(\ln \frac{2}{5}\right) = \frac{2}{5}.}$$

En cambio, $\ln(-7)$ no es una expresión real definida en C18, porque Definición 17.1 tiene dominio $(0, \infty)$. Por tanto $\exp(\ln(-7))$ ni siquiera llega a formar una composición real bien tipada.

Solución B6. Primero probemos el límite en $+\infty$. Sea $M \in \mathbb{R}$. Si $M \leq 0$, como $\exp x > 0$, se tiene automáticamente $\exp x > M$ para todo x.

Si $M > 0$, $\ln M$ está definido. Tomemos $R = \ln M$.

Para $x > R$, el crecimiento estricto de exp da $\exp x > \exp(\ln M) = M$. Ésta es la definición de

$$\boxed{\lim_{x \to +\infty} \exp x = +\infty}.$$

Ahora sea $\varepsilon > 0$. Definimos $L = \ln \varepsilon$.

Si $x < L$, la monotonía de exp produce

$$0 < \exp x < \exp(\ln \varepsilon) = \varepsilon.$$

Así, para todo $\varepsilon > 0$, los valores de $\exp x$ quedan entre 0 y ε cuando x es suficientemente negativo. Por tanto,

$$\boxed{\lim_{x\to-\infty} \exp x = 0}.$$

Solo hemos usado inversión y monotonía.

Soluciones del nivel C. **Solución C1.** Si $n \geq 1$, iterando Corolario 17.2, $e^{nx} = e^{x+\cdots+x} = (e^x)^n$.

Para $n = 0$,

$$e^{0x} = e^0 = 1 = (e^x)^0.$$

Si $n < 0$, escribimos $n = -m$ con $m > 0$. Entonces

$$e^{nx} = e^{-mx} = \frac{1}{e^{mx}} = \frac{1}{(e^x)^m} = (e^x)^{-m} = (e^x)^n.$$

Por tanto,

$$\boxed{e^{nx} = (e^x)^n}$$

para todo $n \in \mathbb{Z}$.

Solución C2. Por la ley exponencial, $e^{\ln 5 - \ln 2} = \frac{e^{\ln 5}}{e^{\ln 2}}$.

Las identidades inversas dan

$$e^{\ln 5} = 5, \qquad e^{\ln 2} = 2.$$

Luego

$$\boxed{e^{\ln 5 - \ln 2} = \frac{5}{2}.}$$

Solución C3. Como $7 > 0$, podemos aplicar ln a ambos miembros: $\ln(e^{2x-1}) = \ln 7$.

Por la identidad inversa, $2x - 1 = \ln 7$.

Por tanto,

$$\boxed{x = \frac{1 + \ln 7}{2}.}$$

La función $x \mapsto e^{2x-1}$ es estrictamente creciente porque $2x - 1$ lo es y exp también. En consecuencia, la solución es única. Equivalentemente, la aplicación de la inyectividad de ln ya impide dos soluciones distintas.

Solución C4. Como $3/2 \in \mathbb{Q}$, Corolario 17.2 garantiza que el nuevo $e^{3/2}$ coincide con la potencia racional de C09. Por las leyes racionales,

$$e^{3/2} = e^{1+1/2} = e\, e^{1/2} = e\sqrt{e}.$$

Así,

$$\boxed{e^{3/2} = e\sqrt{e}}.$$

Asimismo, $e^{2/3} > 0$ y la potencia cúbica es una potencia entera. Por C1,

$$\left(e^{2/3}\right)^3 = e^{3(2/3)} = \boxed{e^2}.$$

La compatibilidad con C09 es lo que asegura que símbolos como $\sqrt{e}$ y $e^{1/2}$ nombran exactamente el mismo número positivo.

Solución C5. A5 afirma que para todo $t > 0$, $\ln t \le t - 1$, con igualdad solo en $t = 1$.
Tomamos $t = e^x > 0$.
Entonces $\ln(e^x) \le e^x - 1$.
Como $\ln(e^x) = x$, $x \le e^x - 1$, es decir,

$$\boxed{e^x \ge 1 + x}.$$

La igualdad ocurre exactamente cuando $e^x = 1$. Por la inyectividad de exp y $\exp 0 = 1$,

$$\boxed{x = 0}.$$

Solución C6. Por linealidad y Teorema 17.3,

$$h'(x) = 3(e^x)' = 3e^x = h(x).$$

Sin embargo, $h(0) = 3e^0 = 3$, mientras que $\exp(0) = 1$.

Por tanto h y exp son funciones distintas aunque ambas satisfacen la identidad diferencial $y' = y$. Esto demuestra directamente que la ecuación diferencial, tomada sin una normalización adicional, no caracteriza por sí sola a la exponencial de C18. En este capítulo exp fue construida primero como $\ln^{-1}$; la igualdad $\exp' = \exp$ apareció después como teorema.

Soluciones del nivel D. **Solución D1.** En C09, $27^{1/3} = 3$, y por tanto

$$27^{2/3} = (27^{1/3})^2 = 3^2 = 9.$$

Con la definición de C18,

$$27^{2/3} = \exp\left(\frac{2}{3}\ln 27\right).$$

Como $27 = 3^3$, Corolario 17.1 da $\ln 27 = 3\ln 3$.
Luego $27^{2/3} = \exp(2\ln 3)$.
Por la ley exponencial,

$$\exp(2\ln 3) = (\exp(\ln 3))^2 = 3^2 = 9.$$

Así ambos caminos dan

$$\boxed{27^{2/3} = 9}.$$

No es coincidencia: Proposición 17.4 demuestra que la definición real nueva prolonga exactamente las potencias racionales positivas de C09.

Solución D2. Por Definición 17.4, $(ab)^x = \exp(x \ln(ab))$.
La ecuación funcional Proposición 17.1 da $\ln(ab) = \ln a + \ln b$.
Por tanto,

$$\begin{aligned}(ab)^x &= \exp(x \ln a + x \ln b) \\ &= \exp(x \ln a) \exp(x \ln b) \\ &= \boxed{a^x b^x}.\end{aligned}$$

Para la segunda ley, $(a^x)^y = \exp(y \ln(a^x))$.
Como $\ln(a^x) = x \ln a$, obtenemos

$$(a^x)^y = \exp(yx \ln a) = \boxed{a^{xy}}.$$

La ecuación funcional de ln se usa explícitamente en la ley del producto de bases; la segunda ley usa la identidad $\ln(a^x) = x \ln a$, que a su vez procede de la definición y de la inversión entre ln y exp.

Solución D3. Como $2 > 1$, Proposición 17.4 implica que $f(x) = 2^x$ es estrictamente creciente. Como $0 < \frac{1}{2} < 1$, la función

$$g(x) = \left(\frac{1}{2}\right)^x$$

es estrictamente decreciente.
Además, $\frac{1}{2} = 2^{-1}$, y las leyes de potencias reales dan

$$g(x) = (2^{-1})^x = 2^{-x} = f(-x).$$

Como $2^0 = 1$ y f es creciente,

$$2^x > 1 \quad \Longleftrightarrow \quad x > 0.$$

Como g es decreciente y $g(0) = 1$,

$$\left(\frac{1}{2}\right)^x > 1 \quad \Longleftrightarrow \quad x < 0.$$

Solución D4. Ambos miembros son positivos. Aplicamos ln: $\ln(9^x) = \ln 27$.
Por Proposición 17.4, $x \ln 9 = \ln 27$.
Como

$$\ln 9 = 2 \ln 3, \qquad \ln 27 = 3 \ln 3,$$

y $\ln 3 > 0$, podemos dividir por $\ln 3$: $2x = 3$.

Por tanto,

$$\boxed{x = \frac{3}{2}}.$$

La función $x \mapsto 9^x$ es estrictamente creciente porque $9 > 1$, así que no puede tomar el valor 27 en dos puntos distintos. La solución es única.

Solución D5. Por cambio de base, $\log_4 32 = \frac{\ln 32}{\ln 4}$.
Como $32 = 2^5$ y $4 = 2^2$,

$$\log_4 32 = \frac{5\ln 2}{2\ln 2} = \boxed{\frac{5}{2}}.$$

Para la segunda expresión, $\log_{1/3} 27 = \frac{\ln 27}{\ln(1/3)}$.
Ahora

$$\ln 27 = 3\ln 3, \qquad \ln(1/3) = -\ln 3,$$

de modo que

$$\boxed{\log_{1/3} 27 = -3.}$$

El signo negativo concuerda con la geometría: la base $1/3$ está en $(0, 1)$, por lo que su logaritmo es una función estrictamente decreciente; un argumento $27 > 1$ debe corresponder a un exponente negativo.

Solución D6. C09 permite raíces cúbicas reales de números negativos. Como $(-2)^3 = -8$, tenemos

$$\boxed{(-8)^{1/3} = -2}.$$

Luego

$$(-8)^{2/3} = \left((-8)^{1/3}\right)^2 = (-2)^2 = \boxed{4}.$$

Esto no contradice Definición 17.4, porque esa definición no pretende reemplazar las potencias racionales de base negativa: construye una potencia para **todo exponente real** solo cuando la base es positiva. En particular, C18 define $a^x = \exp(x \ln a)$ únicamente para $a > 0$. Como $\ln(-8)$ no existe como número real en esta teoría,

$$\boxed{(-8)^{\sqrt{2}}}$$

no ha sido definido en C18.

Soluciones del nivel E. **Solución E1.** Aplicamos Teorema 17.6 con base constante $a = 3$ y la regla de la cadena Teorema 7.4 al argumento $u(x) = x^2$, diferenciable en toda $\mathbb{R}$, incluido $x = 0$: $\frac{d}{dx}3^{x^2} = (\ln 3)3^{x^2}(2x)$.

Por tanto,

$$\boxed{f'(x) = 2x(\ln 3)3^{x^2}.}$$

Como $3 > 0$, la función está definida para todo exponente real x^2, y el dominio es toda $\mathbb{R}$.

Solución E2. Para todo $x \in \mathbb{R}$, $x^2 + 1 > 0$.

Además, la base 5 satisface $5 > 0$ y $5 \neq 1$. Por tanto $\log_5(x^2 + 1)$ está definida para todo real. Aplicando Corolario 17.4, $g'(x) = \frac{2x}{(x^2+1)\ln 5}$.

Así,

$$\boxed{g'(x) = \frac{2x}{(x^2 + 1)\ln 5} \qquad (x \in \mathbb{R}).}$$

Solución E3. La construcción uniforme de Definición 17.4 aplicada a una base variable exige base positiva. Por eso el dominio canónico es

$$\boxed{x > 0}.$$

Como $\sqrt{2} \in \mathbb{R}$, Teorema 17.6 da

$$\boxed{h'(x) = \sqrt{2}\, x^{\sqrt{2}-1} \qquad (x > 0).}$$

No se ha demostrado una fórmula sobre $x < 0$ porque C18 no ha definido una potencia real general x^α con base negativa y exponente real arbitrario. El hecho de que algunos exponentes racionales admitan otros dominios pertenece a C09 y no extiende automáticamente este caso irracional.

Solución E4. Para todo x, $x^2 + 1 > 0$.

Por tanto la base de la potencia es positiva en toda $\mathbb{R}$ y Definición 17.6 define F globalmente.

Tomamos

$$u(x) = x^2 + 1, \qquad v(x) = x.$$

Entonces

$$u'(x) = 2x, \qquad v'(x) = 1.$$

Por Proposición 17.5,

$$\begin{aligned} F'(x) &= (x^2 + 1)^x \left[\ln(x^2 + 1) + x\frac{2x}{x^2 + 1}\right] \\ &= \boxed{(x^2 + 1)^x \left[\ln(x^2 + 1) + \frac{2x^2}{x^2 + 1}\right]}. \end{aligned}$$

Solución E5. En $x > 0$ tomamos

$$u(x) = x, \qquad v(x) = x^2.$$

Entonces

$$u'(x) = 1, \qquad v'(x) = 2x.$$

Por Proposición 17.5,

$$\begin{aligned} G'(x) &= x^{x^2}\left[2x\ln x + x^2\frac{1}{x}\right] \\ &= \boxed{x^{x^2}(2x\ln x + x)}. \end{aligned}$$

También puede factorizarse como $G'(x) = x^{x^2+1}(2\ln x + 1)$, pero la primera forma muestra directamente los dos términos procedentes de variar el exponente y la base.

Solución E6. La función está definida cuando $|x^2 - 1| > 0$, es decir, cuando $x^2 - 1 \neq 0$. Por tanto,

$$\mathrm{Dom}(H) = \boxed{\mathbb{R} \setminus \{-1, 1\}}.$$

La función $\phi(t) = \ln|t|$ es diferenciable para $t \neq 0$ y, por Corolario 17.4, $\phi'(t) = \frac{1}{t}$. Aplicando la regla de la cadena a $t = x^2 - 1$,

$$\boxed{H'(x) = \frac{2x}{x^2 - 1} \qquad (x \neq \pm 1).}$$

No podemos escribir globalmente $H = \ln(x^2 - 1)$ porque para $|x| < 1$ se tiene $x^2 - 1 < 0$, y $\ln(x^2 - 1)$ no está definido como logaritmo real. El valor absoluto mantiene las tres componentes del dominio original: $(-\infty, -1)$, $(-1, 1)$ y $(1, \infty)$.

Soluciones del nivel F. **Solución F1.** Por Definición 17.7,

$$\cosh x = \frac{e^x + e^{-x}}{2}, \qquad \sinh x = \frac{e^x - e^{-x}}{2}.$$

Entonces

$$\begin{aligned} \cosh^2 x - \sinh^2 x &= \frac{(e^x + e^{-x})^2 - (e^x - e^{-x})^2}{4} \\ &= \frac{4e^x e^{-x}}{4} \\ &= 1. \end{aligned}$$

Por tanto,

$$\boxed{\cosh^2 x = 1 + \sinh^2 x \geq 1}.$$

En particular $\cosh^2 x > 0$, así que

$$\boxed{\cosh x \neq 0}$$

para todo x. De hecho Definición 17.7 muestra incluso $\cosh x > 0$.

Solución F2. Como

$$e^{\ln 3} = 3, \qquad e^{-\ln 3} = \frac{1}{3},$$

obtenemos

$$\sinh(\ln 3) = \frac{3 - \frac{1}{3}}{2} = \boxed{\frac{4}{3}},$$

$$\cosh(\ln 3) = \frac{3 + \frac{1}{3}}{2} = \boxed{\frac{5}{3}},$$

y

$$\tanh(\ln 3) = \frac{4/3}{5/3} = \boxed{\frac{4}{5}}.$$

La identidad fundamental se verifica:

$$\left(\frac{5}{3}\right)^2 - \left(\frac{4}{3}\right)^2 = \frac{25 - 16}{9} = 1.$$

Solución F3. Por F2, $\sinh(\ln 3) = \frac{4}{3}$.
Como $\sinh : \mathbb{R} \to \mathbb{R}$ es biyectiva,

$$\boxed{\operatorname{arsinh}\left(\frac{4}{3}\right) = \ln 3.}$$

Para la segunda ruta usamos Corolario 17.5:

$$\begin{aligned}\operatorname{arsinh}\left(\frac{4}{3}\right) &= \ln\left(\frac{4}{3} + \sqrt{\frac{16}{9} + 1}\right) \\ &= \ln\left(\frac{4}{3} + \frac{5}{3}\right) \\ &= \ln 3.\end{aligned}$$

El argumento del logaritmo es $3 > 0$, como requiere su dominio.

Solución F4. Como $-1 < \frac{1}{2} < 1$, el número $1/2$ pertenece al dominio de artanh.
Por Corolario 17.5,

$$\begin{aligned}\operatorname{artanh}\left(\frac{1}{2}\right) &= \frac{1}{2}\ln\left(\frac{1 + \frac{1}{2}}{1 - \frac{1}{2}}\right) \\ &= \frac{1}{2}\ln\left(\frac{3/2}{1/2}\right) \\ &= \boxed{\frac{1}{2}\ln 3}.\end{aligned}$$

Solución F5. Como $4 > 1$, $\ln 4 > 0$, de modo que $\ln 4$ pertenece a la rama $[0, \infty)$ sobre la que se invierte cosh.

Además,

$$e^{\ln 4} = 4, \qquad e^{-\ln 4} = \frac{1}{4}.$$

Entonces

$$\cosh(\ln 4) = \frac{4 + \frac{1}{4}}{2} = \frac{17/4}{2} = \frac{17}{8}.$$

Como la restricción de cosh a $[0, \infty)$ es inyectiva,

$$\boxed{\operatorname{arcosh}\left(\frac{17}{8}\right) = \ln 4.}$$

Solución F6. Por paridad, $\cosh(-x) = \cosh x$.

Si $x \neq 0$, entonces $x \neq -x$ pero ambas entradas tienen la misma imagen. Por tanto cosh no es inyectiva en $\mathbb{R}$ y no puede poseer allí una inversa funcional global.

En cambio, Teorema 17.7 demuestra que cosh es estrictamente creciente en $[0, \infty)$ y que su imagen allí es $[1, \infty)$. Por eso $\cosh : [0, \infty) \to [1, \infty)$ sí es una biyección y puede invertirse.

Para $y > 1$, si $x = \operatorname{arcosh} y$, entonces $x > 0$ y $\cosh'(x) = \sinh x > 0$.

El teorema de la derivada de la inversa Teorema 8.1 se aplica y produce

$$(\operatorname{arcosh} y)' = \frac{1}{\sinh x} = \frac{1}{\sqrt{y^2 - 1}}.$$

En $y = 1$, el antecedente es $x = 0$ y $\cosh'(0) = \sinh 0 = 0$.

Falla precisamente la hipótesis de derivada no nula de Teorema 8.1, y además la fórmula escrita tendría denominador cero. Por eso el enunciado diferencial se restringe a

$$\boxed{y > 1}.$$

Soluciones del nivel G. **Solución G1.** Fijemos $y > 0$ y definamos, para $t > 0$,

$$F_y(t) = \ln(ty), \qquad G(t) = \ln t.$$

Como $ty > 0$, ambas funciones están bien definidas. Por Teorema 17.1 y la regla de la cadena Teorema 7.4, $F_y'(t) = \frac{1}{ty}\, y = \frac{1}{t}$.

También $G'(t) = \frac{1}{t}$.

Ahora fijemos además $x > 0$. Elegimos números

$$0 < a < \min\{1, x\}, \qquad b > \max\{1, x\}.$$

Ambas funciones son diferenciables en el abierto $(0, \infty)$ que contiene a $[a, b]$; sus derivadas relativas en los extremos coinciden con las ordinarias. Por tanto, sus restricciones a $[a, b]$ son primitivas de la misma función $1/t$. Por Proposición 16.2 existe una constante C tal que

$$F_y(t) = G(t) + C \qquad (t \in [a, b]).$$

Evaluando en $t = 1$,

$$C = F_y(1) - G(1) = \ln y - \ln 1 = \ln y.$$

Evaluando ahora en $t = x$,

$$\boxed{\ln(xy) = \ln x + \ln y.}$$

La prueba no usa sustitución integral. La definición integral de ln ya había dado su derivada mediante el FTC; la regla de la cadena compara funciones; la unicidad de primitivas controla la diferencia; la normalización $\ln 1 = 0$ determina la constante. Ningún paso presupone la propia ley del producto.

Solución G2. Por definición, $a^x = \exp(x \ln a)$.

Aplicando ln y usando la identidad inversa,

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

Como $a \neq 1$, Proposición 17.2 implica $\ln a \neq 0$.

Si $a^{x_1} = a^{x_2}$, aplicamos ln: $x_1 \ln a = x_2 \ln a$.

Dividiendo por $\ln a \neq 0$, $x_1 = x_2$.

Así, $x \mapsto a^x$ es inyectiva.

Sea ahora $u > 0$. Buscamos x tal que $a^x = u$. Aplicar ln obliga a $x \ln a = \ln u$, y por tanto el único candidato es

$$\boxed{x = \frac{\ln u}{\ln a}.}$$

Este candidato funciona:

$$\begin{aligned} a^{\ln u/\ln a} &= \exp\left(\frac{\ln u}{\ln a} \ln a\right) \\ &= \exp(\ln u) \\ &= u. \end{aligned}$$

Por inyectividad es el único antecedente. Por definición de función inversa,

$$\boxed{\log_a u = \frac{\ln u}{\ln a}.}$$

Solución G3. La hipótesis $u(x) > 0$ para todo $x \in I$ garantiza que $\ln u(x)$ está definido y que podemos escribir, por Definición 17.6, $w(x) = u(x)^{v(x)} = \exp(h(x))$, donde $h(x) = v(x) \ln u(x)$.

Como u es positiva y diferenciable, Corolario 17.4 da

$$(\ln u(x))' = \frac{u'(x)}{u(x)}.$$

Aplicamos la regla del producto Teorema 7.2 a h:

$$h'(x) = v'(x) \ln u(x) + v(x) \frac{u'(x)}{u(x)}.$$

Finalmente, la regla de la cadena Teorema 7.4 y Teorema 17.3 producen

$$\begin{aligned} w'(x) &= \exp(h(x))h'(x) \\ &= u(x)^{v(x)}\left[v'(x)\ln u(x) + v(x)\frac{u'(x)}{u(x)}\right]. \end{aligned}$$

Por tanto,

$$\boxed{w'(x) = u(x)^{v(x)}\left[v'(x)\ln u(x) + v(x)\frac{u'(x)}{u(x)}\right].}$$

La positividad hace posible la representación logarítmica; el producto diferencia $v\ln u$; la cadena diferencia primero $\ln\circ u$ y después $\exp\circ h$.

Solución G4. Sea $y \in (-1,1)$ y supongamos $y = \tanh x = \frac{e^{2x}-1}{e^{2x}+1}$.
Como $e^{2x}+1>0$, podemos multiplicar: $y(e^{2x}+1) = e^{2x}-1$.
Reordenando, $(1-y)e^{2x} = 1+y$.
Como $y \in (-1,1)$,

$$1-y>0, \qquad 1+y>0.$$

Por tanto, $e^{2x} = \frac{1+y}{1-y} > 0$.
Aplicando ln,

$$2x = \ln\left(\frac{1+y}{1-y}\right),$$

y así

$$\boxed{x = \frac{1}{2}\ln\left(\frac{1+y}{1-y}\right).}$$

La positividad de $1\pm y$ demuestra que la expresión logarítmica está bien definida. Para comprobar que el candidato funciona, fijemos $y \in (-1,1)$ y pongamos $r = (1+y)/(1-y) > 0$ y $x = \frac{1}{2}\ln r$. Entonces $e^{2x} = e^{\ln r} = r$ y

$$\tanh x = \frac{r-1}{r+1} = \frac{2y/(1-y)}{2/(1-y)} = y.$$

Hemos construido y verificado un antecedente real para cada $y \in (-1,1)$, de modo que tanh es sobreyectiva sobre ese intervalo. Por Teorema 17.7, tanh es estrictamente creciente, luego inyectiva. Por tanto, $\tanh : \mathbb{R} \to (-1,1)$ es biyectiva.

Sea ahora $x = \operatorname{artanh} y$.
Entonces $y = \tanh x$ y, por Teorema 17.7, $\tanh'(x) = 1-\tanh^2 x = 1-y^2$.
Como $|y|<1$, $1-y^2>0$.
Se cumplen las hipótesis de Teorema 8.1, así que

$$\boxed{(\operatorname{artanh} y)' = \frac{1}{1-y^2}.}$$

La fórmula logarítmica y la derivada de la inversa han sido reconstruidas sin usar Corolario 17.5 como premisa.

Capítulo 18

Técnicas de integración

18.1. Del teorema fundamental al problema inverso

Las funciones circulares, sus identidades de adición, las derivadas y las ramas inversas utilizadas aquí fueron construidas en el laboratorio de Sección 17.10. Las técnicas de este capítulo se apoyan en esa construcción; cada sustitución conservará su dominio y su elección de rama. El teorema fundamental del cálculo cambió el problema de evaluar una integral definida. Si conocemos una función F cuya derivada es f, entonces Teorema 16.3 permite escribir

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

La dificultad se desplaza así hacia una pregunta anterior:

> **¿cómo encontramos una primitiva de f cuando no está inmediatamente a la vista?**

Hasta ahora hemos trabajado sobre todo en la dirección

$$\boxed{F \longmapsto F'}$$

porque derivar una función conocida es un problema gobernado por reglas directas. Sabemos derivar sumas, productos, cocientes y composiciones; C18 añadió además un repertorio amplio de funciones elementales —logaritmos, exponenciales, potencias reales e hiperbólicas— cuyas derivadas ya forman parte de nuestra teoría. En este capítulo recorreremos el camino contrario:

$$\boxed{f \longmapsto \text{buscar } F \text{ tal que } F' = f.}$$

La derivación parte de una expresión conocida y aplica reglas a su estructura. La búsqueda de una primitiva exige reconocer qué estructura puede haber producido el integrando y comprobar la propuesta mediante derivación.

Tres preguntas que no deben confundirse. Antes de aprender ninguna técnica conviene separar tres problemas distintos. El primero es la existencia de una integral definida. Si f es Riemann integrable en $[a, b]$, entonces

$$\int_a^b f(x)\,dx$$

ya está definido como un número real, independientemente de que sepamos calcularlo mediante una fórmula cerrada. El segundo es la existencia de primitivas. Para una función continua, C17 mostró que la función de acumulación

$$A_f(x) = \int_a^x f(t)\,dt$$

proporciona una primitiva. Por tanto, en ese caso no estamos intentando demostrar que alguna primitiva exista: ya sabemos construir una. El tercer problema es diferente:

> ¿podemos expresar una primitiva mediante funciones que ya conocemos y manipularla de una forma útil para evaluar integrales concretas?

Éste es el problema propio de las **técnicas de integración**.

Existir no es lo mismo que poder escribir una fórmula conocida
Para una función continua f, la función

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

es una primitiva rigurosamente construida. Sin embargo, esa expresión puede no simplificarse de inmediato en términos de las funciones elementales que ya conocemos. Las técnicas de integración no crean la existencia de la primitiva: intentan **reconocerla, reconstruirla o transformarla** en una forma más explícita.

Esta distinción evita un error frecuente. Cuando una integral resulta difícil de calcular, no debemos concluir que «la integral no existe». Lo único que sabemos es que todavía no hemos encontrado una representación conveniente.

Un caso en el que no hay misterio. Consideremos $f(x) = 3x^2 - 2x + 1$.

Las reglas de derivación de potencias permiten reconocer término a término que $F(x) = x^3 - x^2 + x$ satisface $F'(x) = 3x^2 - 2x + 1$. Aquí el problema inverso es casi transparente: cada término del integrando conserva la forma de una derivada elemental conocida.

Podemos escribir, por ejemplo,

$$\int_0^2 (3x^2 - 2x + 1)\,dx = \left[x^3 - x^2 + x\right]_0^2 = 8 - 4 + 2 = 6.$$

No necesitamos todavía una «técnica» especial. Basta leer el integrando como una suma de derivadas reconocibles. La situación cambia en cuanto las reglas de derivación mezclan varias estructuras.

La huella de una composición. Examinemos ahora $2xe^{x^2}$.

La presencia de e^{x^2} sugiere mirar la derivada de una composición. Por la regla de la cadena, $\frac{d}{dx}e^{x^2} = e^{x^2} \cdot 2x$.

Por tanto,

$$\boxed{2xe^{x^2} = \frac{d}{dx}e^{x^2}.}$$

El factor $2x$ no es un adorno: es exactamente la derivada de la función interior x^2. La estructura completa del integrando reproduce la huella dejada por la regla de la cadena.

Así,

$$\int 2xe^{x^2}\,dx = e^{x^2} + C.$$

Compárese con e^{x^2}.

Las dos expresiones contienen la misma exponencial compuesta, pero en la segunda ha desaparecido el factor que la regla de la cadena produciría al derivar. La semejanza superficial no basta para aplicar el mismo razonamiento. Este contraste anticipa una idea central del capítulo:

antes de manipular un integrando, hay que identificar qué regla de derivación podría haberlo producido.

La huella de un producto. Consideremos ahora xe^x.

No es la derivada inmediata de xe^x, porque la regla del producto da $\frac{d}{dx}(xe^x) = e^x + xe^x$. Pero esta identidad puede leerse al revés. Despejando el término que nos interesa, $xe^x = \frac{d}{dx}(xe^x) - e^x$. Como e^x tiene una primitiva conocida, obtenemos

$$\int xe^x\,dx = xe^x - e^x + C.$$

La idea no consistió en adivinar la respuesta. Partimos de una regla de derivación ya conocida, reconocimos que el integrando aparecía como **una parte** de esa regla y reorganizamos la identidad hasta aislarlo. La fórmula general se apoyará en el mecanismo que acabamos de utilizar:

regla del producto $\longrightarrow$ identidad reorganizada $\longrightarrow$ primitiva.

Integrar exige elegir una representación. Un mismo integrando puede admitir varias formas algebraicamente equivalentes, pero no todas son igual de útiles. Por ejemplo, $\frac{x^2-1}{x-1}$ parece una función racional que requeriría alguna técnica especial. Sin embargo, en todo punto de su dominio tenemos $x^2 - 1 = (x-1)(x+1)$, y por tanto, para $x \neq 1$, $\frac{x^2-1}{x-1} = x + 1$. Después de simplificar la estructura algebraica, una primitiva es inmediata: $\frac{x^2}{2} + x$. La integración no empieza necesariamente aplicando una fórmula de integración. Con frecuencia empieza **reescribiendo el integrando**. Esto conecta el capítulo con una habilidad que ya apareció repetidamente en álgebra y cálculo diferencial: antes de calcular, hay que decidir qué forma de una expresión deja visible la estructura relevante.

Las técnicas como reglas de derivación leídas al revés. Los dos ejemplos anteriores sugieren la arquitectura básica de las técnicas que estudiaremos. La regla de la cadena tiene la forma $\frac{d}{dx}F(g(x)) = F'(g(x))g'(x)$. Cuando un integrando contiene simultáneamente una función compuesta y la derivada de su parte interior, aparece una firma estructural que podremos explotar sistemáticamente. La regla del producto tiene la forma

$$\frac{d}{dx}(u(x)v(x)) = u'(x)v(x) + u(x)v'(x).$$

Cuando el integrando contiene uno de esos productos pero no toda la derivada, podremos reorganizar la identidad y transferir parte del problema a otra integral. No derivaremos todavía aquí las fórmulas generales. Lo importante es ver que no aparecerán como recetas independientes de la teoría previa:

regla de la cadena $\longleftrightarrow$ cambio de variable, regla del producto $\longleftrightarrow$ integración por partes.

El sentido de las flechas será precisado y demostrado en los microtramos siguientes.

No existe un botón llamado «integrar». Derivar una expresión elemental es, en gran medida, un procedimiento dirigido por la sintaxis: identificamos las operaciones que la construyen y aplicamos reglas conocidas desde fuera hacia dentro. El problema inverso no tiene esa misma forma. Dado un integrando, puede ser necesario:

- simplificarlo antes de hacer nada;
- reconocer una composición oculta;
- fabricar el diferencial que falta mediante una constante;
- reconstruir un producto;
- separar una expresión en partes más manejables;
- elegir entre varias transformaciones posibles;
- abandonar una ruta que vuelve el problema más difícil.

Por eso las técnicas de integración no deben aprenderse como una lista de órdenes del tipo «si ves esto, haz aquello». Una coincidencia visual puede ser accidental. Lo que necesitamos identificar es una relación matemática precisa entre el integrando y alguna regla de derivación ya demostrada.

Una técnica no es un algoritmo universal
Este capítulo desarrollará procedimientos potentes para clases amplias de integrales, pero no afirmará que toda función continua pueda reducirse mediante un número finito de manipulaciones a una fórmula elemental familiar. El teorema fundamental garantiza una primitiva mediante acumulación; otra cuestión es si esa primitiva admite una expresión elemental conveniente.

Una convención de lectura. Cuando escribamos

$$\int f(x)\,dx = F(x) + C,$$

la igualdad significará que $F'(x) = f(x)$ en el dominio considerado y que, en un intervalo, todas las primitivas difieren de F por una constante, conforme a Proposición 16.2. La proposición Proposición 16.2 fue formulada en C17 para un intervalo compacto $[a, b]$, pero su conclusión se extiende inmediatamente a cualquier intervalo I. En efecto, si P y Q son primitivas de la misma función en I y $x < y$ son dos puntos de I, entonces $[x, y] \subseteq I$; al restringir P y Q a ese segmento, Proposición 16.2 muestra que $P - Q$ tiene el mismo valor en x y en y. Como los puntos eran arbitrarios, $P - Q$ es constante en todo I. Usaremos esta extensión elemental sin repetir el argumento en cada integral indefinida.

Esta notación de **integral indefinida** no sustituye la definición riemanniana de integral definida. Son objetos conceptualmente distintos:

$$\int_a^b f(x)\,dx \quad \text{es un número,}$$
$$\int f(x)\,dx \quad \text{representa una familia de primitivas.}$$

Newton–Leibniz es precisamente el puente que permite usar la segunda para calcular la primera cuando disponemos de una primitiva adecuada.

La pregunta que guiará el capítulo. A partir de ahora, frente a un integrando no preguntaremos primero

«¿qué fórmula de integración debo aplicar?»

sino

«¿qué estructura de derivación está escondida aquí, y qué transformación la vuelve visible?»

Ese cambio de pregunta es la diferencia entre memorizar técnicas y comprenderlas.

Antes de seguir
Sin efectuar todavía ninguna sustitución formal, explica por qué $6x(3x^2+1)^4$ tiene la estructura de una regla de la cadena invertida. Identifica la función interior, su derivada y una función exterior cuya derivada produzca la potencia que aparece.
Después compara con $(3x^2+1)^4$.
¿Qué pieza estructural desapareció? ¿Por qué no basta con decir que «se parecen» para aplicar el mismo razonamiento?

18.2. Cambio de variable: invertir la regla de la cadena

El último ejercicio de §18.1 dejó visible la estructura que queremos formalizar. En $6x(3x^2+1)^4$ aparecen simultáneamente una función compuesta, $(3x^2+1)^4$, y la derivada de su parte interior, $\frac{d}{dx}(3x^2+1) = 6x$. No se trata de una coincidencia tipográfica. Si definimos $g(x) = 3x^2+1$ y $H(u) = \frac{u^5}{5}$, entonces $H'(u) = u^4$. La regla de la cadena Teorema 7.4 produce

$$\begin{aligned}\frac{d}{dx}H(g(x)) &= H'(g(x))g'(x)\\ &= (3x^2+1)^4\,6x.\end{aligned}$$

Por tanto,

$$\boxed{\int 6x(3x^2+1)^4\,dx = \frac{(3x^2+1)^5}{5} + C.}$$

El procedimiento que acabamos de ejecutar será llamado **cambio de variable** o **sustitución**. Su fundamento no es una nueva regla misteriosa de integración: es la regla de la cadena leída en sentido inverso.

La identidad básica. La regla de la cadena afirma que, si $F' = f$, entonces $\frac{d}{dx}F(g(x)) = f(g(x))g'(x)$. Leída desde el integrando hacia la primitiva, esta misma igualdad dice que

$$\boxed{\int f(g(x))g'(x)\,dx = F(g(x)) + C.}$$

Para aplicar sustitución debemos reconocer una composición y el factor diferencial de su función interior, aunque una transformación algebraica sea necesaria para hacerlos visibles. En una integral definida, el cambio de variable debe transformar también los extremos. La fórmula correspondiente precisará cómo se relacionan el intervalo original y los valores de la función de sustitución.

Teorema 18.1. **Teorema (cambio de variable en una dimensión).** Sean $I, J \subseteq \mathbb{R}$ intervalos.

Forma para primitivas. Supongamos que

$$g : I \to J$$

es diferenciable y que $F : J \to \mathbb{R}$ es diferenciable con

$$F'(u) = f(u) \qquad (u \in J).$$

Entonces $F \circ g$ es una primitiva de

$$x \longmapsto f(g(x))g'(x)$$

en I. En consecuencia,

$$\boxed{\int f(g(x))g'(x)\,dx = F(g(x)) + C.}$$

Forma para integrales definidas. Sean $a < b$, supongamos que $g : [a, b] \to J$ es continuamente diferenciable y que $f : J \to \mathbb{R}$ es continua. Entonces

$$\boxed{\int_a^b f(g(x))g'(x)\,dx = \int_{g(a)}^{g(b)} f(u)\,du.}$$

Demostración. La primera parte es exactamente la regla de la cadena. Como $F'(u) = f(u)$, Teorema 7.4 da

$$\begin{aligned}(F \circ g)'(x) &= F'(g(x))g'(x)\\ &= f(g(x))g'(x).\end{aligned}$$

Por tanto, $F \circ g$ es una primitiva del integrando. La extensión de Proposición 16.2 justificada en §18.1 garantiza que todas las primitivas difieren de ella por una constante en el intervalo I. Esto prueba la forma indefinida. Para la forma definida, como g es continuamente diferenciable, es continua en $[a, b]$. Su imagen es compacta; sean

$$m := \min_{x\in[a,b]} g(x), \qquad M := \max_{x\in[a,b]} g(x).$$

Como J es un intervalo y $g([a, b]) \subseteq J$, el segmento $[m, M]$ está contenido en J. Si $m = M$, entonces g es constante, $g' = 0$ y ambos miembros de la identidad definida son 0. Supongamos, pues, $m < M$.

Definamos, para $y \in [m, M]$,

$$B(y) = \int_m^y f(u)\,du.$$

La restricción de f a $[m, M]$ es continua. Por Teorema 16.2,

$$B'_{[m,M]}(y) = f(y) \qquad (y \in [m, M]).$$

Como $g([a, b]) \subseteq [m, M]$, la regla de la cadena Teorema 7.4 aplicada a $B \circ g$ da $\frac{d}{dx}B(g(x)) = f(g(x))g'(x)$. El miembro derecho es continuo en $[a, b]$, de modo que Newton–Leibniz Teorema 16.3 permite escribir

$$\begin{aligned}\int_a^b f(g(x))g'(x)\,dx &= B(g(b)) - B(g(a)) \\ &= \int_m^{g(b)} f(u)\,du - \int_m^{g(a)} f(u)\,du \\ &= \int_{g(a)}^{g(b)} f(u)\,du,\end{aligned}$$

por aditividad y orientación de la integral. Por consiguiente,

$$\int_a^b f(g(x))g'(x)\,dx = \int_{g(a)}^{g(b)} f(u)\,du.$$

■

Qué significa escribir $u = g(x)$ y $du = g'(x)\,dx$. La notación tradicional comprime el teorema anterior en dos líneas:

$$u = g(x), \qquad du = g'(x)\,dx.$$

Entonces el patrón $f(g(x))g'(x)\,dx$ se reescribe como $f(u)\,du$.

Así,

$$\int f(g(x))g'(x)\,dx \quad \rightsquigarrow \quad \int f(u)\,du.$$

Esta escritura es extremadamente útil, pero conviene leerla correctamente dentro de la teoría que estamos desarrollando.

El símbolo du no reemplaza la demostración
La igualdad $du = g'(x)\,dx$ es una notación que registra la transformación sugerida por la regla de la cadena. En este capítulo no necesitamos tratar du y dx como números independientes que puedan manipularse arbitrariamente. La justificación matemática del cambio de variable está en Teorema 18.1.

La notación diferencial funciona precisamente porque conserva a la vista las dos piezas estructurales relevantes:

función interior $g(x)$	+	su derivada $g'(x)$.

Figura 18.1 resume las dos formas de Teorema 18.1 y separa con claridad el cambio de representación de cualquier supuesto de invertibilidad.

Primer ejemplo: la sustitución completa el patrón.

Ejemplo 18.1. **Una potencia compuesta.** Calculemos

$$\int 6x(3x^2+1)^4\,dx.$$

Elegimos $u = 3x^2 + 1$.

Entonces $du = 6x\,dx$.

Todo el integrando queda expresado en la nueva variable:

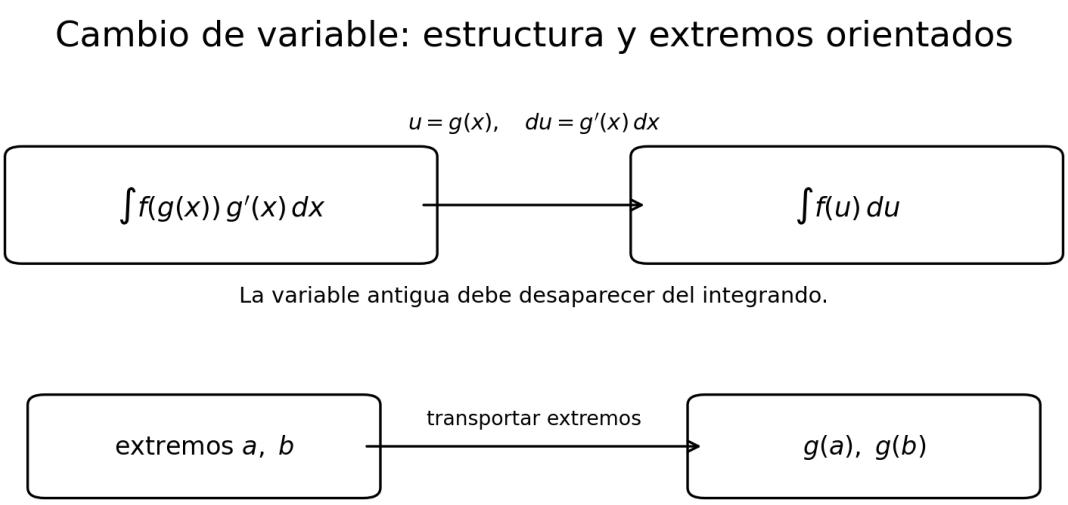


FIGURA 18.1. El cambio de variable traduce una estructura completa: x pasa a $u = g(x)$, el factor $f(g(x))g'(x)\,dx$ pasa a $f(u)\,du$ y, en la integral definida, los extremos a, b pasan a $g(a), g(b)$ conservando la orientación. La identidad no exige despejar x ni invertir g.

$$\int 6x(3x^2+1)^4\,dx = \int u^4\,du.$$

La regla de la potencia da

$$\int u^4\,du = \frac{u^5}{5} + C.$$

Volviendo a la variable original,

$$\boxed{\int 6x(3x^2+1)^4\,dx = \frac{(3x^2+1)^5}{5} + C.}$$

La comprobación debe hacerse siempre que una transformación sea nueva. Derivando,

$$\begin{aligned}\frac{d}{dx}\left[\frac{(3x^2+1)^5}{5}\right] &= \frac{1}{5}\cdot 5(3x^2+1)^4\cdot 6x \\ &= 6x(3x^2+1)^4.\end{aligned}$$

La sustitución no ha producido una fórmula distinta de la regla de la cadena: ha permitido **reconstruir** la composición que la derivación habría desarmado.

Cuando falta solo una constante. El patrón no tiene que aparecer con el coeficiente exacto desde el principio. Consideremos

$$\int xe^{x^2}\,dx.$$

La función interior natural sigue siendo $g(x) = x^2$, pero $g'(x) = 2x$, y el integrando contiene solamente x. La discrepancia es una constante multiplicativa, que podemos fabricar sin cambiar el problema: $x = \frac{1}{2}(2x)$. Por linealidad de la familia de primitivas,

$$\begin{aligned}\int xe^{x^2}\,dx &= \frac{1}{2}\int 2xe^{x^2}\,dx \\ &= \frac{1}{2}e^{x^2} + C.\end{aligned}$$

En notación de sustitución,

$$u = x^2, \qquad du = 2x\,dx,$$

y por tanto

$$\int xe^{x^2}\,dx = \frac{1}{2}\int e^u\,du.$$

Esto sugiere una pregunta de diagnóstico muy útil:

> **¿la derivada de la función interior está presente exactamente, o al menos salvo un factor constante?**

Si la respuesta es sí, el cambio de variable suele cerrar inmediatamente el problema.

Una integral definida: también cambian los extremos. Consideremos ahora

$$\int_0^1 2xe^{x^2}\,dx.$$

Tomamos de nuevo

$$u = x^2, \qquad du = 2x\,dx.$$

Pero ahora debemos transformar también los extremos. Cuando $x = 0$, tenemos $u = 0$, y cuando $x = 1$, tenemos $u = 1$. Por Teorema 18.1,

$$\begin{aligned}\int_0^1 2xe^{x^2}\,dx &= \int_0^1 e^u\,du \\ &= \left[e^u\right]_0^1 \\ &= e - 1.\end{aligned}$$

Una vez transformados los extremos, **no hace falta volver a** x: la integral definida ya es un número.

No mezclar variables y extremos
En una integral definida hay dos procedimientos correctos:

1. hacer la sustitución, transformar los extremos y terminar completamente en la variable nueva;
2. hallar primero una primitiva en la variable nueva, regresar a x y evaluar después en los extremos originales.

Lo que debe evitarse es mezclar ambos procedimientos, por ejemplo conservar límites en x mientras el integrando ya está escrito en u.

La orientación queda incorporada automáticamente. En Teorema 18.1 no hemos supuesto que $g(a) < g(b)$. Si ocurre lo contrario, el extremo transformado izquierdo puede quedar por encima del derecho. La integral orientada de C15 ya sabe manejar esa situación:

$$\int_{g(a)}^{g(b)} f(u)\,du = -\int_{g(b)}^{g(a)} f(u)\,du.$$

Por eso no debemos reordenar los nuevos extremos silenciosamente. Si los intercambiamos, debe aparecer el signo menos correspondiente. Más aún, la demostración tampoco exigió que g fuese monótona o inyectiva en $[a, b]$. La identidad procede de la composición $A \circ g$ y de la regla de la cadena. Si g avanza y luego retrocede, el signo de $g'(x)$ registra esas orientaciones locales y las contribuciones se combinan correctamente.

Cambio de variable no significa necesariamente función inversa
En esta versión del teorema no necesitamos resolver $u = g(x)$ para despejar x como función de u. Lo esencial es que el integrando tenga la forma $f(g(x))g'(x)$. La exigencia de invertibilidad aparecerá en otros contextos, pero no forma parte de la justificación básica de Teorema 18.1.

Una sustitución que no cierra. Volvamos al contraste planteado en §18.1:

$$\int e^{x^2}\,dx.$$

La elección $u = x^2$ produce formalmente $du = 2x\,dx$.

Pero el integrando original no contiene el factor x. A diferencia del ejemplo anterior, la pieza ausente **no es una constante** que podamos introducir y compensar. Si intentáramos escribir $dx = \frac{du}{2x}$, obtendríamos una expresión que todavía contiene x:

$$\int e^{u}\frac{du}{2x}.$$

La variable antigua no ha desaparecido, así que el cambio no ha transformado realmente el problema en una integral en u. Esto no demuestra que la integral «no exista». Tampoco demuestra aquí nada acerca de si posee o no una expresión elemental. Solo demuestra algo mucho más local y preciso:

$$\boxed{u = x^2 \text{ no es, por sí sola, una sustitución que cierre esta integral.}}$$

El fracaso es diagnosticable: falta la derivada de la función interior.

Cómo elegir la función interior. La elección de g puede guiarse examinando si la derivada de una expresión interior aparece como factor en el integrando. Ante un integrando, busquemos una subexpresión $g(x)$ que satisfaga dos condiciones:

1. **aparece dentro de otra función o potencia**, de modo que tratarla como una unidad simplificaría la expresión;
2. **su derivada aparece en el resto del integrando**, exactamente o salvo un factor constante.

Esquemáticamente,

$$\underbrace{f(g(x))}_{\text{estructura exterior}} \qquad \underbrace{g'(x)}_{\text{huella interior}}$$

es la firma más clara de una sustitución directa. Para evaluar la elección $u = g(x)$ debemos comprobar qué expresión queda después del cambio y si facilita el problema:

> después de efectuar el cambio, ¿todo el integrando puede escribirse únicamente con la nueva variable y su diferencial?

Si todavía sobreviven dependencias no eliminadas de x, la sustitución no ha cerrado el problema y debemos reconsiderarla o usar otra técnica.

Verificación: volver a derivar. Toda primitiva obtenida por sustitución admite un control inmediato. Si hemos concluido

$$\int f(x)\,dx = F(x) + C,$$

la comprobación es $F'(x) \stackrel{?}{=} f(x)$.

Este control es especialmente importante porque los errores típicos de sustitución son estructurales:

- olvidar un factor constante;
- omitir la derivada de la función interior;
- regresar incorrectamente de u a x;
- transformar el integrando pero no los extremos de una integral definida;
- cambiar el orden de los extremos sin cambiar el signo.

La derivación de la respuesta detecta varios de estos errores de una sola vez.

Antes de seguir
Calcula

$$\int_0^1 \frac{3x^2}{1+x^3}\,dx$$

mediante un cambio de variable. Identifica la función interior, transforma los extremos y explica qué resultado de C18 permite evaluar la integral resultante.
Respuesta. Tomamos

$$u = 1 + x^3, \qquad du = 3x^2\,dx.$$

Los extremos se transforman como

$$x = 0 \Longrightarrow u = 1, \qquad x = 1 \Longrightarrow u = 2.$$

Por tanto,

$$\begin{aligned}\int_0^1 \frac{3x^2}{1+x^3}\,dx &= \int_1^2 \frac{1}{u}\,du \\ &= \big[\ln u\big]_1^2 \\ &= \ln 2 - \ln 1 \\ &= \boxed{\ln 2},\end{aligned}$$

usando que $(\ln u)' = 1/u$ y $\ln 1 = 0$, resultados establecidos en C18.

El cambio de variable queda así fundamentado como una inversión controlada de la regla de la cadena. La siguiente técnica nacerá de una lectura análoga de otra regla de derivación: la regla del producto.

18.3. Integración por partes: invertir la regla del producto

La integración por partes se obtiene leyendo la regla del producto como una relación entre primitivas, del mismo modo que la sustitución procede de la regla de la cadena. Si u y v son funciones diferenciables, Teorema 7.2 afirma que

$$\frac{d}{dx}\big(u(x)v(x)\big) = u'(x)v(x) + u(x)v'(x).$$

Para integrar no necesitamos inventar una identidad nueva. Basta reorganizar la que ya conocemos:

$$u(x)v'(x) = \frac{d}{dx}\big(u(x)v(x)\big) - u'(x)v(x).$$

Esta igualdad contiene toda la técnica. Al pasar de derivadas a primitivas, obtenemos esquemáticamente

$$\boxed{\int u(x)v'(x)\,dx = u(x)v(x) - \int u'(x)v(x)\,dx.}$$

Una vez justificada la fórmula, tendremos que elegir una descomposición que simplifique la integral restante. La identidad conserva su validez con elecciones diferentes, pero esas elecciones pueden facilitar o dificultar el cálculo.

La fórmula y su demostración.

Teorema 18.2. **Teorema (integración por partes).** Sea $I \subseteq \mathbb{R}$ un intervalo y sean $u, v : I \to \mathbb{R}$ funciones continuamente diferenciables.

Entonces, en el sentido de familias de primitivas,

$$\boxed{\int u(x)v'(x)\,dx = u(x)v(x) - \int u'(x)v(x)\,dx.}$$

Si $a < b$ y $[a, b] \subseteq I$, entonces

$$\boxed{\int_a^b u(x)v'(x)\,dx = \big[u(x)v(x)\big]_a^b - \int_a^b u'(x)v(x)\,dx.}$$

Demostración. Por la regla del producto Teorema 7.2, $(uv)' = u'v + uv'$.

Reordenando, $uv' = (uv)' - u'v$.

Como u y v son continuamente diferenciables, el producto $u'v$ es continuo. Fijemos $x_0 \in I$ y definamos

$$H(x) := \int_{x_0}^x u'(t)v(t)\,dt, \qquad x \in I.$$

La integral está bien definida porque, para cada $x \in I$, el segmento cerrado de extremos x_0 y x está contenido en I y el integrando es continuo allí. En cualquier segmento compacto contenido en I, la función H difiere por una constante de una función de acumulación con base en el extremo izquierdo; por Teorema 16.2, $H'(x) = u'(x)v(x)$. Así, H es una primitiva de $u'v$ en todo I. Entonces

$$\begin{aligned}(uv - H)' &= (uv)' - H' \\ &= (u'v + uv') - u'v \\ &= uv'.\end{aligned}$$

Así, $uv - H$ es una primitiva de uv'. Como H representa una elección de

$$\int u'v\,dx,$$

obtenemos

$$\int uv'\,dx = uv - \int u'v\,dx,$$

con la constante de integración absorbida en la familia resultante. La extensión de Proposición 16.2 establecida en §18.1 garantiza que esta descripción alcanza a todas las primitivas en el intervalo I. Para la forma definida integramos la identidad $uv' = (uv)' - u'v$ sobre $[a, b]$. Por linealidad,

$$\int_a^b uv'\,dx = \int_a^b (uv)'\,dx - \int_a^b u'v\,dx.$$

Newton–Leibniz Teorema 16.3 aplicado al producto uv da

$$\int_a^b (uv)'\,dx = u(b)v(b) - u(a)v(a).$$

Por tanto,

$$\int_a^b u(x)v'(x)\,dx = \big[u(x)v(x)\big]_a^b - \int_a^b u'(x)v(x)\,dx.$$

■

La demostración muestra exactamente de dónde procede el signo menos: aparece al **despejar** uno de los dos términos de la regla del producto. No es una convención añadida al integrar.

La notación $u\,dv$. La fórmula suele escribirse de manera comprimida como

$$\boxed{\int u\,dv = uv - \int v\,du.}$$

La correspondencia con la forma rigurosa es

$$du = u'(x)\,dx, \qquad dv = v'(x)\,dx.$$

Así, $u\,dv = u(x)v'(x)\,dx$ y $v\,du = v(x)u'(x)\,dx$.

La notación no sustituye la estructura
Escribir

$$\int u\,dv = uv - \int v\,du$$

es una abreviación útil de Teorema 18.2. No estamos tratando du, dv y dx como cantidades independientes que puedan separarse arbitrariamente. La técnica es válida porque la regla del producto proporciona una identidad entre derivadas.

Recuperar el ejemplo de §18.1. En §18.1 observamos que

$$\int xe^x\,dx$$

podía reconstruirse a partir de la regla del producto. Ahora podemos expresar ese razonamiento sistemáticamente.

Elegimos

$$u = x, \qquad dv = e^x\,dx.$$

Entonces

$$du = dx, \qquad v = e^x.$$

Aplicando Teorema 18.2,

$$\begin{aligned}\int xe^x\,dx &= xe^x - \int e^x\,dx \\ &= xe^x - e^x + C.\end{aligned}$$

La nueva integral es más sencilla que la original porque diferenciar x lo reduce a la constante 1, mientras integrar e^x no modifica su forma. Podemos verificar el resultado derivando:

$$\begin{aligned}\frac{d}{dx}(xe^x - e^x) &= e^x + xe^x - e^x \\ &= xe^x.\end{aligned}$$

La comprobación reproduce exactamente la cancelación que dio origen a la fórmula.

Un factor que estaba oculto: integrar $\ln x$. La integración por partes puede aplicarse incluso cuando el integrando no parece, a primera vista, un producto. En

$$\int \ln x \, dx, \qquad x > 0,$$

podemos hacer visible el factor constante: $\ln x = \ln x \cdot 1$.

Ejemplo 18.2. **Una primitiva de $\ln x$.** Para $x > 0$,

$$\boxed{\int \ln x \, dx = x \ln x - x + C.}$$

Tomamos

$$u = \ln x, \qquad dv = dx.$$

Los resultados de C18 dan $du = \frac{1}{x}\, dx$, mientras que $v = x$.
Por integración por partes,

$$\begin{aligned} \int \ln x \, dx &= x \ln x - \int x \frac{1}{x}\, dx \\ &= x \ln x - \int 1 \, dx \\ &= x \ln x - x + C. \end{aligned}$$

La simplificación decisiva fue $x \cdot \frac{1}{x} = 1$, válida porque estamos trabajando en el dominio $x > 0$. La derivación confirma inmediatamente el resultado:

$$\frac{d}{dx}(x \ln x - x) = \ln x + 1 - 1 = \ln x.$$

El ejemplo utiliza un producto en el que uno de los factores es 1. Reconocer ese factor permite aplicar partes incluso cuando el integrando no se presenta inicialmente como un producto explícito.

Elegir u y dv es parte del problema. La fórmula

$$\int u \, dv = uv - \int v \, du$$

es simétrica en su origen —proviene de un producto—, pero las dos elecciones posibles no suelen ser igualmente útiles.

La pregunta estratégica es:

> ¿qué factor conviene **diferenciar** y qué factor conviene **integrar** para que la integral restante sea más sencilla?

Dos criterios suelen ser especialmente útiles:

- escoger como u una parte que se simplifique al derivarla;
- escoger como dv una parte cuya primitiva v podamos obtener directamente.

Estos criterios no son un algoritmo infalible. Son una manera de evaluar si la transformación hace progreso real. Consideremos, por ejemplo, la estructura

$$\int x^3 e^x \, dx.$$

Si elegimos

$$u = x^3, \qquad dv = e^x \, dx,$$

entonces

$$du = 3x^2 \, dx, \qquad v = e^x,$$

y la integral restante contiene $x^2 e^x$: el grado del polinomio ha disminuido.

Si elegimos al revés,

$$u = e^x, \qquad dv = x^3 \, dx,$$

entonces

$$du = e^x \, dx, \qquad v = \frac{x^4}{4},$$

y la nueva integral contiene $\frac{x^4}{4} e^x$, que ha aumentado el grado del polinomio. La fórmula sigue siendo correcta, pero la elección ha empeorado el problema.

Correcto no significa útil
Una aplicación de integración por partes puede ser algebraicamente válida y, sin embargo, no constituir progreso. La técnica exige dos controles distintos:

1. **validez:** la descomposición debe satisfacer las hipótesis de Teorema 18.2;
2. **utilidad:** la integral nueva debería ser más manejable que la original.

Figura 18.2 hace visible la distinción entre **validez** y **progreso**: Teorema 18.2 garantiza la identidad, no que toda elección de u y dv simplifique el problema.

La forma definida y el término de borde. En una integral definida la fórmula contiene un término adicional que no debe perderse:

$$\boxed{\int_a^b u(x)v'(x) \, dx = \big[u(x)v(x)\big]_a^b - \int_a^b u'(x)v(x) \, dx.}$$

El producto $\big[u(x)v(x)\big]_a^b = u(b)v(b) - u(a)v(a)$ se denomina habitualmente **término de borde**. Procede directamente de integrar la derivada del producto mediante Newton–Leibniz.

Por ejemplo,

$$\int_0^1 xe^x \, dx$$

se calcula con las mismas elecciones

$$u = x, \qquad dv = e^x \, dx.$$

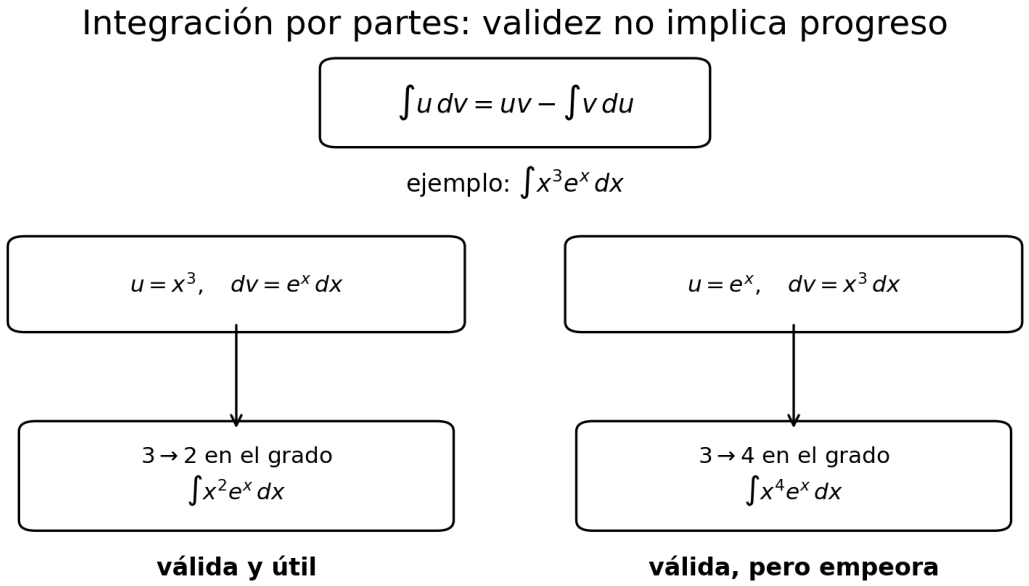


FIGURA 18.2. En integración por partes la identidad es válida en ambas elecciones, pero la estrategia cambia: para $\int x^3 e^x\,dx$, diferenciar x^3 reduce el grado $3 \to 2$, mientras que integrar x^3 y diferenciar e^x lo aumenta $3 \to 4$.

Entonces

$$\begin{aligned}\int_0^1 xe^x\,dx &= \left[xe^x\right]_0^1 - \int_0^1 e^x\,dx \\ &= e - (e-1) \\ &= \boxed{1}.\end{aligned}$$

En una integral indefinida el producto uv forma parte de la primitiva; en una integral definida, ese mismo producto se evalúa en los dos extremos.

Dos errores de lectura que conviene evitar. El primero consiste en olvidar que v debe ser una **primitiva** de la parte elegida como dv. Si escribimos $dv = q(x)\,dx$, la técnica necesita que podamos obtener una función v con $v'(x) = q(x)$. Elegir como dv una expresión más difícil de integrar que la original no aporta nada. El segundo error consiste en pensar que la fórmula garantiza automáticamente una integral más simple. Teorema 18.2 solo garantiza la **identidad**. La disminución de complejidad depende de la elección de los papeles. Podemos resumir el diagnóstico en cuatro preguntas:

1. ¿qué producto veo o puedo hacer visible?
2. ¿qué factor se simplifica al derivarlo?
3. ¿qué factor puedo integrar con facilidad?
4. después de aplicar la fórmula, ¿la integral restante es realmente más sencilla?

Verificación: la regla del producto debe reaparecer. Si una aplicación de partes produce una candidata F, la prueba final sigue siendo derivar. Por ejemplo, $F(x) = x\ln x - x$

debe satisfacer, para $x > 0$, $F'(x) = \ln x$. El control no es accesorio. Como la técnica procede de la regla del producto, una derivación correcta debe reconstruir los dos términos del producto y mostrar la cancelación correspondiente.

Antes de seguir
Calcula

$$\int_1^e \ln x \, dx$$

mediante integración por partes. Identifica u, dv, du y v, y conserva explícitamente el término de borde.
Respuesta. Tomamos

$$u = \ln x, \qquad dv = dx,$$

de modo que

$$du = \frac{1}{x}\,dx, \qquad v = x.$$

Por Teorema 18.2,

$$\begin{aligned}\int_1^e \ln x \, dx &= \big[x \ln x\big]_1^e - \int_1^e 1 \, dx \\ &= (e \ln e - \ln 1) - (e - 1) \\ &= e - (e - 1) \\ &= \boxed{1},\end{aligned}$$

porque C18 estableció

$$\ln e = 1, \qquad \ln 1 = 0.$$

La regla del producto justifica la identidad de integración por partes bajo sus hipótesis. Elegir los factores requiere además examinar la integral resultante y comprobar si el cambio facilita su cálculo.

18.4. Repetir con propósito: partes iteradas y fórmulas de reducción

En §18.3 vimos que una sola aplicación de integración por partes puede sustituir una integral por otra más sencilla. Pero algunos productos no se resuelven en un único paso.

El ejemplo

$$\int x^3 e^x \, dx$$

ya anticipó el mecanismo. Si elegimos

$$u = x^3, \qquad dv = e^x \, dx,$$

entonces

$$du = 3x^2 \, dx, \qquad v = e^x,$$

y Teorema 18.2 produce

$$\int x^3 e^x\,dx = x^3e^x - 3\int x^2e^x\,dx.$$

No hemos terminado, pero sí hemos avanzado: el grado del factor polinómico bajó de 3 a 2. La nueva integral tiene exactamente la misma estructura que la anterior. Podemos aplicar partes otra vez:

$$\int x^2 e^x\,dx = x^2e^x - 2\int xe^x\,dx.$$

Y una tercera aplicación da

$$\int xe^x\,dx = xe^x - \int e^x\,dx = xe^x - e^x + C.$$

Sustituyendo hacia atrás,

$$\begin{aligned}\int x^2e^x\,dx &= x^2e^x - 2(xe^x - e^x) + C\\ &= e^x(x^2 - 2x + 2) + C,\end{aligned}$$

y finalmente

$$\boxed{\int x^3e^x\,dx = e^x(x^3 - 3x^2 + 6x - 6) + C.}$$

Aquí la repetición no fue mecánica. En cada paso conservamos la misma decisión estratégica:

$$\boxed{\text{diferenciar el polinomio} \quad + \quad \text{integrar la exponencial.}}$$

La primera operación reduce el grado; la segunda conserva la forma de e^x. Por eso el proceso termina después de un número finito de pasos.

De un cálculo particular a una fórmula de reducción. El ejemplo anterior pertenece a una familia entera. Para $n \in \mathbb{N}$ consideremos

$$I_n = \int x^n e^x\,dx.$$

Si $n \geq 1$, elegimos

$$u = x^n, \qquad dv = e^x\,dx.$$

Entonces

$$du = nx^{n-1}\,dx, \qquad v = e^x,$$

y la integración por partes da $I_n = x^ne^x - nI_{n-1}$. Esta identidad expresa una integral de orden n mediante otra de orden $n-1$. Una relación de este tipo se llama **fórmula de reducción**.

Proposición 18.1. **Proposición (reducción de $\int x^n e^x\,dx$).** Para todo $n \in \mathbb{N}$ con $n \geq 1$,

$$\boxed{I_n = x^ne^x - nI_{n-1}, \qquad I_n := \int x^n e^x\,dx.}$$

En particular,

$$\boxed{\int x^n e^x\,dx = e^x \sum_{k=0}^{n} (-1)^k \frac{n!}{(n-k)!} x^{n-k} + C.}$$

Demostración. La relación de reducción ya fue obtenida por Teorema 18.2. Falta justificar la fórmula cerrada.

Para $n = 0$,

$$\int e^x\,dx = e^x + C,$$

y la suma del enunciado contiene solamente el término $k = 0$. Supongamos ahora que la fórmula es válida para $n - 1$. Entonces

$$I_{n-1} = e^x \sum_{k=0}^{n-1} (-1)^k \frac{(n-1)!}{(n-1-k)!} x^{n-1-k} + C.$$

Usando la relación de reducción,

$$\begin{aligned} I_n &= x^n e^x - n I_{n-1} \\ &= e^x x^n - e^x \sum_{k=0}^{n-1} (-1)^k \frac{n!}{(n-1-k)!} x^{n-1-k} + C. \end{aligned}$$

En la suma hacemos el cambio de índice $j = k + 1$. Entonces j recorre $1, \ldots, n$ y $-(-1)^k = (-1)^{k+1} = (-1)^j$, mientras $(n-1-k)! = (n-j)!$.

Por tanto,

$$I_n = e^x x^n + e^x \sum_{j=1}^{n} (-1)^j \frac{n!}{(n-j)!} x^{n-j} + C.$$

El término $e^x x^n$ es precisamente el término $j = 0$ de la misma suma, porque $\frac{n!}{n!} = 1$.

Así,

$$I_n = e^x \sum_{j=0}^{n} (-1)^j \frac{n!}{(n-j)!} x^{n-j} + C.$$

La inducción queda cerrada. ■

Qué hace que una reducción sea útil. La palabra *reducción* no significa simplemente que el índice haya cambiado. Debemos comprobar que la nueva integral se encuentre realmente más cerca de una forma ya conocida.

En la familia anterior,

$$I_n \longrightarrow I_{n-1} \longrightarrow \cdots \longrightarrow I_1 \longrightarrow I_0,$$

y

$$I_0 = \int e^x\,dx$$

es inmediata.

Por tanto, existe una **medida de complejidad** que disminuye en cada paso: el grado del polinomio.

Una repetición necesita una dirección de progreso
Aplicar integración por partes varias veces tiene sentido cuando podemos identificar alguna cantidad que disminuye o alguna estructura que se simplifica. En $x^n e^x$, esa cantidad es el exponente n. Si una repetición no disminuye ninguna complejidad reconocible, puede producir una cadena infinita de problemas equivalentes o incluso más difíciles.

Esta idea es más general que la fórmula concreta. Al utilizar una técnica repetidamente conviene poder responder:

¿qué se hace estrictamente más simple después de cada paso?

Si no podemos señalarlo, todavía no hemos justificado que la iteración constituya una estrategia.

Cuando la integral original reaparece. Al repetir integración por partes puede reaparecer la integral original. En ese caso, en lugar de continuar indefinidamente, podemos reunir los términos y resolver la relación algebraica obtenida cuando el coeficiente correspondiente lo permite. Eso no tiene por qué ser un fracaso.

Ejemplo 18.3. **Una integral que vuelve sobre sí misma.** Calculemos

$$\int e^x \cos x \, dx.$$

Denotemos por

$$I = \int e^x \cos x \, dx$$

la familia de primitivas que buscamos. Elegimos

$$u = \cos x, \qquad dv = e^x \, dx.$$

Entonces

$$du = -\sin x \, dx, \qquad v = e^x,$$

y Teorema 18.2 da

$$I = e^x \cos x + \int e^x \sin x \, dx.$$

La nueva integral todavía no es elemental a simple vista. Aplicamos partes una segunda vez, ahora a

$$\int e^x \sin x \, dx.$$

Tomamos

$$u = \sin x, \qquad dv = e^x \, dx,$$

de modo que

$$du = \cos x\,dx, \qquad v = e^x.$$

Así,

$$\int e^x \sin x\,dx = e^x \sin x - I.$$

Sustituyendo en la primera identidad, $I = e^x \cos x + e^x \sin x - I$. Como las integrales indefinidas representan familias de primitivas, las constantes aditivas que aparecen en los dos pasos se absorben en una única constante. Reuniendo los dos términos que representan a I, $2I = e^x(\sin x + \cos x) + C$, y por tanto

$$\boxed{\int e^x \cos x\,dx = \frac{e^x}{2}(\sin x + \cos x) + C.}$$

La verificación elimina cualquier ambigüedad sobre las constantes:

$$\begin{aligned}\frac{d}{dx}\left[\frac{e^x}{2}(\sin x + \cos x)\right] &= \frac{e^x}{2}(\sin x + \cos x) + \frac{e^x}{2}(\cos x - \sin x)\\ &= e^x \cos x.\end{aligned}$$

La integral original reapareció con coeficiente -1. Precisamente por eso pudimos trasladarla al otro miembro y resolver una ecuación algebraica para la familia de primitivas.

Dos tipos de cierre. Los ejemplos de esta sección exhiben dos formas distintas en que una iteración de partes puede terminar.

En la primera,

$$\int x^n e^x\,dx,$$

cada aplicación reduce el grado del polinomio hasta llegar a una integral conocida.

En la segunda,

$$\int e^x \cos x\,dx,$$

el proceso entra en un ciclo y recupera la integral inicial. El ciclo resulta útil porque produce una ecuación lineal en la integral desconocida. Podemos resumir ambos mecanismos así:

reducción finita o retorno controlado a la integral inicial.

Figura 18.3 compara la reducción de Proposición 18.1 con el ciclo de Ejemplo 18.3: repetir una técnica tiene sentido cuando existe descenso o cierre, no por mera reiteración.

La repetición sin ninguno de estos mecanismos no debe considerarse automáticamente progreso.

No perder las constantes al trabajar con integrales indefinidas. En una cadena corta solemos escribir

$$\int f(x)\,dx = F(x) + C$$

y absorber las constantes intermedias en una sola constante final. Este procedimiento es legítimo porque estamos describiendo una **familia** de primitivas. Sin embargo, cuando la integral

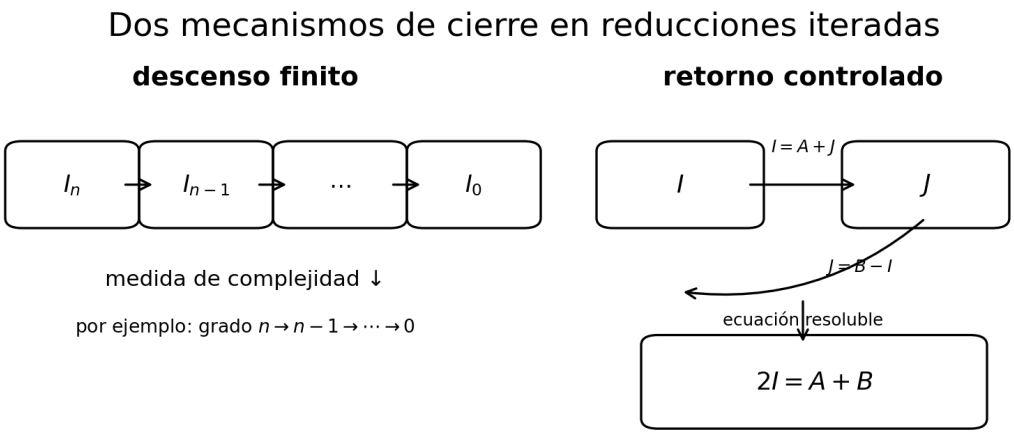


FIGURA 18.3. Dos mecanismos permiten cerrar una repetición de partes: un descenso finito $I_n \to I_{n-1} \to \cdots \to I_0$, o un ciclo controlado $I \to J \to I$ que devuelve la integral original con un coeficiente y permite resolver una ecuación algebraica.

original reaparece en ambos miembros, conviene recordar qué significa realmente la notación. No estamos manipulando un número llamado $\int f(x)\,dx$; estamos relacionando funciones primitivas que son únicas salvo constantes. Por eso hay dos formas seguras de cerrar un cálculo cíclico:

1. mantener explícitas las constantes hasta el final y absorberlas entonces en una sola;
2. obtener una candidata mediante el álgebra formal y **verificarla derivando**.

En este libro utilizaremos con frecuencia la segunda, porque hace visible de inmediato si la identidad final es correcta.

Una ecuación con la integral no autoriza a dividir por cero
Si una iteración conduce a una relación de la forma $I = A + \lambda I$, solo podemos despejar I cuando $1 - \lambda \neq 0$. Si $\lambda = 1$, la relación no determina I; simplemente reproduce una identidad. La aparición de la integral original es útil únicamente cuando el coeficiente permite aislarla.

Antes de seguir.

Antes de seguir
Calcula mediante integración por partes repetida

$$\int x^2 \cos x\,dx.$$

En cada paso explica qué cantidad disminuye.

Respuesta. Tomamos primero

$$u = x^2, \qquad dv = \cos x\,dx,$$

de modo que

$$du = 2x\,dx, \qquad v = \sin x.$$

Entonces

$$\int x^2 \cos x\,dx = x^2 \sin x - 2\int x \sin x\,dx.$$

Para la integral restante elegimos

$$u = x, \qquad dv = \sin x\,dx,$$

por lo que

$$du = dx, \qquad v = -\cos x.$$

Así,

$$\int x \sin x\,dx = -x\cos x + \int \cos x\,dx = -x\cos x + \sin x + C.$$

Sustituyendo,

$$\boxed{\int x^2 \cos x\,dx = x^2 \sin x + 2x\cos x - 2\sin x + C.}$$

En las dos aplicaciones la cantidad que disminuye es el grado del factor polinómico: $2 \longrightarrow 1 \longrightarrow 0$. La derivación de la respuesta recupera $x^2 \cos x$ y cierra la comprobación.

La integración por partes repetida no añade una técnica distinta de Teorema 18.2. Añade una forma de **organizar una secuencia de aplicaciones**: identificar una medida de complejidad que disminuye, una fórmula de reducción que acerca a un caso base o un ciclo que devuelve la integral inicial con un coeficiente algebraicamente resoluble.

18.5. Potencias trigonométricas: reescribir antes de sustituir

Hasta ahora las dos técnicas principales del capítulo nacieron leyendo al revés reglas de derivación: la regla de la cadena produjo el cambio de variable y la regla del producto produjo la integración por partes. Hay, sin embargo, integrandos en los que ninguna de esas estructuras aparece de inmediato. En esos casos el primer paso puede ser **cambiar la representación algebraica del integrando**. Las potencias de seno y coseno constituyen un laboratorio especialmente claro. Consideremos la familia

$$\int \sin^m x\, \cos^n x\,dx, \qquad m, n \in \mathbb{N}.$$

La pregunta decisiva no será todavía «¿qué fórmula integro?», sino

¿qué identidad trigonométrica vuelve visible una sustitución elemental?

Las tres identidades que organizan la familia. Utilizaremos $\sin^2 x + \cos^2 x = 1$,

$$\sin(2x) = 2\sin x \cos x,$$

y $\cos(2x) = \cos^2 x - \sin^2 x$.

Combinando la primera y la tercera obtenemos las fórmulas de reducción de potencia

$$\boxed{\cos^2 x = \frac{1+\cos(2x)}{2},}$$

y

$$\boxed{\sin^2 x = \frac{1-\cos(2x)}{2}.}$$

Estas identidades no integran nada por sí mismas. Su función es **reescribir** el integrando en una forma donde aparezca la huella de una derivada conocida. La paridad de los exponentes indicará qué identidad conviene utilizar.

PROPOSICIÓN 18.2. **Proposición (estrategia para potencias de seno y coseno).** Sean $m, n \in \mathbb{N}$.

Para estudiar

$$\int \sin^m x \, \cos^n x \, dx,$$

pueden aplicarse los principios siguientes.

1. Si n es impar, se separa un factor $\cos x$ y se transforma el resto mediante
$$\cos^2 x = 1 - \sin^2 x.$$
La sustitución natural es
$$u = \sin x, \qquad du = \cos x \, dx.$$
2. Si m es impar, se separa un factor $\sin x$ y se transforma el resto mediante
$$\sin^2 x = 1 - \cos^2 x.$$
La sustitución natural es
$$u = \cos x, \qquad du = -\sin x \, dx.$$
3. Si m y n son ambos pares, las identidades
$$\sin^2 x = \frac{1-\cos(2x)}{2}, \qquad \cos^2 x = \frac{1+\cos(2x)}{2}$$
reducen las potencias. En casos sencillos esta reescritura cierra directamente la integral; para potencias más altas pueden quedar productos o potencias de cosenos con argumentos múltiples, que exigirán una descomposición algebraica adicional.

Justificación. Supongamos primero que n es impar. Escribimos $n = 2r + 1$.

Entonces

$$\cos^n x = \cos^{2r} x \cos x = (\cos^2 x)^r \cos x = (1 - \sin^2 x)^r \cos x.$$

Por tanto,

$$\sin^m x \cos^n x\, dx = \sin^m x(1-\sin^2 x)^r \cos x\, dx.$$

Con

$$u = \sin x, \qquad du = \cos x\, dx,$$

todo el integrando queda convertido en una expresión polinómica en u. El caso m impar es análogo. Si $m = 2r+1$, entonces $\sin^m x = (1-\cos^2 x)^r \sin x$, y la sustitución $u = \cos x$ absorbe el factor $\sin x\, dx$.

Si ambos exponentes son pares, ninguna de las dos potencias contiene un factor aislado $\sin x$ o $\cos x$ que pueda funcionar directamente como diferencial. En ese caso las fórmulas de reducción de potencia disminuyen los exponentes y hacen visible la estructura de ángulos múltiples. En ejemplos sencillos, como los que siguen, la reducción termina directamente en una suma integrable. Para potencias mayores pueden aparecer todavía productos de cosenos con argumentos distintos; §18.6 construirá precisamente la herramienta producto-a-suma que permite continuar esa descomposición. Por tanto, la afirmación de esta sección es de **progreso estructural**, no de cierre universal usando solo las dos fórmulas de reducción de potencia. ■

Exponente impar: reservar exactamente el factor que falta.

Ejemplo 18.4. **Ejemplo.** Calculemos

$$\int \sin^3 x \cos^2 x\, dx.$$

El exponente del seno es impar. Separamos un factor $\sin x$:

$$\sin^3 x = \sin^2 x \sin x = (1-\cos^2 x)\sin x.$$

Así,

$$\int \sin^3 x \cos^2 x\, dx = \int (1-\cos^2 x)\cos^2 x \sin x\, dx.$$

Tomamos

$$u = \cos x, \qquad du = -\sin x\, dx.$$

Entonces

$$\begin{aligned}\int \sin^3 x \cos^2 x\, dx &= -\int (1-u^2)u^2\, du \\ &= -\int (u^2-u^4)\, du \\ &= -\frac{u^3}{3} + \frac{u^5}{5} + C.\end{aligned}$$

Volviendo a x,

$$\boxed{\int \sin^3 x \cos^2 x\, dx = -\frac{\cos^3 x}{3} + \frac{\cos^5 x}{5} + C.}$$

El exponente impar permitió separar el factor $\sin x\,dx$, que corresponde a la derivada de $\cos x$ con el signo adecuado. La identidad del cuadrado transforma los factores restantes en una expresión de la nueva variable.

Si ambos exponentes son impares, hay más de una ruta. Consideremos

$$\int \sin^3 x \cos^3 x\,dx.$$

Podemos reservar un factor $\sin x$ y sustituir $u = \cos x$, o reservar un factor $\cos x$ y sustituir $u = \sin x$. Ambas rutas son correctas porque ambos exponentes son impares. La elección no cambia la familia de primitivas; solo cambia la representación intermedia.

Por ejemplo, separando $\cos x$,

$$\sin^3 x \cos^3 x = \sin^3 x(1 - \sin^2 x)\cos x,$$

y con $u = \sin x$ obtenemos una integral polinómica.

Una técnica puede admitir varias representaciones igualmente buenas
No siempre existe una única sustitución natural. Cuando dos estructuras distintas cierran completamente el cambio de variable, podemos elegir la que produzca la expresión más simple o la que facilite mejor la verificación posterior.

Ambos exponentes pares: reducir potencias. Cuando los dos exponentes son pares no queda disponible un factor aislado de seno o coseno para una sustitución inmediata. El caso más sencillo ya muestra el mecanismo:

$$\int \sin^2 x\,dx.$$

Usamos $\sin^2 x = \frac{1-\cos(2x)}{2}$.

Entonces

$$\begin{aligned}\int \sin^2 x\,dx &= \frac{1}{2}\int 1\,dx - \frac{1}{2}\int \cos(2x)\,dx \\ &= \frac{x}{2} - \frac{1}{4}\sin(2x) + C.\end{aligned}$$

La segunda integral utiliza solamente la sustitución lineal $u = 2x$. El patrón se vuelve más visible en un producto de dos potencias pares.

Ejemplo 18.5. **Ejemplo.** Calculemos

$$\int \sin^2 x \cos^2 x\,dx.$$

Primero utilizamos $2\sin x\cos x = \sin(2x)$.

Por tanto, $\sin^2 x\cos^2 x = \frac{1}{4}\sin^2(2x)$.

Aplicamos ahora la reducción de potencia al ángulo $2x$: $\sin^2(2x) = \frac{1-\cos(4x)}{2}$.

Así, $\sin^2 x\cos^2 x = \frac{1}{8}(1 - \cos(4x))$, y

$$\int \sin^2 x \cos^2 x \, dx = \frac{1}{8}\int 1 \, dx - \frac{1}{8}\int \cos(4x)\, dx$$
$$= \boxed{\frac{x}{8} - \frac{\sin(4x)}{32} + C}.$$

Aquí no apareció una nueva regla de integración. El trabajo decisivo fue **convertir una potencia trigonométrica en una suma de términos integrables**.

Reescribir no es simplificar a ciegas. Una identidad es útil solo si acerca el integrando a una estructura que ya sabemos manejar.

Por ejemplo, ante

$$\int \sin^5 x \cos^2 x \, dx,$$

sería posible comenzar a reemplazar todas las potencias mediante fórmulas de ángulo múltiple. Pero eso destruiría la huella más simple: el exponente impar de $\sin x$ ya nos entrega un factor $\sin x\, dx$ para la sustitución $u = \cos x$. La estrategia preferible es, por tanto,

preservar primero la derivada que necesitamos; reescribir después el resto.

Esto prolonga el principio de §18.2. Una sustitución no se elige por parecido visual, sino porque transforma **todo** el integrando en la nueva variable.

Un árbol de decisión, no una tabla para memorizar. Para la familia

$$\int \sin^m x \cos^n x \, dx$$

podemos resumir el diagnóstico así:

n impar $\Longrightarrow$ reservar $\cos x\, dx$, $\quad u = \sin x$,
m impar $\Longrightarrow$ reservar $\sin x\, dx$, $\quad u = \cos x$,
m, n pares $\Longrightarrow$ reducir potencias con ángulo doble.

Figura 18.4 representa Proposición 18.2 como una decisión reconstruible desde las derivadas de seno y coseno, no como una tabla mnemotécnica autónoma.

La tabla no reemplaza la comprensión. Cada fila debe poder reconstruirse desde las derivadas

$$(\sin x)' = \cos x, \qquad (\cos x)' = -\sin x$$

y desde la identidad $\sin^2 x + \cos^2 x = 1$.

Verificación y control de la representación. Después de obtener una primitiva debemos volver a derivar, como en las secciones anteriores. En integrales trigonométricas este control cumple además otra función: dos respuestas aparentemente distintas pueden representar la misma familia de primitivas porque las identidades trigonométricas permiten escribir una misma función de varias maneras. Por tanto, si dos expresiones difieren en apariencia, no debemos concluir de inmediato que una es incorrecta. Podemos:

1. derivar ambas y comprobar que producen el mismo integrando;

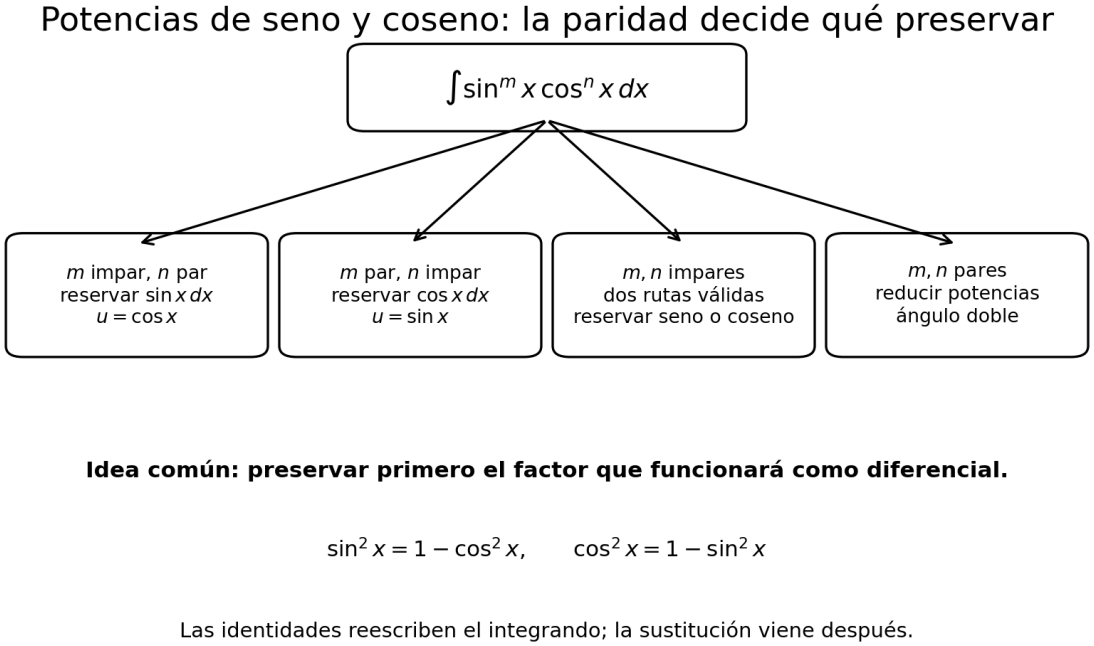


FIGURA 18.4. Árbol de decisión para $\int \sin^m x \cos^n x\, dx$: la paridad determina qué factor conviene reservar como diferencial; si ambos exponentes son impares aparecen dos rutas igualmente válidas, y si ambos son pares conviene reducir potencias.

2. o restarlas y utilizar identidades trigonométricas para verificar que la diferencia es constante.

La derivada sigue siendo el criterio decisivo.

Antes de seguir
Calcula

$$\int \sin^3 x \cos^4 x\, dx.$$

Explica por qué conviene preservar un factor $\sin x$ y no reducir todas las potencias desde el comienzo.
Respuesta. Como el exponente del seno es impar, $\sin^3 x = (1 - \cos^2 x)\sin x$.
Entonces

$$\int \sin^3 x \cos^4 x\, dx = \int (1 - \cos^2 x)\cos^4 x \sin x\, dx.$$

Tomamos

$$u = \cos x, \qquad du = -\sin x\, dx.$$

Por tanto,

$$\begin{aligned}\int \sin^3 x \cos^4 x\, dx &= -\int (1-u^2)u^4\, du \\ &= -\int (u^4 - u^6)\, du \\ &= -\frac{u^5}{5} + \frac{u^7}{7} + C.\end{aligned}$$

Volviendo a x,

$$\boxed{\int \sin^3 x \cos^4 x\, dx = -\frac{\cos^5 x}{5} + \frac{\cos^7 x}{7} + C.}$$

Reservar $\sin x\, dx$ conserva exactamente la derivada de $\cos x$; reducir todas las potencias desde el comienzo ocultaría esa estructura y alargaría innecesariamente el cálculo.

Las potencias trigonométricas muestran una tercera clase de trabajo dentro de las técnicas de integración. No hemos añadido una nueva operación: primero **reescribimos** mediante identidades, después reconocemos una de las técnicas ya demostradas y finalmente verificamos derivando. La representación adecuada puede ser tan importante como la fórmula de integración que se aplica después.

18.6. Productos trigonométricos de frecuencias distintas: convertir productos en sumas

La sección anterior estudió potencias de seno y coseno con el mismo argumento. Allí la paridad de los exponentes nos indicaba qué factor convenía preservar para fabricar un diferencial.

Aparece ahora otra familia:

$$\sin(ax)\cos(bx), \qquad \cos(ax)\cos(bx), \qquad \sin(ax)\sin(bx),$$

con $a, b > 0$.

Para el diagnóstico estratégico no perdemos generalidad esencial con esta convención: un signo negativo se absorbe mediante la paridad de seno y coseno, mientras que una frecuencia nula deja un solo factor trigonométrico y se integra directamente. Cuando a y b son distintos, ninguna sustitución simple del tipo

$$u = \sin(ax) \qquad \text{o} \qquad u = \cos(bx)$$

elimina por sí sola el otro factor. La dificultad no está en hallar una función interior adecuada, sino en que el integrando contiene **dos oscilaciones con escalas distintas**. La estrategia será cambiar nuevamente de representación:

$$\boxed{\text{producto trigonométrico} \longrightarrow \text{suma de funciones trigonométricas simples.}}$$

Después de esa transformación, la linealidad de la integral permite trabajar término a término.

Las fórmulas de adición contienen la transformación. Partimos de las identidades

$$\sin(\alpha+\beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta,$$

$$\sin(\alpha-\beta) = \sin\alpha\cos\beta - \cos\alpha\sin\beta,$$

$$\cos(\alpha+\beta)=\cos\alpha\cos\beta-\sin\alpha\sin\beta,$$

y

$$\cos(\alpha-\beta)=\cos\alpha\cos\beta+\sin\alpha\sin\beta.$$

La transformación que necesitamos no es una identidad adicional independiente. Se obtiene sumando o restando estas cuatro relaciones.

PROPOSICIÓN 18.3. **Proposición (identidades producto-a-suma).** Para cualesquiera $\alpha,\beta\in\mathbb{R}$,

$$\boxed{\sin\alpha\cos\beta=\frac{1}{2}\big[\sin(\alpha+\beta)+\sin(\alpha-\beta)\big],}$$

$$\boxed{\cos\alpha\cos\beta=\frac{1}{2}\big[\cos(\alpha+\beta)+\cos(\alpha-\beta)\big],}$$

y

$$\boxed{\sin\alpha\sin\beta=\frac{1}{2}\big[\cos(\alpha-\beta)-\cos(\alpha+\beta)\big].}$$

Demostración. Sumando las dos fórmulas del seno obtenemos

$$\sin(\alpha+\beta)+\sin(\alpha-\beta)=2\sin\alpha\cos\beta,$$

lo que prueba la primera identidad. Sumando las dos fórmulas del coseno,

$$\cos(\alpha+\beta)+\cos(\alpha-\beta)=2\cos\alpha\cos\beta,$$

y queda probada la segunda.

Finalmente, restando la fórmula de $\cos(\alpha+\beta)$ de la de $\cos(\alpha-\beta)$,

$$\cos(\alpha-\beta)-\cos(\alpha+\beta)=2\sin\alpha\sin\beta.$$

Esto prueba la tercera. ■

Las identidades muestran una estructura importante. Un producto con frecuencias a y b se descompone en términos con frecuencias

$$a+b\qquad\text{y}\qquad a-b.$$

No hemos integrado todavía; solo hemos reemplazado una representación multiplicativa por una aditiva.

Primer ejemplo: dos frecuencias distintas.

EJEMPLO 18.6. **Ejemplo.** Calculemos

$$\int \sin(3x)\cos(2x)\,dx.$$

Aplicamos Proposición 18.3 con

$$\alpha = 3x, \qquad \beta = 2x.$$

Entonces

$$\sin(3x)\cos(2x) = \frac{1}{2}\big[\sin(5x) + \sin x\big].$$

Por linealidad,

$$\begin{aligned}\int \sin(3x)\cos(2x)\,dx &= \frac{1}{2}\int \sin(5x)\,dx + \frac{1}{2}\int \sin x\,dx \\ &= -\frac{1}{10}\cos(5x) - \frac{1}{2}\cos x + C.\end{aligned}$$

Por tanto,

$$\boxed{\int \sin(3x)\cos(2x)\,dx = -\frac{\cos(5x)}{10} - \frac{\cos x}{2} + C.}$$

Cada término final se integra mediante una sustitución lineal elemental. La dificultad original estaba enteramente en el **producto**, no en las funciones resultantes.

El caso resonante: cuando las frecuencias coinciden. La fórmula producto-a-suma sigue siendo válida cuando $a = b$, pero uno de los nuevos argumentos se vuelve cero.

Por ejemplo,

$$\cos(ax)\cos(ax) = \frac{1}{2}\big[\cos(2ax) + \cos 0\big].$$

Como $\cos 0 = 1$, obtenemos

$$\boxed{\cos^2(ax) = \frac{1}{2}\big[1 + \cos(2ax)\big],}$$

que no es otra cosa que la fórmula de reducción de potencia utilizada en §18.5.

De modo análogo, $\sin(ax)\sin(ax) = \frac{1}{2}\big[1 - \cos(2ax)\big]$.

El caso $a = b$ no exige una teoría aparte. Simplemente produce una **frecuencia cero**, es decir, un término constante.

La reducción de potencia es un caso particular
Las fórmulas de §18.5 y las identidades producto-a-suma de esta sección no son técnicas desconectadas. La reducción $\sin^2 x = \frac{1-\cos(2x)}{2}$ es exactamente el caso $\alpha = \beta = x$ de la identidad para $\sin\alpha\sin\beta$.

Un segundo ejemplo: el producto de dos senos.

Ejemplo 18.7. **Ejemplo.** Calculemos

$$\int \sin(4x)\sin x\,dx.$$

Por Proposición 18.3,

$$\sin(4x)\sin x = \frac{1}{2}\big[\cos(3x) - \cos(5x)\big].$$

Entonces

$$\begin{aligned}\int \sin(4x)\sin x\,dx &= \frac{1}{2}\int \cos(3x)\,dx - \frac{1}{2}\int \cos(5x)\,dx \\ &= \frac{1}{6}\sin(3x) - \frac{1}{10}\sin(5x) + C.\end{aligned}$$

Así,

$$\boxed{\int \sin(4x)\sin x\,dx = \frac{\sin(3x)}{6} - \frac{\sin(5x)}{10} + C.}$$

La verificación por derivación recupera el producto original después de recombinar las funciones mediante la misma identidad producto-a-suma.

La frecuencia como medida de la transformación. La notación

$$\sin(ax), \qquad \cos(ax)$$

muestra explícitamente una escala a. No necesitamos todavía desarrollar una teoría de Fourier para aprovechar esta observación. Al multiplicar dos oscilaciones, las identidades anteriores generan las combinaciones

$$a+b \qquad \text{y} \qquad a-b.$$

Por eso un producto que parece más complicado puede convertirse en una suma de términos más sencillos.

Figura 18.5 visualiza la transformación de Proposición 18.3 sin introducir teoría de Fourier: solo registra las combinaciones $a+b$ y $a-b$ ya demostradas algebraicamente.

Este fenómeno también explica por qué algunas integrales definidas sobre intervalos completos se cancelan.

Consideremos, por ejemplo,

$$\int_0^{2\pi} \sin(3x)\sin x\,dx.$$

Producto-a-suma da

$$\sin(3x)\sin x = \frac{1}{2}\big[\cos(2x) - \cos(4x)\big].$$

Por tanto,

$$\begin{aligned}\int_0^{2\pi} \sin(3x)\sin x\,dx &= \frac{1}{2}\int_0^{2\pi} \cos(2x)\,dx - \frac{1}{2}\int_0^{2\pi} \cos(4x)\,dx \\ &= 0.\end{aligned}$$

No necesitamos extraer todavía una teoría general de ortogonalidad. El cálculo concreto basta para mostrar que las oscilaciones completas pueden cancelar sus contribuciones sobre un período común.

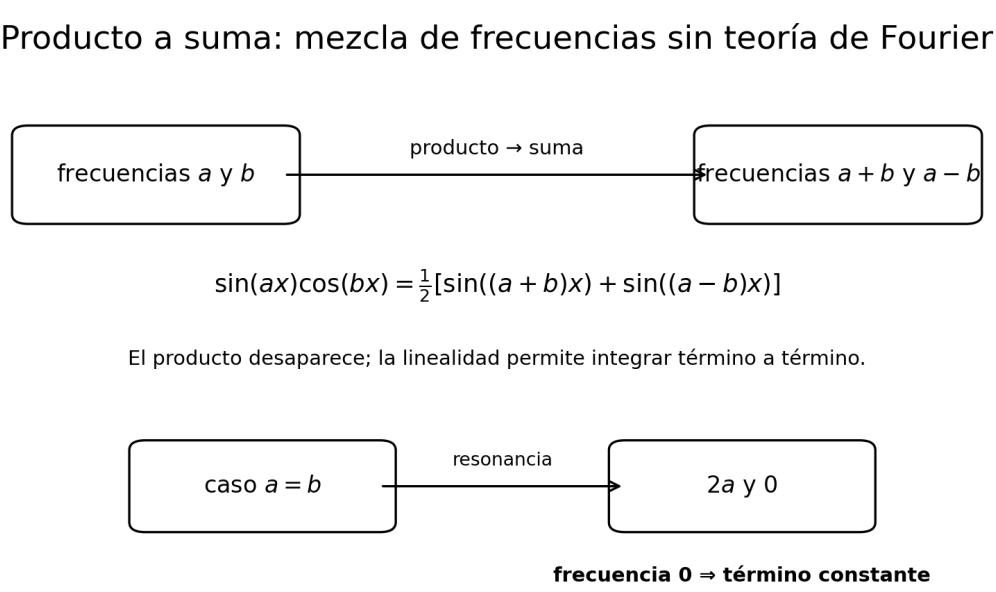


FIGURA 18.5. Las identidades producto-a-suma convierten dos frecuencias a y b en los modos $a+b$ y $a-b$; cuando $a=b$, la diferencia produce frecuencia 0 y aparece un término constante.

No confundir descomposición con sustitución
En §18.2 transformábamos la variable porque el integrando tenía la forma $f(g(x))g'(x)$. Aquí el paso decisivo es distinto: **reemplazamos una identidad por otra algebraicamente equivalente** y solo después integramos cada término. No toda transformación útil de un integrando es un cambio de variable.

Cuándo conviene usar producto-a-suma. Ante un producto trigonométrico, conviene distinguir dos situaciones. Si tenemos potencias del mismo argumento, $\sin^m x \cos^n x$, la paridad de los exponentes y las identidades de §18.5 suelen ofrecer la ruta más corta. Si tenemos funciones de argumentos distintos,

$$\sin(ax)\cos(bx), \qquad \cos(ax)\cos(bx), \qquad \sin(ax)\sin(bx),$$

producto-a-suma suele eliminar el producto en un solo paso. Podemos resumir el diagnóstico como

mismo argumento con potencias $\Longrightarrow$ paridad y reducción de potencia,
argumentos lineales distintos $\Longrightarrow$ producto-a-suma.

La regla no es una prohibición absoluta. Es un criterio para elegir la representación que produce menos trabajo.

Verificación: recomponer lo que descompusimos. Las respuestas obtenidas en esta sección pueden verificarse de dos maneras complementarias. La primera es derivar la candidata

final y comprobar que reaparece el integrando original. La segunda es derivar y luego utilizar Proposición 18.3 para recomponer la suma resultante en el producto inicial. En ambos casos, la identidad trigonométrica y la derivación sirven como controles mutuamente independientes de los signos y de los factores de frecuencia.

Antes de seguir
Calcula

$$\int \sin(5x)\cos(2x)\,dx.$$

Explica por qué una sustitución directa no elimina ambos factores y utiliza producto-a-suma para convertir el problema en dos integrales elementales.
Respuesta. Por Proposición 18.3,

$$\sin(5x)\cos(2x) = \frac{1}{2}\big[\sin(7x) + \sin(3x)\big].$$

Entonces

$$\begin{aligned}\int \sin(5x)\cos(2x)\,dx &= \frac{1}{2}\int \sin(7x)\,dx + \frac{1}{2}\int \sin(3x)\,dx \\ &= -\frac{1}{14}\cos(7x) - \frac{1}{6}\cos(3x) + C.\end{aligned}$$

Por tanto,

$$\boxed{\int \sin(5x)\cos(2x)\,dx = -\frac{\cos(7x)}{14} - \frac{\cos(3x)}{6} + C.}$$

Una sustitución como $u = \sin(5x)$ absorbería un factor $\cos(5x)\,dx$, no $\cos(2x)\,dx$; de modo análogo, $u = \cos(2x)$ exigiría un factor $\sin(2x)\,dx$. Ninguna de esas derivadas está presente. Producto-a-suma elimina el obstáculo estructural.

Los productos trigonométricos de frecuencias distintas refuerzan así el principio general del capítulo: antes de buscar una primitiva, debemos elegir una representación que deje actuar a las reglas ya conocidas. En este caso no cambiamos la variable ni repetimos partes; **descomponemos un producto en una suma y explotamos la linealidad**.

18.7. Funciones racionales I: dividir y descomponer en factores lineales simples

Las secciones anteriores mostraron que una representación adecuada puede convertir una integral difícil en otra ya conocida. Las funciones racionales ofrecen un ejemplo algebraico especialmente sistemático.

Consideremos

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

donde P, Q son polinomios reales y Q no es el polinomio nulo. Antes de buscar una primitiva debemos responder dos preguntas:

1. ¿es la fracción **propia**, es decir, satisface $\deg P < \deg Q$?
2. si ya es propia, ¿cómo está factorizado el denominador?

La primera pregunta se resuelve mediante división polinómica; la segunda, en el caso que estudiaremos aquí, mediante **fracciones parciales**.

Primero separar la parte polinómica. Si $\deg P \geq \deg Q$, la división polinómica permite escribir

$$P(x) = S(x)Q(x) + T(x), \qquad \deg T < \deg Q.$$

Por tanto, en todo punto donde $Q(x) \neq 0$,

$$\boxed{\frac{P(x)}{Q(x)} = S(x) + \frac{T(x)}{Q(x)}.}$$

La integral queda separada en una parte polinómica y una función racional propia. Por ejemplo, $x^2 + 1 = (x-1)(x+1) + 2$, de modo que, para $x \neq 1$, $\frac{x^2+1}{x-1} = x + 1 + \frac{2}{x-1}$. Así, sobre cualquier intervalo que no contenga 1,

$$\begin{aligned}\int \frac{x^2+1}{x-1}\,dx &= \int (x+1)\,dx + 2\int \frac{dx}{x-1} \\ &= \frac{x^2}{2} + x + 2\ln|x-1| + C.\end{aligned}$$

El valor absoluto no es opcional: Corolario 17.4 estableció que

$$\frac{d}{dx}\ln|x-a| = \frac{1}{x-a} \qquad (x \neq a).$$

El problema propio con raíces simples. Supongamos ahora que $\deg P < r$ y que el denominador tiene la forma

$$Q(x) = \prod_{j=1}^{r}(x - a_j),$$

donde $a_1, \ldots, a_r$ son números reales distintos. Queremos reemplazar la fracción racional por una suma de términos elementales:

$$\frac{P(x)}{Q(x)} \stackrel{?}{=} \sum_{j=1}^{r} \frac{A_j}{x - a_j}.$$

La igualdad, si existe, no se adivina. Los coeficientes quedan determinados por los valores de P en las raíces del denominador.

Proposición 18.4. **Proposición (fracciones parciales para factores lineales simples).** Sean $a_1, \ldots, a_r \in \mathbb{R}$ distintos y sea

$$Q(x) = \prod_{j=1}^{r}(x - a_j).$$

Si $P \in \mathbb{R}[x]$ satisface $\deg P < r$, entonces existen únicos números reales $A_1, \ldots, A_r$ tales que

$$\boxed{\frac{P(x)}{Q(x)} = \sum_{j=1}^{r} \frac{A_j}{x - a_j}}$$

para todo

$$x \notin \{a_1, \dots, a_r\}.$$

Además,

$$\boxed{A_j = \frac{P(a_j)}{\displaystyle\prod_{k \neq j}(a_j - a_k)} \qquad (1 \leq j \leq r).}$$

Demostración. Multiplicar la identidad buscada por $Q(x)$ conduce a

$$P(x) \stackrel{?}{=} \sum_{j=1}^{r} A_j \prod_{k \neq j}(x - a_k).$$

Si esta identidad es verdadera, al evaluar en $x = a_j$ todos los sumandos salvo uno se anulan y obtenemos necesariamente

$$P(a_j) = A_j \prod_{k \neq j}(a_j - a_k).$$

Como los a_j son distintos,

$$a_j - a_k \neq 0 \qquad (k \neq j),$$

de modo que el producto del denominador es no nulo y

$$A_j = \frac{P(a_j)}{\displaystyle\prod_{k \neq j}(a_j - a_k)}.$$

Esto prueba la **unicidad** de los posibles coeficientes. Para probar la existencia, definamos esos números A_j mediante la fórmula anterior y consideremos

$$H(x) := P(x) - \sum_{j=1}^{r} A_j \prod_{k \neq j}(x - a_k).$$

Cada sumando tiene grado a lo sumo $r - 1$, y $\deg P < r$; por tanto, $\deg H \leq r - 1$ salvo que H sea el polinomio nulo. Por la elección de los coeficientes,

$$H(a_j) = 0 \qquad (1 \leq j \leq r).$$

Así, H tiene al menos r raíces reales distintas. Un polinomio no nulo de grado a lo sumo $r - 1$ no puede tener r raíces distintas. Luego $H \equiv 0$, y la identidad polinómica queda demostrada. Dividiendo por $Q(x)$ en su dominio obtenemos la descomposición racional. ■

La proposición convierte una sola fracción en una suma de términos cuya primitiva ya conocemos por Corolario 17.4. En efecto,

$$\boxed{\int \frac{P(x)}{\prod_{j=1}^{r}(x - a_j)}\, dx = \sum_{j=1}^{r} A_j \ln|x - a_j| + C}$$

sobre cada intervalo que no contenga ninguna de las raíces a_j.

Un ejemplo con dos polos reales.

Ejemplo 18.8. **Ejemplo.** Calculemos

$$\int \frac{3x+5}{(x-1)(x+2)}\,dx.$$

La fracción ya es propia. Buscamos $\frac{3x+5}{(x-1)(x+2)} = \frac{A}{x-1} + \frac{B}{x+2}$. Por Proposición 18.4, para $a_1 = 1$,

$$A = \frac{3(1)+5}{1-(-2)} = \frac{8}{3},$$

y para $a_2 = -2$, $B = \frac{3(-2)+5}{-2-1} = \frac{1}{3}$.
Por tanto, $\frac{3x+5}{(x-1)(x+2)} = \frac{8}{3(x-1)} + \frac{1}{3(x+2)}$.
Integramos término a término:

$$\boxed{\int \frac{3x+5}{(x-1)(x+2)}\,dx = \frac{8}{3}\ln|x-1| + \frac{1}{3}\ln|x+2| + C.}$$

La verificación puede hacerse diferenciando la respuesta y recombinando las dos fracciones sobre el denominador común.

El dominio no desaparece al descomponer. La función racional del ejemplo no está definida en

$$x = -2 \qquad \text{ni en} \qquad x = 1.$$

Su dominio es $(-\infty,-2)\cup(-2,1)\cup(1,\infty)$.

En cada uno de esos intervalos, todas las primitivas difieren por una constante. Pero las constantes de intervalos distintos son **independientes**.

Un solo símbolo C se interpreta intervalo por intervalo
Cuando escribimos $\frac{8}{3}\ln|x-1| + \frac{1}{3}\ln|x+2| + C$, la expresión describe la familia de primitivas sobre un intervalo contenido en el dominio. Sobre todo el dominio desconectado no existe una razón para imponer la misma constante en las tres componentes. La unicidad «salvo constante» de Proposición 16.2 es un resultado sobre intervalos.

Este detalle también explica por qué una simplificación algebraica nunca debe hacernos olvidar los puntos excluidos del dominio original.

Coeficientes por sustitución, no por magia. En el caso de dos factores es frecuente escribir

$$\frac{P(x)}{(x-a)(x-b)} = \frac{A}{x-a} + \frac{B}{x-b}$$

y luego «tapar» un factor para encontrar A o B. El procedimiento es legítimo porque, después de multiplicar por el denominador, $P(x) = A(x-b) + B(x-a)$, y sustituir $x = a$ o $x = b$ anula exactamente uno de los dos términos.

El método de cobertura es evaluación de una identidad polinómica
No estamos cancelando informalmente un factor en un punto donde la fracción racional no está definida. Primero multiplicamos y obtenemos una identidad entre **polinomios**, válida para todo x; después evaluamos esa identidad en las raíces del denominador.

La distinción importa porque evita introducir una sustitución ilegítima en la función racional original.

Un caso con tres factores.

Ejemplo 18.9. **Ejemplo.** Descompongamos

$$\frac{2x^2+3x-2}{(x-1)(x+1)(x+2)}.$$

Por Proposición 18.4,

$$A=\frac{2(1)^2+3(1)-2}{(1+1)(1+2)}=\frac{1}{2},$$

para el factor $x-1$;

$$B=\frac{2(-1)^2+3(-1)-2}{(-1-1)(-1+2)}=\frac{3}{2},$$

para el factor $x+1$; y

$$D=\frac{2(-2)^2+3(-2)-2}{(-2-1)(-2+1)}=0$$

para el factor $x+2$.
Así,

$$\frac{2x^2+3x-2}{(x-1)(x+1)(x+2)}=\frac{1}{2(x-1)}+\frac{3}{2(x+1)}.$$

Que uno de los coeficientes resulte cero no constituye una excepción: significa simplemente que ese término no contribuye a la descomposición. Por tanto, sobre cada componente del dominio,

$$\boxed{\int \frac{2x^2+3x-2}{(x-1)(x+1)(x+2)}\,dx=\frac{1}{2}\ln|x-1|+\frac{3}{2}\ln|x+1|+C.}$$

Qué hemos resuelto y qué queda fuera. La combinación de división polinómica y Proposición 18.4 resuelve una clase completa:

función racional
$\Downarrow$
división si es impropia
$\Downarrow$
fracción propia con raíces reales simples
$\Downarrow$
suma de $\dfrac{A_j}{x-a_j}$
$\Downarrow$
logaritmos $\ln|x-a_j|$.

No hemos tratado todavía dos situaciones diferentes:

- factores lineales **repetidos**, como $(x-a)^m$ con $m \geq 2$;
- factores cuadráticos reales **sin raíces reales**.

La primera exigirá ampliar el patrón de fracciones parciales; la segunda requerirá estudiar primitivas que no se reducen únicamente a logaritmos mediante la teoría disponible hasta aquí.

Antes de seguir
Calcula

$$\int \frac{2x+1}{(x-2)(x+1)}\,dx.$$

Determina primero los coeficientes de las fracciones parciales y especifica el dominio sobre el que se interpreta cada familia de primitivas.
Respuesta. Buscamos $\frac{2x+1}{(x-2)(x+1)} = \frac{A}{x-2} + \frac{B}{x+1}$.
Por Proposición 18.4,

$$A = \frac{2(2)+1}{2-(-1)} = \frac{5}{3},$$

y $B = \frac{2(-1)+1}{-1-2} = \frac{1}{3}$.
Así, $\frac{2x+1}{(x-2)(x+1)} = \frac{5}{3(x-2)} + \frac{1}{3(x+1)}$, y por tanto

$$\boxed{\int \frac{2x+1}{(x-2)(x+1)}\,dx = \frac{5}{3}\ln|x-2| + \frac{1}{3}\ln|x+1| + C.}$$

El dominio es $(-\infty,-1)\cup(-1,2)\cup(2,\infty)$, y el símbolo C puede tomar un valor independiente en cada componente.

Las fracciones parciales muestran otra forma de reconocer estructura. No estamos invirtiendo directamente una regla de derivación: **desarmamos algebraicamente una función racional hasta convertirla en una suma de derivadas logarítmicas conocidas**. El siguiente paso será ampliar esta descomposición cuando los factores lineales se repiten, sin abandonar todavía el repertorio de primitivas ya construido.

18.8. Funciones racionales II: factores lineales repetidos

La sección anterior resolvió el caso en que el denominador propio se descompone como producto de factores lineales **distintos**. Allí cada raíz a_j producía un único término

$$\frac{A_j}{x-a_j}.$$

Cuando un factor se repite, esta forma ya no contiene suficientes grados de libertad. Si aparece, por ejemplo, $(x-a)^2$, no basta en general con escribir $\frac{A}{(x-a)^2}$. Debemos permitir también el término de orden inferior $\frac{B}{x-a}$. Para una potencia $(x-a)^m$, el patrón completo será

$$\boxed{\frac{A_1}{x-a}+\frac{A_2}{(x-a)^2}+\cdots+\frac{A_m}{(x-a)^m}.}$$

La repetición del factor en el denominador produce, por tanto, una **cadena completa de potencias recíprocas**.

Por qué aparecen todas las potencias. Consideremos primero un solo factor repetido. Sea P un polinomio con $\deg P < m$. Como cualquier polinomio puede escribirse en potencias de $x-a$, existen únicos coeficientes $c_0, c_1, \ldots, c_{m-1}$ tales que

$$P(x) = c_0 + c_1(x-a) + \cdots + c_{m-1}(x-a)^{m-1}.$$

Dividiendo por $(x-a)^m$ obtenemos

$$\frac{P(x)}{(x-a)^m} = \frac{c_{m-1}}{x-a} + \frac{c_{m-2}}{(x-a)^2} + \cdots + \frac{c_0}{(x-a)^m}.$$

Así, incluso antes de combinar distintos factores, la forma de la descomposición ya viene impuesta por la expansión polinómica alrededor de la raíz repetida.

La descomposición general cuando todas las raíces son reales. La idea anterior se combina con §18.7 cuando aparecen varias raíces con multiplicidades distintas.

Proposición 18.5. **Proposición (fracciones parciales con factores lineales repetidos).** Sean

$$a_1, \ldots, a_r \in \mathbb{R}$$

distintos, sean

$$m_1, \ldots, m_r \in \mathbb{N}, \qquad m_j \geq 1,$$

y definamos

$$Q(x) = \prod_{j=1}^{r} (x-a_j)^{m_j}.$$

Si $P \in \mathbb{R}[x]$ satisface

$$\deg P < \deg Q,$$

entonces existen únicos números reales $A_{j,k}$, con

$$1 \le j \le r, \qquad 1 \le k \le m_j,$$

tales que

$$\boxed{\frac{P(x)}{Q(x)} = \sum_{j=1}^{r} \sum_{k=1}^{m_j} \frac{A_{j,k}}{(x-a_j)^k}}$$

para todo x del dominio de la función racional.

Demostración. Sea

$$N := \deg Q = \sum_{j=1}^{r} m_j.$$

Para cada j escribimos

$$S_j(x) := \frac{Q(x)}{(x-a_j)^{m_j}} = \prod_{i \neq j} (x-a_i)^{m_i}.$$

Como las raíces son distintas, $S_j(a_j) \neq 0$. La identidad racional buscada equivale, después de multiplicar por $Q(x)$, a la identidad polinómica

$$P(x) = \sum_{j=1}^{r} \sum_{k=1}^{m_j} A_{j,k} \frac{Q(x)}{(x-a_j)^k}.$$

Para abreviar, definamos

$$\psi_{j,k}(x) := \frac{Q(x)}{(x-a_j)^k} = (x-a_j)^{m_j-k} S_j(x).$$

Fijemos una raíz a_j. Todo término $\psi_{i,k}$ con $i \neq j$ contiene el factor $(x-a_j)^{m_j}$, de modo que él y sus primeras $m_j - 1$ derivadas se anulan en $x = a_j$. Dentro de la familia correspondiente al mismo j, el término $\psi_{j,m_j}(x) = S_j(x)$ no se anula en a_j. Por tanto, la condición de que

$$H(x) := P(x) - \sum_{j=1}^{r} \sum_{k=1}^{m_j} A_{j,k} \psi_{j,k}(x)$$

satisfaga $H(a_j) = 0$ determina de manera única A_{j,m_j}. Después, la condición $H'(a_j) = 0$ determina de manera única A_{j,m_j-1}: el coeficiente con el que aparece ese término es $1!\, S_j(a_j) \neq 0$, mientras que los coeficientes de orden superior ya fueron fijados y los de orden inferior todavía contienen al menos $(x-a_j)^2$. Continuando de este modo, para $\ell = 0, 1, \ldots, m_j - 1$, la condición $H^{(\ell)}(a_j) = 0$ determina sucesivamente y de manera única $A_{j,m_j}, A_{j,m_j-1}, \ldots, A_{j,1}$. El procedimiento puede realizarse independientemente en cada raíz porque los términos asociados a las demás raíces se anulan allí con multiplicidad al menos m_j. Una vez elegidos todos los coeficientes, tenemos

$$H^{(\ell)}(a_j) = 0 \qquad (0 \le \ell < m_j)$$

para cada j. Por Corolario 11.1, si escribimos H en potencias centradas en a_j, las igualdades anteriores anulan exactamente sus primeros m_j coeficientes. Por tanto, $(x-a_j)^{m_j}$ divide a $H(x)$.

Como esto ocurre para cada j y los factores $(x-a_j)^{m_j}$ asociados a raíces distintas son coprimos, su producto $Q(x)$ divide a $H(x)$.

Por otra parte, $\deg H < N = \deg Q$, porque $\deg P < N$ y cada $\psi_{j,k}$ tiene grado a lo sumo $N-1$. La única posibilidad es, por tanto, $H \equiv 0$. Esto prueba la existencia de la descomposición. La construcción sucesiva de los coeficientes prueba al mismo tiempo su unicidad. ■

Integrar la cadena no exige nuevas funciones. La ventaja del resultado anterior es que cada término tiene una primitiva ya disponible.

Para $k = 1$,

$$\int \frac{A_{j,1}}{x-a_j}\,dx = A_{j,1}\ln|x-a_j| + C.$$

Para $k \geq 2$, la regla de la potencia y el cambio lineal dan

$$\int \frac{A_{j,k}}{(x-a_j)^k}\,dx = -\frac{A_{j,k}}{(k-1)(x-a_j)^{k-1}} + C.$$

Por tanto, sobre cada intervalo contenido en el dominio,

$$\boxed{\int \frac{P(x)}{Q(x)}\,dx = \sum_{j=1}^{r}\left[A_{j,1}\ln|x-a_j| - \sum_{k=2}^{m_j}\frac{A_{j,k}}{(k-1)(x-a_j)^{k-1}}\right] + C.}$$

Los factores repetidos no introducen, pues, una nueva clase de funciones elementales: añaden términos racionales junto a los logaritmos ya presentes en el caso simple.

Un primer factor doble.

EJEMPLO 18.10. **Ejemplo.** Calculemos

$$\int \frac{2x+3}{(x-1)^2}\,dx.$$

En lugar de montar un sistema desde cero, escribimos el numerador en potencias de $x-1$: $2x+3 = 2(x-1)+5$.

Por tanto, $\frac{2x+3}{(x-1)^2} = \frac{2}{x-1} + \frac{5}{(x-1)^2}$.

Integramos término a término:

$$\boxed{\int \frac{2x+3}{(x-1)^2}\,dx = 2\ln|x-1| - \frac{5}{x-1} + C.}$$

La función racional está definida en $(-\infty,1)\cup(1,\infty)$, y, como antes, la constante puede elegirse independientemente en cada componente.

Un factor repetido junto a otro factor simple.

EJEMPLO 18.11. **Ejemplo.** Calculemos

$$\int \frac{dx}{(x-1)^2(x+2)}.$$

La presencia del factor doble $(x-1)^2$ exige dos términos asociados a la raíz 1: $\frac{1}{(x-1)^2(x+2)} = \frac{A}{x-1} + \frac{B}{(x-1)^2} + \frac{C}{x+2}$. Multiplicando por el denominador común obtenemos la identidad polinómica

$$1 = A(x-1)(x+2) + B(x+2) + C(x-1)^2.$$

Evaluando en $x = 1$, $1 = 3B$, de modo que $B = \frac{1}{3}$.

Evaluando en $x = -2$, $1 = 9C$, y por tanto $C = \frac{1}{9}$.

Para hallar A podemos comparar los coeficientes de x^2. El miembro izquierdo no contiene término cuadrático, mientras que el derecho aporta $A + C$.

Así, $A + C = 0$, y entonces $A = -\frac{1}{9}$.

La descomposición es

$$\frac{1}{(x-1)^2(x+2)} = -\frac{1}{9(x-1)} + \frac{1}{3(x-1)^2} + \frac{1}{9(x+2)}.$$

Integrando,

$$\boxed{\int \frac{dx}{(x-1)^2(x+2)} = -\frac{1}{9}\ln|x-1| - \frac{1}{3(x-1)} + \frac{1}{9}\ln|x+2| + C.}$$

El dominio sigue siendo $(-\infty,-2)\cup(-2,1)\cup(1,\infty)$, y la constante se interpreta por componente.

Qué cambia en el método de cobertura. En §18.7, cuando todas las raíces eran simples, evaluar la identidad polinómica en cada raíz determinaba inmediatamente el único coeficiente asociado a ella. Con una raíz repetida, la evaluación directa determina solamente el coeficiente del término de **mayor orden**. En el ejemplo anterior, sustituir $x = 1$ produjo $B = \frac{1}{3}$, que corresponde a $\frac{B}{(x-1)^2}$, pero no determinó A.

Una raíz repetida exige más de una condición
Si $(x-a)^m$ aparece en el denominador, hay m coeficientes asociados a esa raíz. La evaluación en $x = a$ proporciona una sola condición. Los coeficientes restantes pueden obtenerse mediante comparación de coeficientes, mediante otros valores convenientes o, de forma sistemática, mediante las derivadas sucesivas utilizadas en la demostración de Proposición 18.5.

Por eso el procedimiento de «tapar factores» deja de ser suficiente por sí solo cuando aparecen multiplicidades.

El orden de trabajo para una función racional totalmente escindida. Podemos reunir §18.7 y §18.8 en un único protocolo.

Ante

$$\frac{P(x)}{Q(x)},$$

con $P, Q \in \mathbb{R}[x]$:

1. si $\deg P \geq \deg Q$, efectuar primero la división polinómica;
2. factorizar el denominador de la parte propia;
3. para cada factor lineal $(x-a)^m$, introducir **todos** los términos $\frac{A_1}{x-a} + \frac{A_2}{(x-a)^2} + \cdots + \frac{A_m}{(x-a)^m}$;

4. determinar los coeficientes a partir de la identidad polinómica resultante;
5. integrar término a término;
6. conservar explícitamente los puntos excluidos del dominio.

Si Q se descompone completamente en factores lineales reales, este procedimiento reduce toda función racional a polinomios, logaritmos y potencias enteras negativas.

Antes de seguir
Calcula

$$\int \frac{x+3}{(x-2)^2(x+1)}\,dx.$$

Escribe primero la cadena completa asociada al factor repetido y determina los tres coeficientes.
Respuesta. Buscamos $\frac{x+3}{(x-2)^2(x+1)} = \frac{A}{x-2} + \frac{B}{(x-2)^2} + \frac{C}{x+1}$.
Multiplicando por $(x-2)^2(x+1)$,

$$x+3 = A(x-2)(x+1) + B(x+1) + C(x-2)^2.$$

En $x=2$,

$$5 = 3B, \qquad B = \frac{5}{3}.$$

En $x=-1$,

$$2 = 9C, \qquad C = \frac{2}{9}.$$

Comparando los coeficientes de x^2, $A + C = 0$, de modo que $A = -\frac{2}{9}$.
Así,

$$\frac{x+3}{(x-2)^2(x+1)} = -\frac{2}{9(x-2)} + \frac{5}{3(x-2)^2} + \frac{2}{9(x+1)},$$

y por tanto

$$\boxed{\int \frac{x+3}{(x-2)^2(x+1)}\,dx = -\frac{2}{9}\ln|x-2| - \frac{5}{3(x-2)} + \frac{2}{9}\ln|x+1| + C.}$$

El dominio es $(-\infty,-1)\cup(-1,2)\cup(2,\infty)$, y el símbolo C se interpreta independientemente en cada componente.

Con los factores lineales repetidos queda resuelta toda función racional cuyo denominador se escinde completamente en factores lineales reales. Si una factorización real disponible contiene además un **factor cuadrático irreducible**, aparece un bloque que las técnicas anteriores no cubren. En este capítulo no demostraremos el teorema algebraico general de factorización de polinomios reales ni la existencia general de la descomposición en fracciones parciales; nuestro objetivo será construir, de manera autosuficiente, las primitivas de los bloques lineales y cuadráticos una vez que esa descomposición algebraica esté dada.

18.9. Funciones racionales III: cuadráticos irreducibles y la aparición de arctan

Las dos secciones anteriores resolvieron todos los denominadores que se descomponen completamente en factores lineales reales. Queda ahora el primer caso que no puede reducirse de ese modo: un factor cuadrático sin raíces reales.

El ejemplo mínimo es x^2+1.

No existen $a, b \in \mathbb{R}$ tales que $x^2+1=(x-a)(x-b)$, porque la ecuación $x^2+1=0$ no tiene soluciones reales. En consecuencia, una fracción como $\frac{1}{x^2+1}$ no puede descomponerse sobre $\mathbb{R}$ en una suma de términos de la forma $A/(x-a)$. Esto señala una diferencia importante con §18.7–§18.8: el obstáculo ya no es hallar más coeficientes de fracciones parciales. Falta una **nueva primitiva elemental**.

Construir la nueva primitiva antes de usarla. Podríamos introducir una fórmula de memoria, pero eso rompería la lógica del libro. En lugar de ello construiremos la función que necesitamos a partir del teorema fundamental del cálculo.

Definición 18.1. **Definición (arco tangente por acumulación).** Para $x \in \mathbb{R}$ definimos

$$\boxed{\arctan x := \int_0^x \frac{dt}{1+t^2}.}$$

La función integrando $t \longmapsto \frac{1}{1+t^2}$ es continua en toda la recta, porque $1+t^2>0$ para todo $t \in \mathbb{R}$. Por tanto la integral está bien definida para cada x.

Aquí arctan **se construye analíticamente**
La notación arctan anticipa la función tradicionalmente llamada arco tangente, pero en este punto no necesitamos utilizarla como inversa de una función tangente previamente construida. Lo único que necesitaremos en este capítulo es la definición por acumulación y la derivada que se deduce de ella. Así evitamos introducir una propiedad inversa que todavía no forma parte de la cadena canónica de resultados.

Por el teorema fundamental, la nueva función trae consigo exactamente la derivada que necesitamos.

Proposición 18.6. **Proposición (primitiva cuadrática básica y forma escalada).** La función definida en Definición 18.1 es diferenciable en $\mathbb{R}$ y satisface

$$\boxed{\frac{d}{dx}\arctan x = \frac{1}{1+x^2}.}$$

Además, si $a>0$ y $b \in \mathbb{R}$, entonces

$$\boxed{\frac{d}{dx}\left[\frac{1}{a}\arctan\left(\frac{x-b}{a}\right)\right] = \frac{1}{(x-b)^2+a^2}.}$$

En consecuencia,

$$\boxed{\int \frac{dx}{(x-b)^2+a^2} = \frac{1}{a}\arctan\left(\frac{x-b}{a}\right)+C.}$$

Demostración. Fijemos $c \in \mathbb{R}$. Elegimos $R > |c| + 1$ y trabajamos en el intervalo compacto $[-R, R]$. Definamos

$$A_R(x) := \int_{-R}^{x} \frac{dt}{1+t^2}.$$

La función integrando es continua en $[-R, R]$, así que Teorema 16.2 da

$$A_R'(x) = \frac{1}{1+x^2} \qquad (x \in (-R, R)).$$

Por aditividad orientada Proposición 14.3, para todo $x \in [-R, R]$,

$$\arctan x = \int_0^x \frac{dt}{1+t^2} = A_R(x) - A_R(0).$$

Como $c \in (-R, R)$, el término $A_R(0)$ es constante respecto de x y, por tanto, $(\arctan)'(c) = A_R'(c) = \frac{1}{1+c^2}$. El punto c fue arbitrario, de modo que la identidad vale en toda $\mathbb{R}$. Para la forma escalada definimos

$$F(x) = \frac{1}{a} \arctan\left(\frac{x-b}{a}\right).$$

Por la regla de la cadena Teorema 7.4 y la primera parte,

$$\begin{aligned} F'(x) &= \frac{1}{a} \frac{1}{1 + \left(\frac{x-b}{a}\right)^2} \frac{1}{a} \\ &= \frac{1}{a^2 + (x-b)^2}. \end{aligned}$$

Esto prueba la identidad y, por la extensión de Proposición 16.2 establecida en §18.1, la fórmula integral sobre cualquier intervalo. ■

Figura 18.6 representa la construcción de Definición 18.1 y la derivada de Proposición 18.6 mediante el teorema fundamental, sin usar todavía arctan como inversa de la tangente.

Completar el cuadrado convierte el denominador en la forma básica. Consideremos ahora un cuadrático mónico $D(x) = x^2 + px + r$, con discriminante negativo: $\Delta = p^2 - 4r < 0$.

Completar el cuadrado da

$$x^2 + px + r = \left(x + \frac{p}{2}\right)^2 + \left(r - \frac{p^2}{4}\right).$$

Como $\Delta < 0$, $r - \frac{p^2}{4} > 0$.

Podemos escribir, por tanto,

$$a^2 := r - \frac{p^2}{4}, \qquad a > 0,$$

y obtenemos

$$\boxed{D(x) = \left(x + \frac{p}{2}\right)^2 + a^2.}$$

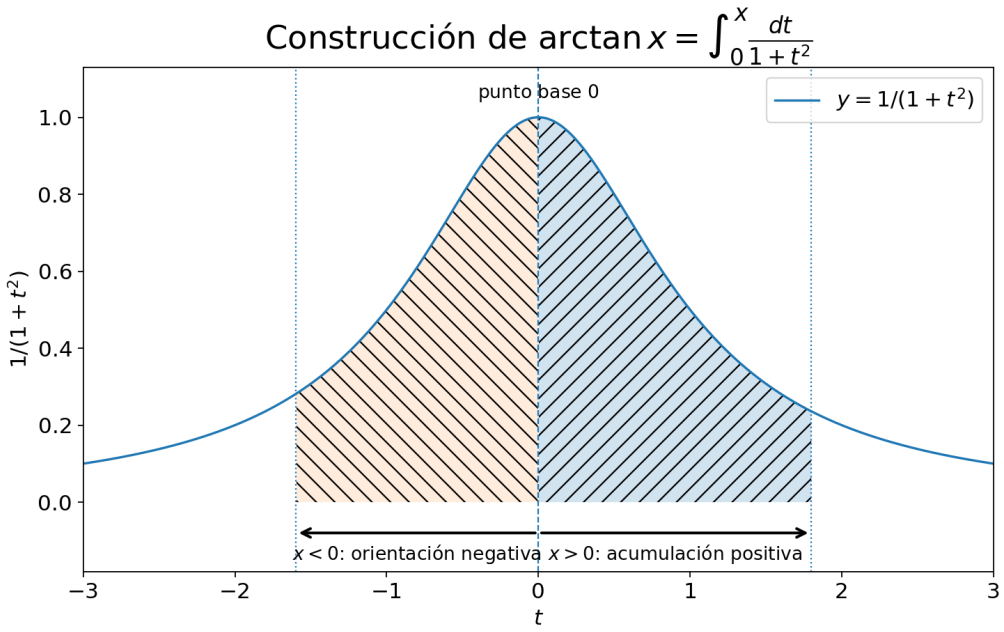


FIGURA 18.6. La función $\arctan x$ se construye aquí como acumulación de $1/(1+t^2)$ desde el punto base 0: para $x > 0$ la acumulación avanza hacia la derecha, para $x < 0$ la integral orientada cambia de signo y para $x = 0$ vale 0.

Ésta es exactamente la forma de Proposición 18.6.

El numerador lineal se separa en dos piezas. En una fracción parcial asociada a un cuadrático irreducible aparece un numerador de grado menor que 2: $mx + n$. La derivada del denominador es $D'(x) = 2x + p$.

Por eso conviene escribir $mx + n = \frac{m}{2}(2x + p) + \left(n - \frac{mp}{2}\right)$.

La primera parte produce un logaritmo; la segunda produce la nueva primitiva cuadrática.

PROPOSICIÓN 18.7. **Proposición (numerador lineal sobre cuadrático irreducible mónico).** Sean $m, n, p, r \in \mathbb{R}$ y supongamos

$$p^2 - 4r < 0.$$

Definamos

$$a := \sqrt{r - \frac{p^2}{4}} > 0.$$

Entonces

$$\boxed{\int \frac{mx+n}{x^2+px+r}\,dx = \frac{m}{2}\ln(x^2+px+r) + \frac{n-\frac{mp}{2}}{a}\arctan\left(\frac{x+\frac{p}{2}}{a}\right) + C.}$$

Demostración. Como $x^2+px+r=\left(x+\frac{p}{2}\right)^2+a^2>0$, el denominador no se anula en $\mathbb{R}$. Además, $mx+n=\frac{m}{2}(2x+p)+\left(n-\frac{mp}{2}\right)$.

Por linealidad,

$$\int \frac{mx+n}{x^2+px+r}\,dx=\frac{m}{2}\int \frac{2x+p}{x^2+px+r}\,dx + \left(n-\frac{mp}{2}\right)\int \frac{dx}{\left(x+\frac{p}{2}\right)^2+a^2}.$$

El primer término es $\frac{m}{2}\ln(x^2+px+r)$, porque el numerador es la derivada del denominador. El segundo se evalúa mediante Proposición 18.6. Sustituyendo ambas primitivas obtenemos la fórmula anunciada. ■

La fórmula no debe memorizarse como una entrada de tabla. Sus dos términos corresponden a dos huellas diferentes:

derivada del denominador $\Longrightarrow \ln D(x)$,
residuo constante después de separar esa derivada $\Longrightarrow \arctan$.

Un cuadrático irreducible con numerador no ajustado.

Ejemplo 18.12. **Ejemplo.** Calculemos

$$\int \frac{2x+3}{x^2+2x+5}\,dx.$$

El denominador se completa como $x^2+2x+5=(x+1)^2+4$.

Su derivada es $2x+2$.

Separamos el numerador: $2x+3=(2x+2)+1$.

Por tanto,

$$\int \frac{2x+3}{x^2+2x+5}\,dx=\int \frac{2x+2}{x^2+2x+5}\,dx+\int \frac{dx}{(x+1)^2+4} = \ln(x^2+2x+5)+\frac{1}{2}\arctan\left(\frac{x+1}{2}\right)+C.$$

Así,

$$\boxed{\int \frac{2x+3}{x^2+2x+5}\,dx=\ln(x^2+2x+5)+\frac{1}{2}\arctan\left(\frac{x+1}{2}\right)+C.}$$

La derivación de la respuesta reconstruye por separado $\frac{2x+2}{x^2+2x+5}$ y $\frac{1}{x^2+2x+5}$, cuya suma devuelve el integrando original.

Cómo entra el cuadrático irreducible en fracciones parciales. Si un denominador contiene un factor lineal y un cuadrático irreducible, la forma correcta ya no es $\frac{A}{x-a}+\frac{B}{q(x)}$. El numerador sobre el cuadrático debe tener grado menor que 2, por lo que necesitamos

$$\boxed{\frac{A}{x-a}+\frac{Bx+C}{q(x)}.}$$

La razón es la misma que apareció con los factores lineales repetidos: la descomposición debe contener suficientes grados de libertad para reconstruir cualquier numerador propio compatible con el denominador.

Ejemplo 18.13. **Ejemplo.** Calculemos

$$\int \frac{dx}{(x-1)(x^2+1)}.$$

Buscamos $\frac{1}{(x-1)(x^2+1)} = \frac{A}{x-1} + \frac{Bx+C}{x^2+1}$.

Multiplicando por el denominador común, $1 = A(x^2+1) + (Bx+C)(x-1)$.

Evaluando en $x = 1$,

$$1 = 2A, \qquad A = \frac{1}{2}.$$

Comparando coeficientes de x^2,

$$A + B = 0, \qquad B = -\frac{1}{2}.$$

Comparando coeficientes de x,

$$C - B = 0, \qquad C = -\frac{1}{2}.$$

Así,

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

Integramos término a término:

$$\boxed{\int \frac{dx}{(x-1)(x^2+1)} = \frac{1}{2}\ln|x-1| - \frac{1}{4}\ln(x^2+1) - \frac{1}{2}\arctan x + C.}$$

El único punto excluido del dominio es $x = 1$, porque $x^2 + 1 > 0$ para todo $x \in \mathbb{R}$. Por tanto, la constante puede elegirse independientemente en

$$(-\infty, 1) \qquad \text{y} \qquad (1, \infty).$$

Un factor cuadrático general puede normalizarse. Si aparece

$$ax^2 + bx + c, \qquad a \neq 0,$$

con discriminante $b^2 - 4ac < 0$, podemos extraer el coeficiente principal:

$$ax^2 + bx + c = a\left(x^2 + \frac{b}{a}x + \frac{c}{a}\right).$$

El factor constante a se absorbe en los coeficientes de la fracción parcial. Por eso basta dominar el caso mónico desarrollado arriba.

El cuadrático irreducible introduce una función nueva por una razón estructural
En los factores lineales, toda primitiva parcial se reducía a logaritmos y potencias negativas. Un cuadrático irreducible deja inevitablemente una pieza proporcional a $\frac{1}{u^2+a^2}$, que no es la derivada logarítmica de un polinomio real. La aparición de arctan no es un artificio de tabla: registra una estructura de derivación nueva.

Qué queda todavía fuera. Con esta sección podemos integrar fracciones parciales asociadas a **factores cuadráticos irreducibles simples**. Sin embargo, una función racional general puede contener potencias

$$q(x)^m, \qquad m \geq 2,$$

de un cuadrático irreducible.

En ese caso aparecen términos de la forma $\frac{A_k x+B_k}{q(x)^k}$, y para $k \geq 2$ necesitaremos una reducción adicional. Ese será el siguiente microtramo.

Antes de seguir
Calcula

$$\int \frac{3x+1}{x^2-4x+8}\,dx.$$

Completa el cuadrado y separa del numerador un múltiplo de la derivada del denominador.
Respuesta. Tenemos $x^2-4x+8=(x-2)^2+4$, y $3x+1=\frac{3}{2}(2x-4)+7$.
Por tanto,

$$\begin{aligned}\int \frac{3x+1}{x^2-4x+8}\,dx &= \frac{3}{2}\int \frac{2x-4}{x^2-4x+8}\,dx + 7\int \frac{dx}{(x-2)^2+4}\\ &= \frac{3}{2}\ln(x^2-4x+8)+\frac{7}{2}\arctan\left(\frac{x-2}{2}\right)+C.\end{aligned}$$

Así,

$$\boxed{\int \frac{3x+1}{x^2-4x+8}\,dx = \frac{3}{2}\ln(x^2-4x+8)+\frac{7}{2}\arctan\left(\frac{x-2}{2}\right)+C.}$$

Los cuadráticos irreducibles completan una pieza que faltaba en el repertorio de fracciones parciales: una parte del numerador sigue produciendo un logaritmo, mientras el residuo introduce la primitiva de $1/(1+x^2)$. La técnica continúa siendo la misma en espíritu: **reescribir hasta que cada término coincida con una derivada ya construida**.

18.10. Funciones racionales IV: cuadráticos irreducibles repetidos y reducción finita

La sección anterior resolvió el caso de un factor cuadrático irreducible simple. Queda por estudiar qué ocurre cuando ese mismo factor aparece con multiplicidad mayor que uno. Después de completar el cuadrado, el bloque esencial tiene la forma

$$\frac{Ax+B}{\left((x-b)^2+a^2\right)^n}, \qquad a>0, \qquad n \geq 2.$$

La dificultad nueva no está en el numerador lineal: como en §18.9, podemos separar una parte proporcional a la derivada del denominador. La pieza verdaderamente nueva es la familia

$$\int \frac{dx}{\left((x-b)^2+a^2\right)^n}.$$

El objetivo será reducir el exponente n hasta llegar a $n=1$, caso ya resuelto por Proposición 18.6.

Centrar primero el cuadrático. Escribamos $u = x - b$.
Entonces

$$(x-b)^2 + a^2 = u^2 + a^2, \qquad du = dx.$$

Por tanto basta entender, para $a > 0$,

$$J_n(u;a) := \int \frac{du}{(u^2+a^2)^n}.$$

Para $n = 1$, Proposición 18.6 da

$$\boxed{J_1(u;a) = \frac{1}{a}\arctan\!\left(\frac{u}{a}\right) + C.}$$

La pregunta es cómo relacionar J_n con J_{n-1} cuando $n \geq 2$.

La derivada que contiene la reducción. Consideremos la función $R_n(u) = \frac{u}{(u^2+a^2)^{n-1}}$.
Su derivada es

$$R_n'(u) = \frac{1}{(u^2+a^2)^{n-1}} - \frac{2(n-1)u^2}{(u^2+a^2)^n}.$$

Usamos ahora $u^2 = (u^2+a^2) - a^2$.
Entonces

$$\begin{aligned} R_n'(u) &= \frac{1}{(u^2+a^2)^{n-1}} - 2(n-1)\left[\frac{1}{(u^2+a^2)^{n-1}} - \frac{a^2}{(u^2+a^2)^n}\right] \\ &= -\frac{2n-3}{(u^2+a^2)^{n-1}} + \frac{2(n-1)a^2}{(u^2+a^2)^n}. \end{aligned}$$

La última identidad puede despejarse para aislar exactamente el integrando de J_n.

Proposición 18.8. **Proposición (reducción para potencias de u^2+a^2).** Sean $a > 0$ y $n \in \mathbb{N}$ con $n \geq 2$. Entonces

$$\boxed{J_n(u;a) = \frac{u}{2(n-1)a^2(u^2+a^2)^{n-1}} + \frac{2n-3}{2(n-1)a^2}J_{n-1}(u;a),}$$

donde

$$J_n(u;a) := \int \frac{du}{(u^2+a^2)^n}.$$

Demostración. De la identidad anterior obtenemos

$$\frac{1}{(u^2+a^2)^n} = \frac{1}{2(n-1)a^2} R_n'(u) + \frac{2n-3}{2(n-1)a^2} \frac{1}{(u^2+a^2)^{n-1}}.$$

Integramos ambos miembros. Como $R_n(u) = \frac{u}{(u^2+a^2)^{n-1}}$, resulta

$$J_n(u;a) = \frac{u}{2(n-1)a^2(u^2+a^2)^{n-1}} + \frac{2n-3}{2(n-1)a^2} J_{n-1}(u;a).$$

La constante queda absorbida en la familia de primitivas. ■

La reducción tiene una dirección inequívoca:

$$J_n \longrightarrow J_{n-1} \longrightarrow \cdots \longrightarrow J_2 \longrightarrow J_1.$$

La medida de complejidad que disminuye es el exponente del cuadrático. Por eso el proceso termina después de un número finito de pasos.

El primer caso repetido.

Ejemplo 18.14. **Ejemplo.** Calculemos

$$\int \frac{dx}{(x^2+1)^2}.$$

Aquí $a = 1$, $u = x$ y $n = 2$. Por Proposición 18.8, $J_2(x;1) = \frac{x}{2(x^2+1)} + \frac{1}{2} J_1(x;1)$. Como Proposición 18.6 da $J_1(x;1) = \arctan x + C$, obtenemos

$$\boxed{\int \frac{dx}{(x^2+1)^2} = \frac{x}{2(x^2+1)} + \frac{1}{2}\arctan x + C.}$$

La respuesta ya muestra el patrón general: al repetir el cuadrático aparecen términos racionales adicionales, mientras la cadena termina finalmente en la misma función arctan del caso simple.

El numerador lineal vuelve a dividirse en dos huellas. Sea ahora

$$q(x) = (x-b)^2 + a^2, \qquad a > 0.$$

Entonces $q'(x) = 2(x-b)$.

Todo numerador lineal puede escribirse como $mx + n = \frac{m}{2} q'(x) + (n + mb)$.

Por tanto, para $k \geq 2$,

$$\begin{aligned} \int \frac{mx+n}{q(x)^k}\,dx &= \frac{m}{2}\int \frac{q'(x)}{q(x)^k}\,dx + (n+mb)\int \frac{dx}{q(x)^k} \\ &= -\frac{m}{2(k-1)q(x)^{k-1}} + (n+mb)J_k(x-b;a) + C. \end{aligned}$$

No aparece, por tanto, una tercera familia de primitivas. La parte proporcional a q' se integra por una potencia; el residuo constante se reduce mediante Proposición 18.8 hasta J_1.

Separar primero la derivada del cuadrático
Ante $\frac{mx+n}{q(x)^k}$, con q cuadrático irreducible, conviene repetir exactamente la decisión de §18.9:

1. aislar un múltiplo de $q'(x)$ en el numerador;
2. integrar esa parte inmediatamente;
3. aplicar la reducción solo al residuo constante.

La fórmula de reducción no sustituye el diagnóstico estructural anterior: actúa después de él.

Un cuadrático desplazado y repetido.

Ejemplo 18.15. **Ejemplo.** Calculemos

$$\int \frac{2x+6}{(x^2-2x+5)^2}\,dx.$$

Completamos el cuadrado: $x^2-2x+5=(x-1)^2+4$.
Definamos $q(x)=x^2-2x+5$.
Entonces $q'(x)=2x-2$, y el numerador se descompone como $2x+6=(2x-2)+8$.
Por linealidad,

$$\int \frac{2x+6}{q(x)^2}\,dx=\int \frac{q'(x)}{q(x)^2}\,dx+8\int \frac{dx}{q(x)^2}.$$

El primer término es $-\frac{1}{q(x)}$.
Para el segundo tomamos

$$u=x-1, \qquad a=2.$$

Por Proposición 18.8,

$$\begin{aligned} J_2(u;2) &= \frac{u}{8(u^2+4)}+\frac{1}{8}J_1(u;2)\\ &= \frac{u}{8(u^2+4)}+\frac{1}{16}\arctan\!\left(\frac{u}{2}\right)+C. \end{aligned}$$

Multiplicando por 8 y regresando a x,

$$8J_2(x-1;2)=\frac{x-1}{q(x)}+\frac{1}{2}\arctan\!\left(\frac{x-1}{2}\right)+C.$$

Por tanto,

$$\boxed{\int \frac{2x+6}{(x^2-2x+5)^2}\,dx=\frac{x-2}{x^2-2x+5}+\frac{1}{2}\arctan\!\left(\frac{x-1}{2}\right)+C.}$$

Derivar la respuesta recupera exactamente el integrando y comprueba simultáneamente el término racional y el factor de escala del arco tangente.

Cómo aparece un cuadrático repetido en fracciones parciales. Si un factor cuadrático irreducible $q(x)$ aparece con multiplicidad m, la cadena correspondiente tiene la forma

$$\boxed{\frac{A_1x+B_1}{q(x)}+\frac{A_2x+B_2}{q(x)^2}+\cdots+\frac{A_mx+B_m}{q(x)^m}.}$$

Cada nivel necesita un numerador de grado menor que 2. Esta condición es el análogo cuadrático de la cadena $\frac{C_1}{x-a}+\frac{C_2}{(x-a)^2}+\cdots+\frac{C_m}{(x-a)^m}$ estudiada en §18.8. Una vez determinados algebraicamente los coeficientes, cada término cuadrático se resuelve mediante el mismo esquema:

$$\boxed{\begin{array}{c}\dfrac{A_kx+B_k}{q(x)^k}\\ \Downarrow\\ \text{separar un múltiplo de } q'(x)\\ \Downarrow\\ \text{potencia de } q \quad + \quad \text{constante}\cdot J_k\\ \Downarrow\\ J_k\to J_{k-1}\to\cdots\to J_1.\end{array}}$$

La parte algebraica y la parte analítica son problemas distintos
En este tomo no demostramos ni el teorema algebraico general de factorización de polinomios reales en factores lineales y cuadráticos irreducibles ni el teorema general de existencia y unicidad de una descomposición completa en fracciones parciales. Esos datos algebraicos deben estar disponibles antes de aplicar la taxonomía que sigue. El resultado de §§18.7–18.10 es otro y es estrictamente analítico: **una vez dada una descomposición en esos bloques lineales y cuadráticos, cada bloque posee una primitiva construida con resultados ya establecidos en este capítulo**.

Cierre analítico de las funciones racionales. Las §§18.7–18.10 permiten organizar todos los bloques reales de una descomposición racional:

$$\boxed{\begin{array}{lcl}\dfrac{A}{x-a} & \longrightarrow & \ln|x-a|,\\ \dfrac{A}{(x-a)^k},\ k\ge 2 & \longrightarrow & \text{potencias negativas},\\ \dfrac{Ax+B}{q(x)} & \longrightarrow & \ln q(x)+\arctan,\\ \dfrac{Ax+B}{q(x)^k},\ k\ge 2 & \longrightarrow & \text{términos racionales}+J_k\longrightarrow\arctan.\end{array}}$$

Aquí q denota un cuadrático irreducible real. **Condicionalmente a que la parte propia ya haya sido descompuesta en los bloques anteriores**, el cierre analítico es completo: no queda una nueva familia de primitivas racionales por construir y los exponentes repetidos se reducen finitamente a los casos base.

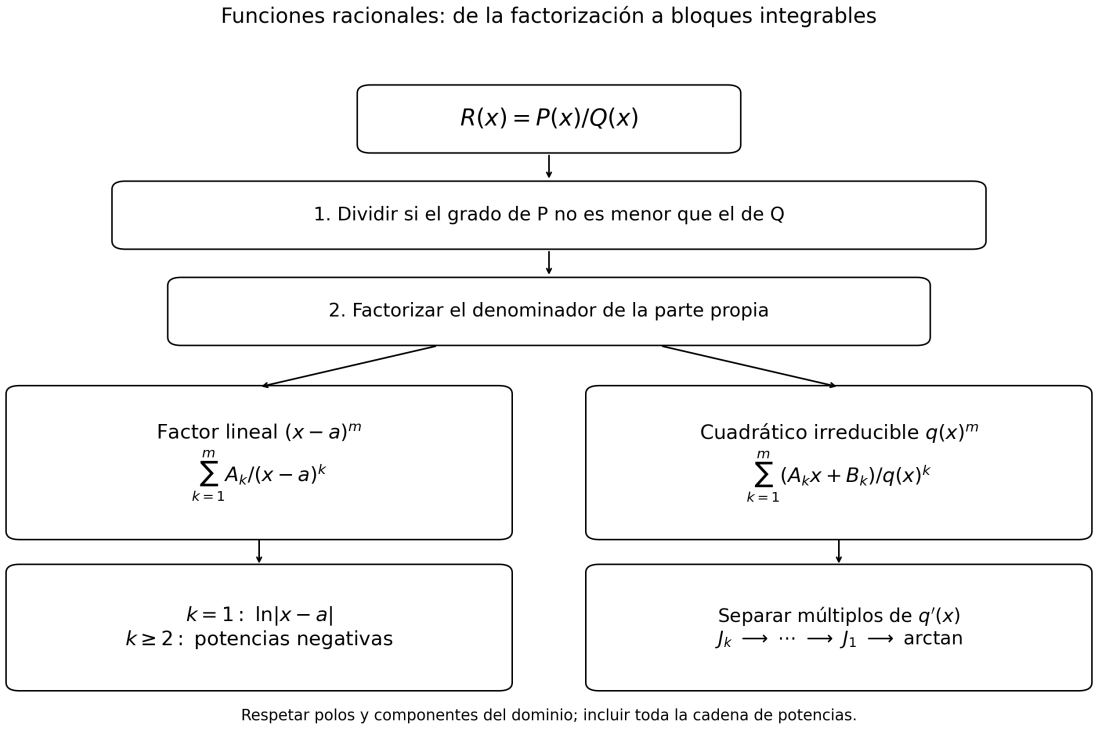


FIGURA 18.7. Mapa de las funciones racionales: una fracción impropia se divide primero; la parte propia se descompone según factores lineales o cuadráticos irreducibles, simples o repetidos. Cada bloque conduce a logaritmos, potencias negativas, términos racionales o arctan, y las potencias cuadráticas descienden mediante $J_k \to \cdots \to J_1$. Los polos lineales separan las componentes del dominio real.

Figura 18.7 reúne Proposición 18.4, Proposición 18.5, Proposición 18.6, Proposición 18.7 y Proposición 18.8 en una sola taxonomía y mantiene separadas la **descomposición algebraica** y la **integración analítica** de los bloques.

Antes de seguir
Usa Proposición 18.8 para calcular

$$\int \frac{dx}{(x^2+1)^3}.$$

Identifica explícitamente el descenso $J_3 \to J_2 \to J_1$.
Respuesta. Para $a = 1$ y $n = 3$, $J_3(x;1) = \frac{x}{4(x^2+1)^2} + \frac{3}{4}J_2(x;1)$.
Por Ejemplo 18.14,

$$J_2(x;1) = \frac{x}{2(x^2+1)} + \frac{1}{2}\arctan x + C.$$

Sustituyendo,

$$\boxed{\int \frac{dx}{(x^2+1)^3} = \frac{x}{4(x^2+1)^2} + \frac{3x}{8(x^2+1)} + \frac{3}{8}\arctan x + C.}$$

La derivación de la expresión final devuelve $1/(x^2+1)^3$.

Los factores cuadráticos irreducibles repetidos completan los bloques de fracciones parciales estudiados. Para combinar las técnicas tendremos que reconocer la estructura del integrando, elegir una representación que facilite el cálculo y verificar la primitiva obtenida.

18.11. Estrategia global: reconocer estructura, elegir representación y verificar

Las secciones anteriores desarrollaron técnicas diferentes, pero todas respondieron a una misma dificultad: el integrando rara vez anuncia por sí solo qué primitiva lo produjo. Integrar exige **leer estructura**, y esa lectura depende con frecuencia de la representación elegida. La elección del método reúne las decisiones anteriores:

La primera pregunta ocurre antes de integrar. Frente a

$$\int f(x)\,dx,$$

la pregunta inicial no debería ser

> «¿qué técnica aplico?»

sino

> **«¿en qué forma conviene escribir el integrando para que alguna estructura conocida se vuelva visible?»**

A veces la respuesta es puramente algebraica. En §18.7 dividimos una función racional impropia antes de descomponerla; en §§18.5–18.6 usamos identidades trigonométricas antes de efectuar ninguna integración; en §§18.9–18.10 completamos cuadrados antes de reconocer los bloques elementales. Así, el primer movimiento puede ser

$$\boxed{\text{simplificar, factorizar, dividir, descomponer o reescribir}}$$

sin que todavía hayamos aplicado una técnica de integración propiamente dicha.

Un protocolo de decisión. El recorrido construido en este capítulo puede organizarse mediante las preguntas siguientes.

1. ¿Hay una simplificación algebraica previa? Antes de introducir nuevas variables o integrar por partes, conviene buscar cancelaciones legítimas, división polinómica, factorizaciones o identidades que reduzcan la complejidad visible. Toda simplificación debe conservar el dominio relevante.

2. ¿Aparece la huella de una composición? Si podemos reconocer una estructura $f(g(x))g'(x)$, exacta o salvo un factor constante, Teorema 18.1 convierte la regla de la cadena en un cambio de variable. El control decisivo es que, después de la sustitución, **desaparezca la variable antigua**.

3. ¿Aparece una estructura de producto? Si el integrando puede escribirse como $u(x)v'(x)$ y diferenciar u reduce alguna complejidad mientras una primitiva de v' es accesible, Teorema 18.2 puede transferir el problema a una integral más sencilla. Si la nueva integral es más complicada, la aplicación es correcta pero estratégicamente mala.

4. ¿La dificultad es trigonométrica y está en la representación? Para potencias del mismo argumento, Proposición 18.2 usa la paridad para preservar el factor que funcionará como diferencial. Para productos de frecuencias lineales distintas, Proposición 18.3 transforma el producto en una suma. En ambos casos las identidades preparan el integrando para técnicas ya conocidas.

5. ¿El integrando es racional? Entonces el orden de trabajo es estructuralmente distinto:

1. dividir primero si la fracción es impropia;
2. factorizar el denominador sobre $\mathbb{R}$;
3. descomponer la parte propia en fracciones parciales;
4. distinguir factores lineales y cuadráticos irreducibles, con sus multiplicidades;
5. integrar cada bloque mediante Proposición 18.4, Proposición 18.5, Proposición 18.7 y Proposición 18.8.

La descomposición algebraica determina **qué bloques aparecen**; los resultados analíticos del capítulo determinan **cómo integrar cada bloque**. Podemos condensar estas preguntas en el esquema

integrando
$\Downarrow$
elegir una representación útil
$\Downarrow$
reconocer una estructura demostrada
$\Downarrow$
aplicar una transformación que haga progreso
$\Downarrow$
verificar por derivación.

Figura 18.8 sintetiza §18.11 como un protocolo con retroalimentación: orienta decisiones, pero no afirma que exista un algoritmo universal que produzca una primitiva elemental.

Una técnica necesita una medida de progreso. Que una transformación sea válida no basta. También debemos saber por qué nos acerca a una primitiva conocida. A lo largo del capítulo aparecieron varias medidas de progreso:

- en $\int x^n e^x\,dx$, disminuye el grado del polinomio;
- en $J_n(u;a)$, disminuye el exponente n;
- en una descomposición en fracciones parciales, una fracción racional global se sustituye por bloques elementales independientes;
- en producto-a-suma desaparece un producto de oscilaciones y aparece una suma de términos simples;
- en una sustitución bien elegida desaparece la composición original y queda una integral completamente expresada en una variable nueva;
- en un cálculo cíclico por partes, la integral original reaparece con un coeficiente que permite aislarla algebraicamente.

Esta observación proporciona una regla de control:

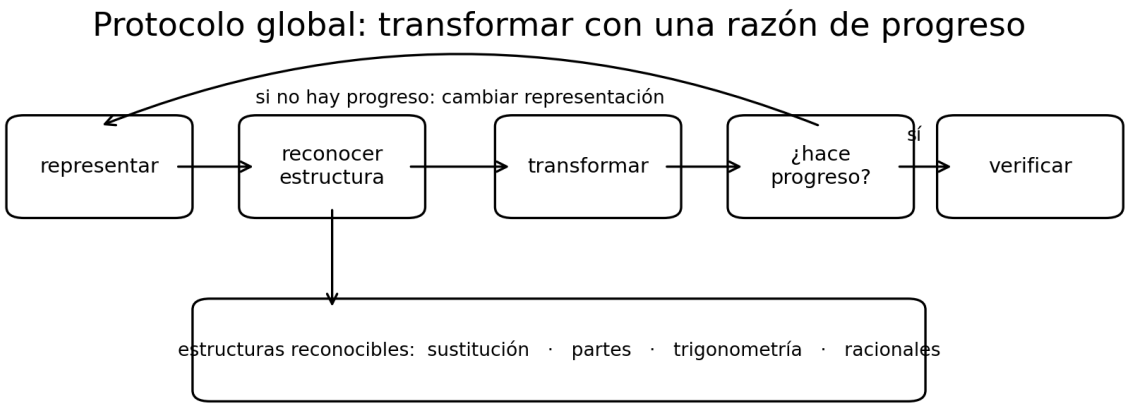


FIGURA 18.8. Mapa global de estrategia para integrar: representar, reconocer estructura, transformar, medir progreso y verificar. Las ramas conducen a sustitución, partes, reescritura trigonométrica o fracciones parciales; si una transformación no hace progreso, el flujo regresa explícitamente a cambiar la representación. El dominio funciona como control transversal.

Antes de repetir una técnica, identifica qué se hace más simple
Si no podemos señalar una cantidad que disminuye, una estructura que desaparece o un ciclo que se cierra, todavía no tenemos una razón matemática para esperar que la repetición progrese. La pregunta «¿qué mejora después de este paso?» es parte del método, no un comentario posterior.

Saber abandonar una ruta también es parte de integrar. §18.2 mostró un ejemplo elemental: la sustitución $u = x^2$ no cerraba

$$\int e^{x^2}\,dx$$

porque la variable antigua sobrevivía. El diagnóstico del fracaso fue tan importante como un cálculo exitoso. La misma actitud se aplica a las demás técnicas. Debemos reconsiderar una ruta cuando:

- una sustitución deja simultáneamente variables antiguas y nuevas;
- una aplicación de partes aumenta de manera persistente la complejidad sin producir un ciclo controlado;
- una identidad trigonométrica destruye precisamente el factor que necesitábamos preservar;

- una descomposición racional se escribe sin respetar multiplicidades o sin usar numeradores del grado adecuado;
- una manipulación ignora puntos excluidos del dominio.

Abandonar una transformación que no hace progreso no significa que la integral sea imposible. Significa únicamente que **esa representación o esa técnica no ha cerrado el problema**.

Un mismo integrando puede admitir más de una ruta. Consideremos

$$\int \frac{x^3}{x^2+1}\,dx.$$

Una primera lectura es algebraica. La división da $\frac{x^3}{x^2+1} = x - \frac{x}{x^2+1}$, y por tanto

$$\begin{aligned}\int \frac{x^3}{x^2+1}\,dx &= \int x\,dx - \int \frac{x}{x^2+1}\,dx \\ &= \frac{x^2}{2} - \frac{1}{2}\ln(x^2+1) + C.\end{aligned}$$

Pero también puede cerrarse mediante

$$u = x^2+1, \qquad du = 2x\,dx,$$

porque $x^2 = u - 1$.

Entonces

$$\int \frac{x^3}{x^2+1}\,dx = \frac{1}{2}\int \frac{u-1}{u}\,du,$$

que produce la misma familia de primitivas, salvo una constante.

El protocolo no exige una ruta única
Dos métodos distintos pueden ser correctos y cerrar completamente una integral. La elección puede hacerse por economía, transparencia o facilidad de verificación. El criterio no es que una técnica sea «la técnica oficial», sino que cada paso esté justificado y conduzca a una estructura más controlada.

Este ejemplo también muestra por qué las técnicas no deben memorizarse como categorías disjuntas. La división algebraica y la sustitución pueden revelar dos lecturas diferentes del mismo integrando.

La verificación forma parte de la solución. Una primitiva candidata F queda validada cuando $F'(x) = f(x)$ en el dominio en que se afirma la igualdad. La derivación final cumple varias funciones a la vez: controla signos, constantes multiplicativas, escalas de argumentos, identidades trigonométricas y términos racionales producidos por reducciones. En una integral definida hay controles adicionales. Si se efectuó un cambio de variable, debemos comprobar también los extremos transformados y su orientación. Si el integrando tiene singularidades, debemos verificar que el intervalo considerado permanezca dentro de su dominio. Para integrales indefinidas sobre dominios desconectados, el símbolo C continúa interpretándose **por componente conexa**, como vimos en las fracciones racionales. Podemos resumir la última etapa como

una expresión plausible no es todavía una primitiva demostrada; la derivada es el control final.

Lo que estas técnicas no dicen. El capítulo ha cerrado familias amplias: composiciones reconocibles, productos adecuados para partes, varias familias trigonométricas y, una vez disponible la factorización real, todos los bloques analíticos que aparecen en una función racional. Eso no convierte la integración en un algoritmo universal para funciones continuas. La advertencia de §18.1 permanece intacta. Si f es continua, C17 garantiza una primitiva mediante acumulación; otra cuestión es que podamos expresarla mediante un número finito de las funciones y operaciones elementales que conocemos. La ausencia de una técnica que cierre un ejemplo concreto no prueba, por sí sola, que no exista una representación elemental. Una afirmación de esa naturaleza requiere una teoría distinta y resultados adicionales.

Antes de cerrar el capítulo

Para cada integrando siguiente, **no calcules todavía la primitiva completa**. Identifica el primer movimiento justificado y explica qué progreso esperas obtener.

1. $\displaystyle\int \frac{x}{(1+x^2)^2}\,dx.$
2. $\displaystyle\int x^2 e^x\,dx.$
3. $\displaystyle\int \sin(4x)\cos x\,dx.$
4. $\displaystyle\int \frac{x^2+1}{x-1}\,dx.$

Diagnóstico.

1. Conviene tomar $u = 1+x^2$: el factor $x\,dx$ contiene, salvo la constante $1/2$, la derivada de la función interior, y la sustitución reduce el problema a una potencia de u.
2. Conviene usar integración por partes con el polinomio como factor que se deriva: su grado disminuye $2 \to 1 \to 0$ mientras e^x conserva su forma.
3. Los argumentos tienen frecuencias distintas; Proposición 18.3 convierte el producto en una suma de senos de frecuencias 5 y 3.
4. La fracción es impropia. El primer paso es la división polinómica; solo después debe decidirse si queda algún bloque racional que requiera otra técnica.

La elección de una representación debe hacer visible una estructura de derivación. Después de transformar la integral comprobaremos si el problema se ha simplificado y, al obtener una primitiva, la verificaremos derivándola en su dominio. La pregunta rectora, por tanto, no es

> «¿qué fórmula de integración debo recordar?»

sino

> **«¿qué estructura está escondida aquí, qué representación la hace visible y por qué el paso siguiente constituye progreso?»**

Ejercicios y soluciones

El banco reservado `0716--0755` se materializará por niveles para conservar el régimen modular del capítulo. Este primer bloque trabaja **reconocimiento estructural y cambio de variable**: antes de calcular, hay que identificar qué subexpresión cumple el papel de función interior, si su derivada está realmente presente y qué ocurre con el dominio o con la orientación en integrales definidas.

Nivel A — Reconocimiento estructural y cambio de variable. Ejercicio A1. La semejanza visual no basta. Para cada integral, decide si una sustitución directa basada en la regla de la cadena cierra el problema. Si la respuesta es afirmativa, indica una función interior $g(x)$ y señala dónde aparece $g'(x)$, quizá salvo un factor constante. Si la respuesta es negativa, explica exactamente qué pieza falta. **No calcules las primitivas.**

1. $\displaystyle\int x^2 e^{x^3}\,dx$;
2. $\displaystyle\int \frac{x}{(1+x^2)^2}\,dx$;
3. $\displaystyle\int e^{x^2}\,dx$;
4. $\displaystyle\int \frac{x}{1+x^3}\,dx$.

Ejercicio A2. Fabricar únicamente la constante que falta. Calcula

$$\int \frac{x}{(1+x^2)^3}\,dx$$

mediante un cambio de variable. Identifica la discrepancia constante entre el factor presente y la derivada de la función interior, y verifica la respuesta por derivación.

Ejercicio A3. Extremos que se invierten. Evalúa

$$\int_0^1 \frac{2x}{2-x^2}\,dx$$

usando la sustitución $u = 2 - x^2$. Transforma los extremos antes de integrar y explica por qué el cambio de orden de los nuevos límites queda compensado por el signo del diferencial.

Ejercicio A4. Una sustitución no necesita ser inyectiva. Calcula

$$\int_{-1}^2 2xe^{x^2}\,dx$$

mediante $u = x^2$. Explica por qué el hecho de que $x \mapsto x^2$ no sea monótona en $[-1, 2]$ no impide aplicar Teorema 18.1.

Ejercicio A5. La variable antigua puede desaparecer después de reescribir. Calcula

$$\int \frac{x^3}{(1+x^2)^2}\,dx$$

con la sustitución $u = 1 + x^2$. El paso decisivo consiste en reescribir el factor x^2 restante únicamente en términos de u. Verifica al final que la derivada recupera el integrando original.

Ejercicio A6. Una misma fórmula, constantes distintas por componente. Considera

$$f(x) = \frac{2x-1}{x^2-x-2}, \qquad D = \mathbb{R} \setminus \{-1, 2\}.$$

1. Obtén una primitiva de f en cada componente conexa de D.
2. Describe la familia completa de primitivas sobre D.

3. Explica por qué las constantes de integración de $(-\infty,-1)$, $(-1,2)$ y $(2,\infty)$ no tienen por qué coincidir.

Soluciones desarrolladas.

Soluciones del nivel A. **Solución A1.** El criterio de Teorema 18.1 no es que aparezca una expresión compuesta, sino que el resto del integrando reproduzca la derivada de su parte interior, salvo una constante.

1. En
$$\int x^2 e^{x^3}\,dx$$
tomamos
$$g(x)=x^3, \qquad g'(x)=3x^2.$$
El factor x^2 es exactamente $\frac{1}{3}g'(x)$. La sustitución directa **sí cierra**.

2. En
$$\int \frac{x}{(1+x^2)^2}\,dx$$
tomamos
$$g(x)=1+x^2, \qquad g'(x)=2x.$$
El numerador es $\frac{1}{2}g'(x)$. La sustitución directa **sí cierra**.

3. En
$$\int e^{x^2}\,dx,$$
la función interior evidente es $g(x)=x^2$, pero $g'(x)=2x$ y el integrando no contiene ningún factor proporcional a x. La pieza ausente no es una constante. Por tanto $u=x^2$ **no cierra** esta integral.

4. En
$$\int \frac{x}{1+x^3}\,dx,$$
la subexpresión natural $g(x)=1+x^3$ tiene derivada $g'(x)=3x^2$, mientras el numerador contiene solo x. De nuevo la discrepancia depende de x y no puede corregirse mediante un factor constante. Esa sustitución directa **no cierra**.

Los cuatro casos muestran que una coincidencia de forma exterior no basta: debe aparecer también la huella diferencial correspondiente.

Solución A2. Elegimos

$$u=1+x^2, \qquad du=2x\,dx.$$

Como $x\,dx=\frac{1}{2}\,du$, obtenemos

$$\begin{aligned}\int \frac{x}{(1+x^2)^3}\,dx &= \frac{1}{2}\int u^{-3}\,du\\ &= \frac{1}{2}\left(-\frac{1}{2}u^{-2}\right)+C\\ &= -\frac{1}{4u^2}+C.\end{aligned}$$

Volviendo a x,

$$\boxed{\int \frac{x}{(1+x^2)^3}\,dx = -\frac{1}{4(1+x^2)^2} + C.}$$

La discrepancia era únicamente el factor constante $1/2$. Para verificar,

$$\frac{d}{dx}\left[-\frac{1}{4}(1+x^2)^{-2}\right] = -\frac{1}{4}(-2)(1+x^2)^{-3}(2x) = \frac{x}{(1+x^2)^3}.$$

Solución A3. Tomamos

$$u = 2 - x^2, \qquad du = -2x\,dx.$$

Los extremos se transforman en

$$x = 0 \Longrightarrow u = 2, \qquad x = 1 \Longrightarrow u = 1.$$

Por tanto,

$$\begin{aligned}\int_0^1 \frac{2x}{2-x^2}\,dx &= -\int_2^1 \frac{du}{u} \\ &= \int_1^2 \frac{du}{u} \\ &= \big[\ln u\big]_1^2 \\ &= \boxed{\ln 2}.\end{aligned}$$

No se han reordenado silenciosamente los extremos. La sustitución produce límites 2 y 1; el signo menos procedente de $du = -2x\,dx$ permite escribir después la integral equivalente con orientación $1 \to 2$.

Solución A4. Sea

$$u = x^2, \qquad du = 2x\,dx.$$

Aunque $x \mapsto x^2$ no es inyectiva en $[-1, 2]$, Teorema 18.1 solo exige, en esta situación, que la función interior sea continuamente diferenciable y que la función exterior sea continua en un intervalo que contenga su imagen.

Los extremos son

$$x = -1 \Longrightarrow u = 1, \qquad x = 2 \Longrightarrow u = 4.$$

Así,

$$\begin{aligned}\int_{-1}^2 2xe^{x^2}\,dx &= \int_1^4 e^u\,du \\ &= \big[e^u\big]_1^4 \\ &= \boxed{e^4 - e}.\end{aligned}$$

La falta de monotonía no introduce ninguna ambigüedad: la prueba de Teorema 18.1 se apoya en la composición de una primitiva con g, no en invertir a g.

Solución A5. Elegimos

$$u = 1 + x^2, \qquad du = 2x\,dx.$$

Todavía queda un factor x^2, pero la propia ecuación de sustitución permite escribir $x^2 = u - 1$. Además, $x^3\,dx = x^2(x\,dx) = \frac{1}{2}(u-1)\,du$.
Por tanto,

$$\begin{aligned}\int \frac{x^3}{(1+x^2)^2}\,dx &= \frac{1}{2}\int \frac{u-1}{u^2}\,du \\ &= \frac{1}{2}\int \left(\frac{1}{u} - \frac{1}{u^2}\right) du \\ &= \frac{1}{2}\ln u + \frac{1}{2u} + C.\end{aligned}$$

Como $u = 1 + x^2 > 0$ para todo x, no hace falta valor absoluto en el logaritmo. Regresando a la variable original,

$$\boxed{\int \frac{x^3}{(1+x^2)^2}\,dx = \frac{1}{2}\ln(1+x^2) + \frac{1}{2(1+x^2)} + C.}$$

La verificación es

$$\begin{aligned}\frac{d}{dx}\left[\frac{1}{2}\ln(1+x^2) + \frac{1}{2(1+x^2)}\right] &= \frac{x}{1+x^2} - \frac{x}{(1+x^2)^2} \\ &= \frac{x[(1+x^2)-1]}{(1+x^2)^2} \\ &= \frac{x^3}{(1+x^2)^2}.\end{aligned}$$

Solución A6. Factorizamos el denominador: $x^2 - x - 2 = (x-2)(x+1)$.
Por tanto su dominio natural es $D = (-\infty, -1) \cup (-1, 2) \cup (2, \infty)$.
Si $q(x) = x^2 - x - 2$, entonces $q'(x) = 2x - 1$.
En cualquiera de las tres componentes del dominio, Corolario 17.4 y la regla de la cadena dan

$$\frac{d}{dx}\ln|q(x)| = \frac{q'(x)}{q(x)} = \frac{2x-1}{x^2-x-2}.$$

Así, en cada componente, una primitiva es $\ln|x^2 - x - 2|$. La familia completa de primitivas sobre el dominio desconectado es, por tanto,

$$F(x) = \begin{cases} \ln|x^2-x-2| + C_1, & x < -1, \\ \ln|x^2-x-2| + C_2, & -1 < x < 2, \\ \ln|x^2-x-2| + C_3, & x > 2, \end{cases}$$

con $C_1, C_2, C_3 \in \mathbb{R}$ independientes.

La unicidad de primitivas hasta una constante se aplica **en un intervalo**. Como D tiene tres componentes conexas y no hay puntos del dominio que conecten una con otra, ninguna condición diferencial obliga a que $C_1 = C_2 = C_3$.

Nivel B — Integración por partes y reducciones directas. Ejercicio B1. Elegir los papeles para que haya progreso. Calcula

$$\int x^2 e^x \, dx$$

mediante integración por partes. Explica por qué conviene diferenciar el polinomio e integrar la exponencial, y verifica la primitiva obtenida.

Ejercicio B2. Un producto que mezcla álgebra y logaritmo. Para $x > 0$, calcula

$$\int x \ln x \, dx$$

mediante una sola aplicación de Teorema 18.2. Haz explícitos u, dv, du y v.

Ejercicio B3. El término de borde no desaparece. Evalúa

$$\int_0^1 x^2 e^x \, dx$$

mediante integración por partes iterada. Conserva explícitamente los términos de borde en cada paso y comprueba que el resultado coincide con la evaluación de una primitiva.

Ejercicio B4. Cuando la integral original reaparece. Calcula

$$\int e^x \sin x \, dx$$

aplicando integración por partes dos veces. Cuando reaparezca la integral inicial, aíslala algebraicamente. Verifica el resultado por derivación.

Ejercicio B5. Usar una reducción ya demostrada. Sin repetir cuatro veces la fórmula de integración por partes, usa Proposición 18.1 para obtener una primitiva de

$$\int x^4 e^x \, dx.$$

Escribe el polinomio final de manera explícita y verifica que su derivada recupera el integrando.

Ejercicio B6. Reducir una potencia del logaritmo. Para $x > 0$, calcula

$$\int (\ln x)^2 \, dx.$$

Haz visible el factor 1, aplica integración por partes y reduce el problema a la primitiva de $\ln x$ ya obtenida en §18.3.

Soluciones del nivel B. **Solución B1.** Elegimos

$$u = x^2, \qquad dv = e^x\,dx.$$

Entonces

$$du = 2x\,dx, \qquad v = e^x.$$

La elección hace progreso porque el grado del polinomio disminuye de 2 a 1, mientras la exponencial conserva su forma. Por Teorema 18.2,

$$\int x^2 e^x\,dx = x^2e^x - 2\int xe^x\,dx.$$

Aplicamos partes una vez más a la integral restante:

$$\int xe^x\,dx = xe^x - e^x + C.$$

Por tanto,

$$\boxed{\int x^2 e^x\,dx = e^x(x^2 - 2x + 2) + C.}$$

La elección inversa, $u = e^x$ y $dv = x^2\,dx$, produciría una integral proporcional a x^3e^x: aumentaría el grado y no constituiría progreso.

La verificación es inmediata:

$$\begin{aligned}\frac{d}{dx}\left[e^x(x^2 - 2x + 2)\right] &= e^x(x^2 - 2x + 2) + e^x(2x - 2)\\ &= x^2e^x.\end{aligned}$$

Solución B2. Para $x > 0$ tomamos

$$u = \ln x, \qquad dv = x\,dx.$$

Entonces

$$du = \frac{1}{x}\,dx, \qquad v = \frac{x^2}{2}.$$

Por Teorema 18.2,

$$\begin{aligned}\int x\ln x\,dx &= \frac{x^2}{2}\ln x - \int \frac{x^2}{2}\frac{1}{x}\,dx\\ &= \frac{x^2}{2}\ln x - \frac{1}{2}\int x\,dx\\ &= \boxed{\frac{x^2}{2}\ln x - \frac{x^2}{4} + C}.\end{aligned}$$

Derivando,

$$\frac{d}{dx}\left(\frac{x^2}{2}\ln x - \frac{x^2}{4}\right) = x\ln x + \frac{x}{2} - \frac{x}{2} = x\ln x.$$

Solución B3. Aplicamos partes con

$$u = x^2, \qquad dv = e^x\,dx,$$

de modo que $du = 2x\,dx$ y $v = e^x$. Entonces

$$\int_0^1 x^2 e^x\,dx = \left[x^2 e^x\right]_0^1 - 2\int_0^1 x e^x\,dx.$$

Para la integral restante usamos nuevamente partes:

$$\int_0^1 x e^x\,dx = \left[x e^x\right]_0^1 - \int_0^1 e^x\,dx.$$

Por tanto,

$$\begin{aligned}\int_0^1 x^2 e^x\,dx &= e - 2\,(e - (e-1)) \\ &= e - 2 \\ &= \boxed{e-2}.\end{aligned}$$

La primitiva de B1 es $F(x) = e^x(x^2 - 2x + 2)$.
Así, $F(1) - F(0) = e - 2$, como exige Newton–Leibniz.

Solución B4. Escribamos

$$I = \int e^x \sin x\,dx.$$

Primero tomamos

$$u = \sin x, \qquad dv = e^x\,dx,$$

de modo que $du = \cos x\,dx$ y $v = e^x$. Entonces

$$I = e^x \sin x - \int e^x \cos x\,dx.$$

Sea

$$J = \int e^x \cos x\,dx.$$

Aplicamos partes a J con $u = \cos x$ y $dv = e^x\,dx$:

$$J = e^x \cos x + \int e^x \sin x\,dx = e^x \cos x + I.$$

Sustituyendo en la primera identidad, $I = e^x \sin x - e^x \cos x - I$.
Por tanto, $2I = e^x(\sin x - \cos x)$, y así

$$\boxed{\int e^x \sin x\,dx = \frac{e^x}{2}(\sin x - \cos x) + C.}$$

La verificación reproduce el cierre del ciclo:

$$\begin{aligned}\frac{d}{dx}\left[\frac{e^x}{2}(\sin x - \cos x)\right] &= \frac{e^x}{2}(\sin x - \cos x) + \frac{e^x}{2}(\cos x + \sin x)\\ &= e^x \sin x.\end{aligned}$$

Solución B5. Proposición 18.1 afirma que

$$\int x^n e^x\,dx = e^x \sum_{k=0}^{n} (-1)^k \frac{n!}{(n-k)!} x^{n-k} + C.$$

Para $n = 4$ obtenemos

$$\int x^4 e^x\,dx = e^x\left(x^4 - 4x^3 + 12x^2 - 24x + 24\right) + C.$$

Por tanto,

$$\boxed{\int x^4 e^x\,dx = e^x(x^4 - 4x^3 + 12x^2 - 24x + 24) + C.}$$

Si

$$P(x) = x^4 - 4x^3 + 12x^2 - 24x + 24,$$

entonces $P'(x) = 4x^3 - 12x^2 + 24x - 24$, y por construcción $P(x) + P'(x) = x^4$. Así, $\frac{d}{dx}\big(e^x P(x)\big) = e^x(P + P') = x^4 e^x$.

Solución B6. Hacemos visible el factor constante:

$$(\ln x)^2 = (\ln x)^2 \cdot 1, \qquad x > 0.$$

Elegimos

$$u = (\ln x)^2, \qquad dv = dx.$$

Entonces

$$du = \frac{2\ln x}{x}\,dx, \qquad v = x.$$

Por integración por partes,

$$\begin{aligned}\int (\ln x)^2\,dx &= x(\ln x)^2 - \int x\frac{2\ln x}{x}\,dx\\ &= x(\ln x)^2 - 2\int \ln x\,dx.\end{aligned}$$

§18.3 estableció

$$\int \ln x\,dx = x\ln x - x + C.$$

Sustituyendo,

$$\boxed{\int (\ln x)^2\,dx = x\big[(\ln x)^2 - 2\ln x + 2\big] + C, \qquad x > 0.}$$

La derivación final da

$$\begin{aligned}\frac{d}{dx}\Big\{x\big[(\ln x)^2 - 2\ln x + 2\big]\Big\} &= (\ln x)^2 - 2\ln x + 2 + 2\ln x - 2\\ &= (\ln x)^2.\end{aligned}$$

Nivel C — Potencias y productos trigonométricos. Ejercicio C1. Preservar el diferencial adecuado. Calcula

$$\int \sin^5 x\cos^2 x\,dx.$$

Explica por qué el exponente impar del seno sugiere reservar un factor $\sin x\,dx$ y usar $u = \cos x$.

Ejercicio C2. La paridad contraria cambia la sustitución. Calcula

$$\int \sin^4 x\cos^3 x\,dx.$$

Identifica qué factor debe reservarse y qué identidad convierte el resto en un polinomio de la nueva variable.

Ejercicio C3. Dos rutas igualmente válidas. Calcula

$$\int \sin^3 x\cos^3 x\,dx$$

de dos maneras: primero con $u = \sin x$ y después con $u = \cos x$. Comprueba que las dos primitivas obtenidas difieren solo en una constante.

Ejercicio C4. Cuando no hay factor impar que reservar. Calcula

$$\int \sin^4 x\,dx$$

usando reducción de potencia. No introduzcas una sustitución antes de haber eliminado la cuarta potencia.

Ejercicio C5. Frecuencias distintas. Calcula

$$\int \cos(5x)\cos(2x)\,dx$$

mediante Proposición 18.3. Explica por qué producto-a-suma elimina el obstáculo estructural en un solo paso.

Ejercicio C6. Cancelación sobre un período común. Evalúa

$$\int_0^{2\pi} \sin(5x)\sin(2x)\,dx$$

mediante producto-a-suma. Justifica la cancelación directamente, sin invocar una teoría general de ortogonalidad.

Soluciones del nivel C. **Solución C1.** Como el exponente del seno es impar,

$$\sin^5 x = \sin^4 x \sin x = (1-\cos^2 x)^2 \sin x.$$

Así,

$$\int \sin^5 x \cos^2 x\,dx = \int (1-\cos^2 x)^2 \cos^2 x \sin x\,dx.$$

Tomamos

$$u = \cos x, \qquad du = -\sin x\,dx.$$

Entonces

$$\begin{aligned}\int \sin^5 x \cos^2 x\,dx &= -\int (1-u^2)^2 u^2\,du \\ &= -\int (u^2 - 2u^4 + u^6)\,du \\ &= -\frac{u^3}{3} + \frac{2u^5}{5} - \frac{u^7}{7} + C.\end{aligned}$$

Por tanto,

$$\boxed{\int \sin^5 x \cos^2 x\,dx = -\frac{\cos^3 x}{3} + \frac{2\cos^5 x}{5} - \frac{\cos^7 x}{7} + C.}$$

El factor reservado $\sin x\,dx$ es precisamente el que permite usar la derivada de $\cos x$; el resto se convierte en un polinomio en u.

Solución C2. Ahora el exponente impar es el del coseno. Escribimos

$$\cos^3 x = \cos^2 x \cos x = (1-\sin^2 x)\cos x.$$

Así,

$$\int \sin^4 x \cos^3 x\,dx = \int \sin^4 x (1-\sin^2 x)\cos x\,dx.$$

Tomamos

$$u = \sin x, \qquad du = \cos x\,dx.$$

Entonces

$$\begin{aligned}\int \sin^4 x \cos^3 x\,dx &= \int u^4(1-u^2)\,du \\ &= \int (u^4 - u^6)\,du \\ &= \frac{u^5}{5} - \frac{u^7}{7} + C.\end{aligned}$$

Por tanto,

$$\boxed{\int \sin^4 x \cos^3 x\,dx = \frac{\sin^5 x}{5} - \frac{\sin^7 x}{7} + C.}$$

La identidad decisiva fue $\cos^2 x = 1 - \sin^2 x$: conserva el factor $\cos x\,dx$ y expresa todo lo demás mediante $u = \sin x$.

Solución C3. Como ambos exponentes son impares, Proposición 18.2 deja dos rutas abiertas.
Primera ruta: $u = \sin x$. Separamos un factor $\cos x$: $\cos^3 x = (1 - \sin^2 x)\cos x$.
Entonces

$$\begin{aligned}\int \sin^3 x \cos^3 x\,dx &= \int \sin^3 x(1 - \sin^2 x)\cos x\,dx \\ &= \int (u^3 - u^5)\,du \\ &= \frac{u^4}{4} - \frac{u^6}{6} + C.\end{aligned}$$

Así, $F_1(x) = \frac{\sin^4 x}{4} - \frac{\sin^6 x}{6}$.
Segunda ruta: $u = \cos x$. Separamos un factor $\sin x$: $\sin^3 x = (1 - \cos^2 x)\sin x$, con $du = -\sin x\,dx$. Por tanto,

$$\begin{aligned}\int \sin^3 x \cos^3 x\,dx &= -\int (1-u^2)u^3\,du \\ &= -\frac{u^4}{4} + \frac{u^6}{6} + C,\end{aligned}$$

y podemos tomar $F_2(x) = -\frac{\cos^4 x}{4} + \frac{\cos^6 x}{6}$.
Ambas derivan al mismo integrando. Además, usando $\sin^2 x = 1 - \cos^2 x$ se obtiene $F_1(x) - F_2(x) = \frac{1}{12}$, una constante. Las dos rutas describen, por tanto, la misma familia de primitivas.

Solución C4. Como no hay un factor impar de seno o coseno que podamos reservar, usamos reducción de potencia: $\sin^2 x = \frac{1-\cos(2x)}{2}$.
Al cuadrar,

$$\sin^4 x = \frac{1}{4}\big(1 - 2\cos(2x) + \cos^2(2x)\big).$$

Reducimos nuevamente la potencia: $\cos^2(2x) = \frac{1+\cos(4x)}{2}$.
Por tanto, $\sin^4 x = \frac{3}{8} - \frac{1}{2}\cos(2x) + \frac{1}{8}\cos(4x)$.
Integramos término a término:

$$\boxed{\int \sin^4 x\,dx = \frac{3x}{8} - \frac{\sin(2x)}{4} + \frac{\sin(4x)}{32} + C.}$$

La reducción de potencia fue el paso estructural previo; las integrales restantes solo requieren sustituciones lineales.

Solución C5. Por Proposición 18.3,

$$\cos(5x)\cos(2x) = \frac{1}{2}\big[\cos(3x) + \cos(7x)\big].$$

Entonces

$$\begin{aligned}\int \cos(5x)\cos(2x)\,dx &= \frac{1}{2}\int \cos(3x)\,dx + \frac{1}{2}\int \cos(7x)\,dx \\ &= \frac{\sin(3x)}{6} + \frac{\sin(7x)}{14} + C.\end{aligned}$$

Así,

$$\boxed{\int \cos(5x)\cos(2x)\,dx = \frac{\sin(3x)}{6} + \frac{\sin(7x)}{14} + C.}$$

Una sustitución directa basada en una sola de las dos frecuencias no elimina el otro factor. Producto-a-suma reemplaza el producto por una suma con frecuencias $5-2=3$ y $5+2=7$, y el problema queda reducido a integrales elementales.

Solución C6. Por Proposición 18.3,

$$\sin(5x)\sin(2x) = \frac{1}{2}\big[\cos(3x) - \cos(7x)\big].$$

Por tanto,

$$\begin{aligned}\int_0^{2\pi} \sin(5x)\sin(2x)\,dx &= \frac{1}{2}\int_0^{2\pi} \cos(3x)\,dx - \frac{1}{2}\int_0^{2\pi} \cos(7x)\,dx \\ &= \frac{1}{6}\big[\sin(3x)\big]_0^{2\pi} - \frac{1}{14}\big[\sin(7x)\big]_0^{2\pi} \\ &= 0.\end{aligned}$$

Luego

$$\boxed{\int_0^{2\pi} \sin(5x)\sin(2x)\,dx = 0.}$$

La cancelación se ha demostrado directamente porque $\sin(6\pi) = \sin(14\pi) = \sin 0 = 0$. No hace falta invocar una teoría general de ortogonalidad para este cálculo concreto.

Nivel D — Funciones racionales: división y factores lineales. Ejercicio D1. Dos polos simples. Calcula

$$\int \frac{4x+1}{(x-1)(x+3)}\,dx.$$

Determina los coeficientes mediante la identidad polinómica asociada y especifica el dominio de la función racional.

Ejercicio D2. Tres factores lineales simples. Calcula

$$\int \frac{x^2+1}{(x-1)(x+1)(x-2)}\,dx.$$

Usa Proposición 18.4 y comprueba al final que la recomposición de las fracciones parciales recupera el integrando.

Ejercicio D3. Dividir antes de descomponer. Calcula

$$\int \frac{x^3+2x^2+1}{x^2-1}\,dx.$$

Explica por qué comenzar directamente con fracciones parciales sería una mala representación del problema y conserva los puntos excluidos del dominio original.

Ejercicio D4. Un factor repetido sin término logarítmico. Calcula

$$\int \frac{3x-1}{(x+2)^3}\,dx.$$

Reescribe el numerador en potencias de $x+2$ y explica por qué en este caso no aparece ningún término proporcional a $1/(x+2)$.

Ejercicio D5. Un factor doble y otro simple. Calcula

$$\int \frac{2x+5}{(x-1)^2(x+2)}\,dx.$$

Escribe la cadena completa de fracciones parciales exigida por Proposición 18.5, determina los tres coeficientes e integra término a término.

Ejercicio D6. Diagnosticar una descomposición incompleta. Un estudiante propone $\frac{1}{(x+1)^2(x-2)} = \frac{A}{(x+1)^2} + \frac{B}{x-2}$.

1. Demuestra que, en general, esa forma no contiene suficientes grados de libertad para este denominador y verifica directamente que aquí no puede funcionar.
2. Escribe la descomposición correcta.
3. Calcula una primitiva sobre cada componente del dominio.

Soluciones del nivel D. **Solución D1.** Buscamos $\frac{4x+1}{(x-1)(x+3)} = \frac{A}{x-1} + \frac{B}{x+3}$.
Multiplicando por el denominador común, $4x + 1 = A(x + 3) + B(x - 1)$.
En $x = 1$,

$$5 = 4A, \qquad A = \frac{5}{4}.$$

En $x = -3$,

$$-11 = -4B, \qquad B = \frac{11}{4}.$$

Por tanto, $\frac{4x+1}{(x-1)(x+3)} = \frac{5}{4(x-1)} + \frac{11}{4(x+3)}$.
Así, sobre cualquier intervalo contenido en el dominio,

$$\boxed{\int \frac{4x+1}{(x-1)(x+3)}\,dx = \frac{5}{4}\ln|x-1| + \frac{11}{4}\ln|x+3| + C.}$$

El dominio es $(-\infty, -3) \cup (-3, 1) \cup (1, \infty)$, y la constante puede elegirse independientemente en cada componente.

Solución D2. Por Proposición 18.4,

$$\frac{x^2+1}{(x-1)(x+1)(x-2)} = \frac{A}{x-1} + \frac{B}{x+1} + \frac{C}{x-2}.$$

Los coeficientes se obtienen evaluando la identidad polinómica en las tres raíces. Para $x = 1$,

$$A = \frac{1^2+1}{(1+1)(1-2)} = -1.$$

Para $x = -1$,

$$B = \frac{(-1)^2+1}{(-1-1)(-1-2)} = \frac{1}{3}.$$

Para $x = 2$,

$$C = \frac{2^2+1}{(2-1)(2+1)} = \frac{5}{3}.$$

Luego

$$\frac{x^2+1}{(x-1)(x+1)(x-2)} = -\frac{1}{x-1} + \frac{1}{3(x+1)} + \frac{5}{3(x-2)}.$$

Por tanto,

$$\boxed{\int \frac{x^2+1}{(x-1)(x+1)(x-2)}\,dx = -\ln|x-1| + \frac{1}{3}\ln|x+1| + \frac{5}{3}\ln|x-2| + C.}$$

La recomposición da

$$-\frac{1}{x-1}+\frac{1}{3(x+1)}+\frac{5}{3(x-2)}=\frac{x^2+1}{(x-1)(x+1)(x-2)},$$

como exige la descomposición.

Solución D3. La fracción es impropia porque el grado del numerador es mayor que el del denominador. La división polinómica da

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

Así,

$$\frac{x^3+2x^2+1}{x^2-1}=x+2+\frac{x+3}{(x-1)(x+1)}.$$

Para la parte propia buscamos $\frac{x+3}{(x-1)(x+1)}=\frac{A}{x-1}+\frac{B}{x+1}$. En $x=1$ obtenemos $4=2A$, luego $A=2$; en $x=-1$, $2=-2B$, luego $B=-1$. Por tanto, $\frac{x^3+2x^2+1}{x^2-1}=x+2+\frac{2}{x-1}-\frac{1}{x+1}$.

Integramos término a término:

$$\boxed{\int\frac{x^3+2x^2+1}{x^2-1}\,dx=\frac{x^2}{2}+2x+2\ln|x-1|-\ln|x+1|+C.}$$

El dominio original excluye $x=\pm1$. Dividir primero separa la parte polinómica y deja una fracción propia, que es la forma a la que se aplica la teoría de fracciones parciales.

Solución D4. Escribimos el numerador en potencias de $x+2$: $3x-1=3(x+2)-7$.

Entonces $\frac{3x-1}{(x+2)^3}=\frac{3}{(x+2)^2}-\frac{7}{(x+2)^3}$.

No aparece un término proporcional a $1/(x+2)$ porque la expansión del numerador alrededor de la raíz -2 no contiene ningún término cuadrático $(x+2)^2$.

Integramos:

$$\begin{aligned}\int\frac{3x-1}{(x+2)^3}\,dx&=3\int(x+2)^{-2}\,dx-7\int(x+2)^{-3}\,dx\\&=-\frac{3}{x+2}+\frac{7}{2(x+2)^2}+C.\end{aligned}$$

Por tanto,

$$\boxed{\int\frac{3x-1}{(x+2)^3}\,dx=-\frac{3}{x+2}+\frac{7}{2(x+2)^2}+C.}$$

El dominio tiene las dos componentes $(-\infty,-2)$ y $(-2,\infty)$.

Solución D5. Como $(x-1)^2$ es un factor doble, Proposición 18.5 exige $\frac{2x+5}{(x-1)^2(x+2)}=\frac{A}{x-1}+\frac{B}{(x-1)^2}+\frac{C}{x+2}$.

Multiplicando por el denominador común,

$$2x+5=A(x-1)(x+2)+B(x+2)+C(x-1)^2.$$

En $x=1$,

$$7 = 3B, \qquad B = \frac{7}{3}.$$

En $x = -2$,

$$1 = 9C, \qquad C = \frac{1}{9}.$$

Comparando los coeficientes de x^2, $A + C = 0$, de modo que $A = -\frac{1}{9}$.

Así,

$$\frac{2x+5}{(x-1)^2(x+2)} = -\frac{1}{9(x-1)} + \frac{7}{3(x-1)^2} + \frac{1}{9(x+2)}.$$

Integramos término a término:

$$\boxed{\int \frac{2x+5}{(x-1)^2(x+2)}\,dx = -\frac{1}{9}\ln|x-1| - \frac{7}{3(x-1)} + \frac{1}{9}\ln|x+2| + C.}$$

Solución D6. La raíz $x = -1$ tiene multiplicidad 2, así que una descomposición general debe incluir **dos** términos asociados a ella: $\frac{A}{x+1} + \frac{B}{(x+1)^2}$.

La propuesta del enunciado omite el primero. Si multiplicamos su identidad por $(x+1)^2(x-2)$, obtendríamos $1 = A(x-2)+B(x+1)^2$. El coeficiente de x^2 obliga a $B = 0$; entonces el coeficiente de x obliga a $A = 0$, lo que hace imposible obtener el término constante 1. La forma propuesta no puede funcionar.

La descomposición correcta es $\frac{1}{(x+1)^2(x-2)} = \frac{A}{x+1} + \frac{B}{(x+1)^2} + \frac{C}{x-2}$.

Multiplicando por el denominador común,

$$1 = A(x+1)(x-2) + B(x-2) + C(x+1)^2.$$

En $x = -1$,

$$1 = -3B, \qquad B = -\frac{1}{3}.$$

En $x = 2$,

$$1 = 9C, \qquad C = \frac{1}{9}.$$

Comparando los coeficientes de x^2,

$$A + C = 0, \qquad A = -\frac{1}{9}.$$

Por tanto,

$$\frac{1}{(x+1)^2(x-2)} = -\frac{1}{9(x+1)} - \frac{1}{3(x+1)^2} + \frac{1}{9(x-2)}.$$

Integramos:

$$\boxed{\int \frac{dx}{(x+1)^2(x-2)} = -\frac{1}{9}\ln|x+1| + \frac{1}{3(x+1)} + \frac{1}{9}\ln|x-2| + C.}$$

El dominio es $(-\infty,-1)\cup(-1,2)\cup(2,\infty)$, y la constante de integración puede elegirse independientemente en cada componente.

Nivel E — Cuadráticos irreducibles: arctan **y reducción. Ejercicio E1. La forma cuadrática básica escalada.** Calcula

$$\int \frac{dx}{x^2+4}.$$

Obtén la respuesta a partir de Proposición 18.6 y verifica el factor de escala derivando.

Ejercicio E2. Separar la derivada del denominador. Calcula

$$\int \frac{4x+1}{x^2-2x+5}\,dx.$$

Completa el cuadrado, escribe el numerador como un múltiplo de la derivada del denominador más un residuo constante y explica por qué aparecen simultáneamente un logaritmo y un término de arco tangente.

Ejercicio E3. Un polo lineal y un cuadrático irreducible. Calcula

$$\int \frac{2x+1}{(x-1)(x^2+1)}\,dx.$$

Usa la forma $\frac{A}{x-1}+\frac{Bx+C}{x^2+1}$ y conserva explícitamente el único punto excluido del dominio real.

Ejercicio E4. El primer cuadrático irreducible repetido. Calcula

$$\int \frac{dx}{\left((x-2)^2+9\right)^2}$$

aplicando Proposición 18.8. Identifica los parámetros u, a y el descenso $J_2 \to J_1$.

Ejercicio E5. Numerador lineal sobre un cuadrático repetido. Calcula

$$\int \frac{2x+1}{(x^2+4x+8)^2}\,dx.$$

Separa primero un múltiplo de la derivada de $q(x)=x^2+4x+8$ y aplica la reducción solo al residuo constante.

Ejercicio E6. Cadena completa con cuadrático repetido. Calcula

$$\int \frac{dx}{(x-1)(x^2+1)^2}.$$

1. Justifica por qué la descomposición debe tener la forma $\frac{A}{x-1}+\frac{Bx+C}{x^2+1}+\frac{Dx+E}{(x^2+1)^2}$.
2. Determina los cinco coeficientes.
3. Integra cada bloque usando únicamente los resultados de §§18.7–18.10 y especifica las componentes del dominio.

Soluciones del nivel E. **Solución E1.** Escribimos $x^2+4 = x^2+2^2$.
En Proposición 18.6 tomamos $a=2$ y $b=0$. Entonces

$$\boxed{\int \frac{dx}{x^2+4} = \frac{1}{2}\arctan\!\left(\frac{x}{2}\right) + C.}$$

La verificación del factor de escala es

$$\begin{aligned}\frac{d}{dx}\left[\frac{1}{2}\arctan\!\left(\frac{x}{2}\right)\right] &= \frac{1}{2}\,\frac{1}{1+(x/2)^2}\,\frac{1}{2}\\ &= \frac{1}{x^2+4}.\end{aligned}$$

El denominador es positivo en toda la recta, así que la familia de primitivas se considera sobre $\mathbb{R}$.

Solución E2. Definimos $q(x) = x^2-2x+5 = (x-1)^2+4$, por lo que $q'(x) = 2x-2$.
El numerador se separa como

$$4x+1 = 2(2x-2)+5 = 2q'(x)+5.$$

Por linealidad,

$$\begin{aligned}\int \frac{4x+1}{q(x)}\,dx &= 2\int \frac{q'(x)}{q(x)}\,dx + 5\int \frac{dx}{(x-1)^2+4}\\ &= 2\ln q(x) + \frac{5}{2}\arctan\!\left(\frac{x-1}{2}\right) + C.\end{aligned}$$

Así,

$$\boxed{\int \frac{4x+1}{x^2-2x+5}\,dx = 2\ln(x^2-2x+5) + \frac{5}{2}\arctan\!\left(\frac{x-1}{2}\right) + C.}$$

El logaritmo procede del múltiplo de q'/q; el arco tangente procede del residuo constante sobre el cuadrático completado. Como $q(x)>0$ para todo x, no hay puntos reales excluidos.

Solución E3. Buscamos $\frac{2x+1}{(x-1)(x^2+1)} = \frac{A}{x-1} + \frac{Bx+C}{x^2+1}$.
Multiplicando por el denominador común,

$$2x+1 = A(x^2+1) + (Bx+C)(x-1).$$

En $x=1$,

$$3 = 2A, \qquad A = \frac{3}{2}.$$

Comparando el coeficiente de x^2,

$$A + B = 0, \qquad B = -\frac{3}{2},$$

y el coeficiente de x da

$$-B + C = 2, \qquad C = \frac{1}{2}.$$

Por tanto,

$$\frac{2x+1}{(x-1)(x^2+1)} = \frac{3}{2(x-1)} - \frac{3x}{2(x^2+1)} + \frac{1}{2(x^2+1)}.$$

Integramos término a término:

$$\boxed{\int \frac{2x+1}{(x-1)(x^2+1)}\,dx = \frac{3}{2}\ln|x-1| - \frac{3}{4}\ln(x^2+1) + \frac{1}{2}\arctan x + C.}$$

Como $x^2+1>0$ en $\mathbb{R}$, el único polo real es $x=1$. El dominio tiene las componentes

$$(-\infty, 1) \qquad \text{y} \qquad (1, \infty),$$

y la constante puede elegirse independientemente en cada una.

Solución E4. Tomamos

$$u = x-2, \qquad a = 3,$$

de modo que $\big((x-2)^2+9\big)^2 = (u^2+a^2)^2$.
Para $n=2$, Proposición 18.8 da

$$J_2(u;a) = \frac{u}{2a^2(u^2+a^2)} + \frac{1}{2a^2}J_1(u;a).$$

Con $a=3$, $J_2(u;3) = \frac{u}{18(u^2+9)} + \frac{1}{18}J_1(u;3)$.
Por Proposición 18.6, $J_1(u;3) = \frac{1}{3}\arctan\big(\frac{u}{3}\big) + C$.
Regresando a x,

$$\boxed{\int \frac{dx}{\big((x-2)^2+9\big)^2} = \frac{x-2}{18\big((x-2)^2+9\big)} + \frac{1}{54}\arctan\left(\frac{x-2}{3}\right) + C.}$$

El descenso fue exactamente $J_2 \to J_1$.

Solución E5. Sea $q(x) = x^2+4x+8 = (x+2)^2+4$.
Entonces $q'(x) = 2x+4$ y $2x+1 = q'(x) - 3$.
Así,

$$\int \frac{2x+1}{q(x)^2}\,dx = \int \frac{q'(x)}{q(x)^2}\,dx - 3\int \frac{dx}{q(x)^2}.$$

El primer término es $-\frac{1}{q(x)}$.

Para el segundo tomamos $u = x+2$ y $a=2$. Por Proposición 18.8, $J_2(u;2) = \frac{u}{8(u^2+4)} + \frac{1}{8}J_1(u;2)$, y por Proposición 18.6, $J_1(u;2) = \frac{1}{2}\arctan\big(\frac{u}{2}\big) + C$.

Por tanto,

$$J_2(u;2) = \frac{u}{8(u^2+4)} + \frac{1}{16}\arctan\!\left(\frac{u}{2}\right) + C.$$

Regresando a x,

$$\boxed{\begin{aligned}\int \frac{2x+1}{(x^2+4x+8)^2}\,dx = &-\frac{1}{x^2+4x+8}\\ &-\frac{3(x+2)}{8(x^2+4x+8)} - \frac{3}{16}\arctan\!\left(\frac{x+2}{2}\right) + C.\end{aligned}}$$

La reducción se aplicó únicamente al residuo constante -3; la parte proporcional a q' se integró de inmediato.

Solución E6. El factor $x-1$ es lineal simple y x^2+1 es cuadrático irreducible de multiplicidad 2. Por eso la cadena completa debe ser

$$\frac{1}{(x-1)(x^2+1)^2} = \frac{A}{x-1} + \frac{Bx+C}{x^2+1} + \frac{Dx+E}{(x^2+1)^2}.$$

Multiplicando por $(x-1)(x^2+1)^2$,

$$1 = A(x^2+1)^2 + (Bx+C)(x-1)(x^2+1) + (Dx+E)(x-1).$$

En $x=1$ obtenemos

$$1 = 4A, \qquad A = \frac{1}{4}.$$

Al comparar coeficientes en la identidad polinómica resultan sucesivamente

$$A + B = 0, \qquad -B + C = 0,$$

$$2A + B - C + D = 0, \qquad -B + C - D + E = 0.$$

Con $A = 1/4$ se obtiene

$$B = -\frac{1}{4}, \qquad C = -\frac{1}{4}, \qquad D = -\frac{1}{2}, \qquad E = -\frac{1}{2}.$$

Por tanto,

$$\frac{1}{(x-1)(x^2+1)^2} = \frac{1}{4(x-1)} - \frac{x+1}{4(x^2+1)} - \frac{x+1}{2(x^2+1)^2}.$$

Integramos cada bloque. Para la parte simple,

$$-\frac{1}{4}\int \frac{x+1}{x^2+1}\,dx = -\frac{1}{8}\ln(x^2+1) - \frac{1}{4}\arctan x.$$

Para el bloque repetido,

$$-\frac{1}{2}\int \frac{x}{(x^2+1)^2}\,dx = \frac{1}{4(x^2+1)},$$

mientras que Ejemplo 18.14 da

$$-\frac{1}{2}\int \frac{dx}{(x^2+1)^2} = -\frac{x}{4(x^2+1)} - \frac{1}{4}\arctan x.$$

Reuniendo términos,

$$\boxed{\int \frac{dx}{(x-1)(x^2+1)^2} = \frac{1}{4}\ln|x-1| - \frac{1}{8}\ln(x^2+1) + \frac{1-x}{4(x^2+1)} - \frac{1}{2}\arctan x + C.}$$

El cuadrático x^2+1 nunca se anula; el único punto excluido es $x=1$. Por tanto las componentes del dominio son

$$(-\infty, 1) \qquad \text{y} \qquad (1, \infty),$$

y la constante puede elegirse independientemente en cada componente.

Nivel F — Síntesis estratégica: elegir, combinar y descartar rutas. Ejercicio F1. El primer movimiento debe estar justificado. Para cada integral, identifica **el primer movimiento que elegirías** y explica qué estructura del integrando lo justifica. No calcules todavía la primitiva completa.

1. $\displaystyle\int x\cos(x^2)\,dx$;
2. $\displaystyle\int x^3 e^x\,dx$;
3. $\displaystyle\int \sin^2(2x)\cos(3x)\,dx$;
4. $\displaystyle\int \frac{x^2+3x+1}{x^2-1}\,dx$.

En cada caso indica también qué medida de progreso esperas obtener después de ese primer paso.

Ejercicio F2. Dos rutas correctas para la misma integral. Calcula

$$\int \frac{x^3}{x^2+4}\,dx$$

de dos maneras:

1. mediante división algebraica previa;
2. mediante la sustitución $u = x^2+4$.

Compara las dos respuestas y demuestra que representan la misma familia de primitivas.

Ejercicio F3. Una sustitución que prepara integración por partes. Calcula

$$\int x^3 e^{x^2}\,dx.$$

Usa primero $u = x^2$ para transformar completamente la integral y explica por qué la nueva integral se resuelve entonces por integración por partes. Verifica la respuesta final derivando.

Ejercicio F4. Dos reescrituras trigonométricas consecutivas. Calcula

$$\int \sin^2(2x)\cos(3x)\,dx.$$

Reduce primero la potencia de $\sin(2x)$ y utiliza después producto-a-suma únicamente donde siga apareciendo un producto. Identifica las frecuencias finales antes de integrar.

Ejercicio F5. Ser racional no obliga a usar fracciones parciales. Calcula

$$\int \frac{2x^3+2x}{x^4+2x^2+5}\,dx.$$

Antes de factorizar el denominador, compara el numerador con su derivada. Explica por qué una sustitución directa constituye una ruta más estructural que intentar una descomposición racional.

Ejercicio F6. Una primitiva local no autoriza Newton–Leibniz a través de un polo. Un cálculo apresurado podría escribir

$$\int_0^2 \frac{dx}{x-1} \stackrel{?}{=} \big[\ln|x-1|\big]_0^2 = 0.$$

Analiza este argumento.

1. Decide si la integral de Riemann del miembro izquierdo está definida en $[0,2]$.
2. Identifica exactamente qué hipótesis necesaria para aplicar Newton–Leibniz falla.
3. Explica qué afirmación correcta puede hacerse sobre las primitivas de $1/(x-1)$ a cada lado del punto $x=1$ sin introducir todavía teoría de integrales impropias.

Soluciones del nivel F. **Solución F1.**

1. En
$$\int x\cos(x^2)\,dx$$
aparece la composición $\cos(x^2)$ y el factor $x\,dx$ es, salvo la constante $1/2$, la derivada de x^2. El primer movimiento natural es
$$u=x^2, \qquad du=2x\,dx.$$
El progreso esperado es que desaparezca la variable x y quede una integral elemental en u.

2. En
$$\int x^3e^x\,dx$$
no hay una composición interior cuya derivada aparezca como factor. La estructura útil es un producto. Conviene aplicar integración por partes con
$$u=x^3, \qquad dv=e^x\,dx,$$
porque el grado del polinomio disminuye $3\to 2$ mientras la exponencial conserva su forma.

3. En
$$\int \sin^2(2x)\cos(3x)\,dx$$
la potencia par impide reservar directamente un factor para una sustitución. El primer movimiento es la reducción $\sin^2(2x) = \frac{1-\cos(4x)}{2}$. El progreso consiste en disminuir la potencia y convertir el problema en una suma donde solo queda un producto trigonométrico, que podrá tratarse con Proposición 18.3.

4. En
$$\int \frac{x^2+3x+1}{x^2-1}\,dx$$
la fracción es impropia porque numerador y denominador tienen el mismo grado. El primer movimiento es la división polinómica. El progreso consiste en separar una parte polinómica y dejar una fracción racional propia, que sí puede analizarse mediante factorización y fracciones parciales.

En los cuatro casos la técnica se elige por una estructura demostrada y por una medida explícita de progreso, no por semejanza visual.

Solución F2. Por división algebraica, $\frac{x^3}{x^2+4} = x - \frac{4x}{x^2+4}$.
Entonces
$$\begin{aligned}\int \frac{x^3}{x^2+4}\,dx &= \int x\,dx - 4\int \frac{x}{x^2+4}\,dx \\ &= \frac{x^2}{2} - 2\ln(x^2+4) + C.\end{aligned}$$
Como $x^2+4>0$ para todo x, no hace falta valor absoluto.
Ahora usamos directamente
$$u = x^2+4, \qquad du = 2x\,dx.$$
La relación $x^2 = u-4$ permite escribir $x^3\,dx = x^2(x\,dx) = \frac{1}{2}(u-4)\,du$.
Así,
$$\begin{aligned}\int \frac{x^3}{x^2+4}\,dx &= \frac{1}{2}\int \frac{u-4}{u}\,du \\ &= \frac{1}{2}u - 2\ln u + C \\ &= \frac{x^2+4}{2} - 2\ln(x^2+4) + C.\end{aligned}$$
Las dos candidatas difieren en la constante 2:
$$\left[\frac{x^2+4}{2} - 2\ln(x^2+4)\right] - \left[\frac{x^2}{2} - 2\ln(x^2+4)\right] = 2.$$
Por tanto describen la misma familia de primitivas.

Solución F3. Tomamos

$$u = x^2, \qquad du = 2x\,dx.$$

Como $x^3\,dx = x^2(x\,dx) = \frac{1}{2}u\,du$, la integral se transforma completamente en

$$\int x^3 e^{x^2}\,dx = \frac{1}{2}\int ue^u\,du.$$

Ahora aparece un producto para el cual integración por partes reduce el factor algebraico. Con

$$U = u, \qquad dV = e^u\,du,$$

tenemos $dU = du$ y $V = e^u$. Por Teorema 18.2,

$$\int ue^u\,du = ue^u - e^u + C = e^u(u-1) + C.$$

Por tanto,

$$\boxed{\int x^3 e^{x^2}\,dx = \frac{1}{2}e^{x^2}(x^2-1) + C.}$$

La verificación es

$$\begin{aligned}\frac{d}{dx}\left[\frac{1}{2}e^{x^2}(x^2-1)\right] &= \frac{1}{2}e^{x^2}(2x)(x^2-1) + \frac{1}{2}e^{x^2}(2x)\\ &= x^3e^{x^2}.\end{aligned}$$

La primera técnica eliminó la composición; la segunda redujo el factor lineal que apareció en la variable nueva.

Solución F4. Primero reducimos la potencia: $\sin^2(2x) = \frac{1-\cos(4x)}{2}$.
Entonces

$$\sin^2(2x)\cos(3x) = \frac{1}{2}\cos(3x) - \frac{1}{2}\cos(4x)\cos(3x).$$

Solo el segundo término conserva un producto. Por Proposición 18.3,

$$\cos(4x)\cos(3x) = \frac{1}{2}[\cos(7x) + \cos x].$$

Por tanto las frecuencias finales son 3, 7 y 1, y

$$\sin^2(2x)\cos(3x) = \frac{1}{2}\cos(3x) - \frac{1}{4}\cos(7x) - \frac{1}{4}\cos x.$$

Integramos término a término:

$$\boxed{\int \sin^2(2x)\cos(3x)\,dx = \frac{1}{6}\sin(3x) - \frac{1}{28}\sin(7x) - \frac{1}{4}\sin x + C.}$$

Derivando la respuesta obtenemos $\frac{1}{2}\cos(3x) - \frac{1}{4}\cos(7x) - \frac{1}{4}\cos x$, que se recompone mediante producto-a-suma en $\sin^2(2x)\cos(3x)$.

Solución F5. Sea $q(x) = x^4 + 2x^2 + 5$.
Entonces $q'(x) = 4x^3 + 4x = 2(2x^3 + 2x)$.
Por tanto el numerador es exactamente la mitad de $q'(x)$: $2x^3 + 2x = \frac{1}{2}q'(x)$.
La sustitución

$$u = q(x), \qquad du = q'(x)\,dx$$

da

$$\begin{aligned}\int \frac{2x^3 + 2x}{x^4 + 2x^2 + 5}\,dx &= \frac{1}{2}\int \frac{du}{u} \\ &= \frac{1}{2}\ln u + C.\end{aligned}$$

Como $q(x) = (x^2 + 1)^2 + 4 > 0$, obtenemos

$$\boxed{\int \frac{2x^3 + 2x}{x^4 + 2x^2 + 5}\,dx = \frac{1}{2}\ln(x^4 + 2x^2 + 5) + C.}$$

La forma racional del integrando no obliga a empezar con fracciones parciales. Antes de factorizar conviene comprobar si el numerador ya contiene la derivada del denominador; aquí esa observación cierra el problema en un paso.

Solución F6. La función $f(x) = \frac{1}{x-1}$ no está definida en $x = 1$ y, además, es no acotada en cualquier intervalo que contenga ese punto. Por tanto no es una función Riemann integrable en $[0, 2]$ en el sentido desarrollado en C15. El símbolo

$$\int_0^2 \frac{dx}{x-1}$$

no representa aquí una integral de Riemann definida. El uso de Newton–Leibniz falla porque no existe una función G continua en todo $[0, 2]$, diferenciable en $(0, 2)$ y con $G'(x) = \frac{1}{x-1}$ para todo $x \in (0, 2)$: el integrando ni siquiera está definido en $x = 1$. En particular, la función $\ln|x-1|$ es una primitiva únicamente en cada componente del dominio,

$$(-\infty, 1) \qquad \text{y} \qquad (1, \infty),$$

no una candidata válida para atravesar el polo dentro de una sola aplicación de Teorema 16.3. La afirmación correcta, dentro de la teoría disponible, es que

$$\boxed{\int \frac{dx}{x-1} = \ln|x-1| + C_- \quad \text{en } (-\infty, 1),}$$

y

$$\boxed{\int \frac{dx}{x-1} = \ln|x-1| + C_+ \quad \text{en } (1, \infty),}$$

con constantes independientes. El cálculo $\big[\ln|x-1|\big]_0^2 = 0$ solo evalúa una expresión en dos puntos; no verifica las hipótesis que permitirían identificar esa diferencia con una integral de Riemann sobre $[0, 2]$. La teoría de integrales impropias queda fuera de este capítulo.

Nivel G — Transferencia, generalización y síntesis final. Ejercicio G1. Del ejemplo particular a una familia. Sean m, n enteros positivos. Demuestra, usando únicamente Proposición 18.3 y primitivas trigonométricas ya conocidas, que

$$\int_0^{2\pi} \sin(mx)\sin(nx)\,dx = \begin{cases} 0, & m \neq n, \\ \pi, & m = n. \end{cases}$$

No invoques teoría de Fourier ni ningún principio general de ortogonalidad: deriva el resultado directamente de las identidades producto-a-suma. Explica por qué el caso $m = n$ debe separarse del caso $m \neq n$ antes de dividir por una frecuencia.

Ejercicio G2. Una reducción con parámetro. Sea $a > 0$. A partir de Proposición 18.8 y del caso base construido en §18.9, calcula

$$\int \frac{dx}{(x^2 + a^2)^3}.$$

La respuesta debe quedar expresada únicamente mediante funciones racionales de x y a y el término $\arctan(x/a)$. Indica en cada paso qué exponente disminuye y verifica el resultado por derivación.

Ejercicio G3. Factor lineal, cuadrático irreducible y dominio. Calcula

$$\int \frac{x^2 + x + 1}{x(x^2 + 4)}\,dx.$$

1. Determina la descomposición en fracciones parciales.
2. Integra cada bloque usando solamente resultados construidos en este capítulo.
3. Describe las componentes del dominio y escribe correctamente la familia completa de primitivas, permitiendo constantes independientes cuando corresponda.
4. Verifica la respuesta derivando en cada componente.

Ejercicio G4. Una integral que cambia de naturaleza durante el cálculo. Calcula

$$\int \ln(x^2 + 1)\,dx.$$

Empieza por integración por partes. Después de ese paso, no uses fracciones parciales: reescribe algebraicamente la función racional restante de modo que aparezcan una constante y el bloque $1/(x^2 + 1)$. Explica por qué cada transformación constituye progreso y verifica la primitiva final por derivación.

Soluciones del nivel G. **Solución G1.** Si $m \neq n$, Proposición 18.3 da

$$\sin(mx)\sin(nx) = \frac{1}{2}\big[\cos((m-n)x) - \cos((m+n)x)\big].$$

Como m y n son enteros positivos y distintos, tanto $m - n$ como $m + n$ son enteros no nulos. Por tanto,

$$\begin{aligned}\int_0^{2\pi} \sin(mx)\sin(nx)\,dx &= \frac{1}{2}\left[\frac{\sin((m-n)x)}{m-n} - \frac{\sin((m+n)x)}{m+n}\right]_0^{2\pi}\\ &= 0,\end{aligned}$$

porque $\sin(2\pi k) = 0$ para todo entero k.

Si $m = n$, la fórmula anterior contiene la frecuencia $m - n = 0$, de modo que no podemos escribir el primer término como $\sin((m-n)x)/(m-n)$. Debemos separar este caso antes de dividir. Usamos entonces la reducción de potencia: $\sin^2(mx) = \frac{1-\cos(2mx)}{2}$.

Así,

$$\begin{aligned}\int_0^{2\pi} \sin^2(mx)\,dx &= \frac{1}{2}\int_0^{2\pi} 1\,dx - \frac{1}{2}\int_0^{2\pi} \cos(2mx)\,dx\\ &= \pi - \frac{1}{4m}\big[\sin(2mx)\big]_0^{2\pi}\\ &= \pi.\end{aligned}$$

Por tanto,

$$\boxed{\int_0^{2\pi} \sin(mx)\sin(nx)\,dx = \begin{cases} 0, & m \neq n,\\ \pi, & m = n.\end{cases}}$$

El resultado surge aquí de una identidad trigonométrica finita y de evaluación directa; no se ha utilizado ninguna teoría general adicional.

Solución G2. Escribimos

$$J_n(x;a) = \int \frac{dx}{(x^2+a^2)^n}.$$

Por Proposición 18.8,

$$J_n = \frac{x}{2(n-1)a^2(x^2+a^2)^{n-1}} + \frac{2n-3}{2(n-1)a^2}J_{n-1}.$$

Para $n = 3$, $J_3 = \frac{x}{4a^2(x^2+a^2)^2} + \frac{3}{4a^2}J_2$.

Aplicamos la misma reducción a J_2: $J_2 = \frac{x}{2a^2(x^2+a^2)} + \frac{1}{2a^2}J_1$.

El caso base de §18.9 es $J_1 = \frac{1}{a}\arctan\left(\frac{x}{a}\right) + C$. Sustituyendo primero en J_2 y luego en J_3 obtenemos

$$\boxed{\begin{aligned}\int \frac{dx}{(x^2+a^2)^3} = {}& \frac{x}{4a^2(x^2+a^2)^2} + \frac{3x}{8a^4(x^2+a^2)}\\ &+ \frac{3}{8a^5}\arctan\left(\frac{x}{a}\right) + C.\end{aligned}}$$

La medida de progreso es el exponente del cuadrático: $3 \longrightarrow 2 \longrightarrow 1$. Derivando la expresión encerrada se obtiene exactamente $\frac{1}{(x^2+a^2)^3}$, lo que verifica la reducción completa.

Solución G3. Como x^2+4 no tiene raíces reales y el factor lineal es x, buscamos

$$\frac{x^2+x+1}{x(x^2+4)} = \frac{A}{x} + \frac{Bx+C}{x^2+4}.$$

Multiplicando por $x(x^2+4)$,

$$x^2+x+1 = A(x^2+4) + x(Bx+C).$$

Comparando coeficientes,

$$4A = 1, \qquad C = 1, \qquad A+B = 1.$$

Por tanto,

$$A = \frac{1}{4}, \qquad B = \frac{3}{4}, \qquad C = 1,$$

y

$$\frac{x^2+x+1}{x(x^2+4)} = \frac{1}{4x} + \frac{3x}{4(x^2+4)} + \frac{1}{x^2+4}.$$

Integramos término a término:

$$\int \frac{dx}{4x} = \frac{1}{4}\ln|x|,$$

$$\int \frac{3x}{4(x^2+4)}\,dx = \frac{3}{8}\ln(x^2+4),$$

y, por la forma escalada de Proposición 18.6,

$$\int \frac{dx}{x^2+4} = \frac{1}{2}\arctan\left(\frac{x}{2}\right).$$

Así, en cualquier intervalo que no cruce $x=0$,

$$\boxed{\frac{1}{4}\ln|x| + \frac{3}{8}\ln(x^2+4) + \frac{1}{2}\arctan\left(\frac{x}{2}\right) + C}$$

es una primitiva.
El dominio es $\mathbb{R}\setminus\{0\} = (-\infty,0)\cup(0,\infty)$.
Por tanto la familia completa debe permitir constantes independientes:

$$F(x) = \frac{1}{4}\ln|x| + \frac{3}{8}\ln(x^2+4) + \frac{1}{2}\arctan\left(\frac{x}{2}\right) + \begin{cases} C_-, & x<0, \\ C_+, & x>0. \end{cases}$$

Derivando en cualquiera de las dos componentes se recupera

$$\frac{1}{4x} + \frac{3x}{4(x^2+4)} + \frac{1}{x^2+4} = \frac{x^2+x+1}{x(x^2+4)}.$$

Solución G4. Hacemos visible el producto $\ln(x^2+1)\cdot 1$ y aplicamos Teorema 18.2 con

$$u = \ln(x^2+1), \qquad dv = dx.$$

Entonces

$$du = \frac{2x}{x^2+1}\,dx, \qquad v = x,$$

y por tanto

$$\int \ln(x^2+1)\,dx = x\ln(x^2+1) - 2\int \frac{x^2}{x^2+1}\,dx.$$

La integración por partes ha transformado el logaritmo en una función racional. Ahora reescribimos $\frac{x^2}{x^2+1} = 1 - \frac{1}{x^2+1}$.

Así,

$$\begin{aligned}\int \ln(x^2+1)\,dx &= x\ln(x^2+1) - 2\int 1\,dx + 2\int \frac{dx}{x^2+1}\\ &= x\ln(x^2+1) - 2x + 2\arctan x + C.\end{aligned}$$

Por tanto,

$$\boxed{\int \ln(x^2+1)\,dx = x\ln(x^2+1) - 2x + 2\arctan x + C.}$$

Cada paso tiene una medida de progreso distinta: partes elimina el logaritmo del integrando restante; la reescritura racional separa una constante de un bloque base ya construido. Finalmente,

$$\begin{aligned}\frac{d}{dx}\left[x\ln(x^2+1) - 2x + 2\arctan x\right] &= \ln(x^2+1) + \frac{2x^2}{x^2+1} - 2 + \frac{2}{x^2+1}\\ &= \ln(x^2+1),\end{aligned}$$

porque $\frac{2x^2+2}{x^2+1} = 2$.

Capítulo 19

Aplicaciones geométricas y cuantitativas de la integral

19.1. De una cantidad local a una integral: aditividad y densidad

El área, el volumen, la masa de una varilla y el trabajo de una fuerza requieren modelos diferentes. Para obtener la cantidad total a partir de contribuciones locales, en cada caso tendremos que responder una misma pregunta:

> **¿cómo reconstruimos una cantidad global cuando conocemos la manera en que esa cantidad se distribuye localmente?**

Supongamos que una variable x recorre un intervalo $[a, b]$ y que una función $q(x)$ describe una **densidad local** respecto de esa variable. Aquí la palabra *densidad* se usa en un sentido amplio: $q(x)$ expresa cuánta cantidad corresponde, aproximadamente, a una unidad de longitud en torno a x. No tiene por qué tratarse de densidad de masa; según el problema, q podrá representar una altura, un área de sección, una fuerza, una tasa u otra magnitud local.

Dividamos $[a, b]$ mediante una partición

$$a = x_0 < x_1 < \cdots < x_n = b$$

y escribamos $\Delta x_k = x_k - x_{k-1}$.

Si escogemos un punto $\xi_k \in [x_{k-1}, x_k]$, la contribución correspondiente a un subintervalo pequeño tiene típicamente la forma $q(\xi_k)\,\Delta x_k$. La cantidad total se aproxima entonces mediante la suma de esas contribuciones:

$$\sum_{k=1}^{n} q(\xi_k)\,\Delta x_k.$$

La expresión depende de una elección de modelo: hay que identificar la contribución local y la variable que mide el ancho de cada pieza. La suma de esas contribuciones conducirá a una integral cuando las estimaciones y las hipótesis permitan justificar el paso al límite. La arquitectura que gobernará todo el capítulo será, por tanto,

$$\boxed{\text{cantidad local} \longrightarrow \text{suma finita} \longrightarrow \text{integral} \longrightarrow \text{interpretación y control.}}$$

Para reunir esa justificación estableceremos un criterio: si una cantidad es aditiva sobre subintervalos y sus contribuciones admiten el control local indicado mediante una densidad integrable, la integral determina la cantidad global.

Teorema 19.1. **Teorema (cantidad aditiva determinada por una densidad).** Sea $q \in \mathcal{R}[a, b]$. Supongamos que a cada subintervalo cerrado $[\alpha, \beta] \subseteq [a, b]$, con $\alpha \le \beta$, se le asigna un número real $Q(\alpha, \beta)$ con las siguientes propiedades.

1. **Aditividad.** Si $a \le \alpha \le \beta \le \gamma \le b$, entonces
$$Q(\alpha, \gamma) = Q(\alpha, \beta) + Q(\beta, \gamma).$$

2. **Control local por la densidad.** Para todo $a \le \alpha \le \beta \le b$,

$$\inf_{[\alpha,\beta]} q\,(\beta-\alpha) \le Q(\alpha,\beta) \le \sup_{[\alpha,\beta]} q\,(\beta-\alpha).$$

Entonces

$$\boxed{Q(a,b) = \int_a^b q(x)\,dx.}$$

La primera hipótesis permite reconstruir una cantidad sobre un intervalo a partir de sus partes. La segunda afirma que, en cada subintervalo, la contribución real queda atrapada entre las contribuciones que producirían las densidades mínima y máxima allí presentes. La integrabilidad de q será precisamente lo que permita hacer desaparecer esa diferencia entre cotas inferiores y superiores.

Demostración. Sea

$$P: \quad a = x_0 < x_1 < \cdots < x_n = b$$

una partición arbitraria de $[a,b]$. Para cada k, escribamos

$$m_k := \inf_{[x_{k-1},x_k]} q, \qquad M_k := \sup_{[x_{k-1},x_k]} q.$$

Por el control local, $m_k \Delta x_k \le Q(x_{k-1},x_k) \le M_k \Delta x_k$.

Sumando estas desigualdades para $k = 1,\dots,n$ y usando repetidamente la aditividad de Q, obtenemos

$$\sum_{k=1}^{n} m_k \Delta x_k \le Q(a,b) \le \sum_{k=1}^{n} M_k \Delta x_k.$$

Los extremos son precisamente las sumas inferior y superior de Darboux asociadas a P. Por tanto,

$$\boxed{L(q,P) \le Q(a,b) \le U(q,P).}$$

Como $q \in \mathcal{R}[a,b]$, su integral inferior y su integral superior coinciden. Además, para toda partición P,

$$L(q,P) \le \int_a^b q(x)\,dx \le U(q,P).$$

Si $Q(a,b)$ fuese menor que la integral, la definición de integral inferior permitiría escoger una partición P con $Q(a,b) < L(q,P)$, contradiciendo el encierro anterior. Del mismo modo, si $Q(a,b)$ fuese mayor que la integral, la definición de integral superior daría una partición P con $U(q,P) < Q(a,b)$, nuevamente imposible. En consecuencia,

$$Q(a,b) = \int_a^b q(x)\,dx.$$

Esto demuestra el teorema. $\square$

Qué hace —y qué no hace— el teorema. Teorema 20.1 no sustituye el trabajo de modelización. Antes de aplicarlo debemos identificar dos objetos distintos:

- una cantidad $Q(\alpha, \beta)$ asociada a cada subintervalo;
- una función $q(x)$ que describa la cantidad aportada por unidad de la variable x.

Una vez identificados, todavía hay que justificar que las contribuciones de intervalos contiguos se suman y que, sobre cada subintervalo, la cantidad queda encerrada por los valores extremos de la densidad. Solo entonces el teorema permite concluir que la acumulación global está representada por una integral. Así, en las aplicaciones que siguen cambiarán los objetos geométricos o físicos, pero no la estructura matemática. Una altura puede actuar como densidad de área respecto de x; un área de sección, como densidad de volumen; una densidad lineal, como densidad de masa; una fuerza, como densidad de trabajo respecto del desplazamiento. En cada caso la pregunta decisiva será la misma:

> **¿qué cantidad aporta un intervalo pequeño y qué función mide esa contribución por unidad de la variable elegida?**

Esta perspectiva también proporciona un control dimensional inmediato. Si q tiene unidades de «cantidad por unidad de x», entonces $q(\xi_k)\,\Delta x_k$ tiene las unidades de la cantidad que se acumula, y lo mismo debe ocurrir con

$$\int_a^b q(x)\,dx.$$

Por eso una integral de aplicación no queda justificada solo porque pueda calcularse: su integrando, su diferencial y sus unidades deben corresponder al modelo que se pretende representar.

Antes de seguir. En una aplicación concreta, ¿qué papel desempeñan Q y q? ¿Qué hipótesis expresa que las contribuciones de piezas contiguas pueden sumarse? ¿Por qué comprobar las unidades puede detectar una integral mal planteada incluso antes de evaluarla?

19.2. Área entre gráficas: diferencia, signo y valor absoluto

La primera aplicación geométrica del principio anterior exige distinguir dos cantidades que se parecen algebraicamente, pero no representan lo mismo. Sean $f, g : [a, b] \longrightarrow \mathbb{R}$ funciones continuas. Para cada $x \in [a, b]$, la diferencia $h(x) := f(x) - g(x)$ mide una separación vertical **con signo** entre sus gráficas. Si $h(x) > 0$, la gráfica de f está por encima de la de g; si $h(x) < 0$, ocurre lo contrario; y si $h(x) = 0$, ambas se encuentran en ese punto. Esta observación revela una primera dificultad. La integral

$$\int_a^b (f(x) - g(x))\,dx$$

acumula esas separaciones conservando el signo. Las regiones donde $f > g$ contribuyen positivamente y las regiones donde $g > f$ contribuyen negativamente. Por tanto, contribuciones geométricamente distintas pueden cancelarse entre sí. El **área geométrica**, en cambio, no admite esa cancelación. Una franja vertical estrecha situada cerca de x tiene como altura la distancia vertical entre las dos gráficas, y una distancia no puede ser negativa. La magnitud local relevante es entonces $|f(x) - g(x)|$. Así aparecen dos objetos diferentes:

$$\boxed{\text{separación orientada} = f - g}$$

frente a

$$\boxed{\text{separación geométrica} = |f - g|.}$$

Confundirlos conduce a uno de los errores más frecuentes en problemas de área. Que una integral definida sea negativa, positiva o incluso nula no determina por sí solo el área de la región comprendida entre dos gráficas: primero hay que decidir si el problema pide una acumulación orientada o una magnitud geométrica no negativa. Antes de integrar debemos responder una pregunta de modelización:

> **¿qué función representa realmente la altura de cada franja vertical de la región?**

Si una misma gráfica permanece arriba en todo el intervalo, la respuesta puede escribirse sin cambiar de expresión. Si las gráficas se cruzan, el signo de $f - g$ cambia y la descripción de la altura exige tener en cuenta esos puntos de cruce. Ese será el paso siguiente.

Cuando una gráfica permanece arriba. Supongamos ahora que

$$f(x) \geq g(x) \qquad \text{para todo } x \in [a, b].$$

Entonces la separación vertical entre las gráficas es simplemente $q(x) := f(x) - g(x) \geq 0$. Sea $A(\alpha, \beta)$ el área geométrica de la región comprendida entre ambas gráficas sobre un subintervalo $[\alpha, \beta] \subseteq [a, b]$. Esta cantidad es aditiva: si $\alpha \leq \beta \leq \gamma$, las regiones sobre $[\alpha, \beta]$ y $[\beta, \gamma]$ tienen interiores disjuntos y forman conjuntamente la región sobre $[\alpha, \gamma]$, de modo que $A(\alpha, \gamma) = A(\alpha, \beta) + A(\beta, \gamma)$. Además, sobre $[\alpha, \beta]$ la altura de la región está comprendida entre $\inf_{[\alpha,\beta]}(f - g)$ y $\sup_{[\alpha,\beta]}(f - g)$. Por tanto, comparando la región con los rectángulos de base $\beta - \alpha$ y esas alturas extremas, obtenemos

$$\inf_{[\alpha,\beta]}(f - g)\,(\beta - \alpha) \leq A(\alpha, \beta) \leq \sup_{[\alpha,\beta]}(f - g)\,(\beta - \alpha).$$

Como f y g son continuas, $f - g$ es continua y, por tanto, Riemann integrable. Se cumplen así exactamente las hipótesis de Teorema 20.1 con

$$Q = A, \qquad q = f - g.$$

En consecuencia,

$$\boxed{A = \int_a^b \big(f(x) - g(x)\big)\,dx.}$$

Cuando $f \geq g$, la altura vertical $f - g$ es la densidad de área respecto de x. El principio de aditividad y control por esa densidad justifica la fórmula para la región entre ambas gráficas. Si la desigualdad se invierte en todo el intervalo, el mismo razonamiento da

$$A = \int_a^b \big(g(x) - f(x)\big)\,dx.$$

El caso realmente nuevo aparece cuando ninguna de las dos gráficas permanece arriba en todo $[a, b]$. Entonces la función que representa la altura cambia de expresión al atravesar los puntos de cruce.

Cuando las gráficas se cruzan. Para obtener una formulación válida sin imponer restricciones sobre el número de intersecciones, definimos directamente la altura geométrica $q(x) := |f(x) - g(x)|$. Como f y g son continuas, también lo es q, y por tanto $q \in \mathcal{R}[a, b]$. Sea $A(\alpha, \beta)$ el área geométrica comprendida entre ambas gráficas sobre un subintervalo $[\alpha, \beta]$. Esta cantidad es aditiva sobre intervalos contiguos. Además, toda franja vertical sobre $[\alpha, \beta]$ tiene altura entre $\text{ínf}_{[\alpha,\beta]}\, q$ y $\sup_{[\alpha,\beta]} q$. La comparación con rectángulos da

$$\inf_{[\alpha,\beta]} q\,(\beta - \alpha) \le A(\alpha, \beta) \le \sup_{[\alpha,\beta]} q\,(\beta - \alpha).$$

Se cumplen así las hipótesis de Teorema 20.1 con

$$Q = A, \qquad q = |f - g|.$$

En consecuencia,

$$\boxed{A = \int_a^b |f(x) - g(x)|\, dx.}$$

Esta demostración no requiere que las gráficas tengan un número finito de puntos de intersección. El valor absoluto aparece antes de integrar porque la cantidad local que se acumula es una distancia vertical y debe ser no negativa. Cuando $f - g$ cambia de signo solo un número finito de veces, puede ser conveniente localizar esos cambios y escribir la integral por tramos. Si

$$a = c_0 < c_1 < \cdots < c_m = b$$

separa intervalos donde el signo de $f - g$ permanece fijo, entonces

$$\int_a^b |f - g| = \sum_{j=1}^{m} \int_{c_{j-1}}^{c_j} |f - g|.$$

En cada tramo, el valor absoluto se elimina usando el signo correspondiente. Esta descomposición es una herramienta de cálculo; la fórmula general del área ya quedó justificada por la densidad local $|f - g|$. Si se integra directamente $f - g$, se obtiene una acumulación orientada y pueden producirse cancelaciones entre regiones situadas a lados opuestos del eje de signo.

Ejemplo 19.1. **Cancelación orientada frente a área geométrica.** Consideremos en $[-1, 1]$ las funciones

$$f(x) = x, \qquad g(x) = 0.$$

Las gráficas se cruzan en $x = 0$. La diferencia es $f(x) - g(x) = x$, de modo que

$$\int_{-1}^{1} (f(x) - g(x))\, dx = \int_{-1}^{1} x\, dx = 0.$$

Ese cero no significa que la región comprendida entre las gráficas tenga área nula. Solo expresa que la contribución negativa de $[-1, 0]$ cancela exactamente la contribución positiva de $[0, 1]$. Para el área geométrica debemos integrar la separación no negativa:

$$A = \int_{-1}^{1} |x|\, dx.$$

Partiendo en el punto de cruce,

$$\begin{aligned} A &= \int_{-1}^{0} (-x)\,dx + \int_{0}^{1} x\,dx \\ &= \frac{1}{2} + \frac{1}{2} \\ &= 1. \end{aligned}$$

Por tanto,

$$\boxed{\int_{-1}^{1} (f-g) = 0 \qquad \text{mientras que} \qquad \int_{-1}^{1} |f-g| = 1.}$$

La diferencia entre ambas cantidades no es computacional, sino conceptual: una conserva orientación y permite cancelación; la otra mide una magnitud geométrica.

Antes de seguir. Si dos gráficas se cruzan varias veces, ¿qué debe determinarse antes de integrar? ¿Por qué una integral orientada puede ser cero aunque el área encerrada sea positiva? ¿Qué representa localmente la función $|f-g|$?

19.3. Volumen por secciones: cortar antes de integrar

Para calcular un volumen mediante una integral, el primer problema no es encontrar una primitiva. Es decidir **cómo cortar el sólido**. Supongamos que un sólido se extiende a lo largo del eje x entre $x = a$ y $x = b$. Para cada $x \in [a, b]$, consideremos la sección obtenida al cortar el sólido con un plano perpendicular al eje x, y denotemos por $A(x)$ el área de esa sección transversal. Dividamos ahora $[a, b]$ mediante una partición

$$a = x_0 < x_1 < \cdots < x_n = b.$$

Sobre un subintervalo pequeño $[x_{k-1}, x_k]$, de espesor $\Delta x_k = x_k - x_{k-1}$, escogemos un punto ξ_k. Si el área transversal cambia poco dentro de esa rebanada, podemos aproximarla por un prisma cuya base tiene área $A(\xi_k)$ y cuyo espesor es Δx_k. Su volumen aproximado es entonces $A(\xi_k)\,\Delta x_k$. Sumando las contribuciones de todas las rebanadas obtenemos

$$\sum_{k=1}^{n} A(\xi_k)\,\Delta x_k.$$

La contribución de volumen se obtiene de un **área transversal** multiplicada por un espesor. Para aplicar el principio de §19.1 debemos justificar que esas secciones describen el sólido y que la función de áreas satisface las hipótesis del modelo. Por eso, antes de escribir ninguna integral, debemos identificar tres elementos:

1. la dirección de corte;
2. el área $A(x)$ de cada sección;
3. el espesor correspondiente a la variable elegida.

La fórmula de volumen aparecerá solo después de que estos tres datos hayan sido fijados correctamente.

Principio de volumen por secciones. La suma anterior sugiere la fórmula correcta, pero podemos justificarla con el resultado general de §19.1.

PROPOSICIÓN 19.1. **Proposición (volumen por secciones).** Sea $A \in \mathcal{R}[a,b]$ y supongamos que a cada subintervalo cerrado $[\alpha,\beta] \subseteq [a,b]$ se le asocia el volumen $V(\alpha,\beta)$ de la parte del sólido comprendida entre los planos $x=\alpha$ y $x=\beta$. Supongamos además que:

1. V es aditivo sobre intervalos contiguos;
2. para todo $a \le \alpha \le \beta \le b$,

$$\inf_{[\alpha,\beta]} A\,(\beta-\alpha) \le V(\alpha,\beta) \le \sup_{[\alpha,\beta]} A\,(\beta-\alpha).$$

Entonces

$$\boxed{V(a,b) = \int_a^b A(x)\,dx.}$$

Demostración. Basta aplicar Teorema 20.1 con

$$Q(\alpha,\beta) := V(\alpha,\beta) \qquad \text{y} \qquad q(x) := A(x).$$

La aditividad de V proporciona la primera hipótesis del teorema, mientras que el encierro entre prismas de alturas extremas proporciona la segunda. Por tanto,

$$V(a,b) = \int_a^b A(x)\,dx.$$

□

La fórmula

$$\boxed{V = \int_a^b A(x)\,dx}$$

no debe leerse como «volumen igual a integral» sin más. El integrando $A(x)$ tiene una interpretación geométrica precisa: es el área de la sección perpendicular a la variable de integración. El diferencial dx representa el espesor de las rebanadas que se están acumulando. Así, el modelo local es

$$\text{área transversal} \times \text{espesor}$$

y las unidades lo confirman:

$$[\text{longitud}^2]\,[\text{longitud}] = [\text{longitud}^3].$$

La dificultad de los problemas concretos estará, por tanto, en construir correctamente $A(x)$. Una vez conocida esa función, la integral aparece como consecuencia de la acumulación de secciones.

EJEMPLO 19.2. **Secciones cuadradas sobre una base triangular.** La base de un sólido es la región del plano comprendida entre

$$y = x \qquad \text{y} \qquad y = 2 - x,$$

para $0 \leq x \leq 1$. Las secciones perpendiculares al eje x son cuadrados. Antes de integrar debemos construir el área de una sección. Para un valor fijo de x, el lado del cuadrado es la distancia vertical entre las dos rectas: $s(x) = (2-x) - x = 2(1-x)$. Por tanto, el área transversal es $A(x) = s(x)^2 = 4(1-x)^2$. El principio de Proposición 20.1 da entonces

$$V = \int_0^1 4(1-x)^2\,dx.$$

Evaluando,

$$\begin{aligned} V &= 4\int_0^1 (1 - 2x + x^2)\,dx \\ &= 4\left[x - x^2 + \frac{x^3}{3}\right]_0^1 \\ &= \frac{4}{3}. \end{aligned}$$

Así,

$$\boxed{V = \frac{4}{3}.}$$

El paso decisivo no fue la integración. Fue la cadena

$$\boxed{\text{distancia entre las fronteras} \longrightarrow \text{lado del cuadrado} \longrightarrow A(x) \longrightarrow \int_0^1 A(x)\,dx.}$$

Si se hubiera integrado directamente la distancia $2(1-x)$, se habría acumulado una longitud por un espesor y el resultado tendría dimensiones de área, no de volumen.

Antes de seguir. ¿Qué parte del problema determina la forma de $A(x)$? ¿Por qué la distancia entre las fronteras de la base no es todavía el integrando de volumen? ¿Qué cambiaría si las secciones fueran triángulos equiláteros en vez de cuadrados?

19.4. Sólidos de revolución I: discos y arandelas

El método de discos no introduce una nueva teoría de volumen. Es el principio de secciones de §19.3 aplicado a un sólido cuyas secciones transversales son círculos.

Discos como secciones transversales. Sea $f : [a,b] \longrightarrow [0,\infty)$ una función continua. Consideremos la región del plano comprendida entre la gráfica de f, el eje x y las rectas $x = a$ y $x = b$, y hagámosla girar alrededor del eje x. Para un valor fijo de x, el corte perpendicular al eje de rotación es un disco. Su radio no se obtiene de una fórmula nueva: es simplemente la distancia vertical desde el eje x hasta la gráfica, es decir, $R(x) = f(x)$. Por tanto, el área transversal es $A(x) = \pi R(x)^2 = \pi f(x)^2$. Al sustituir esta función de área en el principio de volumen por secciones obtenemos el caso de discos.

Corolario 19.1. **Corolario (volumen por discos).** Sea $f : [a,b] \to [0,\infty)$ continua. El volumen del sólido obtenido al hacer girar alrededor del eje x la región

$$\{(x,y) : a \leq x \leq b,\ 0 \leq y \leq f(x)\}$$

es

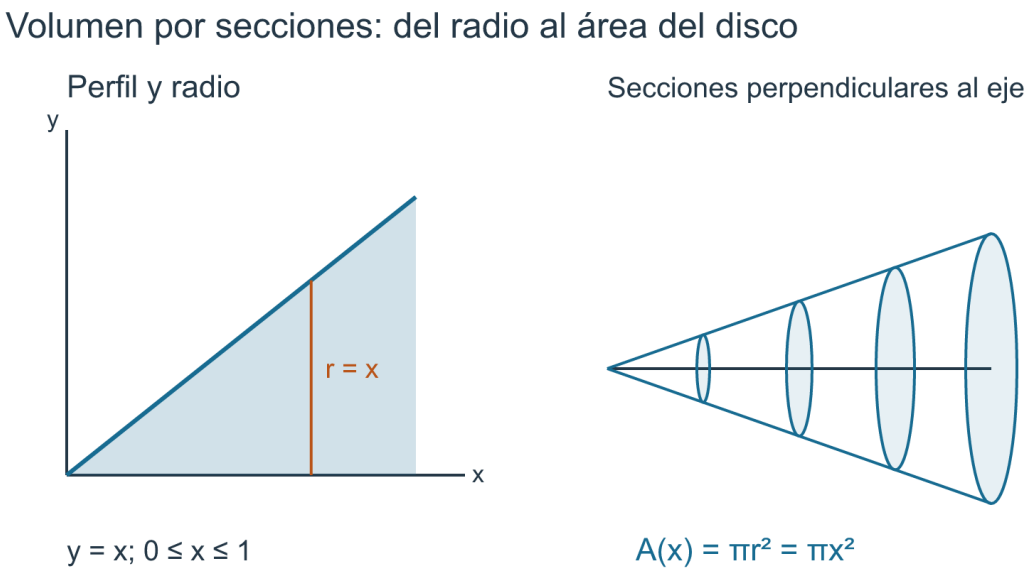


FIGURA 19.1. Al girar $y = x$, $0 \leq x \leq 1$, alrededor del eje horizontal, el radio de la sección es $r(x) = x$ y su área es $A(x) = \pi x^2$. La integral de secciones usa esta área, no solo el radio.

$$\boxed{V = \pi \int_a^b f(x)^2 \, dx.}$$

Demostración. Cada sección perpendicular al eje x es un disco de radio $f(x)$, luego $A(x) = \pi f(x)^2$.

Como f es continua, A también lo es y, en particular, es Riemann integrable. Aplicando Proposición 20.1,

$$V = \int_a^b A(x)\, dx = \int_a^b \pi f(x)^2 \, dx = \pi \int_a^b f(x)^2 \, dx.$$

□

La estructura del razonamiento es, por tanto,

$$\boxed{\text{radio } R(x) \longrightarrow \text{área } \pi R(x)^2 \longrightarrow \text{volumen } \int \pi R(x)^2 \, dx.}$$

La integral aparece al final. La decisión geométrica esencial sigue siendo identificar la sección transversal correcta y su radio respecto del eje de rotación.

Arandelas: un disco menos un disco interior. Si la región que gira alrededor del eje x deja un hueco entre ella y el eje, cada sección perpendicular tendrá la forma de una **arandela**. Su área se obtiene restando el disco interior al exterior. Sean $R, r : [a, b] \longrightarrow [0, \infty)$ funciones continuas tales que

$$0 \le r(x) \le R(x) \qquad (x \in [a,b]).$$

Interpretamos $R(x)$ como la distancia desde el eje de rotación hasta la frontera exterior de la región y $r(x)$ como la distancia hasta la frontera interior. Al girar la sección vertical alrededor del eje x, obtenemos una arandela de radio exterior $R(x)$ y radio interior $r(x)$. Su área es la diferencia entre las áreas de dos discos: $A(x) = \pi R(x)^2 - \pi r(x)^2$.

Por tanto,

$$\boxed{A(x) = \pi(R(x)^2 - r(x)^2).}$$

COROLARIO 19.2. **Corolario (volumen por arandelas).** Sean $R, r : [a,b] \to [0,\infty)$ continuas y supongamos que

$$0 \le r(x) \le R(x) \qquad (x \in [a,b]).$$

Si las secciones perpendiculares al eje x del sólido son arandelas de radio exterior $R(x)$ y radio interior $r(x)$, entonces

$$\boxed{V = \pi \int_a^b (R(x)^2 - r(x)^2)\, dx.}$$

Demostración. El área transversal es $A(x) = \pi(R(x)^2 - r(x)^2)$.

Como R y r son continuas, A es Riemann integrable. Aplicando Proposición 20.1,

$$\begin{aligned} V &= \int_a^b A(x)\, dx \\ &= \int_a^b \pi(R(x)^2 - r(x)^2)\, dx \\ &= \pi \int_a^b (R(x)^2 - r(x)^2)\, dx. \end{aligned}$$

□

La condición $0 \le r(x) \le R(x)$ garantiza que los radios sean distancias admisibles, que el interior no exceda al exterior y que la diferencia de áreas sea no negativa. También conviene observar que los radios son **distancias al eje de rotación**, no necesariamente valores de las funciones que delimitan la región. Cuando el eje de rotación no es el eje x, identificar correctamente esas distancias forma parte del modelo.

EJEMPLO 19.3. **Arandelas con eje de rotación desplazado.** Consideremos la región comprendida entre

$$y = x \qquad \text{y} \qquad y = x^2,$$

para $0 \le x \le 1$, y hagámosla girar alrededor de la recta $y = 2$. En $[0,1]$ se tiene $x^2 \le x$, de modo que ambas curvas están por debajo del eje de rotación. La frontera más alejada de $y = 2$ es $y = x^2$, mientras que la más cercana es $y = x$. Por tanto, $R(x) = 2 - x^2$ y $r(x) = 2 - x$. El área transversal es entonces $A(x) = \pi\left[(2 - x^2)^2 - (2 - x)^2\right]$. Aplicando Corolario 20.2,

$$V = \pi \int_0^1 \left[(2 - x^2)^2 - (2 - x)^2\right] dx.$$

Simplificando el integrando,

$$(2-x^2)^2-(2-x)^2 = 4x-5x^2+x^4.$$

Así,

$$\begin{aligned} V &= \pi\int_0^1 \left(4x-5x^2+x^4\right)\,dx \\ &= \pi\left[2x^2-\frac{5}{3}x^3+\frac{1}{5}x^5\right]_0^1 \\ &= \pi\left(2-\frac{5}{3}+\frac{1}{5}\right) \\ &= \frac{8\pi}{15}. \end{aligned}$$

Por tanto,

$$\boxed{V=\frac{8\pi}{15}.}$$

Obsérvese que habría sido incorrecto tomar $R(x)=x$ y $r(x)=x^2$. Esas son alturas medidas desde el eje x, pero el sólido gira alrededor de $y=2$. Los radios deben medirse siempre como distancias al eje real de rotación.

Antes de seguir. ¿Cómo se decide cuál radio es exterior cuando el eje de rotación no coincide con un eje coordenado? ¿Por qué no basta con mirar qué función tiene mayor valor? ¿Qué control geométrico permite detectar un intercambio accidental entre R y r?

19.5. Sólidos de revolución II: cascarones y elección de representación

El método de cascarones parte de una elección geométrica distinta. En vez de cortar el sólido con secciones perpendiculares al eje de rotación, comenzamos con tiras **paralelas** a ese eje y seguimos lo que ocurre cuando cada una gira. Sea $f:[a,b]\longrightarrow[0,\infty)$ continua, con $0\le a<b$, y consideremos la región

$$\{(x,y): a\le x\le b,\ 0\le y\le f(x)\}.$$

Al hacerla girar alrededor del eje y, una tira vertical estrecha situada cerca de x ya no produce un disco. Produce un **cascarón cilíndrico**. Si la tira tiene espesor Δx, entonces los tres datos locales son:

- radio aproximado: x;
- altura aproximada: $f(x)$;
- espesor: Δx.

Al desplegar lateralmente un cascarón muy delgado, su volumen se aproxima por (circunferencia)(altura)(espesor). Como la circunferencia correspondiente al radio x es $2\pi x$, la contribución local toma la forma

$$\boxed{2\pi x\, f(x)\,\Delta x.}$$

Con una partición

$$a = x_0 < x_1 < \cdots < x_n = b$$

y puntos $\xi_k \in [x_{k-1}, x_k]$, la suma de cascarones delgados queda modelada por

$$\sum_{k=1}^{n} 2\pi \xi_k f(\xi_k)\, \Delta x_k.$$

La estructura local es ahora

$$\boxed{\text{circunferencia} \times \text{altura} \times \text{espesor.}}$$

Obsérvese la diferencia con discos y arandelas. Allí el corte perpendicular al eje producía directamente un **área transversal**. Aquí una tira paralela al eje genera una superficie cilíndrica delgada, y el volumen local se construye a partir de su circunferencia, su altura y su espesor. Hemos identificado la contribución de una tira vertical al girar. Para obtener una fórmula global por cascarones tendremos que justificar la acumulación de esas contribuciones y sus límites de integración.

De los cascarones delgados a la integral. La expresión local anterior puede justificarse sin introducir una nueva teoría de integración. Consideremos una partición

$$P: \quad a = x_0 < x_1 < \cdots < x_n = b$$

y, en cada intervalo $[x_{k-1}, x_k]$, escribamos

$$m_k := \inf_{[x_{k-1}, x_k]} f, \qquad M_k := \sup_{[x_{k-1}, x_k]} f.$$

La parte del sólido comprendida entre los cilindros de radios x_{k-1} y x_k queda encerrada entre dos cascarones cilíndricos de alturas m_k y M_k. Por tanto, su volumen V_k satisface

$$\pi(x_k^2 - x_{k-1}^2) m_k \le V_k \le \pi(x_k^2 - x_{k-1}^2) M_k.$$

Si definimos $c_k := \frac{x_{k-1} + x_k}{2}$, entonces $x_k^2 - x_{k-1}^2 = 2c_k\, \Delta x_k$.

Así, al sumar sobre todos los intervalos,

$$\sum_{k=1}^{n} 2\pi c_k m_k\, \Delta x_k \le V \le \sum_{k=1}^{n} 2\pi c_k M_k\, \Delta x_k.$$

Entre ambas cotas se encuentra la suma etiquetada

$$\sum_{k=1}^{n} 2\pi c_k f(c_k)\, \Delta x_k.$$

Además,

$$0 \le \sum_{k=1}^{n} 2\pi c_k (M_k - m_k)\, \Delta x_k \le 2\pi b \sum_{k=1}^{n} (M_k - m_k)\, \Delta x_k.$$

Como f es continua en $[a, b]$, es Riemann integrable y, al refinar suficientemente la partición, la última cantidad puede hacerse tan pequeña como se quiera. Al mismo tiempo, las sumas etiquetadas de $x \longmapsto 2\pi x f(x)$ convergen a su integral.

PROPOSICIÓN 19.2. **Proposición (volumen por cascarones cilíndricos).** Sea $f : [a, b] \to [0, \infty)$ continua, con $0 \le a < b$. El volumen del sólido obtenido al hacer girar alrededor del eje y la región

$$\{(x, y) : a \le x \le b,\ 0 \le y \le f(x)\}$$

es

$$\boxed{V = 2\pi \int_a^b x f(x)\, dx.}$$

La demostración anterior muestra de dónde procede cada factor:

$$\boxed{2\pi x \ \times\ f(x)\ \times\ dx}$$

corresponde, respectivamente, a la circunferencia del cascarón, su altura y su espesor. La fórmula global no se memoriza aparte: es el límite de esa construcción local.

Dos representaciones del mismo volumen. La elección entre arandelas y cascarones depende de cómo describamos la misma región antes de hacerla girar.

EJEMPLO 19.4. **El mismo sólido con arandelas y cascarones.** Consideremos la región

$$0 \le x \le 1, \qquad 0 \le y \le x,$$

y hagámosla girar alrededor del eje y. Con **tiras verticales**, una tira situada en x genera un cascarón de radio x y altura x. Por Proposición 20.2,

$$V = 2\pi \int_0^1 x^2\, dx = \frac{2\pi}{3}.$$

La misma región puede describirse horizontalmente. Para un valor fijo de $y \in [0, 1]$, $y \le x \le 1$. Al girar esa tira horizontal alrededor del eje y, se obtiene una arandela de radio exterior $R(y) = 1$ y radio interior $r(y) = y$. Por Corolario 20.2,

$$\begin{aligned} V &= \pi \int_0^1 (1 - y^2)\, dy \\ &= \pi \left[y - \frac{y^3}{3} \right]_0^1 \\ &= \frac{2\pi}{3}. \end{aligned}$$

Ambas integrales representan el mismo sólido:

$$\boxed{2\pi \int_0^1 x^2\, dx = \pi \int_0^1 (1 - y^2)\, dy = \frac{2\pi}{3}.}$$

Este ejemplo muestra que el método se elige **antes** del cálculo. Una representación trabaja con tiras paralelas al eje de rotación y produce cascarones; otra trabaja con tiras perpendiculares y produce discos o arandelas. El criterio útil consiste en comparar las descripciones locales que genera cada elección:

- qué variable recorre la región con mayor naturalidad;
- qué fronteras deben escribirse como funciones de esa variable;
- cuántos tramos exige la descripción;
- qué radios, alturas o áreas transversales aparecen.

Una representación puede exigir despejar funciones o dividir el intervalo en varios tramos, mientras otra mantiene una sola expresión. La integral más conveniente surge de la geometría de la región y del eje de rotación.

Antes de seguir. Para una región dada y un eje de rotación fijo, ¿qué tipo de tira produce arandelas y cuál produce cascarones? ¿Por qué dos integrales con integrandos distintos pueden representar exactamente el mismo volumen? ¿Qué conviene examinar antes de decidir la variable de integración?

19.6. Longitud de la gráfica: de polígonos a $\int \sqrt{1+(f')^2}$

El área y el volumen se construyeron acumulando cantidades locales que ya tenían una interpretación directa: alturas, áreas transversales o cascarones. La longitud de una gráfica exige un paso distinto. Antes de disponer de una densidad que podamos integrar, debemos decidir cómo aproximar la curva mediante objetos cuya longitud sí sepamos calcular exactamente.

Sea $f \in C^1([a,b])$ y consideremos una partición

$$P: \quad a = x_0 < x_1 < \cdots < x_n = b.$$

Sobre la gráfica de f tomamos los puntos

$$P_k := (x_k, f(x_k)), \qquad k = 0, \ldots, n.$$

Uniendo cada par consecutivo P_{k-1} y P_k mediante un segmento recto obtenemos una poligonal inscrita en la gráfica. La longitud del segmento correspondiente al subintervalo $[x_{k-1}, x_k]$ es, por el teorema de Pitágoras,

$$\sqrt{(x_k - x_{k-1})^2 + \big(f(x_k) - f(x_{k-1})\big)^2}.$$

Por tanto, la longitud total de la poligonal es la suma de esas longitudes.

Definición 19.1. **Definición (longitud poligonal asociada a una partición).** Sea $f : [a,b] \to \mathbb{R}$ y sea

$$P: \quad a = x_0 < x_1 < \cdots < x_n = b$$

una partición. Definimos la longitud poligonal de la gráfica de f asociada a P por

$$\boxed{L_P(f) := \sum_{k=1}^{n} \sqrt{(x_k - x_{k-1})^2 + \big(f(x_k) - f(x_{k-1})\big)^2}.}$$

Esta definición no supone todavía que exista una longitud límite de la gráfica. Solo asigna una longitud exacta a cada aproximación poligonal determinada por una partición. La pregunta siguiente será cómo relacionar cada término de $L_P(f)$ con la derivada de f. Esa relación permitirá convertir la geometría de los segmentos en una suma de Riemann.

Del segmento a una suma etiquetada. Como $f \in C^1([a,b])$, podemos aplicar Teorema 9.3 en cada subintervalo $[x_{k-1}, x_k]$. Existe entonces un punto $\xi_k \in (x_{k-1}, x_k)$ tal que

$$\frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}} = f'(\xi_k).$$

Como $\Delta x_k := x_k - x_{k-1} > 0$, se sigue que $f(x_k) - f(x_{k-1}) = f'(\xi_k)\,\Delta x_k$.

Sustituyendo esta identidad en la longitud de la cuerda correspondiente,

$$\begin{aligned}\sqrt{(\Delta x_k)^2 + \big(f(x_k) - f(x_{k-1})\big)^2} &= \sqrt{(\Delta x_k)^2 + f'(\xi_k)^2 (\Delta x_k)^2} \\ &= \sqrt{1 + f'(\xi_k)^2}\,\Delta x_k.\end{aligned}$$

Por tanto, la longitud poligonal puede escribirse exactamente como

$$\boxed{L_P(f) = \sum_{k=1}^{n} \sqrt{1 + f'(\xi_k)^2}\,\Delta x_k.}$$

Esta ya es una suma etiquetada de la función $g(x) := \sqrt{1 + f'(x)^2}$. La aparición de g no es una conjetura geométrica añadida después: surge directamente de aplicar el teorema del valor medio a cada cuerda de la poligonal. Como f' es continua en $[a,b]$, también lo es g, y por tanto g es Riemann integrable. Queda por justificar que, al refinar las particiones, las longitudes poligonales convergen a una cantidad bien definida y que ese límite coincide con la integral de g.

Refinar una poligonal solo puede aumentar su longitud.

LEMA 19.1. **Lema (monotonía bajo refinamiento).** Si Q es un refinamiento de una partición P de $[a,b]$, entonces

$$L_P(f) \le L_Q(f).$$

Demostración. Basta considerar el caso en que Q se obtiene insertando un único punto $c \in (x_{k-1}, x_k)$ en P. La cuerda que une

$$(x_{k-1}, f(x_{k-1})) \quad \text{con} \quad (x_k, f(x_k))$$

queda reemplazada por dos segmentos que pasan por $(c, f(c))$. Por la desigualdad triangular para distancias euclidianas, la suma de las longitudes de esos dos segmentos es al menos la longitud de la cuerda original. Los demás segmentos no cambian. Por tanto, $L_P(f) \le L_Q(f)$. Insertando sucesivamente todos los puntos adicionales de Q obtenemos el caso general. □

Este lema organiza las aproximaciones poligonales de manera monótona bajo refinamiento y motiva definir la longitud mediante el supremo de todas ellas.

DEFINICIÓN 19.2. **Definición (longitud de la gráfica).** Sea $f : [a,b] \to \mathbb{R}$. Si el conjunto de longitudes poligonales

$$\{L_P(f) : P \text{ partición de } [a,b]\}$$

está acotado superiormente, definimos la longitud de la gráfica de f por

$$\boxed{L(f) := \sup_P L_P(f).}$$

Para funciones C^1 la acotación requerida en esta definición viene dada precisamente por la integral que estamos a punto de identificar.

TEOREMA 19.2. **Teorema (longitud de una gráfica C^1).** Si $f \in C^1([a,b])$, entonces la gráfica de f tiene longitud finita y

$$\boxed{L(f) = \int_a^b \sqrt{1 + f'(x)^2}\, dx.}$$

Demostración. Definamos $g(x) := \sqrt{1 + f'(x)^2}$.

Como f' es continua, también lo es g, y por tanto $g \in \mathcal{R}[a,b]$. Fijemos una partición cualquiera P. Construyamos una sucesión de refinamientos $P \preceq P_1 \preceq P_2 \preceq \cdots$ cuya malla tienda a cero. Por Lema 20.1,

$$L_P(f) \leq L_{P_m}(f) \qquad (m \geq 1).$$

Por el teorema del valor medio aplicado en cada subintervalo de P_m, existen etiquetas $\xi_{m,k}$ tales que

$$L_{P_m}(f) = \sum_k g(\xi_{m,k})\, \Delta x_{m,k}.$$

Como la malla de P_m tiende a cero y g es Riemann integrable, Teorema 14.2 implica

$$L_{P_m}(f) \longrightarrow \int_a^b g(x)\, dx.$$

Pasando al límite en la desigualdad anterior,

$$L_P(f) \leq \int_a^b g(x)\, dx.$$

Como P era arbitraria, todas las longitudes poligonales están acotadas por esa integral. En particular, Definición 20.2 es aplicable y

$$L(f) \leq \int_a^b g(x)\, dx.$$

Para la desigualdad opuesta, tomemos las particiones uniformes

$$U_n : \quad a = x_0^{(n)} < \cdots < x_n^{(n)} = b.$$

Su malla tiende a cero. De nuevo, por el teorema del valor medio, $L_{U_n}(f)$ es una suma etiquetada de g, de modo que Teorema 14.2 da

$$L_{U_n}(f) \longrightarrow \int_a^b g(x)\, dx.$$

Pero, por definición de supremo, $L_{U_n}(f) \leq L(f)$ para todo n. Al tomar límites,

$$\int_a^b g(x)\,dx \le L(f).$$

Combinando ambas desigualdades,

$$L(f) = \int_a^b g(x)\,dx = \int_a^b \sqrt{1 + f'(x)^2}\,dx.$$

Esto demuestra el teorema. □

La fórmula de longitud surge así de tres ingredientes que ya pertenecían al libro: geometría euclidiana de las cuerdas, teorema del valor medio y convergencia de sumas etiquetadas de Riemann.

Antes de seguir. ¿Por qué no basta observar que una sola longitud poligonal es una suma etiquetada? ¿Qué papel cumple la monotonía bajo refinamiento? ¿Dónde se usa exactamente la hipótesis $f \in C^1([a,b])$?

19.7. Densidad lineal y cantidad total

En una varilla o un cable, la masa puede distribuirse de manera desigual a lo largo de su longitud. La **densidad lineal** describe la contribución local que debemos acumular para obtener la masa total. Sea una varilla situada sobre el intervalo $[a,b]$. Una función $\rho : [a,b] \longrightarrow [0,\infty)$ describe su densidad lineal si $\rho(x)$ representa la cantidad de masa por unidad de longitud cerca del punto x. La densidad y la masa total no son la misma magnitud. Si x se mide en metros y la masa en kilogramos, entonces $[\rho] = \frac{\text{kg}}{\text{m}}$, mientras que una cantidad de masa tiene unidades de kg. Consideremos un subintervalo pequeño $[x_{k-1}, x_k]$ de longitud $\Delta x_k = x_k - x_{k-1}$.

Si escogemos un punto ξ_k dentro de ese tramo y la densidad varía poco allí, la masa contenida en la pequeña pieza se aproxima por

$$\boxed{\Delta M_k \approx \rho(\xi_k)\,\Delta x_k.}$$

La interpretación dimensional confirma la estructura: $\frac{\text{kg}}{\text{m}} \cdot \text{m} = \text{kg}$. Así, al dividir la varilla mediante una partición, la masa total se aproxima por

$$\sum_{k=1}^{n} \rho(\xi_k)\,\Delta x_k.$$

La expresión tiene exactamente la forma que apareció en §19.1. En este caso,

$$\boxed{\text{densidad lineal} \times \text{longitud} = \text{cantidad local.}}$$

La misma estructura puede modelar otras cantidades distribuidas a lo largo de una dimensión: carga eléctrica por unidad de longitud, cantidad de material por unidad de longitud o cualquier magnitud cuya contribución sobre un tramo pequeño quede controlada por una densidad lineal. El paso siguiente será convertir esta aproximación local en una fórmula exacta para la cantidad total, usando el principio de cantidad aditiva determinada por una densidad.

De la densidad a la cantidad total. Sea ahora

$$\rho \in \mathcal{R}[a,b], \qquad \rho(x) \geq 0 \quad (x \in [a,b]),$$

y sea $M(\alpha,\beta)$ la masa contenida en el tramo $[\alpha,\beta]$ de la varilla. La masa es aditiva sobre intervalos contiguos: $M(\alpha,\gamma) = M(\alpha,\beta) + M(\beta,\gamma)$ siempre que

$$a \leq \alpha \leq \beta \leq \gamma \leq b.$$

Además, si en un subintervalo $[\alpha,\beta]$ la densidad está comprendida entre sus valores extremo inferior y superior, la masa de ese tramo queda encerrada entre las masas que tendrían dos varillas uniformes con esas densidades constantes:

$$\inf_{[\alpha,\beta]} \rho\,(\beta-\alpha) \leq M(\alpha,\beta) \leq \sup_{[\alpha,\beta]} \rho\,(\beta-\alpha).$$

Estamos exactamente en la situación de Teorema 20.1.

Corolario 19.3. **Corolario (cantidad total a partir de una densidad lineal).** Sea $\rho \in \mathcal{R}[a,b]$ una densidad lineal no negativa y sea $M(\alpha,\beta)$ una cantidad aditiva asociada a cada subintervalo, sometida al control local

$$\inf_{[\alpha,\beta]} \rho\,(\beta-\alpha) \leq M(\alpha,\beta) \leq \sup_{[\alpha,\beta]} \rho\,(\beta-\alpha).$$

Entonces

$$\boxed{M(a,b) = \int_a^b \rho(x)\,dx.}$$

Demostración. Basta aplicar Teorema 20.1 con

$$Q(\alpha,\beta) := M(\alpha,\beta) \qquad \text{y} \qquad q(x) := \rho(x).$$

□

En el caso de una varilla, podemos escribir simplemente

$$\boxed{M = \int_a^b \rho(x)\,dx.}$$

La fórmula expresa que la masa total es la acumulación de masa por unidad de longitud. Su lectura dimensional es

$$\left[\frac{\text{masa}}{\text{longitud}}\right][\text{longitud}] = [\text{masa}].$$

La misma construcción sirve para otras cantidades unidimensionales. Si $\lambda(x)$ representa carga eléctrica por unidad de longitud, por ejemplo, la carga total es

$$Q = \int_a^b \lambda(x)\,dx,$$

siempre que el modelo satisfaga las correspondientes hipótesis de aditividad e integrabilidad. El símbolo que recibe la densidad puede cambiar, y también puede cambiar la naturaleza física de la cantidad acumulada. La estructura matemática permanece:

$$\boxed{\text{cantidad total} = \int \text{densidad respecto de la variable}}.$$

En la sección siguiente la densidad dejará de usarse solo para calcular cuánto hay. La emplearemos además como **peso** para determinar dónde se encuentra distribuida esa cantidad.

Ejemplo 19.5. **Una varilla con densidad creciente.** Una varilla ocupa el intervalo $0 \le x \le 2$ metros. Su densidad lineal viene dada numéricamente por $\rho(x) = 1+2x$ kilogramos por metro; el coeficiente de x incorpora las unidades necesarias para que ρ tenga siempre unidades de masa por longitud.

La masa total es

$$M = \int_0^2 (1 + 2x)\, dx.$$

Por tanto,

$$\begin{aligned} M &= \Big[x + x^2\Big]_0^2 \\ &= 2 + 4 \\ &= 6. \end{aligned}$$

Así,

$$\boxed{M = 6 \text{ kg}.}$$

El valor $\rho(2) = 5$ no es la masa de la varilla: es una densidad lineal, medida en kg/m, en el extremo derecho. Del mismo modo, la longitud total 2 m tampoco determina por sí sola la masa, porque la densidad no es constante.

El control dimensional es

$$\left(\frac{\text{kg}}{\text{m}}\right)(\text{m}) = \text{kg},$$

exactamente las unidades del resultado.

Antes de seguir. ¿Qué información aporta $\rho(2)$ y qué información aporta $\int_0^2 \rho(x)\, dx$? ¿Por qué multiplicar la longitud total por una densidad tomada en un solo punto no determina en general la masa? ¿Qué unidades debe tener siempre la integral de una densidad lineal respecto de la longitud?

19.8. Promedio ponderado, primer momento y centro de masa de una varilla

La densidad ρ permite responder primero cuánto hay en total:

$$M = \int_a^b \rho(x)\, dx.$$

Ahora queremos usar la misma función para describir **dónde** está distribuida esa masa. La herramienta necesaria ya fue construida en C16. Según Definición 15.2, si una función w actúa como peso y

$$\int_a^b w(x)\,dx \neq 0,$$

podemos formar el promedio integral de una función f ponderado por w. En una varilla, la función que queremos promediar es la propia posición, $p(x) := x$, y el peso natural es la densidad, $w(x) := \rho(x)$. Esto expresa una idea sencilla: una posición donde hay más masa debe influir más en la posición media de la varilla que otra donde hay menos. Para que ese promedio esté definido necesitamos que la masa total sea positiva:

$$M = \int_a^b \rho(x)\,dx > 0.$$

Esta hipótesis merece quedar visible. De $\rho \geq 0$ y $\rho \not\equiv 0$ no se sigue automáticamente, para una función meramente Riemann-integrable, que su integral sea positiva; C16 ya separó cuidadosamente ambas condiciones. Bajo la hipótesis $M > 0$, Definición 15.2 permite considerar $\mathrm{Avg}_{\rho,[a,b]}(p)$. Este objeto es el promedio ponderado de la posición: cada punto x interviene con un peso proporcional a la cantidad de masa situada cerca de él. La estructura conceptual es, por tanto,

$$\boxed{\text{posición} + \text{densidad como peso} \longrightarrow \text{posición media ponderada.}}$$

En el siguiente paso identificaremos el numerador que aparece en este promedio, le daremos su interpretación como **primer momento** y obtendremos de ahí la fórmula del centro de masa.

Primer momento y centro de masa. El numerador del promedio ponderado de la posición es

$$\int_a^b x\rho(x)\,dx.$$

Cada contribución local tiene la forma $x\,\rho(x)\,dx$. La densidad aporta la cantidad de masa por unidad de longitud y el factor x registra la posición de esa masa respecto del origen.

Definición 19.3. **Definición (primer momento respecto del origen).** Sea $\rho \in \mathcal{R}[a,b]$ una densidad lineal tal que $x\rho(x)$ sea Riemann-integrable. Definimos el **primer momento respecto del origen** por

$$\boxed{\mathcal{M}_0 := \int_a^b x\rho(x)\,dx.}$$

Si x se mide en unidades de longitud y ρ en masa por longitud, entonces

$$[\mathcal{M}_0] = [\text{longitud}]\left[\frac{\text{masa}}{\text{longitud}}\right][\text{longitud}] = [\text{masa}]\,[\text{longitud}].$$

El primer momento no es una posición: combina masa y posición. Para obtener una posición debemos normalizarlo por la masa total.

Corolario 19.4. **Corolario (centro de masa de una varilla).** Sea $\rho \in \mathcal{R}[a,b]$ una densidad lineal no negativa y supongamos

$$M := \int_a^b \rho(x)\,dx > 0.$$

El centro de masa de la varilla es el promedio de la posición ponderado por ρ y viene dado por

$$\boxed{\bar{x} = \frac{\displaystyle\int_a^b x\rho(x)\,dx}{\displaystyle\int_a^b \rho(x)\,dx} = \frac{\mathcal{M}_0}{M}.}$$

Demostración. Aplicamos Definición 15.2 a la función posición $p(x) = x$ con peso $w(x) = \rho(x)$.

Como

$$\int_a^b \rho(x)\,dx = M > 0,$$

el promedio ponderado está definido y

$$\operatorname{Avg}_{\rho,[a,b]}(p) = \frac{\displaystyle\int_a^b x\rho(x)\,dx}{\displaystyle\int_a^b \rho(x)\,dx}.$$

Por definición física del centro de masa como posición media ponderada por la masa, esta cantidad es $\bar{x}$. □

La normalización por M elimina la unidad de masa:

$$\frac{[\text{masa}]\,[\text{longitud}]}{[\text{masa}]} = [\text{longitud}],$$

como corresponde a una posición.

La fórmula muestra también la diferencia entre las dos preguntas de esta parte del capítulo:

$$\int_a^b \rho(x)\,dx$$

responde **cuánta masa hay**, mientras que

$$\frac{\int_a^b x\rho(x)\,dx}{\int_a^b \rho(x)\,dx}$$

responde **dónde está centrada esa masa**.

El centro de masa permanece dentro de la varilla. La teoría ponderada de C16 proporciona inmediatamente una cota geométrica importante. Como la función posición $p(x) = x$ es continua en $[a, b]$, la densidad satisface $\rho \geq 0$ y

$$\int_a^b \rho(x)\,dx = M > 0,$$

podemos aplicar Teorema 15.2. Existe entonces un punto $c \in [a, b]$ tal que $\operatorname{Avg}_{\rho,[a,b]}(p) = p(c)$. Por Corolario 20.4, el miembro izquierdo es $\bar{x}$, mientras que $p(c) = c$. Por tanto, $\bar{x} = c \in [a, b]$.

En consecuencia,

$$\boxed{a \le \bar{x} \le b.}$$

Esta cota no requiere calcular explícitamente ninguna integral: es una consecuencia estructural del hecho de que el centro de masa es un promedio ponderado de posiciones pertenecientes al propio intervalo.

Ejemplo 19.6. **Centro de masa de la varilla con densidad creciente.** Retomemos la varilla de Ejemplo 20.5:

$$0 \le x \le 2, \qquad \rho(x) = 1 + 2x.$$

Ya sabemos que su masa total es $M = 6$. Su primer momento respecto del origen es

$$\begin{aligned} \mathcal{M}_0 &= \int_0^2 x(1+2x)\,dx \\ &= \int_0^2 (x + 2x^2)\,dx \\ &= \left[\frac{x^2}{2} + \frac{2x^3}{3}\right]_0^2 \\ &= 2 + \frac{16}{3} \\ &= \frac{22}{3}. \end{aligned}$$

Por tanto,

$$\bar{x} = \frac{\mathcal{M}_0}{M} = \frac{22/3}{6} = \frac{11}{9}.$$

Así,

$$\boxed{\bar{x} = \frac{11}{9}\ \text{m}.}$$

La cota general se verifica: $0 < \frac{11}{9} < 2$. Además, $\frac{11}{9} > 1$, de modo que el centro de masa queda a la derecha del punto medio geométrico de la varilla. Esto concuerda con el modelo: la densidad aumenta con x, por lo que las posiciones cercanas al extremo derecho reciben mayor peso.

Antes de seguir. ¿Por qué $a \le \bar{x} \le b$ puede deducirse sin calcular las integrales? ¿Qué papel cumple la condición $M > 0$? ¿Por qué una densidad creciente puede desplazar el centro de masa respecto del punto medio geométrico?

19.9. Trabajo de una fuerza variable

Cuando una fuerza constante actúa en la misma dirección que un desplazamiento rectilíneo, el trabajo se calcula como el producto de la fuerza por el desplazamiento. Si la fuerza cambia con la posición, esa fórmula ya no puede aplicarse usando un único valor de la fuerza sobre todo el recorrido. Supongamos que una partícula se mueve a lo largo del eje x desde $x = a$ hasta $x = b$, y que $F : [a, b] \longrightarrow \mathbb{R}$ representa la componente de la fuerza en la dirección positiva del eje. Dividamos el intervalo mediante una partición

$$a = x_0 < x_1 < \cdots < x_n = b.$$

En un subintervalo pequeño $[x_{k-1}, x_k]$, elegimos un punto ξ_k. Si la fuerza varía poco en ese tramo, podemos aproximarla por el valor constante $F(\xi_k)$. El trabajo realizado en la pequeña pieza queda entonces aproximado por

$$\boxed{\Delta W_k \approx F(\xi_k)\,\Delta x_k.}$$

La estructura vuelve a ser la de §19.1:

$$\boxed{\text{fuerza local} \times \text{desplazamiento} = \text{trabajo local.}}$$

Al sumar todos los aportes obtenemos

$$\sum_{k=1}^{n} F(\xi_k)\,\Delta x_k.$$

Aquí el signo forma parte del modelo. Si $F(\xi_k) > 0$, la fuerza favorece el desplazamiento positivo y la contribución al trabajo es positiva. Si $F(\xi_k) < 0$, la fuerza se opone a ese desplazamiento y la contribución es negativa. Por tanto, el trabajo que estamos acumulando es una cantidad **orientada**. Reemplazar F por $|F|$ cambiaría la pregunta: eliminaría la información sobre si la fuerza ayuda o se opone al movimiento. El control dimensional también coincide con la interpretación física. Si la fuerza se mide en newtons y la posición en metros, [trabajo] = [N][m] = [J]. Así, la integral que aparecerá debe tener unidades de energía. El siguiente paso será convertir esta suma local en la fórmula exacta del trabajo de una fuerza variable.

De la fuerza local al trabajo total. Sea ahora $F \in \mathcal{R}[a, b]$, y denote $W(\alpha, \beta)$ el trabajo realizado por la fuerza cuando la partícula se desplaza desde $x = \alpha$ hasta $x = \beta$, con

$$a \le \alpha \le \beta \le b.$$

El trabajo es aditivo sobre intervalos contiguos: $W(\alpha, \gamma) = W(\alpha, \beta) + W(\beta, \gamma)$ si

$$a \le \alpha \le \beta \le \gamma \le b.$$

Además, sobre un tramo $[\alpha, \beta]$, la componente de la fuerza queda comprendida entre $\inf_{[\alpha,\beta]} F$ y $\sup_{[\alpha,\beta]} F$. El modelo físico local se formaliza exigiendo que el trabajo realizado en ese tramo quede sometido al encierro

$$\inf_{[\alpha,\beta]} F\,(\beta - \alpha) \le W(\alpha, \beta) \le \sup_{[\alpha,\beta]} F\,(\beta - \alpha).$$

Esta condición expresa que, sobre un desplazamiento positivo $\beta - \alpha$, el trabajo de una fuerza variable queda entre los trabajos que producirían fuerzas constantes iguales a sus valores extremo inferior y superior en el tramo. Junto con la aditividad, proporciona exactamente las hipótesis necesarias para aplicar Teorema 20.1.

Corolario 19.5. **Corolario (trabajo de una fuerza variable).** Sea $F \in \mathcal{R}[a, b]$ la componente de una fuerza sobre el eje x, positiva en el sentido de los valores crecientes de x, y supongamos que la partícula se desplaza desde $x = a$ hasta $x = b$. Si el trabajo sobre subintervalos es aditivo y satisface el control local anterior, entonces

$$\boxed{W = \int_a^b F(x)\,dx.}$$

Demostración. Aplicamos Teorema 20.1 con

$$Q(\alpha,\beta) := W(\alpha,\beta) \qquad \text{y} \qquad q(x) := F(x).$$

Así,

$$W(a,b) = \int_a^b F(x)\,dx.$$

□

La fórmula conserva el signo de la fuerza. Por ejemplo, si

$$F(x) \le 0 \qquad (x \in [a,b]),$$

mientras el desplazamiento se realiza de a hacia b, entonces por monotonía de la integral $W \le 0$. Esto expresa que la fuerza se opone al movimiento en todo el recorrido. La orientación también importa en los extremos. Con la convención de integral orientada,

$$\int_b^a F(x)\,dx = -\int_a^b F(x)\,dx.$$

Por tanto, recorrer el mismo intervalo en sentido contrario cambia el signo del trabajo asociado al mismo campo unidimensional de fuerzas.

Finalmente, las unidades verifican el modelo:

$$\int_a^b \underbrace{F(x)}_{\mathrm{N}}\underbrace{dx}_{\mathrm{m}}$$

tiene unidades de $\mathrm{N\,m} = \mathrm{J}$.

La integral representa, por tanto, una acumulación orientada de energía transferida por la fuerza a lo largo del desplazamiento.

El resorte de Hooke: qué fuerza realiza el trabajo. Un resorte ideal permite ver por qué el signo debe fijarse antes de integrar. Sea $x \ge 0$ la elongación medida desde la posición de equilibrio y sea $k > 0$ la constante elástica. La fuerza que ejerce el resorte sobre el extremo desplazado es restauradora: $F_{\mathrm{res}}(x) = -kx$. Si el extremo se mueve desde $x = 0$ hasta $x = L > 0$, el trabajo realizado por el propio resorte es

$$W_{\mathrm{res}} = \int_0^L (-kx)\,dx = -\frac{kL^2}{2}.$$

El signo negativo expresa que la fuerza del resorte apunta en sentido opuesto al desplazamiento.

En cambio, si el resorte se estira lentamente mediante una fuerza externa que equilibra en cada posición a la fuerza restauradora, entonces $F_{\mathrm{ext}}(x) = kx$. El trabajo realizado por esa fuerza externa es

$$W_{\mathrm{ext}} = \int_0^L kx\,dx = \frac{kL^2}{2}.$$

Por tanto,

$$\boxed{W_{\text{ext}} = -\,W_{\text{res}} = \frac{kL^2}{2}.}$$

Ejemplo 19.7. **Trabajo para estirar un resorte.** Un resorte ideal tiene constante elástica $k = 40$ N/m.

¿Cuánto trabajo debe realizar una fuerza externa para estirarlo lentamente desde su posición de equilibrio hasta una elongación de $L = 0.30$ m?

La fuerza externa necesaria en la posición x es $F_{\text{ext}}(x) = 40x$.

Luego

$$\begin{aligned} W_{\text{ext}} &= \int_0^{0.30} 40x\,dx \\ &= 20x^2\Big|_0^{0.30} \\ &= 20(0.30)^2 \\ &= 1.8. \end{aligned}$$

Así,

$$\boxed{W_{\text{ext}} = 1.8 \text{ J}.}$$

El resorte realiza simultáneamente trabajo $W_{\text{res}} = -1.8$ J. Las dos respuestas no compiten: corresponden a fuerzas distintas.

El ejemplo ilustra una regla de modelización esencial: antes de escribir la integral hay que identificar **qué fuerza** se está integrando y cuál es la orientación del desplazamiento. La misma trayectoria puede producir trabajos con signos opuestos para fuerzas diferentes.

Antes de seguir. ¿Qué debe identificarse antes de integrar una fuerza variable: la magnitud de la fuerza, su componente en la dirección del movimiento o ambas? ¿Cómo cambia el signo del trabajo si se invierte la orientación del desplazamiento? ¿Por qué el trabajo del resorte y el de la fuerza externa pueden tener igual magnitud y signos opuestos sobre la misma trayectoria?

Una tasa conocida aporta otra contribución local: multiplicada por un pequeño incremento de su variable, permite aproximar un cambio. Su acumulación reconstruirá el cambio neto bajo las hipótesis correspondientes.

19.10. Cambio neto desde una tasa conocida

Muchas cantidades cambian a lo largo del tiempo —o de otra variable— y lo que conocemos directamente no es la cantidad misma, sino su **tasa de cambio**. Supongamos que $Q : [a, b] \longrightarrow \mathbb{R}$ es continua en $[a, b]$ y diferenciable en (a, b), y que su tasa $r(t) := Q'(t)$ es Riemann-integrable. Entonces Teorema 16.3 da

$$\boxed{Q(b) - Q(a) = \int_a^b r(t)\,dt.}$$

Esta igualdad admite una lectura cuantitativa inmediata:

> **integrar una tasa conocida sobre un intervalo recupera el cambio neto de la cantidad cuya tasa estamos acumulando.**

La palabra *neto* es esencial. Si $r(t) > 0$, la cantidad Q aumenta localmente; si $r(t) < 0$, disminuye. La integral

$$\int_a^b r(t)\,dt$$

conserva esos signos y permite que aumentos y disminuciones se compensen. Por eso no debemos confundir

$$\boxed{\int_a^b r(t)\,dt}$$

con

$$\boxed{\int_a^b |r(t)|\,dt.}$$

La primera expresión mide el **cambio neto**. La segunda acumula la magnitud de los cambios locales sin cancelación cuando ese modelo es el que corresponde a la pregunta planteada. Son cantidades distintas y deben elegirse antes de calcular. La estructura local vuelve a coincidir con el principio general del capítulo. En un intervalo pequeño de longitud Δt, $\Delta Q \approx r(\xi)\,\Delta t$. Al sumar esas variaciones locales y pasar al límite aparece

$$\int_a^b r(t)\,dt,$$

mientras que Newton–Leibniz identifica esa acumulación exactamente con $Q(b) - Q(a)$. Aquí no estamos resolviendo una ecuación diferencial. La tasa r se considera **conocida** y la tarea consiste únicamente en acumularla para obtener el cambio de una cantidad entre dos instantes o posiciones.

Movimiento rectilíneo: desplazamiento y distancia. Sea $s : [a, b] \longrightarrow \mathbb{R}$ la posición de una partícula en movimiento rectilíneo y supongamos $s \in C^1([a, b])$.

Su velocidad es $v(t) := s'(t)$.

Por Teorema 16.3,

$$\boxed{s(b) - s(a) = \int_a^b v(t)\,dt.}$$

El miembro izquierdo es el **desplazamiento**: compara únicamente la posición final con la inicial. Por eso conserva la orientación y permite cancelación entre intervalos en los que la partícula se mueve en sentidos opuestos. La **distancia recorrida** responde a otra pregunta. Localmente, la magnitud del desplazamiento sobre un intervalo pequeño se aproxima por $|v(\xi)|\,\Delta t$, porque $|v(t)|$ es la rapidez. Al acumular esas contribuciones no negativas obtenemos

$$\boxed{D = \int_a^b |v(t)|\,dt.}$$

Así,

$$\boxed{\text{desplazamiento} = \int_a^b v(t)\,dt}$$

mientras que

$$\boxed{\text{distancia recorrida} = \int_a^b |v(t)|\,dt.}$$

Si la velocidad no cambia de signo, ambas magnitudes coinciden en valor absoluto. Si cambia de signo, la integral de v cancela recorridos en sentidos opuestos, mientras que la integral de $|v|$ los suma.

En particular,

$$\left|\int_a^b v(t)\,dt\right| \le \int_a^b |v(t)|\,dt,$$

de modo que

$$\boxed{|s(b) - s(a)| \le D.}$$

La desigualdad expresa una propiedad geométrica elemental: la distancia total recorrida no puede ser menor que la distancia entre la posición inicial y la final.

Tasas de entrada y salida: el balance neto. Otra situación frecuente aparece cuando una cantidad recibe aportes y pérdidas simultáneamente. Supongamos que

$$r_{\text{in}}, r_{\text{out}} \in \mathcal{R}[a,b], \qquad r_{\text{in}}(t) \ge 0, \qquad r_{\text{out}}(t) \ge 0,$$

representan, respectivamente, una tasa de entrada y una tasa de salida, medidas en cantidad por unidad de tiempo. La orientación del balance no está en el signo de cada tasa por separado, sino en la resta $r_{\text{in}} - r_{\text{out}}$. La tasa neta de cambio es

$$\boxed{r_{\text{net}}(t) = r_{\text{in}}(t) - r_{\text{out}}(t).}$$

Si $Q(t)$ denota la cantidad presente y el modelo satisface $Q'(t) = r_{\text{net}}(t)$, entonces Teorema 16.3 da

$$Q(b) - Q(a) = \int_a^b \big(r_{\text{in}}(t) - r_{\text{out}}(t)\big)\,dt.$$

Por linealidad,

$$\boxed{Q(b) - Q(a) = \int_a^b r_{\text{in}}(t)\,dt - \int_a^b r_{\text{out}}(t)\,dt.}$$

Esta identidad separa tres cantidades distintas:

$$\underbrace{\int_a^b r_{\text{in}}(t)\,dt}_{\text{entrada acumulada}}, \qquad \underbrace{\int_a^b r_{\text{out}}(t)\,dt}_{\text{salida acumulada}},$$

y

$$\underbrace{Q(b) - Q(a)}_{\text{cambio neto}}.$$

Una cantidad puede tener cambio neto cero aunque durante el intervalo haya entrado y salido una cantidad considerable. Basta que las dos acumulaciones coincidan:

$$\int_a^b r_{\text{in}}(t)\,dt = \int_a^b r_{\text{out}}(t)\,dt.$$

En ese caso, $Q(b) = Q(a)$, pero eso no significa que «no haya ocurrido nada»: solo significa que los aportes totales y las pérdidas totales se compensaron exactamente. El control dimensional vuelve a ser inmediato. Si las tasas se miden, por ejemplo, en litros por minuto, entonces

$$\left[\frac{\text{L}}{\text{min}}\right][\text{min}] = [\text{L}],$$

de modo que cada integral representa una cantidad de volumen y su diferencia representa un cambio neto de volumen. Este balance no constituye una resolución de una ecuación diferencial. Las tasas se consideran dadas y la integral solo acumula sus efectos sobre el intervalo.

Ejemplo 19.8. **Cambio neto nulo con distancia recorrida positiva.** Una partícula se mueve durante tres segundos con velocidad

$$v(t) = \begin{cases} 2, & 0 \le t \le 1, \\ -1, & 1 < t \le 3, \end{cases} \qquad \text{m/s}.$$

Su desplazamiento entre $t = 0$ y $t = 3$ es

$$\begin{aligned} s(3) - s(0) &= \int_0^3 v(t)\,dt \\ &= \int_0^1 2\,dt + \int_1^3 (-1)\,dt \\ &= 2 - 2 \\ &= 0. \end{aligned}$$

La partícula termina, por tanto, en la misma posición en la que comenzó. La distancia recorrida acumula la rapidez:

$$\begin{aligned} D &= \int_0^3 |v(t)|\,dt \\ &= \int_0^1 2\,dt + \int_1^3 1\,dt \\ &= 2 + 2 \\ &= 4. \end{aligned}$$

Así,

$$\boxed{s(3) - s(0) = 0, \qquad D = 4 \text{ m}.}$$

El ejemplo muestra que un cambio neto nulo puede coexistir con una acumulación total positiva: las contribuciones orientadas se cancelan en el desplazamiento, mientras que sus magnitudes se suman en la distancia.

Antes de seguir. Si una tasa cambia de signo, ¿qué información conserva $\int_a^b r(t)\,dt$ y qué información conserva $\int_a^b |r(t)|\,dt$? ¿Cómo puede una cantidad terminar con el mismo valor inicial después de haber experimentado cambios apreciables? En un problema de entradas y salidas, ¿qué debe integrarse para obtener el cambio neto?

Las aplicaciones anteriores comparten una misma decisión de modelización: identificar primero la cantidad local que debe acumularse y determinar si su signo forma parte de la información buscada. La sección siguiente reunirá estas decisiones en problemas donde el desafío principal será construir la integral adecuada antes de evaluarla.

19.11. Laboratorio de modelización: elegir cortes, densidades y controles

Para justificar la integral en cada contexto necesitamos precisar qué contribución estamos acumulando:

> **¿Qué cantidad local se está sumando y por qué su suma finita conduce precisamente a esta integral?**

Una forma práctica de organizar la modelización es la siguiente.

1. **Identificar la cantidad que se busca.** Puede ser área, volumen, longitud, masa, primer momento, trabajo, desplazamiento, distancia o cambio neto.
2. **Elegir la variable de acumulación.** La variable debe parametrizar de manera natural los pequeños elementos que se sumarán: anchura dx, espesor dx o dy, intervalo temporal dt, entre otros.
3. **Construir la contribución local.** Sobre un tramo pequeño, escribir una cantidad del tipo

$$\boxed{\text{factor local} \times \text{incremento de la variable.}}$$

   El factor local puede ser una altura, un área de sección, una densidad, una fuerza, una rapidez o una tasa.
4. **Decidir cómo interviene el signo.** En algunas cantidades el signo conserva información de orientación o balance; en otras, la magnitud local debe ser no negativa. Esta decisión pertenece al modelo y debe tomarse antes de integrar.
5. **Fijar los límites y controlar unidades.** Los extremos deben corresponder al dominio efectivo de acumulación, y las unidades del integrando multiplicadas por las de la variable deben coincidir con las unidades de la cantidad final.

Solo después de estos pasos aparece la integral de acumulación:

$$\boxed{\text{cantidad acumulada} = \int \text{contribución por unidad } d(\text{variable}).}$$

Algunas magnitudes finales se obtienen directamente de esta acumulación —como área, volumen, masa, trabajo o cambio neto— y otras se construyen a partir de varias acumulaciones. El centro de masa, por ejemplo, normaliza el primer momento mediante la masa total. La evaluación simbólica pertenece a una etapa posterior. Las técnicas de C19 entran en juego una vez que el modelo ya ha determinado qué integral o qué combinación de integrales debe calcularse. Este orden permite detectar errores antes de realizar álgebra. Una integral puede estar evaluada correctamente y, sin embargo, responder a una pregunta distinta si se eligieron mal los cortes, la densidad, el signo o la variable de acumulación.

Ejemplo 19.9. **Diagnóstico de representación: ¿disco o cascarón?** Consideremos la región comprendida entre

$$y = x^2 \qquad \text{y} \qquad y = x,$$

para $0 \le x \le 1$, y hagámosla girar alrededor del eje y. Supongamos que decidimos trabajar con tiras verticales de espesor dx. Aparecen entonces dos propuestas:

$$\text{(A)} \qquad 2\pi \int_0^1 x(x - x^2)\, dx,$$

y

$$\text{(B)} \qquad \pi \int_0^1 (x - x^2)^2\, dx.$$

Antes de calcular debemos identificar qué sólido elemental genera una tira vertical. La tira situada en la posición x tiene:

- distancia al eje de rotación: x;
- altura: $x - x^2$;
- espesor: dx.

Al girar alrededor del eje y, esa tira genera un cascarón cilíndrico. Su volumen local es, por tanto,

$$\Delta V \approx 2\pi \underbrace{x}_{\text{radio}} \underbrace{(x - x^2)}_{\text{altura}}\, dx.$$

Esto conduce a

$$\boxed{V = 2\pi \int_0^1 x(x - x^2)\, dx.}$$

La propuesta (B) trata la altura $x - x^2$ como si fuera el radio de un disco. Esa interpretación no corresponde al elemento geométrico producido por una tira vertical al girar alrededor del eje y. También sería posible construir una representación correcta mediante arandelas, pero entonces habría que usar cortes horizontales, expresar las fronteras en función de y y acumular con respecto a dy. El método cambia porque cambia el elemento local. El diagnóstico se decide, por tanto, antes de toda integración:

$$\boxed{\text{tira elegida} \longrightarrow \text{sólido elemental} \longrightarrow \text{contribución local} \longrightarrow \text{integral.}}$$

Ejemplo 19.10. **Diagnóstico de signo: acumulación orientada o área geométrica.** Consideremos

$$f(x) = x^2 - 1, \qquad -2 \le x \le 2.$$

La gráfica corta el eje x en

$$x = -1 \qquad \text{y} \qquad x = 1.$$

Supongamos que queremos calcular el **área geométrica total** comprendida entre la gráfica y el eje x. Dos expresiones pueden parecer plausibles:

$$\text{(A)} \qquad \int_{-2}^{2} (x^2 - 1)\, dx,$$

y

$$\text{(B)} \qquad \int_{-2}^{2} |x^2 - 1|\, dx.$$

La decisión se toma observando la contribución local. El área de una tira vertical debe ser no negativa, de modo que su altura es $|f(x)| = |x^2 - 1|$. Por tanto, el modelo correcto para el área geométrica es

$$\boxed{A = \int_{-2}^{2} |x^2 - 1|\, dx.}$$

Como el signo cambia en $x = \pm 1$, podemos escribir

$$\begin{aligned} A &= 2\int_{1}^{2} (x^2 - 1)\, dx + \int_{-1}^{1} (1 - x^2)\, dx \\ &= 2\left(\frac{4}{3}\right) + \frac{4}{3} \\ &= 4. \end{aligned}$$

En cambio,

$$\int_{-2}^{2} (x^2 - 1)\, dx = \frac{4}{3}.$$

Este segundo valor es la **acumulación orientada** de f: las contribuciones situadas bajo el eje se restan de las situadas sobre él. Responde a una pregunta distinta. Así, las dos integrales son matemáticamente legítimas, pero modelan cantidades diferentes:

$$\boxed{\int_{-2}^{2} f(x)\, dx \quad \text{mide acumulación orientada,}}$$

mientras que

$$\boxed{\int_{-2}^{2} |f(x)|\, dx \quad \text{mide área geométrica total.}}$$

El valor absoluto no se introduce como una corrección posterior. Surge al decidir que cada contribución local de área debe ser no negativa.

Ejemplo 19.11. **Diagnóstico de unidades e información insuficiente.** Una varilla ocupa el intervalo $0 \le x \le 2$ m y tiene masa total $M = 6$ kg.

¿Bastan estos datos para determinar su centro de masa?

Una propuesta podría ser

$$\bar{x} \stackrel{?}{=} \frac{1}{6}\int_{0}^{2} 6x\, dx.$$

Antes de calcular, el control dimensional detecta un problema. El número 6 dado en el enunciado representa una **masa total**, con unidades de kilogramos. Por tanto, $x\,(6\text{ kg})\,dx$

tiene unidades [m] [kg] [m] = [kg m^2], y dividir por 6 kg produciría unidades de [m^2], que no corresponden a una posición. En la fórmula de Corolario 20.4, el factor que multiplica a x debe ser una densidad lineal $\rho(x)$, con unidades de masa por longitud. El dato $M = 6$ kg no proporciona por sí solo esa función. De hecho, dos varillas pueden tener la misma longitud y la misma masa total, pero centros de masa diferentes. Si la densidad es uniforme, $\rho_1(x) = 3$ kg/m, entonces

$$\int_0^2 \rho_1(x)\,dx = 6 \text{ kg}$$

y, por simetría o por Corolario 20.4, $\bar{x}_1 = 1$ m.

En cambio, la densidad de Ejemplo 20.5,

$$\rho_2(x) = 1 + 2x \qquad (\text{kg/m}),$$

también satisface

$$\int_0^2 \rho_2(x)\,dx = 6 \text{ kg},$$

pero Ejemplo 20.6 dio $\bar{x}_2 = \frac{11}{9}$ m.

Así,

$$\boxed{\text{misma longitud} + \text{misma masa total} \not\Longrightarrow \text{mismo centro de masa.}}$$

El modelo requiere información sobre **cómo está distribuida la masa**. Cuando esa información falta, la respuesta correcta es diagnosticar que el problema está subdeterminado, en lugar de inventar una densidad.

Síntesis del laboratorio. Los ejemplos anteriores muestran que modelar una aplicación de la integral exige tomar varias decisiones antes de evaluar:

$$\boxed{\text{cantidad buscada} \longrightarrow \text{variable} \longrightarrow \text{contribución local} \longrightarrow \text{signo} \longrightarrow \text{límites y unidades} \longrightarrow \text{integral.}}$$

Cada etapa controla a la siguiente. La elección de cortes determina la forma de la contribución local; el significado de la cantidad decide si el signo debe conservarse o si corresponde acumular magnitudes; las unidades permiten detectar incompatibilidades; y la información disponible debe bastar para determinar el factor local que se integra. Solo después de construir el modelo entra la etapa de cálculo. En ese punto pueden aplicarse las técnicas desarrolladas en C19, pero la integral que se evalúa ya debe haber quedado justificada por la situación.

Para construir un modelo nuevo debemos identificar la cantidad local y la variable de acumulación, decidir qué información conserva el signo y verificar las unidades. También habrá que comprobar que los datos determinan esa contribución; una integral correctamente evaluada no compensa una densidad que el problema no permite establecer.

Ejercicios y soluciones

El banco reservado `0756--0795` se organiza por niveles. En este primer bloque el objetivo es **modelar antes de integrar**: identificar qué cantidad local se suma, qué variable produce los pequeños incrementos y qué papel cumple el signo. La evaluación simbólica queda en segundo plano.

Nivel A — Modelar antes de integrar. Ejercicio A1. Reconocer la contribución local. En cada situación, identifica la contribución correspondiente a un pequeño incremento y escribe la integral que representa la cantidad total. **No evalúes las integrales.**

1. Una región está comprendida entre $y = f(x)$ y $y = g(x)$ en $[a, b]$, con $f(x) \geq g(x)$.
2. Un sólido tiene área transversal $A(x)$ perpendicular al eje x.
3. Una varilla ocupa $[a, b]$ y tiene densidad lineal $\rho(x)$.
4. Una partícula se desplaza sobre el eje x bajo una fuerza cuya componente en la dirección positiva es $F(x)$.

En cada caso, indica también las unidades de la contribución local.

Ejercicio A2. Dos preguntas, dos integrales. Una cantidad Q tiene tasa conocida

$$r(t) = 3 - 2t, \qquad 0 \leq t \leq 2.$$

1. Escribe una integral para el **cambio neto** $Q(2) - Q(0)$.
2. Escribe una integral para la **cantidad total de cambio acumulado**, entendida como suma de las magnitudes de los cambios locales.
3. Explica por qué las dos integrales deben ser distintas.
4. Evalúa ambas únicamente después de haber fijado el modelo.

Ejercicio A3. De una sección geométrica a una suma finita. La base de un sólido es la región comprendida entre

$$y = x \qquad \text{y} \qquad y = x^2,$$

para $0 \leq x \leq 1$. Las secciones perpendiculares al eje x son semicírculos cuyo diámetro es el segmento vertical contenido en la base.

Sea

$$P = \{0 = x_0 < x_1 < \cdots < x_n = 1\}$$

una partición y sea $\xi_k \in [x_{k-1}, x_k]$.

1. Expresa el diámetro de la sección situada en ξ_k.
2. Obtén el área de esa sección.
3. Escribe una suma finita que aproxime el volumen.
4. Identifica la integral a la que converge esa suma cuando la malla de la partición tiende a cero.

No evalúes la integral.

Ejercicio A4. La fórmula calcula otra cantidad. Una varilla ocupa $[0, 3]$ metros y tiene densidad lineal

$$\rho(x) = 2 + x \qquad (\text{kg/m}).$$

Un estudiante propone calcular su centro de masa mediante

$$\bar{x} \stackrel{?}{=} \int_0^3 x\rho(x)\,dx.$$

Audita la propuesta.

1. Determina qué cantidad representa realmente la integral escrita.
2. Usa las unidades para explicar por qué el resultado no puede ser una posición.
3. Repara el modelo.
4. Calcula la masa total y el centro de masa correctos.

Ejercicio A5. El cruce cambia el modelo. Considera las gráficas

$$f(x) = x \qquad \text{y} \qquad g(x) = x^3$$

en el intervalo $[-1, 1]$.

1. Determina los puntos donde las gráficas se cortan.
2. Explica por qué
$$\int_{-1}^{1} (f(x) - g(x))\, dx$$
no representa el área geométrica total encerrada entre ambas gráficas.
3. Construye una integral correcta para esa área.
4. Evalúala.

Tu solución debe identificar primero qué función está arriba en cada subintervalo.

Ejercicio A6. Las unidades correctas no validan por sí solas el modelo. Una partícula se mueve sobre una recta con velocidad

$$v(t) = t - 1, \qquad 0 \le t \le 2,$$

medida en metros por segundo. Un estudiante propone calcular la distancia recorrida mediante

$$D \stackrel{?}{=} \int_0^2 v(t)\, dt.$$

1. Verifica que la expresión propuesta tiene unidades de longitud.
2. Explica por qué ese control dimensional no basta para concluir que la fórmula calcula distancia.
3. Identifica la cantidad que representa realmente la integral propuesta.
4. Repara el modelo y calcula la distancia recorrida.

Nivel B — Áreas y volúmenes por secciones. Ejercicio B1. De la base a la sección transversal. La base de un sólido es la región comprendida entre

$$y = 4 - x^2 \qquad \text{y} \qquad y = 0,$$

para $-2 \le x \le 2$. Las secciones perpendiculares al eje x son triángulos equiláteros.

1. Determina la longitud del lado de la sección situada en la posición x.
2. Obtén el área transversal $A(x)$.
3. Construye la integral que representa el volumen.
4. Evalúa el volumen.

La solución debe justificar cada paso geométrico antes de integrar.

Ejercicio B2. La misma región con dos variables. Considera la región encerrada por

$$y = x^2 \qquad \text{y} \qquad y = 2x.$$

1. Determina los puntos de intersección.
2. Modela el área usando tiras verticales y una integral respecto de x.
3. Modela la misma área usando tiras horizontales y una integral respecto de y.
4. Evalúa ambas integrales y verifica que producen el mismo resultado.
5. Explica qué información geométrica cambia al pasar de dx a dy.

No se permite escribir la segunda integral sin identificar primero las fronteras izquierda y derecha en función de y.

Ejercicio B3. Secciones rectangulares que cambian de forma. La base de un sólido es la región comprendida entre

$$y = x \qquad \text{y} \qquad y = x^2,$$

para $0 \le x \le 1$. Las secciones perpendiculares al eje x son rectángulos. En la posición x, uno de sus lados coincide con el segmento vertical contenido en la base y el otro mide $(1 + x)$ veces ese segmento.

1. Determina las longitudes de los dos lados de la sección situada en x.
2. Construye el área transversal $A(x)$.
3. Escribe la integral del volumen.
4. Evalúala.
5. Explica por qué las secciones no forman una familia de rectángulos semejantes.

Ejercicio B4. Cuando la sección cambia de descripción. La base de un sólido es el triángulo limitado por

$$y = 0, \qquad y = x, \qquad y = 2 - x.$$

Las secciones perpendiculares al eje x son cuadrados.

1. Determina el intervalo de valores de x que ocupa la base.
2. Explica por qué la longitud del lado del cuadrado no puede describirse mediante una sola de las expresiones x o $2 - x$ en todo el intervalo.
3. Construye el área transversal por tramos.
4. Escribe y evalúa la integral del volumen.

La solución debe identificar el punto en que cambia la frontera superior antes de integrar.

Ejercicio B5. La sección no entrega directamente su área. La base de un sólido es la región comprendida entre

$$y = 2\sqrt{x} \qquad \text{y} \qquad y = x,$$

para $0 \le x \le 4$. Las secciones perpendiculares al eje x son triángulos rectángulos isósceles cuya **hipotenusa** coincide con el segmento vertical contenido en la base.

1. Determina la longitud de la hipotenusa en la posición x.
2. Expresa el área de un triángulo rectángulo isósceles en función de su hipotenusa.
3. Construye el área transversal $A(x)$.
4. Escribe y evalúa la integral del volumen.

Ejercicio B6. Diámetro no es radio. La base de un sólido es la región comprendida entre

$$y = x \qquad \text{y} \qquad y = x^2,$$

para $0 \le x \le 1$. Las secciones perpendiculares al eje x son círculos cuyo **diámetro** es el segmento vertical contenido en la base.

Un estudiante propone

$$V \stackrel{?}{=} \pi \int_0^1 (x - x^2)^2 \, dx.$$

1. Explica qué interpretación geométrica equivocada contiene la propuesta.
2. Corrige el área transversal.
3. Escribe la integral correcta.
4. Evalúa el volumen.
5. Explica por qué el análisis dimensional, por sí solo, no detecta este error.

Nivel C — Revolución y representaciones equivalentes. Ejercicio C1. Un mismo sólido, dos representaciones. Considera la región comprendida entre

$$y = \sqrt{x} \qquad \text{y} \qquad y = \frac{x}{2},$$

y hazla girar alrededor del eje x.

1. Determina los puntos de intersección.
2. Modela el volumen mediante arandelas usando dx.
3. Modela el mismo volumen mediante cascarones usando dy.
4. Evalúa ambas integrales y verifica que producen el mismo resultado.
5. Explica qué cantidad geométrica desempeña el papel de radio y cuál el de altura en cada representación.

Ejercicio C2. El radio es una distancia al eje. Considera la región

$$0 \le x \le 1, \qquad 0 \le y \le x,$$

y hazla girar alrededor de la recta horizontal $y = 2$.

1. Usando tiras verticales, determina el radio exterior y el radio interior de cada arandela.
2. Construye y evalúa la integral del volumen respecto de x.
3. Usando tiras horizontales, determina el radio y la altura de cada cascarón.
4. Construye y evalúa la integral correspondiente respecto de y.
5. Explica por qué usar y directamente como radio en el método de cascarones sería incorrecto.

Ejercicio C3. Elegir la representación más eficiente. Considera la región comprendida entre

$$y = x \qquad \text{y} \qquad y = x^3,$$

para $0 \le x \le 1$, y hazla girar alrededor del eje y.

1. Modela el volumen mediante cascarones usando dx.
2. Modela el mismo volumen mediante arandelas usando dy.
3. Evalúa ambas integrales.
4. Compara las dos representaciones y explica cuál elegirías si el objetivo fuera minimizar el trabajo algebraico.

La comparación debe hacerse después de identificar correctamente radios, alturas y límites.

Ejercicio C4. Una arandela no es un disco con radio igual al espesor radial. La región comprendida entre

$$y = x \qquad \text{y} \qquad y = x^2,$$

para $0 \le x \le 1$, gira alrededor del eje x.

Un estudiante propone

$$V \stackrel{?}{=} \pi \int_0^1 (x - x^2)^2 \, dx.$$

1. Explica qué objeto geométrico está tratando implícitamente como disco.
2. Identifica el radio exterior y el radio interior correctos.
3. Construye la integral correcta.
4. Evalúa tanto la propuesta incorrecta como la integral correcta y compara los resultados.
5. Explica por qué las unidades no bastan para detectar el error.

Ejercicio C5. Cuando una arandela se convierte en disco. Considera la región

$$0 \le x \le 4, \qquad 0 \le y \le \sqrt{x},$$

y hazla girar alrededor de la recta $y = 1$.

Usa tiras verticales.

1. Determina para qué valores de x la tira está completamente por debajo del eje de rotación.
2. Determina para qué valores de x la tira atraviesa el eje.
3. Construye el área transversal por tramos.
4. Escribe y evalúa la integral del volumen.
5. Explica por qué una sola fórmula de arandela con radio interior positivo no describe todo el sólido.

Ejercicio C6. El cascarón respecto del eje equivocado. La región comprendida entre

$$y = x - x^2 \qquad \text{y} \qquad y = 0,$$

para $0 \le x \le 1$, gira alrededor de la recta vertical $x = 2$.

Un estudiante propone

$$V \stackrel{?}{=} 2\pi \int_0^1 x(x - x^2)\, dx.$$

1. Identifica qué distancia está usando el estudiante como radio.
2. Determina el radio correcto del cascarón.
3. Construye y evalúa la integral correcta.
4. Evalúa también la propuesta incorrecta.
5. Explica por qué ambas expresiones tienen unidades de volumen aunque solo una modele el sólido pedido.

Nivel D — Longitud de gráfica. Ejercicio D1. Del polígono a la longitud exacta. Considera

$$f(x) = \frac{2}{3}x^{3/2}, \qquad 0 \le x \le 4,$$

y la partición $P = \{0, 1, 4\}$.

1. Calcula la longitud poligonal $L_P(f)$.
2. Usa el teorema del valor medio en cada subintervalo para explicar por qué cada cuerda puede escribirse en la forma $\sqrt{1 + f'(\xi_k)^2}\,\Delta x_k$ para algún ξ_k del subintervalo correspondiente.
3. Verifica que $f \in C^1([0, 4])$.
4. Usa Teorema 20.2 para obtener la longitud exacta de la gráfica.
5. Compara conceptualmente $L_P(f)$ con la longitud exacta y explica por qué una sola aproximación poligonal no constituye todavía la definición de longitud.

Ejercicio D2. Resultado correcto, justificación incompleta. Sea

$$f(x) = |x|, \qquad -1 \le x \le 1.$$

Un estudiante escribe

$$L \stackrel{?}{=} \int_{-1}^{1} \sqrt{1 + f'(x)^2}\, dx = 2\sqrt{2}$$

y afirma que la fórmula de Teorema 20.2 se aplica directamente.

1. Audita la afirmación: ¿se cumplen las hipótesis de Teorema 20.2 en todo $[-1, 1]$?
2. Explica por qué el problema está concentrado en un solo punto.
3. Repara la demostración dividiendo la gráfica en dos tramos.
4. Calcula rigurosamente la longitud total.
5. Explica por qué obtener el valor numérico correcto no vuelve válida una aplicación incorrecta de un teorema.

Ejercicio D3. Refinar aumenta la longitud poligonal. Considera

$$f(x) = x^2, \qquad 0 \le x \le 1,$$

y las particiones

$$P = \{0, 1\}, \qquad Q = \left\{0, \frac{1}{2}, 1\right\}.$$

1. Calcula $L_P(f)$.
2. Calcula $L_Q(f)$.
3. Verifica numéricamente que $L_P(f) < L_Q(f)$.
4. Explica conceptualmente por qué esta desigualdad es un caso particular de Lema 20.1.
5. ¿Permite este cálculo concluir que $L_Q(f)$ es la longitud exacta de la gráfica? Justifica.

Ejercicio D4. Verificar antes de integrar. Sea

$$f(x) = \frac{1}{3}(2x + 1)^{3/2}, \qquad 0 \le x \le 3.$$

1. Calcula $f'(x)$.
2. Verifica explícitamente que $f \in C^1([0, 3])$.
3. Construye la integral de longitud usando Teorema 20.2.
4. Simplifica el integrando antes de evaluar.
5. Calcula la longitud exacta de la gráfica.

Ejercicio D5. Trasladar verticalmente no cambia la longitud. Sea $f \in C^1([a, b])$ y, para una constante $c \in \mathbb{R}$, define $g(x) = f(x) + c$.

1. Demuestra directamente desde Definición 20.1 que para toda partición P de $[a, b]$, $L_P(g) = L_P(f)$.
2. Deduce desde Definición 20.2 que $L(g) = L(f)$.
3. Verifica la misma conclusión usando Teorema 20.2.
4. Explica geométricamente por qué una traslación vertical conserva la longitud de la gráfica.

Ejercicio D6. Una fórmula plausible con la geometría equivocada. Un estudiante propone que, para una función $f \in C^1([a, b])$, la longitud de su gráfica debería ser

$$L_{\text{inc}} = \int_a^b (1 + |f'(x)|)\, dx.$$

1. Usa la recta
$$f(x) = x, \qquad 0 \le x \le 1,$$
para comparar esta propuesta con la longitud geométrica exacta del segmento.
2. Calcula la longitud mediante Teorema 20.2.
3. Explica, a partir de un pequeño incremento Δx, por qué aparece $\sqrt{1 + f'(x)^2}$ y no $1 + |f'(x)|$.
4. Explica por qué el control dimensional tampoco detecta aquí el error.

Nivel E — Densidad, masa y centro de masa. Ejercicio E1. Recuperar la densidad antes de localizar el centro. Una varilla ocupa $0 \le x \le 2$ m y tiene densidad lineal $\rho(x) = kx$, donde $k > 0$ es una constante. La masa total de la varilla es 12 kg.

1. Determina las unidades de k.
2. Usa la masa total para hallar k.

3. Calcula el primer momento respecto del origen.
4. Determina el centro de masa.
5. Verifica que el resultado pertenece al intervalo ocupado por la varilla y explica por qué queda a la derecha del punto medio.

Ejercicio E2. Simetría antes del cálculo. Una varilla ocupa el intervalo

$$[-a, a], \qquad a > 0,$$

y tiene densidad

$$\rho(x) = \rho_0 \left(1 + \frac{x^2}{a^2}\right), \qquad \rho_0 > 0.$$

1. Explica por qué la densidad es simétrica respecto del origen.
2. Determina, sin calcular todavía la masa total, cuál debe ser el centro de masa.
3. Justifica esa conclusión usando la fórmula del primer momento.
4. Calcula la masa total como control.
5. Explica qué propiedad de la densidad sería necesario romper para desplazar el centro de masa fuera del origen.

Ejercicio E3. Una cota más fina desde el soporte efectivo. Una varilla ocupa el intervalo $0 \le x \le 5$, y tiene densidad lineal

$$\rho \in \mathcal{R}[0, 5], \qquad \rho(x) \ge 0.$$

Supongamos además que $\rho(x) = 0$ para $x \in [0, 1) \cup (4, 5]$, y que la masa total es positiva.

1. Explica por qué la cota general $0 \le \bar{x} \le 5$ puede mejorarse.
2. Demuestra que
$$\boxed{1 \le \bar{x} \le 4.}$$
3. Haz la demostración sin calcular explícitamente ninguna primitiva.
4. Explica qué información geométrica de la distribución de masa permite obtener la cota más fina.

Ejercicio E4. Igual masa, distribución reflejada. Dos varillas ocupan el intervalo $0 \le x \le 2$.

Sus densidades son

$$\rho_L(x) = k(2 - x), \qquad \rho_R(x) = kx,$$

donde $k > 0$.

1. Verifica que ambas varillas tienen la misma masa total.
2. Calcula el centro de masa de cada una.
3. Comprueba que los centros son simétricos respecto del punto medio $x = 1$.
4. Explica por qué la igualdad de masa total no determina por sí sola la posición del centro de masa.
5. Interpreta geométricamente el efecto de reflejar la densidad respecto de $x = 1$.

Ejercicio E5. Cambiar el origen cambia el momento, pero no el punto físico. Sea una varilla sobre $[a, b]$ con densidad lineal no negativa $\rho \in \mathcal{R}[a, b]$, y masa total

$$M = \int_a^b \rho(x)\,dx > 0.$$

Su primer momento respecto del origen es

$$\mathcal{M}_0 = \int_a^b x\rho(x)\,dx.$$

Ahora elegimos como nuevo origen el punto $x = c$ y medimos posiciones mediante $u = x - c$.

1. Define el primer momento respecto del nuevo origen.
2. Demuestra que $\mathcal{M}_c = \mathcal{M}_0 - cM$.
3. Demuestra que la coordenada del centro de masa respecto del nuevo origen es $\bar{u} = \bar{x} - c$.
4. Explica por qué el centro de masa físico no cambia aunque cambien sus coordenadas.

Ejercicio E6. Una función integrable puede ser una mala densidad de masa. En el intervalo $[0, 2]$ se propone $\rho(x) = 1 - x$ como densidad lineal de masa.

1. Determina en qué parte del intervalo ρ es negativa.
2. Calcula
$$\int_0^2 \rho(x)\,dx.$$
3. Explica por qué esta función no satisface el modelo físico de densidad de masa utilizado en §19.7–§19.8.
4. Explica por qué la fórmula
$$\bar{x} = \frac{\int_0^2 x\rho(x)\,dx}{\int_0^2 \rho(x)\,dx}$$
ni siquiera está definida en este caso.
5. Indica qué interpretación matemática sí puede conservar una función con signo cuando el contexto admite contribuciones orientadas.

Nivel F — Trabajo y cambio neto. Ejercicio F1. Misma trayectoria, trabajos opuestos. Un resorte ideal tiene constante elástica $k = 50$ N/m.

Se estira lentamente desde una elongación de $x = 0.10$ m hasta $x = 0.40$ m.

1. Escribe la fuerza ejercida por el resorte.
2. Escribe la fuerza externa necesaria para realizar el estiramiento cuasiestáticamente.
3. Calcula el trabajo realizado por cada fuerza.
4. Verifica que ambos trabajos tienen igual magnitud y signos opuestos.
5. Explica qué información física codifica el signo.

Ejercicio F2. Balance neto y flujo total no son la misma cantidad. Un depósito recibe líquido con tasa $r_{\text{in}}(t) = 5 - t$ y lo pierde con tasa $r_{\text{out}}(t) = 1 + t$, medidas en litros por minuto, para $0 \le t \le 4$.

1. Determina la tasa neta de cambio del volumen.
2. Calcula la entrada acumulada y la salida acumulada durante los cuatro minutos.
3. Calcula el cambio neto del volumen.

4. Calcula el flujo bruto total, definido como entrada acumulada más salida acumulada.
5. Calcula
$$\int_0^4 |r_{\text{in}}(t) - r_{\text{out}}(t)|\, dt$$
e interpreta qué mide.
6. Explica por qué estas tres cantidades —cambio neto, flujo bruto y variación acumulada del volumen— no deben confundirse.

Ejercicio F3. Trabajo neto nulo con transferencia no nula. Una partícula se desplaza sobre el eje x desde $x = 0$ hasta $x = 4$ m, bajo una fuerza cuya componente en la dirección positiva es

$$F(x) = 6 - 3x \qquad \text{(N)}.$$

1. Determina dónde cambia de signo la fuerza.
2. Calcula el trabajo realizado en el tramo donde la fuerza favorece el desplazamiento.
3. Calcula el trabajo realizado en el tramo donde la fuerza se opone al desplazamiento.
4. Calcula el trabajo neto sobre todo el recorrido.
5. Calcula
$$\int_0^4 |F(x)|\, dx$$
e interpreta qué información conserva esta cantidad.
6. Explica cómo puede ser nulo el trabajo neto aunque haya contribuciones locales no nulas.

Ejercicio F4. Desplazamiento, distancia y posición final. Una partícula se mueve sobre una recta con velocidad

$$v(t) = t^2 - 4t + 3 \qquad \text{(m/s)},$$

para $0 \le t \le 4$, y su posición inicial es $s(0) = 5$ m.

1. Determina los instantes en que la partícula cambia de sentido.
2. Calcula el desplazamiento total.
3. Determina la posición final $s(4)$.
4. Calcula la distancia total recorrida.
5. Explica por qué ninguno de esos tres resultados puede sustituirse automáticamente por otro.

Ejercicio F5. Invertir el recorrido invierte el trabajo. Una partícula se mueve sobre el eje x bajo la fuerza

$$F(x) = x^2 + 1 \qquad \text{(N)}.$$

1. Calcula el trabajo realizado por la fuerza cuando la partícula se desplaza desde $x = 0$ hasta $x = 2$ m.
2. Calcula el trabajo realizado por la misma fuerza cuando la partícula recorre el mismo intervalo en sentido contrario, desde $x = 2$ hasta $x = 0$.
3. Explica por qué los dos trabajos tienen igual magnitud y signos opuestos.
4. Explica por qué no corresponde insertar un valor absoluto en la integral para “hacer positivo” el segundo resultado.

Ejercicio F6. Una tasa determina el cambio, no el nivel absoluto. Una cantidad $Q(t)$, medida en kilogramos, tiene tasa neta conocida

$$Q'(t) = 2t - 3 \qquad (\text{kg/min}),$$

para $0 \le t \le 4$.

1. Calcula el cambio neto $Q(4) - Q(0)$.
2. Expresa $Q(4)$ en función de $Q(0)$.
3. Explica por qué la tasa conocida no basta para determinar numéricamente $Q(4)$.
4. Calcula $Q(4)$ si $Q(0) = 10$ kg.
5. Calcula $Q(4)$ si $Q(0) = 25$ kg.
6. Identifica qué información es común a ambos casos y cuál depende del dato inicial.

Nivel G — Transferencia y auditoría. Ejercicio G1. La misma función, cuatro modelos distintos. Considera

$$q(x) = 2 - x, \qquad 0 \le x \le 3.$$

La misma fórmula aparece en cuatro problemas diferentes.

1. $q(x)$ es la diferencia vertical entre una gráfica y el eje x, y se pide el **área geométrica total** comprendida entre ambas.
2. $q(x)$ se propone como **densidad lineal de masa** de una varilla.
3. $q(x)$ es la componente de una **fuerza** en la dirección positiva del eje, y una partícula se desplaza desde $x = 0$ hasta $x = 3$.
4. $q(x)$ es la **velocidad** de una partícula y se piden tanto su desplazamiento como la distancia recorrida.

Para cada contexto:

- decide si debe aparecer q, $|q|$ o si el modelo es inadmisible tal como está planteado;
- escribe la integral correspondiente;
- evalúa las cantidades que estén bien definidas;
- explica qué papel cumple el signo.

Ejercicio G2. Reparar una solución que inventa información. Una varilla ocupa $0 \le x \le 2$ m, tiene masa total 6 kg, y se sabe únicamente que su densidad lineal es una función afín creciente de x.

Un estudiante escribe:

> «Como la masa es 6, tomo $\rho(x) = 3x$ kg/m. Entonces el centro de masa es $\bar{x} = 4/3$ m.»

Audita la solución.

1. Identifica el problema dimensional en la expresión escrita para ρ.
2. Decide si los datos del enunciado determinan una única densidad afín creciente.
3. Escribe la familia general de densidades afines crecientes y no negativas compatibles con la masa total.
4. Calcula el centro de masa en función del parámetro libre.
5. Determina qué valores puede tomar $\bar{x}$ dentro de esa familia.

6. Explica en qué sentido el valor $4/3$ obtenido por el estudiante puede corresponder a un modelo válido sin quedar determinado por los datos.

Ejercicio G3. Una misma función, tres factores locales distintos. Considera

$$q(x) = x(2 - x), \qquad 0 \le x \le 2.$$

1. Interpreta $q(x)$ como altura sobre el eje x y calcula el área bajo la gráfica.
2. Haz girar esa región alrededor del eje y. Modela el volumen mediante cascarones y calcúlalo.
3. Interpreta ahora la misma fórmula $q(x)$ como densidad lineal de una varilla en $[0, 2]$.
4. Calcula la masa total, el primer momento y el centro de masa.
5. Explica por qué en las tres preguntas aparecen, respectivamente,
$$q(x), \qquad xq(x), \qquad \frac{\int xq(x)\,dx}{\int q(x)\,dx}.$$
6. Identifica qué parte del modelo cambia aunque la fórmula algebraica q sea la misma.

Ejercicio G4. Masa y centro de masa tampoco determinan la densidad. Una varilla ocupa $0 \le x \le 2$, tiene masa total $M = 6$ y centro de masa $\bar{x} = 1$.

Un estudiante concluye que necesariamente la densidad es constante: $\rho(x) = 3$.

1. Verifica que la densidad constante propuesta satisface los datos.
2. Construye una densidad no constante, no negativa y Riemann-integrable que tenga la misma masa total y el mismo centro de masa.
3. Verifica explícitamente ambos datos para tu densidad alternativa.
4. Explica por qué conocer
$$\int_0^2 \rho(x)\,dx$$
y
$$\int_0^2 x\rho(x)\,dx$$
no determina, en general, la función ρ.
5. Indica qué error lógico comete el estudiante al pasar de «una densidad compatible» a «la densidad».

Soluciones desarrolladas.

Soluciones del nivel A. **Solución A1.**

1. Una tira vertical de ancho Δx tiene altura aproximada $f(\xi) - g(\xi)$. Su área local es $\Delta A \approx (f(\xi) - g(\xi))\Delta x$, y el área total queda modelada por
$$\boxed{A = \int_a^b (f(x) - g(x))\,dx.}$$
Si x, f y g se miden en unidades de longitud, la contribución local tiene unidades de longitud al cuadrado.

2. Una sección de área $A(\xi)$ y espesor Δx aporta aproximadamente $\Delta V \approx A(\xi)\Delta x$. Por tanto,
$$\boxed{V = \int_a^b A(x)\,dx.}$$
Las unidades son área por longitud, es decir, volumen.

3. Un tramo de longitud Δx aporta masa aproximada $\Delta M \approx \rho(\xi)\Delta x$. Así,
$$\boxed{M = \int_a^b \rho(x)\,dx.}$$
Las unidades son masa por longitud multiplicada por longitud, es decir, masa.

4. Sobre un desplazamiento pequeño Δx, el trabajo local es $\Delta W \approx F(\xi)\Delta x$. Por tanto,
$$\boxed{W = \int_a^b F(x)\,dx.}$$
En el SI, las unidades son $\mathrm{N\,m} = \mathrm{J}$.

Los cuatro modelos tienen la misma forma estructural: una cantidad por unidad de la variable se multiplica por un pequeño incremento de esa variable y luego se acumula.

Solución A2. La tasa cambia de signo cuando $3 - 2t = 0$, es decir, en $t = \frac{3}{2}$.
El cambio neto conserva el signo de la tasa:

$$\boxed{Q(2) - Q(0) = \int_0^2 (3 - 2t)\,dt.}$$

En cambio, la cantidad total de cambio acumulado debe sumar magnitudes locales:

$$\boxed{T = \int_0^2 |3 - 2t|\,dt.}$$

Ahora evaluamos. Para el cambio neto,

$$\begin{aligned} Q(2) - Q(0) &= \left[3t - t^2\right]_0^2 \\ &= 6 - 4 \\ &= 2. \end{aligned}$$

Para la acumulación total separamos en el punto donde cambia el signo:

$$\begin{aligned} T &= \int_0^{3/2} (3 - 2t)\,dt + \int_{3/2}^2 (2t - 3)\,dt \\ &= \frac{9}{4} + \frac{1}{4} \\ &= \frac{5}{2}. \end{aligned}$$

Por tanto,

$$\boxed{Q(2) - Q(0) = 2, \qquad T = \frac{5}{2}.}$$

La diferencia aparece porque la integral orientada permite cancelación, mientras que la integral del valor absoluto acumula todas las magnitudes localmente positivas.

Solución A3. En la posición ξ_k, la distancia vertical entre las dos fronteras de la base es $d_k = \xi_k - \xi_k^2$.

Ese segmento es el diámetro del semicírculo. Por tanto, su radio es $r_k = \frac{\xi_k - \xi_k^2}{2}$. El área transversal correspondiente es

$$\begin{aligned} A(\xi_k) &= \frac{1}{2}\pi r_k^2 \\ &= \frac{1}{2}\pi \left(\frac{\xi_k - \xi_k^2}{2} \right)^2 \\ &= \frac{\pi}{8}(\xi_k - \xi_k^2)^2. \end{aligned}$$

Si $\Delta x_k = x_k - x_{k-1}$, el pequeño bloque asociado al subintervalo tiene volumen aproximado

$$\Delta V_k \approx A(\xi_k)\Delta x_k = \frac{\pi}{8}(\xi_k - \xi_k^2)^2 \Delta x_k.$$

La suma finita es entonces

$$\boxed{\sum_{k=1}^{n} \frac{\pi}{8}(\xi_k - \xi_k^2)^2 \Delta x_k.}$$

Esta es una suma etiquetada para la función $A(x) = \frac{\pi}{8}(x - x^2)^2$. Como A es continua en $[0, 1]$, al hacer tender la malla a cero obtenemos

$$\boxed{V = \frac{\pi}{8} \int_0^1 (x - x^2)^2 \, dx.}$$

La integral aparece después de identificar la sección local: diámetro $\to$ radio $\to$ área $\to$ volumen elemental.

Solución A4. La integral propuesta es

$$\int_0^3 x\rho(x) \, dx.$$

Por Definición 20.3, esta cantidad es el **primer momento respecto del origen**:

$$\mathcal{M}_0 = \int_0^3 x\rho(x) \, dx.$$

Sus unidades son

$$[\mathrm{m}] \left[\frac{\mathrm{kg}}{\mathrm{m}} \right] [\mathrm{m}] = [\mathrm{kg\,m}].$$

Por tanto, el resultado no puede ser una posición, que debe tener unidades de longitud.

La masa total es

$$\begin{aligned} M &= \int_0^3 (2+x)\,dx \\ &= \left[2x + \frac{x^2}{2}\right]_0^3 \\ &= 6 + \frac{9}{2} \\ &= \frac{21}{2}\ \text{kg}. \end{aligned}$$

El primer momento es

$$\begin{aligned} \mathcal{M}_0 &= \int_0^3 x(2+x)\,dx \\ &= \int_0^3 (2x + x^2)\,dx \\ &= \left[x^2 + \frac{x^3}{3}\right]_0^3 \\ &= 9 + 9 \\ &= 18\ \text{kg m}. \end{aligned}$$

Según Corolario 20.4, el centro de masa se obtiene normalizando el primer momento por la masa total:

$$\begin{aligned} \bar{x} &= \frac{\mathcal{M}_0}{M} \\ &= \frac{18}{21/2} \\ &= \frac{12}{7}\ \text{m}. \end{aligned}$$

Así,

$$\boxed{M = \frac{21}{2}\ \text{kg}, \qquad \bar{x} = \frac{12}{7}\ \text{m}.}$$

El error del modelo consistía en identificar directamente un momento con una posición; la división por la masa total es la normalización que recupera las unidades y el significado correctos.

Solución A5. Los puntos de corte satisfacen $x = x^3$, de modo que $x(x^2-1) = 0$.
Por tanto,

$$x = -1, \qquad x = 0, \qquad x = 1.$$

En $(-1, 0)$ se tiene $x^3 > x$, mientras que en $(0, 1)$ ocurre $x > x^3$. La diferencia $f - g = x - x^3$ cambia entonces de signo en $x = 0$. La integral

$$\int_{-1}^{1} (x - x^3)\,dx$$

es una acumulación orientada y permite cancelación entre las dos mitades del intervalo. De hecho, el integrando es impar, por lo que esa integral vale 0. Para el área geométrica, cada contribución local debe ser no negativa. Podemos escribir

$$A = \int_{-1}^{0} (x^3 - x)\,dx + \int_{0}^{1} (x - x^3)\,dx.$$

Por simetría,

$$A = 2\int_{0}^{1} (x - x^3)\,dx.$$

Evaluando,

$$\begin{aligned} A &= 2\left[\frac{x^2}{2} - \frac{x^4}{4}\right]_0^1 \\ &= 2\left(\frac{1}{2} - \frac{1}{4}\right) \\ &= \frac{1}{2}. \end{aligned}$$

Así,

$$\boxed{A = \frac{1}{2}.}$$

Equivalentemente,

$$\boxed{A = \int_{-1}^{1} |x - x^3|\,dx.}$$

El punto decisivo es localizar primero los cruces y determinar qué gráfica ocupa la posición superior en cada tramo.

Solución A6. La velocidad tiene unidades de $\frac{\mathrm{m}}{\mathrm{s}}$, y dt tiene unidades de segundos. Por tanto, $v(t)\,dt$ tiene unidades de metros. La integral propuesta supera, pues, el control dimensional.

Sin embargo, la velocidad es una magnitud orientada. La integral

$$\int_{0}^{2} v(t)\,dt$$

representa el desplazamiento:

$$\begin{aligned} s(2) - s(0) &= \int_{0}^{2} (t - 1)\,dt \\ &= \left[\frac{t^2}{2} - t\right]_0^2 \\ &= 0. \end{aligned}$$

La velocidad cambia de signo en $t = 1$. La distancia recorrida acumula la rapidez:

$$D = \int_0^2 |t-1|\,dt.$$

Separando en el cambio de signo,

$$\begin{aligned} D &= \int_0^1 (1-t)\,dt + \int_1^2 (t-1)\,dt \\ &= \frac{1}{2} + \frac{1}{2} \\ &= 1. \end{aligned}$$

Por tanto,

$$\boxed{s(2) - s(0) = 0, \qquad D = 1 \text{ m}.}$$

El ejemplo muestra que el análisis dimensional es un control necesario, pero no determina por sí solo el significado de la integral. También debe comprobarse qué cantidad local se está acumulando y cómo interviene el signo.

Soluciones del nivel B. **Solución B1.** Para cada $x \in [-2, 2]$, la base contiene un segmento vertical que va desde $y = 0$ hasta $y = 4 - x^2$.

Por tanto, la longitud de ese segmento es $s(x) = 4 - x^2$. Ese segmento es el lado del triángulo equilátero correspondiente. El área de un triángulo equilátero de lado s es $\frac{\sqrt{3}}{4}s^2$, de modo que $A(x) = \frac{\sqrt{3}}{4}(4 - x^2)^2$. El volumen queda modelado por

$$\boxed{V = \frac{\sqrt{3}}{4} \int_{-2}^{2} (4 - x^2)^2\,dx.}$$

La función integranda es par, así que

$$V = \frac{\sqrt{3}}{2} \int_0^2 (16 - 8x^2 + x^4)\,dx.$$

Evaluando,

$$\begin{aligned} V &= \frac{\sqrt{3}}{2} \left[16x - \frac{8x^3}{3} + \frac{x^5}{5} \right]_0^2 \\ &= \frac{\sqrt{3}}{2} \left(32 - \frac{64}{3} + \frac{32}{5} \right) \\ &= \frac{\sqrt{3}}{2} \cdot \frac{256}{15} \\ &= \boxed{\frac{128\sqrt{3}}{15}.} \end{aligned}$$

La cadena de modelización es

$$\boxed{\text{segmento vertical de la base} \longrightarrow \text{lado del triángulo} \longrightarrow A(x) \longrightarrow V.}$$

Solución B2. Los puntos de intersección satisfacen $x^2 = 2x$, de modo que $x(x-2) = 0$. Por tanto,

$$x = 0 \qquad \text{y} \qquad x = 2.$$

Los puntos correspondientes son

$$(0,0) \qquad \text{y} \qquad (2,4).$$

Con tiras verticales, para $0 \le x \le 2$ la recta $y = 2x$ está por encima de la parábola $y = x^2$. La altura local es $2x - x^2$, y por tanto

$$\boxed{A = \int_0^2 (2x - x^2)\, dx.}$$

Evaluando,

$$\begin{aligned} A &= \left[x^2 - \frac{x^3}{3} \right]_0^2 \\ &= 4 - \frac{8}{3} \\ &= \frac{4}{3}. \end{aligned}$$

Para usar tiras horizontales debemos expresar las fronteras en función de y. De $y = 2x$ obtenemos $x = \frac{y}{2}$, y de $y = x^2$ tomamos, en esta región, $x = \sqrt{y}$. Cuando $0 \le y \le 4$, la frontera izquierda es $x = \frac{y}{2}$, y la derecha es $x = \sqrt{y}$. La longitud horizontal es entonces $\sqrt{y} - \frac{y}{2}$, de modo que

$$\boxed{A = \int_0^4 \left(\sqrt{y} - \frac{y}{2} \right) dy.}$$

Evaluando,

$$\begin{aligned} A &= \left[\frac{2}{3} y^{3/2} - \frac{y^2}{4} \right]_0^4 \\ &= \frac{16}{3} - 4 \\ &= \frac{4}{3}. \end{aligned}$$

Así,

$$\boxed{\int_0^2 (2x - x^2)\, dx = \int_0^4 \left(\sqrt{y} - \frac{y}{2} \right) dy = \frac{4}{3}.}$$

La región es la misma, pero el elemento local cambia. Con dx se necesita **altura = superior − inferior**; con dy se necesita **longitud = derecha − izquierda**.

Solución B3. Para $0 \le x \le 1$, la recta $y = x$ está por encima de la parábola $y = x^2$. El segmento vertical de la base tiene longitud $s(x) = x - x^2$.

Uno de los lados del rectángulo mide entonces $s(x) = x - x^2$, y el otro mide $(1 + x)s(x) = (1 + x)(x - x^2)$. Por tanto, el área transversal es

$$\begin{aligned} A(x) &= s(x)\,(1 + x)s(x) \\ &= (1 + x)(x - x^2)^2. \end{aligned}$$

El volumen queda modelado por

$$\boxed{V = \int_0^1 (1 + x)(x - x^2)^2\,dx.}$$

Expandiendo solo en la etapa de cálculo,

$$(1 + x)(x - x^2)^2 = x^2 - x^3 - x^4 + x^5.$$

Por tanto,

$$\begin{aligned} V &= \int_0^1 (x^2 - x^3 - x^4 + x^5)\,dx \\ &= \left[\frac{x^3}{3} - \frac{x^4}{4} - \frac{x^5}{5} + \frac{x^6}{6}\right]_0^1 \\ &= \frac{1}{3} - \frac{1}{4} - \frac{1}{5} + \frac{1}{6} \\ &= \boxed{\frac{1}{20}}. \end{aligned}$$

Las secciones no son semejantes entre sí porque la razón entre sus lados es $1 + x$, que cambia con la posición. El método por secciones solo necesita conocer el área transversal $A(x)$; no exige que todas las secciones tengan la misma forma relativa.

Solución B4. Las rectas

$$y = x \qquad \text{y} \qquad y = 2 - x$$

se cortan cuando $x = 2 - x$, es decir, en $x = 1$. La base se extiende desde $x = 0$ hasta $x = 2$. Para $0 \le x \le 1$, la frontera superior del triángulo es $y = x$, de modo que el lado del cuadrado vale $s(x) = x$. Para $1 \le x \le 2$, la frontera superior es $y = 2 - x$, y entonces $s(x) = 2 - x$.

Por tanto,

$$A(x) = \begin{cases} x^2, & 0 \le x \le 1, \\ (2 - x)^2, & 1 \le x \le 2. \end{cases}$$

El volumen debe escribirse por tramos:

$$\boxed{V = \int_0^1 x^2\,dx + \int_1^2 (2 - x)^2\,dx.}$$

Evaluando,

$$\int_0^1 x^2\,dx = \frac{1}{3}.$$

En la segunda integral, por simetría o mediante el cambio $u = 2 - x$,

$$\int_1^2 (2-x)^2\,dx = \frac{1}{3}.$$

Así,

$$\boxed{V = \frac{2}{3}.}$$

La partición del dominio no es una técnica de integración añadida al final: viene impuesta por la geometría de la base, porque la frontera superior cambia en $x = 1$.

Solución B5. En $0 \le x \le 4$ se tiene $2\sqrt{x} \ge x$, de modo que la longitud del segmento vertical de la base es $h(x) = 2\sqrt{x} - x$.

Ese segmento es la hipotenusa del triángulo rectángulo isósceles. Si la hipotenusa es h, cada cateto mide $\frac{h}{\sqrt{2}}$, porque, por Pitágoras, $h^2 = \ell^2 + \ell^2 = 2\ell^2$. Por tanto, el área de la sección es

$$A = \frac{1}{2}\left(\frac{h}{\sqrt{2}}\right)^2 = \frac{h^2}{4}.$$

Así, $A(x) = \frac{1}{4}(2\sqrt{x} - x)^2$.

El volumen queda modelado por

$$\boxed{V = \frac{1}{4}\int_0^4 (2\sqrt{x} - x)^2\,dx.}$$

Expandiendo, $\frac{1}{4}(2\sqrt{x} - x)^2 = x - x^{3/2} + \frac{x^2}{4}$.

Entonces

$$\begin{aligned} V &= \int_0^4 \left(x - x^{3/2} + \frac{x^2}{4}\right) dx \\ &= \left[\frac{x^2}{2} - \frac{2}{5}x^{5/2} + \frac{x^3}{12}\right]_0^4 \\ &= 8 - \frac{64}{5} + \frac{16}{3} \\ &= \boxed{\frac{8}{15}}. \end{aligned}$$

El paso esencial fue convertir primero la información «hipotenusa» en una fórmula de área transversal.

Solución B6. El segmento vertical de la base tiene longitud $d(x) = x - x^2$.

El enunciado establece que esta longitud es el **diámetro** del círculo. La propuesta del estudiante usa $\pi d(x)^2$, que correspondería al área de un círculo cuyo **radio** fuera $d(x)$.

El radio correcto es $r(x) = \frac{d(x)}{2} = \frac{x-x^2}{2}$.

Por tanto, $A(x) = \pi r(x)^2 = \frac{\pi}{4}(x - x^2)^2$.

La integral correcta es

$$\boxed{V = \frac{\pi}{4} \int_0^1 (x - x^2)^2 \, dx.}$$

Como $(x - x^2)^2 = x^2 - 2x^3 + x^4$, obtenemos

$$\begin{aligned} V &= \frac{\pi}{4} \left[\frac{x^3}{3} - \frac{x^4}{2} + \frac{x^5}{5} \right]_0^1 \\ &= \frac{\pi}{4} \left(\frac{1}{3} - \frac{1}{2} + \frac{1}{5} \right) \\ &= \frac{\pi}{4} \cdot \frac{1}{30} \\ &= \boxed{\frac{\pi}{120}.} \end{aligned}$$

La fórmula incorrecta también tiene unidades de volumen: una longitud al cuadrado multiplicada por dx. El análisis dimensional no distingue entre radio y diámetro; ese control exige interpretar correctamente la geometría de la sección.

Soluciones del nivel C. **Solución C1.** Los puntos de intersección satisfacen $\sqrt{x} = \frac{x}{2}$.

Como $x \geq 0$, al elevar al cuadrado obtenemos $x = \frac{x^2}{4}$, es decir, $x(x - 4) = 0$.

Por tanto,

$$x = 0 \qquad \text{y} \qquad x = 4.$$

En $0 \leq x \leq 4$ se cumple $\sqrt{x} \geq \frac{x}{2}$.

Al girar alrededor del eje x, una tira vertical produce una arandela con

$$R(x) = \sqrt{x}, \qquad r(x) = \frac{x}{2}.$$

Así,

$$\boxed{V = \pi \int_0^4 \left(x - \frac{x^2}{4} \right) \, dx.}$$

Evaluando,

$$\begin{aligned} V &= \pi \left[\frac{x^2}{2} - \frac{x^3}{12} \right]_0^4 \\ &= \pi \left(8 - \frac{16}{3} \right) \\ &= \boxed{\frac{8\pi}{3}.} \end{aligned}$$

Para usar cascarones debemos describir la región horizontalmente. De $y = \sqrt{x}$ obtenemos $x = y^2$, y de $y = \frac{x}{2}$ obtenemos $x = 2y$. Cuando $0 \leq y \leq 2$, la región va desde $x = y^2$ hasta $x = 2y$. Una tira horizontal genera un cascarón de radio y y altura $2y - y^2$.

Por tanto,

$$\boxed{V = 2\pi \int_0^2 y(2y - y^2)\, dy.}$$

Evaluando,

$$\begin{aligned} V &= 2\pi \int_0^2 (2y^2 - y^3)\, dy \\ &= 2\pi \left[\frac{2y^3}{3} - \frac{y^4}{4} \right]_0^2 \\ &= 2\pi \left(\frac{16}{3} - 4 \right) \\ &= \boxed{\frac{8\pi}{3}.} \end{aligned}$$

Las dos integrales describen el mismo sólido mediante elementos locales distintos:

$$\boxed{\text{arandela: área transversal} \times dx, \qquad \text{cascarón: circunferencia} \times \text{altura} \times dy.}$$

Solución C2. La región está situada debajo del eje de rotación $y = 2$.

Con una tira vertical en la posición x, la región va desde $y = 0$ hasta $y = x$. La distancia desde el eje $y = 2$ hasta $y = 0$ es $R(x) = 2$, mientras que la distancia hasta $y = x$ es $r(x) = 2 - x$. Por tanto, la arandela tiene área

$$\pi\big(R(x)^2 - r(x)^2\big) = \pi \left(4 - (2 - x)^2 \right).$$

El volumen es

$$\boxed{V = \pi \int_0^1 \left(4 - (2 - x)^2 \right)\, dx.}$$

Como $4 - (2 - x)^2 = 4x - x^2$, obtenemos

$$V = \pi \left[2x^2 - \frac{x^3}{3} \right]_0^1$$
$$= \boxed{\frac{5\pi}{3}.}$$

Ahora usamos tiras horizontales. Para $0 \le y \le 1$, la región se extiende desde $x = y$ hasta $x = 1$. La altura del cascarón es $1 - y$. Como el eje de rotación es $y = 2$, el radio es la distancia vertical $2 - y$.

Por tanto,

$$\boxed{V = 2\pi \int_0^1 (2 - y)(1 - y)\, dy.}$$

Evaluando,

$$V = 2\pi \int_0^1 (2 - 3y + y^2)\, dy$$
$$= 2\pi \left[2y - \frac{3y^2}{2} + \frac{y^3}{3} \right]_0^1$$
$$= 2\pi \left(2 - \frac{3}{2} + \frac{1}{3} \right)$$
$$= \boxed{\frac{5\pi}{3}.}$$

Usar y como radio mediría la distancia al eje x, no a la recta $y = 2$. En problemas de revolución, el radio se obtiene siempre como **distancia al eje de rotación**.

Solución C3. En $0 \le x \le 1$ se cumple $x \ge x^3$.

Con tiras verticales, el radio del cascarón es x, y su altura es $x - x^3$.

Por tanto,

$$\boxed{V = 2\pi \int_0^1 x(x - x^3)\, dx.}$$

Evaluando,

$$V = 2\pi \int_0^1 (x^2 - x^4)\, dx$$
$$= 2\pi \left(\frac{1}{3} - \frac{1}{5} \right)$$
$$= \boxed{\frac{4\pi}{15}.}$$

Para usar arandelas debemos describir la región horizontalmente. De $y = x$ obtenemos $x = y$, y de $y = x^3$ obtenemos $x = y^{1/3}$.

Para $0 \le y \le 1$ se tiene $y^{1/3} \ge y$.

Así, la arandela tiene radio exterior $R(y) = y^{1/3}$ y radio interior $r(y) = y$.

Por tanto,

$$\boxed{V = \pi \int_0^1 \left(y^{2/3} - y^2\right)\, dy.}$$

Evaluando,

$$\begin{aligned} V &= \pi \left(\frac{3}{5} - \frac{1}{3}\right) \\ &= \boxed{\frac{4\pi}{15}.} \end{aligned}$$

Las dos representaciones son correctas. En este caso, los cascarones producen directamente un integrando polinómico, $x^2 - x^4$, mientras que las arandelas introducen la potencia fraccionaria $y^{2/3}$. Si solo se busca eficiencia algebraica, los cascarones ofrecen aquí la representación más simple.

Solución C4. En una tira vertical, la región se extiende desde $y = x^2$ hasta $y = x$.

Al girar alrededor del eje x, esa tira genera una **arandela**. La propuesta del estudiante utiliza $\pi(x - x^2)^2$, como si la separación vertical entre las dos curvas fuera el radio de un disco macizo.

Los radios correctos son

$$R(x) = x, \qquad r(x) = x^2.$$

Por tanto, el área transversal es $\pi\big(R(x)^2 - r(x)^2\big) = \pi(x^2 - x^4)$.

La integral correcta es

$$\boxed{V = \pi \int_0^1 (x^2 - x^4)\, dx.}$$

Evaluando,

$$\begin{aligned} V &= \pi \left(\frac{1}{3} - \frac{1}{5}\right) \\ &= \boxed{\frac{2\pi}{15}.} \end{aligned}$$

La propuesta incorrecta da

$$\begin{aligned} V_{\text{inc}} &= \pi \int_0^1 (x - x^2)^2\, dx \\ &= \pi \int_0^1 (x^2 - 2x^3 + x^4)\, dx \\ &= \pi \left(\frac{1}{3} - \frac{1}{2} + \frac{1}{5}\right) \\ &= \boxed{\frac{\pi}{30}.} \end{aligned}$$

Los resultados difieren por un factor de cuatro: $\frac{2\pi}{15} = 4 \cdot \frac{\pi}{30}$. Ambas expresiones tienen unidades de volumen, porque en ambos casos se integra una cantidad con dimensión de área respecto

de una longitud. El error solo aparece al interpretar la geometría: una arandela requiere restar **áreas**, $\pi R^2 - \pi r^2$, y esa cantidad no coincide con $\pi(R-r)^2$.

Solución C5. En la posición x, la tira vertical ocupa el intervalo $0 \le y \le \sqrt{x}$.

El eje de rotación es $y = 1$. Si $0 \le x \le 1$, entonces $\sqrt{x} \le 1$, de modo que toda la tira queda por debajo del eje. Al girar se obtiene una arandela con radio exterior $R(x) = 1$ y radio interior $r(x) = 1 - \sqrt{x}$.

Por tanto, $A(x) = \pi\left[1 - (1-\sqrt{x})^2\right]$.

Si $1 \le x \le 4$, entonces $\sqrt{x} \ge 1$, y la tira atraviesa el eje. Además, $0 \le \sqrt{x} - 1 \le 1$. La parte inferior de la tira alcanza una distancia 1 del eje, mientras que la parte superior alcanza como máximo esa misma distancia. Al rotar, la sección llena un disco de radio $R(x) = 1$.

Así,

$$A(x) = \pi \qquad (1 \le x \le 4).$$

El volumen es entonces

$$\boxed{V = \pi \int_0^1 \left[1 - (1-\sqrt{x})^2\right] dx + \pi \int_1^4 1\, dx.}$$

Simplificando el primer integrando, $1 - (1-\sqrt{x})^2 = 2\sqrt{x} - x$.

Por tanto,

$$\begin{aligned} V &= \pi \int_0^1 (2\sqrt{x} - x)\, dx + 3\pi \\ &= \pi \left[\frac{4}{3}x^{3/2} - \frac{x^2}{2}\right]_0^1 + 3\pi \\ &= \pi\left(\frac{4}{3} - \frac{1}{2}\right) + 3\pi \\ &= \frac{5\pi}{6} + 3\pi \\ &= \boxed{\frac{23\pi}{6}.} \end{aligned}$$

El cambio de fórmula ocurre porque, a partir de $x = 1$, el eje de rotación atraviesa la tira. Desde ese punto ya no queda un hueco central: la sección es un disco.

Solución C6. La propuesta del estudiante utiliza x como radio. Esa cantidad es la distancia desde la tira vertical situada en x hasta el eje y.

El eje real de rotación es $x = 2$. Por tanto, el radio correcto es la distancia horizontal $R(x) = 2 - x$. La altura del cascarón viene dada por $h(x) = x - x^2$. Así, el volumen correcto es

$$\boxed{V = 2\pi \int_0^1 (2-x)(x-x^2)\, dx.}$$

Expandimos: $(2-x)(x-x^2) = 2x - 3x^2 + x^3$.

Entonces

$$\begin{aligned} V &= 2\pi \int_0^1 (2x - 3x^2 + x^3)\, dx \\ &= 2\pi \left[x^2 - x^3 + \frac{x^4}{4} \right]_0^1 \\ &= 2\pi \left(\frac{1}{4} \right) \\ &= \boxed{\frac{\pi}{2}.} \end{aligned}$$

La propuesta incorrecta produce

$$\begin{aligned} V_{\text{inc}} &= 2\pi \int_0^1 x(x - x^2)\, dx \\ &= 2\pi \int_0^1 (x^2 - x^3)\, dx \\ &= 2\pi \left(\frac{1}{3} - \frac{1}{4} \right) \\ &= \boxed{\frac{\pi}{6}.} \end{aligned}$$

Ambas expresiones tienen unidades de volumen porque en las dos aparece

$$(\text{longitud})(\text{longitud})(\text{longitud}).$$

El control dimensional no puede determinar cuál eje se está usando. Ese dato pertenece a la geometría del modelo: el radio de un cascarón es siempre la distancia de la tira al eje de rotación indicado.

Soluciones del nivel D. **Solución D1.** Los puntos de la gráfica correspondientes a la partición son

$$(0, 0), \qquad \left(1, \frac{2}{3}\right), \qquad \left(4, \frac{16}{3}\right).$$

La primera cuerda tiene longitud

$$\sqrt{(1-0)^2 + \left(\frac{2}{3} - 0\right)^2} = \frac{\sqrt{13}}{3}.$$

La segunda tiene longitud

$$\sqrt{(4-1)^2 + \left(\frac{16}{3} - \frac{2}{3}\right)^2} = \sqrt{9 + \left(\frac{14}{3}\right)^2} = \frac{\sqrt{277}}{3}.$$

Por tanto,

$$\boxed{L_P(f) = \frac{\sqrt{13} + \sqrt{277}}{3}.}$$

En un subintervalo $[x_{k-1}, x_k]$, el teorema del valor medio Teorema 9.3 proporciona un punto $\xi_k \in (x_{k-1}, x_k)$ tal que $f(x_k) - f(x_{k-1}) = f'(\xi_k)\Delta x_k$. Por consiguiente, la longitud de la cuerda correspondiente es

$$\begin{aligned}\sqrt{(\Delta x_k)^2 + \big(f(x_k) - f(x_{k-1})\big)^2} &= \sqrt{(\Delta x_k)^2 + f'(\xi_k)^2(\Delta x_k)^2} \\ &= \sqrt{1 + f'(\xi_k)^2}\,\Delta x_k.\end{aligned}$$

Para la función dada, $f'(x) = \sqrt{x}$, que es continua en $[0, 4]$. Luego $f \in C^1([0, 4])$. Podemos aplicar Teorema 20.2:

$$\begin{aligned}L(f) &= \int_0^4 \sqrt{1 + f'(x)^2}\,dx \\ &= \int_0^4 \sqrt{1 + x}\,dx \\ &= \left[\frac{2}{3}(1 + x)^{3/2}\right]_0^4 \\ &= \frac{2}{3}\left(5^{3/2} - 1\right).\end{aligned}$$

Así,

$$\boxed{L(f) = \frac{2}{3}(5\sqrt{5} - 1).}$$

La longitud poligonal asociada a P es una aproximación inferior al objeto definido mediante el supremo de todas las longitudes poligonales. El paso desde una partición concreta hasta la longitud exacta exige considerar el proceso de refinamiento y usar el resultado de Teorema 20.2; una sola cuerda poligonal no determina por sí misma la longitud de la gráfica.

Solución D2. La función $f(x) = |x|$ es continua en $[-1, 1]$, pero no es diferenciable en $x = 0$.

Por tanto, $f \notin C^1([-1, 1])$, y Teorema 20.2 no puede aplicarse directamente sobre todo el intervalo con sus hipótesis tal como fue formulado. El defecto está concentrado en un único punto. En los dos subintervalos

$$[-1, 0] \qquad \text{y} \qquad [0, 1],$$

la función es lineal:

$$f(x) = -x \qquad (-1 \le x \le 0),$$

y

$$f(x) = x \qquad (0 \le x \le 1).$$

Cada restricción pertenece a C^1 en su intervalo. Por Teorema 20.2,

$$L_- = \int_{-1}^0 \sqrt{1 + (-1)^2}\,dx = \sqrt{2},$$

y

$$L_+ = \int_0^1 \sqrt{1+1^2}\,dx = \sqrt{2}.$$

La longitud total es aditiva al concatenar ambos segmentos:

$$\boxed{L = L_- + L_+ = 2\sqrt{2}.}$$

El valor escrito por el estudiante es correcto, pero la invocación directa de Teorema 20.2 sobre $[-1,1]$ no está justificada por las hipótesis disponibles. La reparación consiste en partir el dominio en tramos donde la regularidad requerida sí se cumple. Este ejemplo distingue dos preguntas:

$$\boxed{\text{¿el resultado es correcto?} \qquad \text{y} \qquad \text{¿el argumento lo demuestra?}}$$

Una respuesta numérica correcta puede provenir de una justificación incompleta.

Solución D3. Para la partición $P = \{0,1\}$, la aproximación poligonal consiste en una sola cuerda entre

$$(0,0) \qquad \text{y} \qquad (1,1).$$

Por tanto,

$$\boxed{L_P(f) = \sqrt{2}.}$$

Para

$$Q = \left\{0, \frac{1}{2}, 1\right\},$$

aparecen los puntos

$$(0,0), \qquad \left(\frac{1}{2}, \frac{1}{4}\right), \qquad (1,1).$$

La primera cuerda tiene longitud

$$\sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{1}{4}\right)^2} = \frac{\sqrt{5}}{4},$$

y la segunda,

$$\sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{3}{4}\right)^2} = \frac{\sqrt{13}}{4}.$$

Así,

$$\boxed{L_Q(f) = \frac{\sqrt{5}+\sqrt{13}}{4}.}$$

Numéricamente, $L_P(f) \approx 1.4142$, mientras que $L_Q(f) \approx 1.4604$.
Luego $L_P(f) < L_Q(f)$.

Esto coincide con Lema 20.1, porque $P \preceq Q$: la partición Q se obtiene refinando P mediante el punto adicional $x = 1/2$. Geométricamente, sustituir una cuerda por dos segmentos que pasan por un punto intermedio de la curva no puede disminuir la longitud poligonal. Sin embargo, $L_Q(f)$ sigue siendo la longitud asociada a una partición concreta. La longitud de la gráfica está definida mediante el supremo de **todas** las longitudes poligonales: $L(f) = \sup_R L_R(f)$. Por tanto, el refinamiento mejora la aproximación, pero una partición finita aislada no proporciona por sí sola la longitud exacta.

Solución D4. Derivamos: $f(x) = \frac{1}{3}(2x+1)^{3/2}$.
Por la regla de la cadena,

$$\begin{aligned} f'(x) &= \frac{1}{3} \cdot \frac{3}{2}(2x+1)^{1/2} \cdot 2 \\ &= \sqrt{2x+1}. \end{aligned}$$

La función $x \longmapsto \sqrt{2x+1}$ es continua en $[0,3]$. Además, f es continua en ese intervalo. Luego

$$\boxed{f \in C^1([0,3]).}$$

Las hipótesis de Teorema 20.2 están satisfechas. Por tanto,

$$L(f) = \int_0^3 \sqrt{1 + f'(x)^2}\, dx.$$

Como $f'(x)^2 = 2x+1$, tenemos

$$\sqrt{1+f'(x)^2} = \sqrt{2x+2} = \sqrt{2}\,\sqrt{x+1}.$$

Así,

$$\begin{aligned} L(f) &= \sqrt{2}\int_0^3 \sqrt{x+1}\, dx \\ &= \sqrt{2}\left[\frac{2}{3}(x+1)^{3/2}\right]_0^3 \\ &= \frac{2\sqrt{2}}{3}\left(4^{3/2} - 1\right) \\ &= \frac{2\sqrt{2}}{3}(8-1). \end{aligned}$$

Por consiguiente,

$$\boxed{L(f) = \frac{14\sqrt{2}}{3}.}$$

El cálculo comienza con la verificación de regularidad. La fórmula de longitud entra solo después de confirmar que las hipótesis del teorema que la justifica están disponibles.

Solución D5. Sea

$$P = \{a = x_0 < x_1 < \cdots < x_n = b\}.$$

Como $g(x) = f(x) + c$, para cada k se tiene

$$g(x_k) - g(x_{k-1}) = f(x_k) - f(x_{k-1}).$$

Por Definición 20.1,

$$\begin{aligned} L_P(g) &= \sum_{k=1}^{n} \sqrt{(x_k - x_{k-1})^2 + (g(x_k) - g(x_{k-1}))^2} \\ &= \sum_{k=1}^{n} \sqrt{(x_k - x_{k-1})^2 + (f(x_k) - f(x_{k-1}))^2} \\ &= L_P(f). \end{aligned}$$

Por tanto,

$$\boxed{L_P(g) = L_P(f)}$$

para toda partición P.

Tomando supremos sobre todas las particiones y usando Definición 20.2,

$$\boxed{L(g) = L(f).}$$

La misma conclusión se obtiene desde Teorema 20.2. Como $g'(x) = f'(x)$, tenemos

$$\begin{aligned} L(g) &= \int_a^b \sqrt{1 + g'(x)^2}\, dx \\ &= \int_a^b \sqrt{1 + f'(x)^2}\, dx \\ &= L(f). \end{aligned}$$

Geométricamente, sumar una constante desplaza todos los puntos de la gráfica la misma distancia vertical. Las diferencias entre pares de puntos permanecen intactas, de modo que también permanecen intactas todas las longitudes de cuerdas y, en consecuencia, la longitud de la gráfica.

Solución D6. Para

$$f(x) = x, \qquad 0 \le x \le 1,$$

la gráfica es el segmento que une

$$(0, 0) \qquad \text{con} \qquad (1, 1).$$

Su longitud geométrica es

$$\sqrt{(1-0)^2 + (1-0)^2} = \boxed{\sqrt{2}}.$$

La fórmula propuesta por el estudiante da

$$\begin{aligned} L_{\text{inc}} &= \int_0^1 (1 + |1|)\, dx \\ &= \int_0^1 2\, dx \\ &= 2, \end{aligned}$$

que no coincide con $\sqrt{2}$.

En cambio, Teorema 20.2 produce

$$\begin{aligned} L &= \int_0^1 \sqrt{1 + f'(x)^2}\, dx \\ &= \int_0^1 \sqrt{1 + 1^2}\, dx \\ &= \boxed{\sqrt{2}}. \end{aligned}$$

La razón geométrica aparece ya en un pequeño tramo. Si el incremento horizontal es Δx y $\Delta y \approx f'(x)\Delta x$, la longitud local viene dada por Pitágoras:

$$\begin{aligned} \Delta s &\approx \sqrt{(\Delta x)^2 + (\Delta y)^2} \\ &\approx \sqrt{(\Delta x)^2 + f'(x)^2(\Delta x)^2} \\ &= \sqrt{1 + f'(x)^2}\, \Delta x. \end{aligned}$$

La expresión $(1 + |f'(x)|)\Delta x$ suma por separado una variación horizontal y una vertical; corresponde a otra noción de distancia local, no a la distancia euclidiana de la cuerda.

Finalmente, ambos integrandos son adimensionales y ambos se multiplican por dx, de modo que las dos integrales tienen unidades de longitud. El análisis dimensional no puede distinguir aquí la geometría euclidiana correcta de una regla local diferente.

Soluciones del nivel E. **Solución E1.** Como $\rho(x) = kx$ debe tener unidades de masa por longitud y x tiene unidades de longitud, k debe tener unidades

$$\boxed{\text{kg/m}^2.}$$

La masa total es

$$M = \int_0^2 kx\, dx.$$

Por tanto,

$$\begin{aligned} 12 &= k \left[\frac{x^2}{2} \right]_0^2 \\ &= 2k, \end{aligned}$$

de donde

$$\boxed{k = 6 \text{ kg/m}^2.}$$

Así,

$$\rho(x) = 6x \qquad (\text{kg/m}).$$

El primer momento respecto del origen es

$$\begin{aligned}\mathcal{M}_0 &= \int_0^2 x\rho(x)\,dx \\ &= \int_0^2 6x^2\,dx \\ &= 6\left[\frac{x^3}{3}\right]_0^2 \\ &= 16\ \text{kg m}.\end{aligned}$$

Por Corolario 20.4,

$$\begin{aligned}\bar{x} &= \frac{\mathcal{M}_0}{M} \\ &= \frac{16}{12} \\ &= \boxed{\frac{4}{3}\ \text{m}.}\end{aligned}$$

En efecto, $0 < \frac{4}{3} < 2$.

Además, $\frac{4}{3} > 1$, de modo que el centro de masa queda a la derecha del punto medio geométrico. Esto concuerda con el hecho de que la densidad aumenta linealmente con x: los puntos cercanos al extremo derecho reciben mayor peso.

Solución E2. La densidad satisface

$$\rho(-x) = \rho_0\left(1 + \frac{(-x)^2}{a^2}\right) = \rho(x).$$

Por tanto, ρ es una función par: puntos situados a igual distancia del origen tienen la misma densidad. Geométricamente, la distribución de masa es simétrica respecto de $x = 0$, así que esperamos

$$\boxed{\bar{x} = 0.}$$

La fórmula del primer momento confirma esta conclusión. Tenemos

$$\mathcal{M}_0 = \int_{-a}^{a} x\rho(x)\,dx.$$

Como x es impar y ρ es par, el producto $x\rho(x)$ es impar. Por simetría,

$$\boxed{\mathcal{M}_0 = 0.}$$

La masa total es positiva, pues

$$\begin{aligned} M &= \int_{-a}^{a} \rho_0 \left(1 + \frac{x^2}{a^2}\right) dx \\ &= 2\rho_0 \int_0^a \left(1 + \frac{x^2}{a^2}\right) dx \\ &= 2\rho_0 \left[x + \frac{x^3}{3a^2}\right]_0^a \\ &= 2\rho_0 \left(a + \frac{a}{3}\right) \\ &= \frac{8a\rho_0}{3}. \end{aligned}$$

Así, $\bar{x} = \frac{\mathcal{M}_0}{M} = 0$.

Para desplazar el centro de masa fuera del origen habría que romper la simetría de la distribución; por ejemplo, hacer que la densidad en una región a la derecha del origen difiera de la densidad en su punto reflejado a la izquierda.

Solución E3. Como $\rho(x) = 0$ fuera de $[1, 4]$, toda la masa de la varilla está efectivamente distribuida dentro de ese subintervalo. Por tanto,

$$M = \int_0^5 \rho(x)\, dx = \int_1^4 \rho(x)\, dx > 0.$$

Del mismo modo,

$$\mathcal{M}_0 = \int_0^5 x\rho(x)\, dx = \int_1^4 x\rho(x)\, dx.$$

Para $x \in [1, 4]$ se tiene $1 \leq x \leq 4$.

Como $\rho(x) \geq 0$, $\rho(x) \leq x\rho(x) \leq 4\rho(x)$.

Integrando en $[1, 4]$,

$$\int_1^4 \rho(x)\, dx \leq \int_1^4 x\rho(x)\, dx \leq 4 \int_1^4 \rho(x)\, dx.$$

Es decir, $M \leq \mathcal{M}_0 \leq 4M$.

Como $M > 0$, podemos dividir por la masa total: $1 \leq \frac{\mathcal{M}_0}{M} \leq 4$. Por Corolario 20.4, $\bar{x} = \frac{\mathcal{M}_0}{M}$.

Así,

$$\boxed{1 \leq \bar{x} \leq 4.}$$

La mejora respecto de la cota global $0 \leq \bar{x} \leq 5$ proviene de una información adicional del modelo: aunque la varilla geométrica ocupa $[0, 5]$, la densidad se anula fuera de $[1, 4]$. Para localizar el centro de masa importa el intervalo donde realmente hay masa.

Solución E4. Para la primera varilla,

$$\begin{aligned} M_L &= \int_0^2 k(2-x)\,dx \\ &= k\left[2x - \frac{x^2}{2}\right]_0^2 \\ &= 2k. \end{aligned}$$

Para la segunda,

$$\begin{aligned} M_R &= \int_0^2 kx\,dx \\ &= k\left[\frac{x^2}{2}\right]_0^2 \\ &= 2k. \end{aligned}$$

Por tanto,

$$\boxed{M_L = M_R = 2k.}$$

El primer momento de la varilla izquierda es

$$\begin{aligned} \mathcal{M}_L &= \int_0^2 x\,k(2-x)\,dx \\ &= k\int_0^2 (2x - x^2)\,dx \\ &= k\left[x^2 - \frac{x^3}{3}\right]_0^2 \\ &= \frac{4k}{3}. \end{aligned}$$

Así,

$$\boxed{\bar{x}_L = \frac{\mathcal{M}_L}{M_L} = \frac{2}{3}.}$$

Para la varilla derecha,

$$\begin{aligned} \mathcal{M}_R &= \int_0^2 x(kx)\,dx \\ &= k\int_0^2 x^2\,dx \\ &= \frac{8k}{3}. \end{aligned}$$

Por tanto,

$$\boxed{\bar{x}_R = \frac{\mathcal{M}_R}{M_R} = \frac{4}{3}.}$$

Los dos centros satisfacen $\frac{2}{3} + \frac{4}{3} = 2$, de modo que son simétricos respecto del punto medio $x = 1$. La igualdad de masa total solo fija

$$\int_0^2 \rho(x)\,dx.$$

El centro de masa depende además del primer momento

$$\int_0^2 x\rho(x)\,dx,$$

que registra **dónde** se sitúa esa masa.

Finalmente, $\rho_R(x) = \rho_L(2 - x)$.

La segunda distribución es el reflejo de la primera respecto de $x = 1$. Esa reflexión transforma el centro $\bar{x}_L = \frac{2}{3}$ en $2 - \bar{x}_L = \frac{4}{3}$, que coincide con $\bar{x}_R$.

Solución E5. Respecto del nuevo origen situado en $x = c$, la coordenada de un punto es $u = x - c$.

El primer momento correspondiente es

$$\boxed{\mathcal{M}_c = \int_a^b (x - c)\rho(x)\,dx.}$$

Por linealidad,

$$\begin{aligned}\mathcal{M}_c &= \int_a^b x\rho(x)\,dx - c\int_a^b \rho(x)\,dx \\ &= \mathcal{M}_0 - cM.\end{aligned}$$

Así,

$$\boxed{\mathcal{M}_c = \mathcal{M}_0 - cM.}$$

La coordenada del centro de masa respecto del nuevo origen es $\bar{u} = \frac{\mathcal{M}_c}{M}$.

Sustituyendo la relación anterior,

$$\begin{aligned}\bar{u} &= \frac{\mathcal{M}_0 - cM}{M} \\ &= \frac{\mathcal{M}_0}{M} - c \\ &= \bar{x} - c.\end{aligned}$$

Por tanto,

$$\boxed{\bar{u} = \bar{x} - c.}$$

El punto físico no ha cambiado. Solo ha cambiado el sistema de coordenadas usado para describirlo. Si en el sistema original el centro está en $x = \bar{x}$, en el sistema trasladado su coordenada es exactamente la distancia algebraica desde el nuevo origen: $\bar{x} - c$. El mismo fenómeno aparece en el momento: trasladar el origen modifica su valor en la cantidad cM, pero la posición física reconstruida a partir del cociente momento/masa sigue siendo la misma.

Solución E6. La función $\rho(x) = 1 - x$ es negativa cuando $x > 1$.
En particular,

$$\rho(x) < 0 \qquad (1 < x \le 2).$$

Su integral total es

$$\begin{aligned}\int_0^2 (1-x)\,dx &= \left[x - \frac{x^2}{2}\right]_0^2 \\ &= 2 - 2 \\ &= 0.\end{aligned}$$

Por tanto,

$$\boxed{\int_0^2 \rho(x)\,dx = 0.}$$

En el modelo físico de §19.7–§19.8, una densidad de masa debe ser no negativa: una pequeña longitud Δx aporta aproximadamente $\rho(\xi)\Delta x$ unidades de masa, y esa contribución no puede ser negativa. Aquí la función cambia de signo, de modo que describe una acumulación **orientada** más que una distribución ordinaria de masa. Además, el denominador de la fórmula del centro de masa vale cero:

$$\int_0^2 \rho(x)\,dx = 0.$$

Así,

$$\frac{\int_0^2 x\rho(x)\,dx}{\int_0^2 \rho(x)\,dx}$$

no está definido.

Esto no vuelve inútil a la función. En otros contextos, una densidad con signo puede representar contribuciones orientadas; por ejemplo, una densidad lineal de carga eléctrica puede tomar valores positivos y negativos. En ese caso, la integral total representa una carga neta, y la cancelación forma parte del modelo. La lección es que la integrabilidad de una función no basta para convertirla en una densidad física de masa: también deben respetarse las restricciones semánticas del modelo.

Soluciones del nivel F. **Solución F1.** Si x mide la elongación desde la posición de equilibrio, la fuerza restauradora del resorte es $F_{\text{res}}(x) = -kx = -50x$.

Para estirar el resorte lentamente, la fuerza externa equilibra en cada posición a la fuerza restauradora: $F_{\text{ext}}(x) = 50x$. El trabajo de la fuerza externa es

$$\begin{aligned}W_{\text{ext}} &= \int_{0.10}^{0.40} 50x\,dx \\ &= 25x^2\Big|_{0.10}^{0.40} \\ &= 25(0.16 - 0.01) \\ &= \boxed{3.75\text{ J.}}\end{aligned}$$

El trabajo realizado por el resorte sobre la misma trayectoria es

$$\begin{aligned} W_{\text{res}} &= \int_{0.10}^{0.40} (-50x)\,dx \\ &= -25(0.16 - 0.01) \\ &= \boxed{-3.75 \text{ J.}} \end{aligned}$$

Por tanto,

$$\boxed{W_{\text{ext}} = -W_{\text{res}}.}$$

La igualdad de magnitudes refleja que, en el modelo cuasiestático, las fuerzas tienen igual magnitud en cada posición. Los signos distinguen su orientación respecto del desplazamiento: la fuerza externa favorece el estiramiento, mientras que la fuerza del resorte se opone a él.

Solución F2. La tasa neta es

$$\begin{aligned} r_{\text{net}}(t) &= r_{\text{in}}(t) - r_{\text{out}}(t) \\ &= (5 - t) - (1 + t) \\ &= 4 - 2t. \end{aligned}$$

La entrada acumulada es

$$\begin{aligned} I &= \int_0^4 (5 - t)\,dt \\ &= \left[5t - \frac{t^2}{2}\right]_0^4 \\ &= 20 - 8 \\ &= \boxed{12 \text{ L.}} \end{aligned}$$

La salida acumulada es

$$\begin{aligned} O &= \int_0^4 (1 + t)\,dt \\ &= \left[t + \frac{t^2}{2}\right]_0^4 \\ &= 4 + 8 \\ &= \boxed{12 \text{ L.}} \end{aligned}$$

Por tanto, el cambio neto del volumen es

$$\boxed{Q(4) - Q(0) = I - O = 0.}$$

El flujo bruto total suma todo el líquido que cruzó la frontera del sistema en cualquier dirección:

$$\boxed{I + O = 24 \text{ L.}}$$

Ahora, $r_{\text{net}}(t) = 4 - 2t$ cambia de signo en $t = 2$.
Así,

$$\begin{aligned}\int_0^4 |r_{\text{net}}(t)|\,dt &= \int_0^2 (4-2t)\,dt + \int_2^4 (2t-4)\,dt \\ &= 4+4 \\ &= \boxed{8 \text{ L.}}\end{aligned}$$

Esta última cantidad mide la variación total acumulada del volumen descrita por la **tasa neta**: durante los dos primeros minutos el volumen aumenta en 4 litros y durante los dos siguientes disminuye en 4 litros. Las tres cantidades responden a preguntas diferentes:

$$\boxed{\text{cambio neto} = 0, \qquad \text{flujo bruto} = 24 \text{ L}, \qquad \text{variación total} = 8 \text{ L}.}$$

El cambio neto compara solo los estados inicial y final; el flujo bruto cuenta toda entrada y toda salida; la variación total acumula la magnitud de los cambios del contenido del depósito.

Solución F3. La fuerza cambia de signo cuando $6 - 3x = 0$, es decir, en $x = 2$.

Para $0 \le x < 2$, se tiene $F(x) > 0$, de modo que la fuerza favorece el desplazamiento. El trabajo en ese tramo es

$$\begin{aligned}W_+ &= \int_0^2 (6-3x)\,dx \\ &= \left[6x - \frac{3x^2}{2}\right]_0^2 \\ &= 12 - 6 \\ &= \boxed{6 \text{ J.}}\end{aligned}$$

Para $2 < x \le 4$, se tiene $F(x) < 0$, y la fuerza se opone al desplazamiento. El trabajo correspondiente es

$$\begin{aligned}W_- &= \int_2^4 (6-3x)\,dx \\ &= \left[6x - \frac{3x^2}{2}\right]_2^4 \\ &= 0 - 6 \\ &= \boxed{-6 \text{ J.}}\end{aligned}$$

Por tanto, el trabajo neto es

$$\boxed{W = W_+ + W_- = 0.}$$

En cambio,

$$\begin{aligned}\int_0^4 |F(x)|\,dx &= \int_0^2 (6-3x)\,dx + \int_2^4 (3x-6)\,dx \\ &= 6+6 \\ &= \boxed{12 \text{ J.}}\end{aligned}$$

Esta última integral acumula las magnitudes de las contribuciones locales de trabajo sin permitir cancelación entre los tramos de signo opuesto. El trabajo neto puede ser cero porque la transferencia positiva de los primeros dos metros se cancela exactamente con la transferencia negativa de los dos metros siguientes. La ausencia de cambio neto no implica ausencia de interacción durante el recorrido.

Solución F4. Factorizamos la velocidad: $v(t) = t^2 - 4t + 3 = (t-1)(t-3)$.
Por tanto, la partícula cambia de sentido en

$$\boxed{t = 1, \qquad t = 3.}$$

La velocidad es positiva en $[0, 1)$, negativa en $(1, 3)$ y positiva en $(3, 4]$.
El desplazamiento total es

$$\begin{aligned} s(4) - s(0) &= \int_0^4 (t^2 - 4t + 3)\, dt \\ &= \left[\frac{t^3}{3} - 2t^2 + 3t\right]_0^4 \\ &= \frac{64}{3} - 32 + 12 \\ &= \boxed{\frac{4}{3}\ \text{m.}} \end{aligned}$$

Como $s(0) = 5$ m, la posición final es

$$\begin{aligned} s(4) &= s(0) + \frac{4}{3} \\ &= 5 + \frac{4}{3} \\ &= \boxed{\frac{19}{3}\ \text{m.}} \end{aligned}$$

Para la distancia debemos integrar la rapidez. Sea $A(t) = \frac{t^3}{3} - 2t^2 + 3t$.
Entonces

$$A(0) = 0, \qquad A(1) = \frac{4}{3}, \qquad A(3) = 0, \qquad A(4) = \frac{4}{3}.$$

Por tanto,

$$\begin{aligned} D &= \int_0^1 v(t)\, dt - \int_1^3 v(t)\, dt + \int_3^4 v(t)\, dt \\ &= \frac{4}{3} + \frac{4}{3} + \frac{4}{3} \\ &= \boxed{4\ \text{m.}} \end{aligned}$$

Las tres cantidades responden a preguntas distintas:

$$\boxed{\text{desplazamiento} = \frac{4}{3}\ \text{m},}$$

$$\boxed{\text{posición final} = \frac{19}{3}\ \text{m},}$$

y

$$\boxed{\text{distancia recorrida} = 4\ \text{m}.}$$

El desplazamiento mide cambio neto de posición, la posición final incorpora además el dato inicial y la distancia acumula la magnitud de todos los desplazamientos locales.

Solución F5. En el recorrido desde $x = 0$ hasta $x = 2$, el trabajo es

$$\begin{aligned} W_{0\to 2} &= \int_0^2 (x^2 + 1)\, dx \\ &= \left[\frac{x^3}{3} + x\right]_0^2 \\ &= \frac{8}{3} + 2 \\ &= \boxed{\frac{14}{3}\ \text{J}.} \end{aligned}$$

Al recorrer el mismo intervalo en sentido contrario,

$$\begin{aligned} W_{2\to 0} &= \int_2^0 (x^2 + 1)\, dx \\ &= -\int_0^2 (x^2 + 1)\, dx \\ &= \boxed{-\frac{14}{3}\ \text{J}.} \end{aligned}$$

Por tanto,

$$\boxed{W_{2\to 0} = -W_{0\to 2}.}$$

La fuerza tiene componente positiva en todo el intervalo. En el recorrido $0 \to 2$, el desplazamiento elemental está orientado en la dirección positiva y el trabajo es positivo. En el recorrido $2 \to 0$, el desplazamiento elemental tiene orientación opuesta a la fuerza y el trabajo es negativo. Introducir $|F|$ no cambiaría nada aquí porque $F > 0$, pero reemplazar la integral orientada por una acumulación forzada a ser positiva eliminaría precisamente la información sobre la orientación del desplazamiento. El signo del trabajo forma parte de la cantidad física modelada.

Solución F6. Por Teorema 16.3,

$$Q(4) - Q(0) = \int_0^4 (2t - 3)\, dt.$$

Evaluando,

$$\begin{aligned} Q(4) - Q(0) &= \Big[t^2 - 3t\Big]_0^4 \\ &= 16 - 12 \\ &= \boxed{4 \text{ kg.}} \end{aligned}$$

Por tanto,

$$\boxed{Q(4) = Q(0) + 4 \text{ kg.}}$$

La tasa determina cuánto cambia la cantidad durante el intervalo, pero no fija el nivel desde el cual comienza ese cambio. Sin conocer $Q(0)$ no puede obtenerse un valor numérico único para $Q(4)$.

Si $Q(0) = 10$ kg, entonces

$$\boxed{Q(4) = 14 \text{ kg.}}$$

Si $Q(0) = 25$ kg, entonces

$$\boxed{Q(4) = 29 \text{ kg.}}$$

En ambos casos el incremento es el mismo:

$$\boxed{Q(4) - Q(0) = 4 \text{ kg.}}$$

Lo que cambia es el nivel absoluto de la cantidad, porque ese nivel incorpora el dato inicial. La distinción puede resumirse como

$$\boxed{\text{tasa conocida} \longrightarrow \text{cambio neto conocido,}}$$

mientras que

$$\boxed{\text{tasa conocida} + \text{valor inicial} \longrightarrow \text{valor final.}}$$

Soluciones del nivel G. **Solución G1.** La función cambia de signo en $2 - x = 0$, es decir, en $x = 2$.

Primero calculamos las dos acumulaciones básicas que aparecerán en los distintos modelos.

La integral orientada es

$$\begin{aligned} \int_0^3 (2 - x)\, dx &= \left[2x - \frac{x^2}{2}\right]_0^3 \\ &= 6 - \frac{9}{2} \\ &= \boxed{\frac{3}{2}.} \end{aligned}$$

La acumulación de magnitudes es

$$\begin{aligned}\int_0^3 |2-x|\,dx &= \int_0^2 (2-x)\,dx + \int_2^3 (x-2)\,dx \\ &= 2 + \frac{1}{2} \\ &= \boxed{\frac{5}{2}.}\end{aligned}$$

1. Área geométrica. Cada contribución local de área debe ser no negativa. Por tanto,

$$\boxed{A = \int_0^3 |2-x|\,dx = \frac{5}{2}.}$$

2. Densidad de masa. La función propuesta satisface

$$q(x) < 0 \qquad (2 < x \leq 3).$$

Por ello no cumple la condición de no negatividad exigida para una densidad ordinaria de masa en este capítulo. El modelo debe corregirse o reinterpretarse antes de hablar de masa total de una varilla.

3. Trabajo. El signo de la fuerza conserva información sobre si favorece o se opone al desplazamiento. Así,

$$\boxed{W = \int_0^3 (2-x)\,dx = \frac{3}{2}}$$

en las unidades de trabajo correspondientes.

4. Movimiento. Como velocidad, q produce el desplazamiento

$$\boxed{s(3) - s(0) = \int_0^3 (2-x)\,dx = \frac{3}{2},}$$

mientras que la distancia recorrida es

$$\boxed{D = \int_0^3 |2-x|\,dx = \frac{5}{2}.}$$

La misma fórmula algebraica conduce a modelos distintos porque el significado del signo depende de la cantidad local que se está acumulando. El integrando se elige desde el contexto antes de efectuar el cálculo.

Solución G2. La expresión del estudiante $\rho(x) = 3x$ kg/m es dimensionalmente ambigua. Si x representa una longitud, el coeficiente que multiplica a x debe tener unidades de masa por longitud al cuadrado.

Para parametrizar la familia sin mezclar magnitudes de unidades distintas, escribamos $\xi := \frac{x}{1\ \mathrm{m}}$, de modo que ξ es adimensional y recorre $[0, 2]$. Toda densidad afín creciente puede escribirse entonces como $\rho(x) = (\alpha + \beta\xi)\frac{\mathrm{kg}}{\mathrm{m}}$, con $\beta > 0$.

La masa total vale

$$\begin{aligned} 6\ \mathrm{kg} &= \int_0^{2\mathrm{m}} \rho(x)\,dx \\ &= \mathrm{kg}\int_0^{2} (\alpha+\beta\xi)\,d\xi \\ &= (2\alpha+2\beta)\,\mathrm{kg}. \end{aligned}$$

Por tanto, $\alpha+\beta=3$, y así $\alpha=3-\beta$.

Como la densidad es creciente, su mínimo se alcanza en $\xi=0$. La no negatividad exige $\alpha=3-\beta\geq 0$.

En consecuencia,

$$\boxed{0<\beta\leq 3.}$$

La familia completa compatible con los datos es

$$\boxed{\rho_\beta(x)=\left[(3-\beta)+\beta\frac{x}{1\ \mathrm{m}}\right]\frac{\mathrm{kg}}{\mathrm{m}},\qquad 0<\beta\leq 3.}$$

El primer momento respecto del origen es

$$\begin{aligned} \mathcal{M}_0 &= \int_0^{2\mathrm{m}} x\,\rho_\beta(x)\,dx \\ &= \mathrm{kg\,m}\int_0^{2} \xi\big((3-\beta)+\beta\xi\big)\,d\xi \\ &= \mathrm{kg\,m}\left[2(3-\beta)+\frac{8\beta}{3}\right] \\ &= \left(6+\frac{2\beta}{3}\right)\mathrm{kg\,m}. \end{aligned}$$

Como la masa total es 6 kg,

$$\begin{aligned} \bar{x} &= \frac{\mathcal{M}_0}{M} \\ &= \left(1+\frac{\beta}{9}\right)\mathrm{m}. \end{aligned}$$

Así,

$$\boxed{\bar{x}=\left(1+\frac{\beta}{9}\right)\mathrm{m},\qquad 0<\beta\leq 3.}$$

Por tanto,

$$\boxed{1\ \mathrm{m}<\bar{x}\leq\frac{4}{3}\ \mathrm{m}.}$$

El valor $\bar{x}=\frac{4}{3}$ m corresponde al caso $\beta=3$, para el cual $\rho(x)=3\frac{x}{1\ \mathrm{m}}\frac{\mathrm{kg}}{\mathrm{m}}$. Esta es la versión dimensionalmente correcta de la idea que el estudiante intentó expresar con $\rho(x)=3x$. La densidad elegida por el estudiante puede, por tanto, representar un modelo compatible cuando se interpretan correctamente sus unidades, pero los datos del problema admiten infinitas densidades afines crecientes. La falla lógica consiste en convertir una posibilidad en una conclusión.

Solución G3. Como $q(x) = x(2-x) = 2x - x^2$ es no negativa en $[0,2]$, el área bajo la gráfica es

$$\begin{aligned} A &= \int_0^2 q(x)\,dx \\ &= \int_0^2 (2x - x^2)\,dx \\ &= \left[x^2 - \frac{x^3}{3}\right]_0^2 \\ &= 4 - \frac{8}{3} \\ &= \boxed{\frac{4}{3}.} \end{aligned}$$

Al girar la región alrededor del eje y, una tira vertical genera un cascarón de radio x y altura $q(x)$.

Por tanto,

$$\begin{aligned} V &= 2\pi \int_0^2 xq(x)\,dx \\ &= 2\pi \int_0^2 (2x^2 - x^3)\,dx \\ &= 2\pi \left[\frac{2x^3}{3} - \frac{x^4}{4}\right]_0^2 \\ &= 2\pi \left(\frac{16}{3} - 4\right) \\ &= \boxed{\frac{8\pi}{3}.} \end{aligned}$$

Ahora interpretamos q como densidad lineal. La masa total es

$$\boxed{M = \int_0^2 q(x)\,dx = \frac{4}{3}.}$$

El primer momento respecto del origen es

$$\begin{aligned} \mathcal{M}_0 &= \int_0^2 xq(x)\,dx \\ &= \int_0^2 (2x^2 - x^3)\,dx \\ &= \frac{4}{3}. \end{aligned}$$

Por tanto,

$$\boxed{\bar{x} = \frac{\mathcal{M}_0}{M} = 1.}$$

Las tres apariciones de la misma función responden a estructuras locales distintas. Para el área, $q(x)\,dx$ es altura por anchura. Para los cascarones, $2\pi x q(x)\,dx$ es circunferencia por altura por espesor. Para el centro de masa, $xq(x)\,dx$ es posición por masa local, y la división por la masa total normaliza el primer momento:

$$\bar{x} = \frac{\int_0^2 xq(x)\,dx}{\int_0^2 q(x)\,dx}.$$

La fórmula algebraica q permanece igual; lo que cambia es el significado físico o geométrico asignado a $q(x)$ y, en consecuencia, la cantidad local que debe acumularse.

Solución G4. La densidad constante $\rho_1(x) = 3$ tiene masa

$$\int_0^2 3\,dx = 6.$$

Su primer momento es

$$\int_0^2 3x\,dx = 3\left[\frac{x^2}{2}\right]_0^2 = 6,$$

de modo que $\bar{x} = \frac{6}{6} = 1$.

Por tanto, ρ_1 es compatible con los datos. Ahora construimos otra densidad. Definimos

$$\rho_2(x) = \begin{cases} 2, & 0 \le x < \frac{1}{2}, \\ 4, & \frac{1}{2} \le x \le \frac{3}{2}, \\ 2, & \frac{3}{2} < x \le 2. \end{cases}$$

Esta función es no negativa, acotada y Riemann-integrable.

Su masa total es

$$\begin{aligned} M_2 &= 2\left(\frac{1}{2}\right) + 4(1) + 2\left(\frac{1}{2}\right) \\ &= 1 + 4 + 1 \\ &= \boxed{6}. \end{aligned}$$

Además, ρ_2 es simétrica respecto de $x = 1$. Por esa simetría, su centro de masa debe estar en $\bar{x}_2 = 1$.

También podemos verificarlo directamente:

$$\begin{aligned} \mathcal{M}_2 &= \int_0^{1/2} 2x\,dx + \int_{1/2}^{3/2} 4x\,dx + \int_{3/2}^2 2x\,dx \\ &= \frac{1}{4} + 4 + \frac{7}{4} \\ &= 6. \end{aligned}$$

Así,

$$\boxed{\bar{x}_2 = \frac{\mathcal{M}_2}{M_2} = 1.}$$

Hemos obtenido dos densidades distintas, $\rho_1 \neq \rho_2$, con la misma masa y el mismo centro de masa.

Los datos

$$\int_0^2 \rho(x)\,dx = 6$$

y

$$\int_0^2 x\rho(x)\,dx = 6$$

imponen dos restricciones integrales sobre ρ, pero no especifican todos sus valores. Muchas distribuciones diferentes pueden compartir esas dos cantidades globales. El error lógico del estudiante consiste en confundir **existencia de un modelo compatible** con **unicidad del modelo**. La densidad constante es una posibilidad válida, pero no una consecuencia necesaria de los datos.

Capítulo 20

Ecuaciones diferenciales elementales y síntesis Newton–Leibniz

El cálculo diferencial obtiene una tasa a partir de un estado conocido, y el teorema fundamental permite reconstruir el estado desde una tasa continua y un valor de referencia. Si la tasa depende del propio estado buscado, como ocurre en $y' = y$, debemos encontrar una función que satisfaga la relación con su derivada.

Las herramientas construidas permiten resolver algunas clases de estas ecuaciones. En cada caso comprobaremos que la fórmula define una solución en el intervalo indicado y, cuando el argumento lo permita, que el dato inicial determina la única solución. La conclusión de cada prueba quedará limitada a las hipótesis de la clase estudiada.

20.1. Solución sobre un intervalo y dato inicial

La incógnita es una función. Consideremos primero la relación $y' = 2x$. Si sustituimos $y(x) = x^2$, la derivada coincide con el miembro derecho para todo $x \in \mathbb{R}$. Lo mismo sucede con $y(x) = x^2 + 3$ y con $y(x) = x^2 - 5$: la derivación elimina una constante añadida. Por tanto, conocer esta tasa no determina todavía una única función. La notación y' abrevia la derivada de una función que debe buscarse. En una **ecuación diferencial ordinaria**, la incógnita depende de una sola variable y la ecuación relaciona sus valores con alguna de sus derivadas. El orden es el de la derivada más alta que aparece; aquí estudiaremos principalmente ecuaciones de primer orden escritas como $y' = F(x, y)$. Aunque la expresión F reciba dos argumentos, la función incógnita y sigue dependiendo únicamente de x: al sustituirla obtenemos $F(x, y(x))$.

La verificación debe incluir el dominio. En $y' = 1/y$, por ejemplo, los valores de una solución tienen que ser distintos de cero en todos los puntos de su intervalo; de lo contrario el miembro derecho no está definido y la sustitución no permite comprobar la ecuación.

Definición 20.1. **Definición (solución de una ecuación diferencial de primer orden).** Sean $D \subseteq \mathbb{R}^2$ y $F\colon D \to \mathbb{R}$. Una **solución de la ecuación**

$$y' = F(x, y)$$

sobre un intervalo abierto no vacío $I \subseteq \mathbb{R}$ es una función $y\colon I \to \mathbb{R}$ diferenciable en todo I que satisface, para cada $x \in I$, las dos condiciones

$$(x, y(x)) \in D \qquad \text{y} \qquad y'(x) = F(x, y(x)).$$

La primera condición hace admisible la sustitución; la segunda afirma que la ecuación se cumple. Ambas se exigen en **todos** los puntos de I. Comprobar una igualdad en un punto, o incluso en varios puntos escogidos, no equivale a verificar una solución sobre el intervalo. No hemos supuesto que F sea continua ni afirmado que toda ecuación así escrita posea una solución. La definición establece qué significa ser solución. Los resultados de existencia y unicidad necesitarán sus propias hipótesis y demostraciones. Tampoco se exige que D sea un rectángulo: basta que contenga cada par $(x, y(x))$ que la sustitución utiliza.

En este capítulo trabajaremos con intervalos abiertos, que pueden ser acotados, semirrectas o toda $\mathbb{R}$. Así la derivada en cada punto de la solución es bilateral. Un extremo finito que no pertenece al intervalo no adquiere un valor por aparecer en una fórmula; para incorporarlo y continuar más allá habrá que justificar una prolongación, cuestión que estudiaremos posteriormente.

El valor que fija una trayectoria. Volvamos a $y' = 2x$. Si, además de la tasa, conocemos $y(0) = 3$, la función $x^2 + 3$ cumple las dos exigencias, mientras que x^2 y $x^2 - 5$ ya no cumplen el valor prescrito. Este dato no altera la ecuación: selecciona, entre sus soluciones, aquellas que pasan por un punto determinado.

Definición 20.2. **Definición (problema de valor inicial).** Sean $F\colon D \to \mathbb{R}$ y $(x_0, y_0) \in D$. El **problema de valor inicial**

$$y' = F(x, y), \qquad y(x_0) = y_0$$

consiste en encontrar una solución de la ecuación sobre un intervalo abierto I que contenga x_0 y cuyo valor en ese punto sea y_0.

Cuando se fija de antemano el intervalo I, la búsqueda se realiza sobre ese intervalo. Si no se fija, también debemos indicar en qué intervalo hemos construido la solución. La palabra «inicial» procede de problemas donde la variable representa el tiempo, pero matemáticamente x_0 es un punto de referencia y no tiene que ser un extremo izquierdo. Una solución puede estudiarse a ambos lados de él. El dato debe pertenecer a D para que sea admisible. Una vez comprobada esa condición, quedan por resolver la existencia de una solución, su posible unicidad y los intervalos en que puede definirse. El ejemplo siguiente permite estudiar cada cuestión directamente.

Ejemplo 20.1. **Ejemplo (una familia completa y el dato que elige su constante).** Sea I un intervalo abierto no vacío y consideremos $y' = 2x$.

El miembro derecho está definido para todos los pares $(x, y) \in \mathbb{R}^2$. Para cada $C \in \mathbb{R}$, la función

$$y_C\colon I \to \mathbb{R}, \qquad y_C(x) = x^2 + C$$

es diferenciable y satisface $y_C'(x) = 2x$ en todo I. Hemos verificado una familia de soluciones; todavía debemos comprobar que la familia las contiene a todas. Sea y una solución cualquiera sobre I y fijemos $x_0 \in I$. La diferencia $h(x) = y(x) - x^2$ es diferenciable y tiene derivada nula en I. Para cada $x \in I$ distinto de x_0, el segmento cerrado cuyos extremos son x y x_0 está contenido en I. En ese segmento h es continua, y en su interior es diferenciable con derivada cero. Por Proposición 9.2, es constante allí; en particular, $h(x) = h(x_0)$. La misma igualdad vale cuando $x = x_0$, de modo que

$$y(x) = x^2 + C \qquad (x \in I), \qquad C = y(x_0) - x_0^2.$$

Así queda probada la exhaustividad: toda solución sobre I pertenece a la familia encontrada. Si exigimos además $y(x_0) = y_0$, la constante queda determinada por $C = y_0 - x_0^2$. El problema tiene, sobre cualquier intervalo abierto I que contenga x_0, la única solución

$$\boxed{y(x) = x^2 + y_0 - x_0^2.}$$

En particular, para $x_0 = 0$ y $y_0 = 3$, obtenemos $y(x) = x^2 + 3$. La fórmula define también una solución sobre toda $\mathbb{R}$, porque tanto ella como la ecuación están definidas allí.

Control del dominio. La restricción de $x^2 + 3$ a $(0, \infty)$ sigue resolviendo la ecuación en ese intervalo. Sin embargo, esa función restringida no resuelve el problema con dato $y(0) = 3$, pues 0

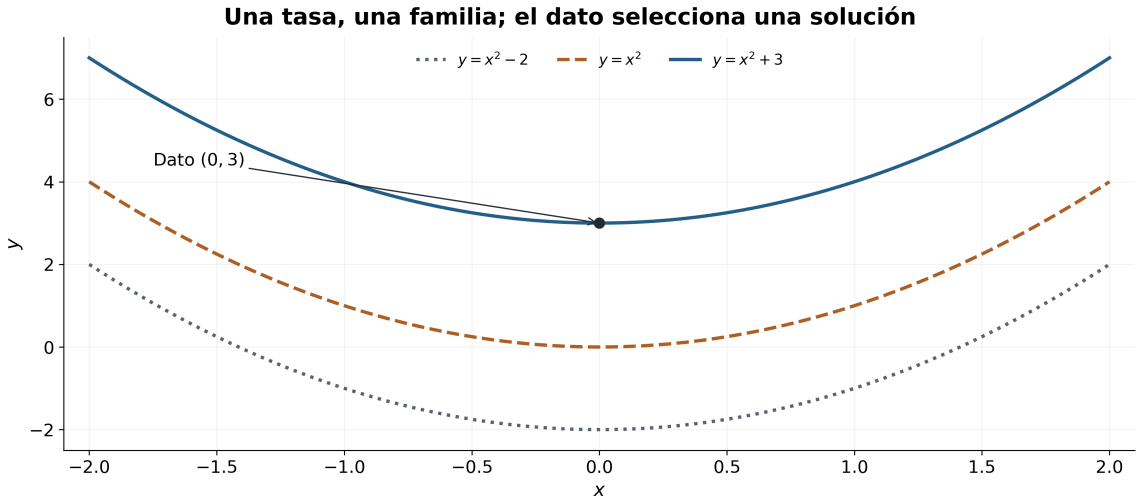


FIGURA 20.1. Las funciones $y = x^2 + C$ tienen la misma derivada; el dato $y(0) = 3$ selecciona $C = 3$.

no pertenece a su dominio. La extensión a $\mathbb{R}$ sí lo resuelve. La coincidencia de las fórmulas no vuelve idénticas dos funciones con dominios distintos.

En Figura 21.1 se distingue la familia completa del miembro seleccionado por el dato. El argumento de derivada nula, y no la inspección del dibujo, prueba la exhaustividad. El ejemplo contiene dos argumentos con trabajos diferentes. Derivar $x^2 + C$ demuestra que cada función propuesta es solución. Examinar una solución arbitraria y y probar que $y - x^2$ es constante demuestra que la familia es completa. Finalmente, evaluar en x_0 selecciona la constante y acredita la unicidad para ese dato sobre un intervalo fijado. Ninguno de estos pasos debe confundirse con los demás.

Verificar antes de generalizar. La función $z(x) = x^2 + x + 3$, definida en toda $\mathbb{R}$, cumple el mismo dato $z(0) = 3$, pero no la ecuación: $z'(x) = 2x + 1$. Pasar por el punto prescrito no basta. En cambio, $w(x) = x^2$ cumple la ecuación en toda la recta, pero falla el dato $w(0) = 3$. Ser solución de la ecuación y ser solución del problema de valor inicial son afirmaciones distintas. La definición tampoco obliga a que una solución tenga segunda derivada o derivadas de todos los órdenes. Solo exige la diferenciabilidad necesaria para la ecuación de primer orden considerada. Si más adelante una ecuación y sus coeficientes permiten deducir regularidad adicional, habrá que demostrarla.

Para verificar una propuesta, el recorrido es ya concreto: especificar el intervalo, comprobar que todos los pares $(x, y(x))$ pertenecen al dominio de la ecuación, derivar y sustituir en cada punto, y evaluar el dato inicial cuando exista. Si el objetivo es encontrar **todas** las soluciones, falta además un argumento de exhaustividad. En la sección siguiente veremos cómo el teorema fundamental del cálculo organiza precisamente esta reconstrucción cuando la tasa depende solo de la variable independiente.

20.2. Reconstruir una función desde una tasa conocida

La ecuación $y' = 2x$ de la sección anterior se resolvió porque reconocimos una función cuya derivada era $2x$. Pero esa identificación no es el fundamento de la existencia de una solución. Si una tasa g es continua, el teorema fundamental del cálculo permite construir una primitiva

mediante una integral, aunque no sepamos escribirla con las funciones elementales conocidas. El dato inicial fija después el valor de esa primitiva en el punto de referencia.

La construcción natural es acumular la tasa desde x_0 hasta x y añadir el valor y_0. La integral orientada conserva el mismo significado a ambos lados de x_0: cuando $x < x_0$, restamos la acumulación desde x hasta x_0. No necesitamos una integral sobre todo el intervalo de solución, que podría ser no acotado; cada evaluación utiliza únicamente un segmento compacto.

Teorema 20.1. **Teorema (reconstrucción desde una tasa continua y un dato inicial).** Sea $I \subseteq \mathbb{R}$ un intervalo abierto no vacío, sea $g\colon I \to \mathbb{R}$ continua y sean $x_0 \in I$ y $y_0 \in \mathbb{R}$. Entonces el problema

$$y'(x) = g(x) \quad (x \in I), \qquad y(x_0) = y_0$$

tiene una única solución sobre I, dada por

$$\boxed{y(x) = y_0 + \int_{x_0}^{x} g(t)\,dt \qquad (x \in I).}$$

La solución pertenece a $C^1(I)$.

Idea de la prueba. Las dos direcciones del teorema fundamental realizan trabajos distintos. Derivar la acumulación construida demuestra que tenemos una solución. Integrar la derivada de una solución cualquiera demuestra que necesariamente coincide con esa construcción. Conviene mantener separados ambos argumentos para no confundir una fórmula propuesta con una descripción exhaustiva.

Demostración. Para cada $x \in I$, el segmento cerrado cuyos extremos son x_0 y x está contenido en I. La continuidad de g implica su integrabilidad en ese segmento; por tanto, la integral orientada de la fórmula está bien definida y es finita. Para $x = x_0$ es cero, de modo que $y(x_0) = y_0$.

Fijemos ahora $c \in I$. Como I es abierto, podemos elegir $a, b \in I$ tales que $a < \min\{c, x_0\}$ y $\max\{c, x_0\} < b$. En $[a, b]$, g es continua. Aplicamos Teorema 16.4 con punto base $p = x_0$ y extremo variable $\phi(x) = x$, restringido a $[a, b]$. La identidad es diferenciable en c, y el resultado da

$$\frac{d}{dx}\int_{x_0}^{x} g(t)\,dt\Big|_{x=c} = g(c).$$

La constante y_0 tiene derivada cero. Así $y'(c) = g(c)$; como c era arbitrario, la fórmula define una solución en todo I. Su derivada es continua porque coincide con g, luego $y \in C^1(I)$. Para demostrar unicidad, sea z cualquier solución del mismo problema sobre I. Si $x > x_0$, la función z es continua en $[x_0, x]$ y diferenciable en su interior, y allí $z' = g$. Newton–Leibniz (Teorema 16.3) proporciona

$$z(x) - z(x_0) = \int_{x_0}^{x} g(t)\,dt.$$

Si $x < x_0$, aplicamos ese mismo teorema en $[x, x_0]$ y cambiamos la orientación:

$$z(x_0) - z(x) = \int_{x}^{x_0} g(t)\,dt, \qquad z(x) - z(x_0) = \int_{x_0}^{x} g(t)\,dt.$$

Para $x = x_0$, la última identidad también vale, con ambos miembros iguales a cero. Como $z(x_0) = y_0$, en todos los casos obtenemos

$$z(x) = y_0 + \int_{x_0}^{x} g(t)\,dt = y(x).$$

Toda solución coincide con la construida, lo que demuestra la unicidad. □

El argumento acredita la solución sobre **todo el intervalo I declarado**. No requiere que g esté acotada en todo I: basta su continuidad y, con ella, la integrabilidad sobre cada segmento compacto utilizado. Tampoco hay aquí un límite en un extremo infinito ni una integral impropia. Si cambiamos el punto de referencia a $x_1 \in I$ y usamos el valor ya obtenido $y(x_1)$, la aditividad de las integrales orientadas devuelve la misma función:

$$y(x_1) + \int_{x_1}^{x} g(t)\,dt = y_0 + \int_{x_0}^{x} g(t)\,dt.$$

El punto base organiza la descripción de la solución; no cambia la trayectoria cuando el dato se transforma de manera compatible.

Ejemplo 20.2. **Ejemplo (una solución definida por una integral).** Consideremos el problema

$$y'(x) = e^{-x^2}, \qquad y(0) = 2, \qquad x \in \mathbb{R}.$$

La función $g(x) = e^{-x^2}$ es continua en toda $\mathbb{R}$, pues es la composición de un polinomio con la exponencial ya construida. Por Teorema 21.1, la única solución sobre toda la recta es

$$\boxed{y(x) = 2 + \int_{0}^{x} e^{-t^2}\,dt.}$$

La integral define un número para cada x real y una función diferenciable de x. No necesitamos encontrar una expresión adicional para tener una solución exacta. Este ejemplo no demuestra que el integrando carezca de primitiva elemental; simplemente muestra que la construcción y la verificación no dependen de disponer de tal expresión. Podemos obtener información sobre la solución sin evaluar la integral. Si $x > 0$, entonces $e^{-x^2} \le e^{-t^2} \le 1$ para $0 \le t \le x$. Por comparación, $2 + xe^{-x^2} \le y(x) \le 2 + x$. Si $x < 0$, en el segmento $[x, 0]$ tenemos las mismas cotas $e^{-x^2} \le e^{-t^2} \le 1$. Como la integral desde 0 hasta x es la negativa de la integral desde x hasta 0, obtenemos $2 + x \le y(x) \le 2 + xe^{-x^2}$.

En particular, $y(x) > 2$ cuando $x > 0$ y $y(x) < 2$ cuando $x < 0$, porque el integrando es positivo en cada segmento no degenerado. La fórmula proporciona así existencia, valor inicial y cotas, sin reducir el problema a una búsqueda de fórmulas conocidas.

Cuando la tasa es independiente del estado, esta sección resuelve el problema mediante una acumulación de una función ya dada. Para una ecuación como $y' = y$, escribir $y(x) = y_0 + \int_{x_0}^{x} y(t)\,dt$ no produce de inmediato una función conocida: la incógnita aparece también dentro de la integral. La siguiente sección utilizará la exponencial para construir y caracterizar las soluciones cuando la tasa es proporcional al estado.

20.3. Cuando la tasa es proporcional al estado

En $y' = a(x)y$, el coeficiente a es conocido, pero la tasa completa no lo es: depende del estado y que buscamos. Integrar el miembro derecho directamente dejaría la incógnita dentro de la integral. La exponencial ofrece una salida porque su derivada reproduce la propia función. Si construimos una primitiva A de a, la regla de la cadena da $(e^A)' = ae^A$ y sugiere buscar

soluciones que sean múltiplos de e^A. Podríamos intentar dividir la ecuación por y y reconocer una derivada logarítmica. Ese paso solo sería admisible donde y no se anula y dejaría fuera, desde el comienzo, la solución $y \equiv 0$. Para construir todas las soluciones utilizaremos otra transformación: multiplicar por una exponencial que nunca se anula.

TEOREMA 20.2. **Teorema (tasa proporcional con coeficiente continuo).** Sea $I \subseteq \mathbb{R}$ un intervalo abierto no vacío, sea $a\colon I \to \mathbb{R}$ continua y sean $x_0 \in I$ y $y_0 \in \mathbb{R}$. Definamos

$$A(x) = \int_{x_0}^{x} a(t)\,dt \qquad (x \in I).$$

Entonces el problema

$$y'(x) = a(x)y(x) \quad (x \in I), \qquad y(x_0) = y_0$$

tiene una única solución sobre I, dada por

$$\boxed{y(x) = y_0 e^{A(x)}.}$$

La solución pertenece a $C^1(I)$. Si $y_0 = 0$, es idénticamente nula; si $y_0 \neq 0$, no se anula en ningún punto de I y conserva el signo de y_0.

Idea de la prueba. La función e^A produce el término ay al derivar. Para demostrar que ninguna otra solución escapa a la fórmula, buscamos una cantidad constante: el producto ye^{-A}. La derivada del segundo factor cancela exactamente la contribución de y'.

Demostración. El teorema Teorema 21.1, aplicado a la tasa a y al dato cero, muestra que A está bien definida en I, que $A(x_0) = 0$ y que $A' = a$. Por la derivada de la exponencial (Teorema 17.3) y la regla de la cadena, la función propuesta satisface $y'(x) = y_0 e^{A(x)} A'(x) = a(x)y(x)$.

Además, $y(x_0) = y_0 e^0 = y_0$. Hemos construido y verificado una solución. Su derivada es continua porque es el producto de las funciones continuas a e y. Sea ahora z una solución cualquiera del mismo problema. Definamos $h(x) = z(x)e^{-A(x)}$. Por las reglas de producto y cadena,

$$h'(x) = e^{-A(x)}z'(x) - z(x)e^{-A(x)}a(x) = e^{-A(x)}\big(z'(x) - a(x)z(x)\big) = 0$$

para todo $x \in I$. Como en la sección inicial, aplicamos Proposición 9.2 en cada segmento cerrado con extremos x_0 y x contenido en I. Concluimos que $h(x) = h(x_0) = y_0$. Multiplicar por $e^{A(x)}$ da

$$z(x) = y_0 e^{A(x)} \qquad (x \in I).$$

Esto demuestra exhaustividad y unicidad. Finalmente, $e^{A(x)} > 0$ en todo I: el producto es nulo en todos los puntos si $y_0 = 0$, y tiene en todos ellos el signo de y_0 en caso contrario. □

La conservación del signo es una conclusión de la fórmula demostrada, no una hipótesis que hayamos impuesto para dividir por la solución. En particular, una solución que toma el valor cero en algún punto es idénticamente nula en todo su intervalo: basta usar ese punto como referencia en el teorema. Para esta ecuación, una trayectoria no nula no puede atravesar el estado cero.

También podemos describir todas las soluciones sin prescribir un dato: fijado $x_0 \in I$, cada solución tiene la forma $y(x) = Ce^{A(x)}$, con $C = y(x_0)$, y todo $C \in \mathbb{R}$ produce una solución. La continuidad de a permite construir A sobre cada segmento compacto; no se exige que a esté acotada en todo I.

Coeficiente constante: crecimiento, decaimiento y signo.

COROLARIO 20.1. **Corolario (tasa proporcional constante).** Sean $k \in \mathbb{R}$, $x_0 \in \mathbb{R}$ y $y_0 \in \mathbb{R}$. La única solución sobre toda $\mathbb{R}$ del problema

$$y' = ky, \qquad y(x_0) = y_0$$

es

$$\boxed{y(x) = y_0 e^{k(x-x_0)}.}$$

Su restricción a cualquier intervalo abierto que contenga x_0 es la única solución sobre ese intervalo con el mismo dato.

Demostración. Aplicamos Teorema 21.2 a $a(x) = k$. La integral de la constante es $A(x) = k(x - x_0)$, lo que proporciona la fórmula y ambas afirmaciones de unicidad. □

Si $y_0 > 0$, un coeficiente $k > 0$ produce una función estrictamente creciente, mientras que $k < 0$ produce una función estrictamente decreciente. Esto puede leerse directamente de la monotonía de la exponencial. Si $y_0 < 0$, las direcciones se invierten al multiplicar por un número negativo: para $k > 0$, el valor de y disminuye aunque su magnitud aumente. Por eso «crecimiento» requiere precisar si nos referimos al valor firmado o a $|y|$. Para $y_0 \neq 0$, la magnitud es $|y(x)| = |y_0|e^{k(x-x_0)}$: aumenta con x cuando $k > 0$ y disminuye cuando $k < 0$. Si $k = 0$, la solución es constante; si $y_0 = 0$, es la solución nula cualquiera sea k. Los casos cero pertenecen al mismo resultado y no requieren una fórmula separada obtenida por división.

EJEMPLO 20.3. **Ejemplo (coeficiente variable y cambio de régimen).** Consideremos

$$y' = 2xy, \qquad y(1) = -3, \qquad x \in \mathbb{R}.$$

El coeficiente $a(x) = 2x$ es continuo en toda la recta, y su acumulación desde el punto inicial es

$$A(x) = \int_1^x 2t\,dt = x^2 - 1.$$

El teorema da la única solución

$$\boxed{y(x) = -3e^{x^2-1}.}$$

La verificación independiente confirma el resultado: $y(1) = -3$ y $y'(x) = -6xe^{x^2-1} = 2xy(x)$ en todo $\mathbb{R}$. La solución es siempre negativa, pero no es monótona en toda la recta. Para $x < 0$, su derivada es positiva; para $x > 0$, es negativa. Crece hasta $x = 0$ y decrece después, alcanzando allí su máximo $y(0) = -3/e$. En cambio, su magnitud tiene allí un mínimo. La tasa proporcional variable puede cambiar de signo sin que la solución cambie de signo. En este ejemplo, $a(0) = 0$ hace que la tasa instantánea se anule en 0, pero no convierte a la solución en constante ni obliga a que su valor sea cero.

Multiplicar por e^{-A} produjo una función de derivada nula. Con un término conocido adicional buscaremos que la derivada del producto sea una función que podamos acumular; así podremos aplicar la reconstrucción de una tasa conocida.

20.4. Ecuaciones lineales y factor integrante

La tasa proporcional se puede acompañar de una entrada conocida. Al escribir la ecuación como $y' + py = q$, el término py recoge la dependencia lineal respecto del estado, mientras que q aporta una tasa que no depende de la incógnita. La transformación de la sección anterior sigue siendo útil: buscamos un multiplicador que convierta el miembro izquierdo en la derivada de un producto.

DEFINICIÓN 20.3. **Definición (ecuación lineal de primer orden).** Sean $p, q : I \to \mathbb{R}$ continuas en un intervalo abierto no vacío I. Llamamos ecuación lineal de primer orden a

$$y'(x) + p(x)y(x) = q(x), \qquad x \in I.$$

Si q es idénticamente nula, la ecuación es **homogénea**. La ecuación homogénea asociada a una ecuación dada conserva p y sustituye q por cero.

La forma normal exige que el coeficiente de y' sea uno. Si una ecuación se presenta como $r(x)y' + s(x)y = t(x)$, dividir por r solo produce esta forma en intervalos donde r no se anula y los cocientes correspondientes son continuos. Un cero de r no desaparece por simplificación algebraica: la ecuación original debe examinarse allí por separado.

TEOREMA 20.3. **Teorema (factor integrante y problema inicial lineal).** Sean p, q continuas en un intervalo abierto I, $x_0 \in I$ e $y_0 \in \mathbb{R}$. Definamos

$$P(x) = \int_{x_0}^{x} p(t)\,dt, \qquad \mu(x) = e^{P(x)}.$$

La única solución en I del problema

$$y' + py = q, \qquad y(x_0) = y_0$$

es

$$y(x) = e^{-P(x)} \left(y_0 + \int_{x_0}^{x} e^{P(t)} q(t)\,dt \right).$$

Esta solución es de clase C^1. La función positiva μ se llama **factor integrante**.

Demostración. Por el teorema de reconstrucción [Teorema 21.1], P es de clase C^1 y $P' = p$. La derivación de la exponencial y la regla de la cadena dan

$$\mu' = p\mu, \qquad \mu > 0, \qquad \mu(x_0) = 1.$$

Por tanto, para cualquier función diferenciable y,

$$(\mu y)' = \mu y' + \mu' y = \mu(y' + py).$$

La ecuación original equivale a $(\mu y)' = \mu q$: la implicación inversa es válida porque μ nunca es cero.

El producto μq es continuo. Aplicando nuevamente [Teorema 21.1], la única función w que satisface

$$w' = \mu q, \qquad w(x_0) = y_0$$

es

$$w(x) = y_0 + \int_{x_0}^{x} \mu(t) q(t)\,dt.$$

Definamos $y = w/\mu = e^{-P} w$. Es de clase C^1, toma el valor y_0 en x_0 y cumple $(\mu y)' = \mu q$; la equivalencia anterior prueba que es una solución.

Si z es cualquier otra solución con el mismo dato inicial, μz satisface el mismo problema de reconstrucción que w. Su unicidad implica $\mu z = w$, y la positividad de μ implica $z = y$. Todas

las integrales utilizadas recorren segmentos compactos contenidos en I, también cuando $x < x_0$; no requieren una integral impropia sobre los extremos de I.

El punto base normaliza el factor integrante, pero no es una condición adicional de la ecuación. Multiplicar μ por una constante no nula produce la misma transformación. La normalización $\mu(x_0) = 1$ hace que el dato inicial entre directamente en la fórmula. Si $q = 0$, obtenemos $y = y_0 e^{-P}$, exactamente la solución de $y' = ay$ con $a = -p$. La ecuación lineal amplía, por tanto, el resultado proporcional sin abandonar su mecanismo de prueba.

Corolario 20.2. **Corolario (familia a partir de una solución particular).** Bajo las hipótesis de [Teorema 21.3], sea y_p una solución de $y' + py = q$ en I. Todas las soluciones de esa ecuación en I son

$$y(x) = y_p(x) + Ce^{-P(x)}, \qquad C \in \mathbb{R},$$

donde $P(x) = \int_{x_0}^{x} p(t)\,dt$.

Demostración. Si y e y_p satisfacen la misma ecuación, su diferencia $h = y - y_p$ cumple $h' + ph = 0$. Por [Teorema 21.2], sus soluciones son $h = Ce^{-P}$. Recíprocamente, sumar cualquiera de estas soluciones homogéneas a y_p da

$$(y_p + h)' + p(y_p + h) = q + 0 = q.$$

La suma de dos soluciones particulares satisface una ecuación con miembro derecho $2q$; no tiene por qué satisfacer la ecuación inicial. Lo que podemos sumar libremente a una solución particular es una solución de la ecuación homogénea asociada.

Ejemplo 20.4. **Ejemplo (entrada exponencial).** Resolvamos en $\mathbb{R}$

$$y' + 2y = e^{x}, \qquad y(0) = 1.$$

Aquí $P(x) = 2x$ y $\mu(x) = e^{2x}$. La fórmula del factor integrante da

$$y(x) = e^{-2x}\left(1 + \int_{0}^{x} e^{3t}\,dt\right) = \frac{1}{3}e^{x} + \frac{2}{3}e^{-2x}.$$

La comprobación directa es

$$y' = \frac{1}{3}e^{x} - \frac{4}{3}e^{-2x}, \qquad y' + 2y = e^{x}, \qquad y(0) = 1.$$

La solución particular $y_p = e^{x}/3$ permite escribir toda la familia como $y = \frac{1}{3}e^{x} + Ce^{-2x}$. El dato inicial fija $C = 2/3$. El término homogéneo decrece cuando x aumenta, pero la entrada produce un término creciente: el signo del coeficiente p por sí solo no determina la evolución de una solución no homogénea.

El factor integrante transforma la ecuación en una tasa conocida. Al separar variables utilizaremos una función del estado para obtener una transformación semejante; su inversión exigirá controlar el dominio y la imagen.

20.5. Separación de variables y rama inversa

La expresión «separar variables» suele resumirse mediante una igualdad entre diferenciales. Para justificarla construiremos una función cuya derivada compense la dependencia respecto de y. El paso decisivo será aplicar la regla de la cadena, y la recuperación del estado se hará mediante una inversa en un intervalo declarado.

DEFINICIÓN 20.4. **Definición (ecuación separable).** Sean I, J intervalos abiertos no vacíos, $a : I \to \mathbb{R}$ y $b : J \to \mathbb{R}$ continuas. Una ecuación de la forma

$$y' = a(x)b(y)$$

se llama **separable**. Una solución considerada en este dominio tiene su intervalo de definición contenido en I y todos sus valores en J.

Esta definición admite ceros de b. El procedimiento siguiente, en cambio, exige trabajar en un intervalo de estados donde b no se anule. Más adelante examinaremos las soluciones que esa restricción puede excluir.

TEOREMA 20.4. **Teorema (separación en un intervalo sin ceros).** Sean I, J intervalos abiertos no vacíos, $a : I \to \mathbb{R}$ y $b : J \to \mathbb{R}$ continuas, y supongamos que $b(v) \neq 0$ para todo $v \in J$. Fijemos $x_0 \in I$ e $y_0 \in J$ y definamos

$$A(x) = \int_{x_0}^{x} a(t)\,dt, \qquad H(v) = \int_{y_0}^{v} \frac{ds}{b(s)}.$$

Entonces H es estrictamente monótona, $H(J)$ es un intervalo abierto y $H^{-1} : H(J) \to J$ existe. Sea

$$U = \{x \in I : A(x) \in H(J)\},$$

y sea K el mayor intervalo abierto que contiene x_0 y está contenido en U. La función

$$y(x) = H^{-1}(A(x)), \qquad x \in K,$$

es una solución de clase C^1 de

$$y' = a(x)b(y), \qquad y(x_0) = y_0.$$

Toda solución de este problema inicial cuyos valores permanezcan en J es una restricción de y. En particular, K es el mayor intervalo de solución para ese dato inicial **con la restricción de valores en** J.

Demostración. Como b es continua y nunca se anula en el intervalo J, el teorema del valor intermedio impide que cambie de signo. La función $1/b$ es continua, y [Teorema 21.1] da

$$H'(v) = \frac{1}{b(v)}.$$

Esta derivada tiene un signo fijo y nunca es cero. Para $v_1 < v_2$ en J, el teorema del valor medio aplicado sobre $[v_1, v_2]$ muestra que $H(v_2) - H(v_1)$ tiene ese mismo signo. Así, H es estrictamente monótona. El resultado sobre inversas continuas en intervalos [Teorema 5.7] asegura que su imagen es un intervalo y que su inversa es continua.

Además, la imagen es abierta. Dado $v \in J$, podemos elegir $v_- < v < v_+$ en J. Por monotonía estricta, $H(v)$ queda estrictamente entre $H(v_-)$ y $H(v_+)$. Por continuidad, todos los valores entre estos dos últimos pertenecen a $H(J)$. Por tanto, cada punto de $H(J)$ tiene un entorno contenido en esa imagen.

La función A es de clase C^1, con $A' = a$. Cada punto $x \in U$ tiene un entorno contenido en I que permanece en U: elegimos un entorno de $A(x)$ contenido en $H(J)$ y usamos la continuidad de A en x. Así, U es abierto. Además, $A(x_0) = H(y_0) = 0$, de modo que $x_0 \in U$. Existe, pues, algún intervalo abierto que contiene x_0 y está contenido en U. La unión de todos esos intervalos es abierta y es un intervalo: todos contienen el mismo punto y, entre dos puntos de la unión, los intervalos que los contienen cubren también los puntos intermedios. Esa unión es el intervalo K del enunciado y contiene cualquier otro intervalo de la misma clase.

El teorema de derivación de la inversa [Teorema 8.1] se aplica a H, pues H' no se anula. En cada $u \in H(J)$,

$$(H^{-1})'(u) = \frac{1}{H'(H^{-1}(u))} = b(H^{-1}(u)).$$

La regla de la cadena da, para $x \in K$,

$$y'(x) = b(H^{-1}(A(x)))A'(x) = b(y(x))a(x).$$

Esta derivada es continua. Asimismo, $y(x_0) = H^{-1}(0) = y_0$.

Finalmente, sea $z : L \to J$ otra solución del mismo problema, donde L es un intervalo abierto contenido en I y contiene x_0. La regla de la cadena implica

$$(H \circ z)'(x) = \frac{z'(x)}{b(z(x))} = a(x).$$

Como $H(z(x_0)) = 0$, la unicidad de reconstrucción [Teorema 21.1] aplicada en L muestra que $H(z(x)) = A(x)$. En particular, $A(x) \in H(J)$ para todo $x \in L$, por lo que $L \subseteq U$. La definición de K implica $L \subseteq K$, y la inyectividad de H da $z = H^{-1} \circ A$ en L.

La restricción de estados forma parte del resultado. Si elegimos un intervalo J menor del necesario, el extremo de K puede reflejar esa elección, en vez de un impedimento para continuar la solución de la ecuación original. Tampoco debemos tomar todo U como intervalo de solución: puede tener partes separadas. El dato inicial selecciona el intervalo que lo contiene.

La escritura habitual

$$\int \frac{dy}{b(y)} = \int a(x)\,dx$$

abrevia la identidad $(H \circ y)' = a$. La justificación es la regla de la cadena, no una cancelación algebraica de símbolos en dy/dx. Las integrales con punto base incorporan el dato inicial y evitan dejar indeterminada la constante.

Ejemplo 20.5. **Ejemplo (crecimiento cuadrático y dominio de la inversa).** Consideremos

$$y' = y^2, \qquad y(0) = c, \qquad c \neq 0.$$

Tomamos $I = \mathbb{R}$ y elegimos $J = (0, \infty)$ si $c > 0$, o $J = (-\infty, 0)$ si $c < 0$. En ambos casos,

$$H(v) = \int_c^v \frac{ds}{s^2} = \frac{1}{c} - \frac{1}{v}, \qquad A(x) = x.$$

De $H(y) = x$ resulta $y(x) = \frac{c}{1-cx}$. Si $c > 0$, la imagen $H(J)$ es $(-\infty, 1/c)$ y $K = (-\infty, 1/c)$. Si $c < 0$, la imagen es $(1/c, \infty)$ y $K = (1/c, \infty)$. En cada caso, K contiene 0, el denominador es positivo en K y la solución conserva el signo de c. La verificación directa da

$$y'(x) = \frac{c^2}{(1-cx)^2} = y(x)^2, \qquad y(0) = c.$$

La fórmula también tiene valores al otro lado de $1/c$, pero ese segundo intervalo no contiene el dato inicial y no puede unirse a K como intervalo de solución: la función no está definida en $1/c$. Al aproximarnos a ese extremo desde K, su magnitud crece sin límite.

El caso $c = 0$ no está cubierto por la elección de J, aunque $y \equiv 0$ satisface la ecuación. Esta solución se reconoce directamente y debe conservarse al estudiar el problema completo.

La comparación de Figura 21.2 muestra cómo la imagen de H restringe los valores admisibles de $A(x) = x$. Las variables de los paneles desempeñan trabajos distintos: v es el estado que se transforma y x es el argumento de la solución.

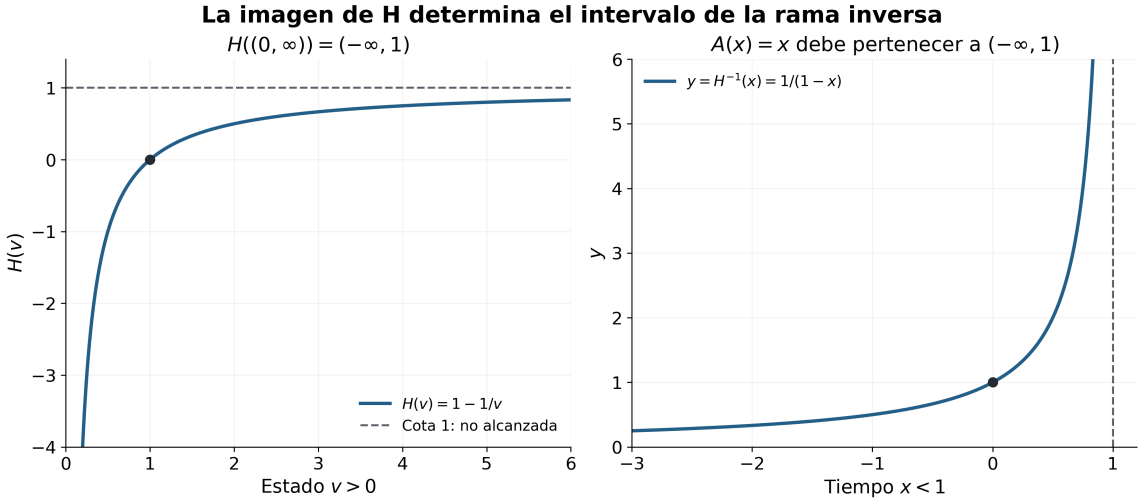


FIGURA 20.2. Para $y' = y^2$ con $y(0) = 1$, $H(v) = 1 - 1/v$ lleva los estados positivos a $(-\infty, 1)$; su inversa construye la solución únicamente para $x < 1$.

20.6. Equilibrios y límites de la unicidad

Dividir por $b(y)$ supone que la trayectoria evita sus ceros. Antes de hacerlo, conviene comprobar las soluciones constantes. También hay que distinguir dos afirmaciones: que un valor constante resuelve la ecuación y que toda solución con ese valor inicial debe permanecer constante. La segunda exige una demostración de unicidad que la primera no proporciona.

DEFINICIÓN 20.5. **Definición (solución de equilibrio).** Una **solución de equilibrio** de una ecuación diferencial es una solución constante $y(x) = c$ sobre su intervalo de definición. El número c se llama **valor de equilibrio** para ese intervalo.

PROPOSICIÓN 20.1. **Proposición (equilibrios de una ecuación separable).** Sean $a : I \to \mathbb{R}$ y $b : J \to \mathbb{R}$ continuas en intervalos abiertos no vacíos. Para $c \in J$, la función constante $y = c$ resuelve $y' = a(x)b(y)$ en I si y solo si

$$a(x)b(c) = 0 \quad \text{para todo } x \in I.$$

En particular:

- si $b(c) = 0$, la función constante $y = c$ es una solución;
- si a no es idénticamente nula en I, los valores de equilibrio son exactamente los ceros de b en J;
- si a es idénticamente nula en I, cada $c \in J$ da una solución constante.

Demostración. La derivada de $y = c$ es cero, por lo que la condición es exactamente la sustitución en la ecuación. Si $b(c) = 0$, el producto se anula para todo x. Si existe $x_1 \in I$ con $a(x_1) \neq 0$, la igualdad en ese punto obliga a $b(c) = 0$. Cuando $a = 0$ en todo I, la condición se cumple para cualquier $c \in J$.

En una ecuación autónoma $y' = b(y)$, es decir, una ecuación cuya tasa depende solo del estado, tenemos $a = 1$. Sus valores de equilibrio son, por tanto, los ceros de b. El adjetivo «equilibrio» describe una solución constante disponible; no asegura que sea la única trayectoria que pase por ese estado.

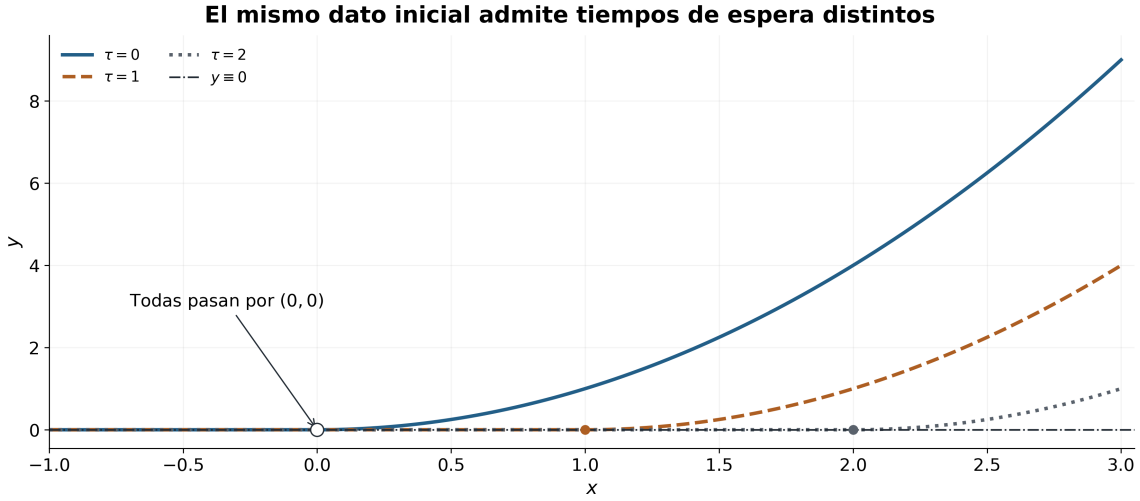


FIGURA 20.3. La ecuación $y' = 2\sqrt{|y|}$ admite soluciones que esperan en cero y parten en tiempos diferentes, todas con $y(0) = 0$. Se representan $\tau = 0, 1, 2$ y la solución nula.

EJEMPLO 20.6. **Ejemplo (espera en un equilibrio y falta de unicidad).** Consideremos en $\mathbb{R}$ el problema

$$y' = 2\sqrt{|y|}, \qquad y(0) = 0.$$

El miembro derecho es continuo para todo estado real, y $y \equiv 0$ es una solución. Para cada $\tau \geq 0$, definamos además

$$y_\tau(x) = \begin{cases} 0, & x \leq \tau, \\ (x-\tau)^2, & x > \tau. \end{cases}$$

Estas funciones permanecen en el equilibrio hasta el instante τ y después toman valores positivos.

Comprobemos el punto de unión. Como $y_\tau(\tau) = 0$, el cociente incremental es

$$\frac{y_\tau(\tau+h) - y_\tau(\tau)}{h} = \begin{cases} 0, & h < 0, \\ h, & h > 0. \end{cases}$$

Ambos límites son cero. Por tanto, $y_\tau'(\tau) = 0$, y en toda la recta

$$y_\tau'(x) = \begin{cases} 0, & x \leq \tau, \\ 2(x-\tau), & x > \tau. \end{cases}$$

La derivada es continua. Para $x \leq \tau$, ambos miembros de la ecuación valen cero. Para $x > \tau$, se cumple

$$2\sqrt{|y_\tau(x)|} = 2\sqrt{(x-\tau)^2} = 2(x-\tau) = y_\tau'(x).$$

Como $\tau \geq 0$, también $y_\tau(0) = 0$. Todas estas funciones resuelven el mismo problema inicial en el mismo intervalo $\mathbb{R}$.

Por ejemplo, $y_0(1/2) = 1/4$, mientras que $y_1(1/2) = 0$. Son soluciones distintas con el mismo dato inicial. La familia expuesta basta para probar falta de unicidad; no se afirma que enumere todas las soluciones.

Las trayectorias de Figura 21.3 comparten el dato inicial y se separan después. La tangente horizontal en cada punto de partida corresponde a la derivada bilateral cero que acabamos de comprobar. La continuidad del miembro derecho no basta para asegurar unicidad de un problema

inicial. Los resultados anteriores sí la probaron para sus clases concretas: reconstrucción de tasa conocida, tasa proporcional, ecuación lineal y separación en un intervalo donde b no se anula. El ejemplo no contradice ninguno de ellos: su dato inicial está precisamente en un cero de $b(y) = 2\sqrt{|y|}$.

Si dividimos esta ecuación por $2\sqrt{|y|}$, excluimos los tramos donde $y = 0$. La separación puede describir una parte positiva de la trayectoria, pero no determina cuánto tiempo permanece en el equilibrio antes de entrar en esa parte. La comprobación en el punto de unión fue indispensable: unir expresiones por tramos solo construye una solución cuando el resultado es diferenciable y satisface la ecuación también en la unión.

Así, antes de separar una ecuación debemos registrar los ceros de b, comprobar las soluciones constantes y declarar el intervalo de estados utilizado. Después debemos verificar la fórmula y su intervalo de definición. Estas precauciones serán parte del análisis de los modelos elementales de la siguiente sección: una expresión explícita tiene valor matemático cuando se acompaña de las condiciones bajo las cuales describe una solución.

20.7. Del balance a la ecuación: enfriamiento y mezcla

Una ecuación diferencial no describe un fenómeno por el solo hecho de tener una fórmula resoluble. Primero debemos elegir el estado, precisar las hipótesis que permiten expresar su tasa y declarar el intervalo donde el modelo tiene sentido. La solución matemática permitirá después examinar las consecuencias de esas hipótesis.

Temperatura respecto de un ambiente constante.

Ejemplo 20.7. **Ejemplo (modelo de enfriamiento).** Sea $T(t)$ la temperatura de un cuerpo y sea T_a la temperatura constante del ambiente. Suponemos que el cuerpo se puede describir por una sola temperatura y que la tasa de cambio es proporcional a la diferencia respecto del ambiente, con signo opuesto:

$$T' = -k(T - T_a), \qquad k > 0, \qquad T(0) = T_0.$$

El coeficiente k tiene unidades de tiempo inverso. Tanto T' como $k(T - T_a)$ tienen unidades de temperatura por unidad de tiempo. La diferencia de temperaturas evita que la elección del cero de la escala altere la ley.

Al introducir $\theta = T - T_a$, usamos que T_a es constante y obtenemos $\theta' = -k\theta$. Por [Corolario 21.1], la única solución es

$$T(t) = T_a + (T_0 - T_a)e^{-kt}, \qquad t \in \mathbb{R}.$$

Su derivada es $-k(T_0 - T_a)e^{-kt}$, que coincide con $-k(T - T_a)$, y el dato inicial se satisface. Si $T_0 > T_a$, la temperatura decrece y permanece por encima del ambiente; si $T_0 < T_a$, aumenta y permanece por debajo. Si $T_0 = T_a$, es constante.

La fórmula prueba que $T(t) \to T_a$ cuando $t \to +\infty$. Para un dato distinto del ambiente, la diferencia no se anula en ningún tiempo finito. Si queremos reducir su magnitud a una fracción $0 < \varepsilon < 1$ de la inicial, resolvemos

$$e^{-kt} = \varepsilon, \qquad t = \frac{\ln(1/\varepsilon)}{k}.$$

En particular, el tiempo de reducción a la mitad es $(\ln 2)/k$, independiente de la magnitud inicial de la diferencia.

La solución matemática está definida en toda la recta. Una experiencia iniciada en $t = 0$ utiliza su parte futura, y la validez física requiere que el ambiente y el coeficiente permanezcan

aproximadamente constantes durante esa experiencia. Si T_a depende del tiempo, la sustitución da $\theta' = T' - T_a'$; la ecuación proporcional anterior ya no se obtiene sin ese término adicional.

El signo del estado relativo θ tiene una interpretación física, pero su conservación procede de la fórmula y del resultado de unicidad ya probado. No hemos supuesto una prohibición universal de atravesar equilibrios.

Cantidad y concentración en un tanque.

EJEMPLO 20.8. **Ejemplo (mezcla con volumen variable).** Un tanque contiene inicialmente un volumen $V_0 > 0$ de líquido y una cantidad de soluto $m_0 \geq 0$. Entra líquido con caudal constante $u > 0$ y concentración constante $c_{\text{in}} \geq 0$; sale mezcla con caudal constante $v > 0$. Suponemos mezcla instantánea y uniforme, y ausencia de producción o pérdida química de soluto.

El balance de volumen es $V(t) = V_0 + (u - v)t$. Mientras $V(t) > 0$, la concentración interna es $m(t)/V(t)$. La entrada de soluto por unidad de tiempo es uc_{in} y la salida es $vm(t)/V(t)$. El balance de cantidad da

$$m' + \frac{v}{V(t)} m = uc_{\text{in}}, \qquad m(0) = m_0.$$

Aquí m tiene unidades de cantidad de soluto, V de volumen y u, v de volumen por tiempo. Cada término de la ecuación tiene unidades de cantidad por tiempo. La tasa de salida usa la concentración interna, no la concentración entrante.

Sea $\delta = u - v$. El mayor intervalo abierto que contiene 0 donde el volumen es positivo y los coeficientes están definidos es

$$I = \begin{cases} (-V_0/\delta, \infty), & \delta > 0, \\ \mathbb{R}, & \delta = 0, \\ (-\infty, V_0/(v-u)), & \delta < 0. \end{cases}$$

En él V es positivo y la ecuación es lineal con coeficientes continuos, de modo que [Teorema 21.3] acredita una única solución.

La función $m_p = c_{\text{in}} V$ es una solución particular: su derivada es $c_{\text{in}}(u - v)$ y el término de salida añade vc_{in}, produciendo uc_{in}. Si $z = m - c_{\text{in}} V$, entonces

$$z' = -\frac{v}{V} z, \qquad z(0) = m_0 - c_{\text{in}} V_0.$$

Para $\delta \neq 0$,

$$\int_0^t \frac{v}{V(s)}\, ds = \frac{v}{\delta} \ln \frac{V(t)}{V_0}.$$

Así,

$$m(t) = c_{\text{in}} V(t) + (m_0 - c_{\text{in}} V_0) \left(\frac{V(t)}{V_0} \right)^{-v/\delta}.$$

Las potencias reales se toman sobre la base positiva $V(t)/V_0$, mediante exponencial y logaritmo. Para $\delta = 0$, la fórmula correspondiente es

$$m(t) = c_{\text{in}} V_0 + (m_0 - c_{\text{in}} V_0) e^{-vt/V_0}.$$

Ambas fórmulas se verifican mediante la ecuación homogénea para z y la solución particular ya comprobada.

Escribiendo $c_0 = m_0/V_0$, la concentración resulta

$$c(t) = \frac{m(t)}{V(t)} = \begin{cases} c_{\text{in}} + (c_0 - c_{\text{in}}) \left(\dfrac{V(t)}{V_0} \right)^{-u/\delta}, & \delta \neq 0, \\ c_{\text{in}} + (c_0 - c_{\text{in}}) e^{-ut/V_0}, & \delta = 0. \end{cases}$$

En el caso de volumen constante, $u = v$. Para $t \geq 0$ dentro del dominio, el factor que multiplica $c_0 - c_{\text{in}}$ pertenece a $(0, 1]$: también puede escribirse como

$$\exp\left(-\int_0^t \frac{u}{V(s)}\, ds\right).$$

Por ello, la concentración permanece entre c_0 y c_{in}, y es no negativa. La cantidad $m = cV$ también es no negativa.

Si $v > u$, el tanque se vacía en el tiempo $t_* = V_0/(v - u)$. La fórmula de concentración tiene límite c_{in} cuando $t \to t_*^-$, mientras que $m(t) \to 0$. Estos límites no definen concentración de un tanque vacío: el cociente m/V y la ecuación dejan de estar definidos en t_*. Si $u > v$, el modelo ideal permite volumen creciente indefinidamente, pero un tanque de capacidad finita introduce un tiempo de desborde que limita su uso físico. Tal capacidad no fue incluida en la ecuación.

Los modelos muestran dos dominios distintos que debemos declarar: el de la ecuación matemática y el de uso de las hipótesis físicas. La restricción temporal $t \geq 0$ selecciona la parte de una solución que interesa a la experiencia; las pruebas de existencia y unicidad anteriores se aplican a la solución sobre su intervalo abierto.

20.8. Restricción, prolongación e intervalo maximal

Toda solución puede restringirse a un intervalo abierto menor. El problema inverso consiste en saber si puede prolongarse conservando sus valores y satisfaciendo la misma ecuación. Una fórmula puede dejar de estar escrita en un punto sin que la solución tenga allí una singularidad; también puede presentar una divergencia que impide cualquier prolongación.

Definición 20.6. **Definición (restricción, prolongación y solución maximal).** Fijemos una ecuación $y' = F(x, y)$ con dominio $D \subseteq \mathbb{R}^2$ y una solución $y : I \to \mathbb{R}$.

- Si $L \subseteq I$ es un intervalo abierto no vacío, la función $y|_L$ es una **restricción** de la solución.
- Una **prolongación** de y es una solución $\widetilde{y} : \widetilde{I} \to \mathbb{R}$ de la misma ecuación, con $I \subsetneq \widetilde{I}$ y $\widetilde{y}|_I = y$, donde $\widetilde{I}$ es un intervalo abierto.
- La solución es **maximal** si no admite tal prolongación.

Si se conserva un problema inicial, los intervalos considerados deben contener su punto base. «Maximal» no significa «definida en toda $\mathbb{R}$».

La restricción conserva diferenciabilidad, grafo admisible y ecuación porque esas condiciones se comprueban punto por punto. Una prolongación exige además admisibilidad en cada punto nuevo. No se afirma aquí que toda ecuación tenga soluciones ni que toda solución posea una prolongación maximal.

Ejemplo 20.9. **Ejemplo (unicidad específica y explosión en tiempo finito).** Volvamos a

$$y' = y^2, \qquad y(0) = c.$$

Primero podemos probar unicidad sin excluir el cero. Si y_1, y_2 son soluciones sobre un intervalo común L y coinciden en algún $s \in L$, su diferencia $d = y_1 - y_2$ satisface

$$d' = y_1^2 - y_2^2 = (y_1 + y_2)d, \qquad d(s) = 0.$$

El coeficiente $y_1 + y_2$ es continuo, pues ambas soluciones son diferenciables. Por [Teorema 21.2], $d = 0$ en L. Para soluciones inicialmente definidas en intervalos distintos que contienen s, aplicamos el argumento en su intersección, que es un intervalo abierto que contiene s.

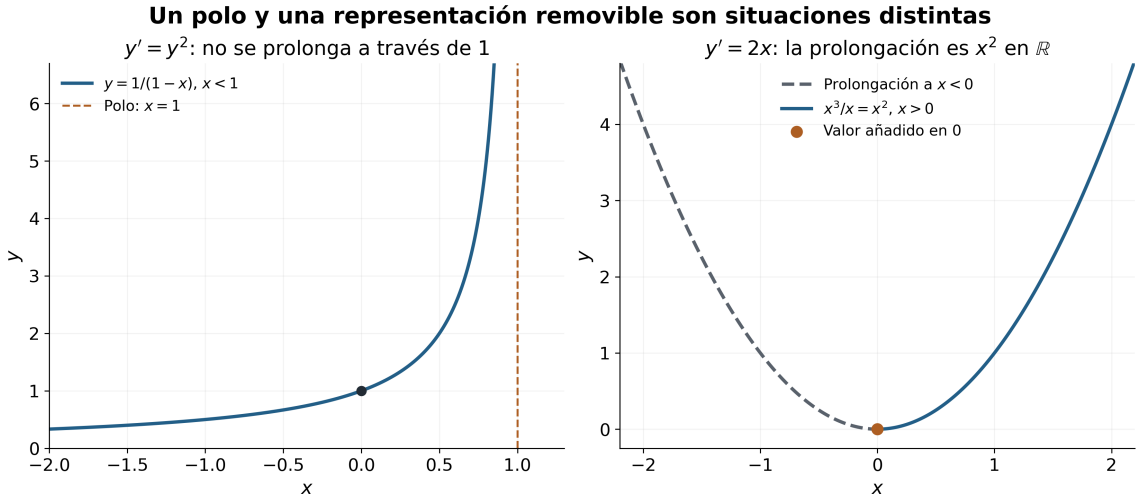


FIGURA 20.4. Un polo impide prolongar la solución $y = 1/(1-x)$ de $y' = y^2$ a través de 1; la omisión de 0 en x^3/x, para otra ecuación, se repara mediante la solución x^2 de $y' = 2x$.

Si $c = 0$, la solución global $y = 0$ existe y el argumento demuestra que cualquier otra solución con ese dato es una restricción de ella. Si $c \neq 0$, [Ejemplo 21.5] construyó $y(x) = \frac{c}{1-cx}$ sobre el intervalo

$$K = \begin{cases} (-\infty, 1/c), & c > 0, \\ (1/c, \infty), & c < 0. \end{cases}$$

Probemos que esta solución es maximal para la ecuación, sin imponer ya una restricción de signo sobre sus posibles prolongaciones.

Sea $p = 1/c$. Para $c > 0$, $y(x) \to +\infty$ cuando $x \to p^-$; para $c < 0$, $y(x) \to -\infty$ cuando $x \to p^+$. Cualquier intervalo abierto estrictamente mayor que K debe contener p. Una prolongación tendría que ser diferenciable, y por tanto continua, en p, con un valor real finito. La divergencia de los valores heredados de y lo impide. Así, K es maximal.

Además, cualquier solución con el dato $(0, c)$ coincide con esta fórmula en la intersección de sus intervalos, por la unicidad ya probada. Si su intervalo atravesara p, esa coincidencia produciría la misma contradicción de continuidad. Por tanto, su intervalo está contenido en K y es una restricción de la solución construida. La expresión racional al otro lado de p resuelve también la ecuación en ese otro intervalo. Es otra solución, que no contiene el punto inicial 0. Dos intervalos separados por un polo no forman un único intervalo de solución.

Una representación defectuosa no es una explosión. Consideremos ahora la ecuación $y' = 2x$, definida para todo $(x, y) \in \mathbb{R}^2$. En $(0, \infty)$, la expresión $y(x) = \frac{x^3}{x}$ es igual a x^2 y resuelve la ecuación. El cociente escrito no está definido en 0, pero la función $\widetilde{y}(x) = x^2$ sí está definida y es diferenciable en toda $\mathbb{R}$, coincide con y en $(0, \infty)$ y satisface $\widetilde{y}' = 2x$. Hemos construido una prolongación de la solución. La omisión del punto era un defecto de la representación, no una singularidad de la trayectoria.

En Figura 21.4 se comparan estas dos obstrucciones aparentes. Solo el segundo panel permite añadir un valor y continuar la misma solución; en ambos casos la decisión exige comprobar la ecuación y su dominio. El dominio de la ecuación sigue siendo decisivo. Si la misma ley $F(x, y) = 2x$ se declara únicamente para $x > 0$, la función x^2 no es una solución de esa ecuación restringida en un intervalo que incluya 0: su grafo saldría del dominio declarado. Tener un

límite finito o una fórmula simplificable no basta para prolongar una solución; también hay que comprobar diferenciabilidad, ecuación y admisibilidad.

Esta distinción se aplica al tanque de la sección anterior. Que la cantidad tenga límite cero al vaciarse no elimina el coeficiente $v/V(t)$ ni aporta por sí solo un modelo válido después del vaciado.

20.9. Signo de la tasa y lectura cualitativa: la logística

Una solución diferenciable permite leer crecimiento y decrecimiento a partir del signo de su derivada. Para una ecuación autónoma $y' = b(y)$, el signo de b informa sobre la dirección del cambio mientras el estado permanece en una región dada. Para concluir que una trayectoria nunca abandona esa región, necesitamos además justificar que no atraviesa su frontera. Los equilibrios y una prueba específica de unicidad pueden proporcionar esa justificación.

Proposición 20.2. **Proposición (unicidad y separación por equilibrios en la logística).** Fijemos $r > 0$ y consideremos

$$y' = ry(1 - y).$$

Dos soluciones que coinciden en un punto coinciden en todo su intervalo común que contiene ese punto. Los valores de equilibrio son 0 y 1. Una solución que toma alguno de esos valores en un punto es constante en todo su intervalo de definición. En consecuencia, una solución no constante permanece enteramente en una de las regiones

$$(-\infty, 0), \qquad (0, 1), \qquad (1, \infty).$$

Es estrictamente decreciente en la primera y en la tercera, y estrictamente creciente en la segunda.

Demostración. Si y_1, y_2 son soluciones y $d = y_1 - y_2$, entonces

$$d' = r(y_1 - y_1^2 - y_2 + y_2^2) = r(1 - y_1 - y_2)d.$$

El coeficiente es continuo. Si $d(s) = 0$ en un punto común, [Teorema 21.2] implica $d = 0$ en el intervalo común que contiene s.

Por [Proposición 21.1], los ceros de $rv(1-v)$, que son 0 y 1, dan exactamente los equilibrios. Si una solución alcanza uno de ellos en s, la comparamos con la solución constante correspondiente, tomando s como punto base. La unicidad anterior implica que ambas coinciden en todo el intervalo de la solución.

Una solución no constante no puede, pues, tomar los valores 0 o 1. Como es continua en un intervalo, el teorema del valor intermedio impide que pase entre las tres regiones sin tomar uno de esos valores. Dentro de $(0, 1)$, su derivada es positiva; fuera de $[0, 1]$, es negativa. Para dos tiempos cualesquiera de su intervalo, el teorema del valor medio sobre el segmento que los une convierte ese signo estricto en la monotonía indicada.

La afirmación de no cruce se obtuvo por una ecuación lineal para la diferencia. El ejemplo con tiempo de espera de [Ejemplo 21.6] muestra por qué no podemos convertirla en una regla para toda ecuación autónoma con segundo miembro continuo.

Ejemplo 20.10. **Ejemplo (fórmula logística y comportamiento futuro).** Sea

$$y' = ry(1 - y), \qquad r > 0, \qquad y(0) = c.$$

Para $0 < c < 1$, podemos separar en $J = (0, 1)$. La identidad $\frac{1}{v(1-v)} = \frac{1}{v} + \frac{1}{1-v}$ da una primitiva $\ln v - \ln(1 - v)$. Por tanto, $\ln \frac{y}{1-y} - \ln \frac{c}{1-c} = rt$, y al despejar obtenemos

$$y(t) = \frac{ce^{rt}}{1 - c + ce^{rt}} = \frac{1}{1 + \dfrac{1 - c}{c} e^{-rt}}.$$

El denominador es positivo para todo $t \in \mathbb{R}$, y la fórmula toma valores en $(0, 1)$. Su derivada, escribiendo $E = e^{rt}$ y $D = 1 - c + cE$, es

$$y' = \frac{rc(1-c)E}{D^2} = r\,\frac{cE}{D}\left(1 - \frac{cE}{D}\right).$$

Además, $y(0) = c$. Esto demuestra existencia en toda la recta, y [Proposición 21.2] demuestra unicidad. La solución crece estrictamente y

$$\lim_{t\to+\infty} y(t) = 1, \qquad \lim_{t\to-\infty} y(t) = 0.$$

No alcanza ninguno de esos equilibrios en un tiempo finito.

Los datos $c = 0$ y $c = 1$ dan las soluciones constantes, únicas por la misma proposición. Para $c > 1$, la fórmula racional anterior sigue siendo una solución donde su denominador no se anula; la comprobación por derivación sigue siendo válida. El intervalo que contiene 0 es

$$K = (t_*, \infty), \qquad t_* = \frac{1}{r}\ln\frac{c-1}{c} < 0.$$

En K, el valor es mayor que 1. Decrece y tiende a 1 cuando $t \to +\infty$. Hacia el extremo pasado t_*^+, tiende a $+\infty$. La continuidad impide prolongarla a través de ese extremo, como en [Ejemplo 21.9]. Así, un dato por encima del equilibrio superior también posee una solución en toda la semirrecta futura, aunque no en toda la recta.

Para $c < 0$, el cero del denominador ocurre en $t_* = \frac{1}{r}\ln\frac{c-1}{c} > 0$, y el intervalo que contiene 0 es $K = (-\infty, t_*)$. La solución es negativa, decrece y tiende a $-\infty$ cuando $t \to t_*^-$. Esta rama es matemáticamente admisible, pero no representa una población no negativa. También muestra que la ecuación logística, considerada para todos los estados reales, no asegura existencia para todo tiempo futuro con cualquier dato.

Si y representa la fracción de una capacidad de referencia, r tiene unidades de tiempo inverso y y es adimensional. Para una población N y una capacidad $M > 0$, la sustitución $y = N/M$ transforma

$$N' = rN\left(1 - \frac{N}{M}\right)$$

en la ecuación estudiada. Para $N(0) > 0$, la solución existe para todo tiempo futuro y tiende a M; si el dato es menor que M, crece, y si es mayor, decrece. Estas conclusiones pertenecen al modelo, que supone r y M constantes.

La figura Figura 21.5 reúne el signo de la tasa y las trayectorias positivas en tiempos futuros. La lectura gráfica acompaña la prueba de no cruce y la verificación de existencia; no las reemplaza. El signo de la tasa permite anticipar la dirección de las trayectorias, pero un límite cuando el tiempo tiende a infinito requiere que la solución exista durante todo ese recorrido. Aquí la fórmula comprobada establece esa existencia para los datos no negativos; después podemos interpretar su comportamiento. La monotonía de una solución definida solo en un intervalo corto no justificaría por sí sola un límite futuro infinito.

Los tres modelos combinaron el balance con construcciones explícitas, pruebas de unicidad específicas y control de los intervalos. La formulación integral permite expresar el mismo problema bajo las hipótesis que estableceremos; la equivalencia, por sí sola, no construye una solución para una ecuación general.

20.10. Síntesis Newton–Leibniz: del dato a la identidad integral

En una ecuación diferencial, la tasa depende de una función que todavía debemos encontrar. Una vez que disponemos de una candidata y, podemos evaluar esa tasa a lo largo de su trayectoria.

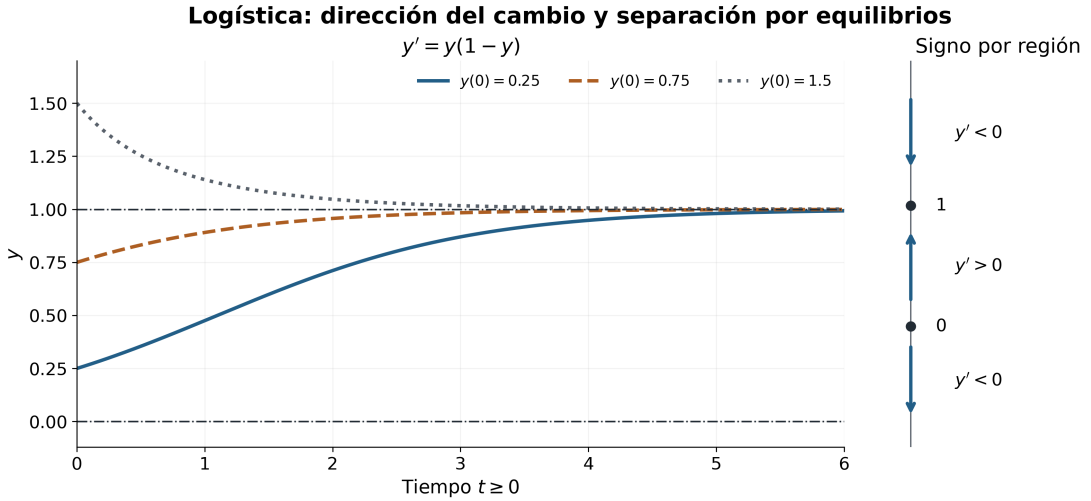


FIGURA 20.5. Para la logística con $r = 1$, los datos positivos permanecen en su región respecto de los equilibrios 0 y 1 y se aproximan a 1 en el futuro. El no cruce se demuestra por unicidad específica.

Si el resultado es continuo, Newton–Leibniz permite reconstruir los valores de y desde el dato inicial. La operación inversa, derivar una acumulación continua, permite recuperar la ecuación. Precisaremos ambas direcciones sin suponer que la identidad integral produzca por sí sola una candidata.

PROPOSICIÓN 20.3. **Proposición (equivalencia diferencial–integral).** Sea $F : D \to \mathbb{R}$, donde $D \subseteq \mathbb{R}^2$, y sean I un intervalo abierto no vacío, $x_0 \in I$ e $y_0 \in \mathbb{R}$. Sea $y : I \to \mathbb{R}$ una función continua cuyo grafo está contenido en D. Supongamos que la función

$$h(t) = F(t, y(t))$$

es continua en I. Entonces son equivalentes las siguientes afirmaciones:

1. La función y es diferenciable en I y satisface
$$y'(x) = F(x, y(x)), \qquad y(x_0) = y_0.$$
2. Para todo $x \in I$ se cumple
$$y(x) = y_0 + \int_{x_0}^{x} F(t, y(t))\, dt.$$

En cualquiera de los dos casos, y es de clase C^1.

Demostración. Supongamos primero la afirmación diferencial. Para la función continua ya determinada h, tenemos $y' = h$ e $y(x_0) = y_0$. El teorema de reconstrucción [Teorema 21.1], que aplica Newton–Leibniz a cualquier solución de ese problema, da

$$y(x) = y_0 + \int_{x_0}^{x} h(t)\, dt.$$

Esta es la identidad integral requerida.

Recíprocamente, supongamos esa identidad. La continuidad de h permite aplicar la dirección constructiva de [Teorema 21.1]: la función del miembro derecho es de clase C^1 y su derivada es $h(x)$. Como coincide con y, resulta

$$y'(x) = h(x) = F(x, y(x)).$$

En $x = x_0$, la integral es cero, por lo que $y(x_0) = y_0$. La admisibilidad del grafo era una hipótesis y se conserva. Finalmente, en ambas direcciones $y' = h$ es continua, lo que acredita la afirmación de clase C^1.

Las integrales del enunciado se toman sobre segmentos compactos contenidos en I, con la orientación correspondiente si $x < x_0$. No se integra hasta un extremo excluido ni se exige una integral impropia. Tampoco se ha invocado una teoría de continuidad multivariable: la hipótesis necesaria se formula directamente para el integrando compuesto $t \mapsto F(t, y(t))$.

Caracterización y construcción. Si $F(x, y) = g(x)$ no depende del estado, la identidad integral entrega inmediatamente la función:

$$y(x) = y_0 + \int_{x_0}^{x} g(t)\,dt.$$

Si F depende de y, la incógnita aparece también dentro de la integral. La expresión

$$y(x) = y_0 + \int_{x_0}^{x} F(t, y(t))\,dt$$

es entonces una condición sobre una función desconocida. Demostrar que una función satisface esta condición es una forma de verificar una solución; demostrar que alguna función la satisface requiere un argumento adicional.

La equivalencia tampoco aporta unicidad por sí sola. Las soluciones con tiempo de espera de [Ejemplo 21.6] son continuas, y su tasa compuesta es continua; todas satisfacen, por tanto, la misma identidad

$$y(x) = \int_{0}^{x} 2\sqrt{|y(t)|}\,dt.$$

Para y_τ, con $\tau \geq 0$, la comprobación es explícita: si $x \leq \tau$, la tasa es cero en todo el segmento entre 0 y x; si $x > \tau$,

$$\int_{0}^{x} 2\sqrt{|y_\tau(t)|}\,dt = \int_{\tau}^{x} 2(t-\tau)\,dt = (x-\tau)^2.$$

La identidad admite esas soluciones distintas. Cambiar de formulación no elimina una falta de unicidad presente en el problema.

Ejemplo 20.11. **Ejemplo (la misma solución en dos lenguajes).** La solución de [Ejemplo 21.4], $y(x) = \frac{1}{3}e^x + \frac{2}{3}e^{-2x}$, resuelve $y' = e^x - 2y$ con $y(0) = 1$. El integrando compuesto es $e^t - 2y(t) = \frac{1}{3}e^t - \frac{4}{3}e^{-2t}$, una función continua. Por integración,

$$\begin{aligned} 1 + \int_{0}^{x} \left(e^t - 2y(t)\right) dt &= 1 + \frac{1}{3}(e^x - 1) + \frac{2}{3}(e^{-2x} - 1) \\ &= \frac{1}{3}e^x + \frac{2}{3}e^{-2x} \\ &= y(x). \end{aligned}$$

La identidad integral registra el valor inicial y la acumulación de la tasa neta. Su derivación recupera la ecuación diferencial.

La existencia de esta función se construyó con el factor integrante, y la unicidad se probó para la ecuación lineal. La equivalencia integral organiza esos resultados, pero no sustituye sus pruebas.

La síntesis Newton–Leibniz consiste aquí en reconocer que derivar una acumulación y acumular una derivada son operaciones inversas bajo las hipótesis declaradas. Cuando la tasa depende del estado, esa relación sigue siendo válida para una trayectoria admisible, aunque encontrar la trayectoria exija una transformación adicional.

20.11. Laboratorio de reconstrucción y cierre del tomo

Resolver una ecuación elemental exige coordinar varias preguntas. ¿Qué función buscamos y en qué dominio puede vivir su grafo? ¿Cuál es el dato inicial? ¿Qué estructura permite construirla? ¿La transformación conserva todas las soluciones consideradas? ¿La fórmula obtenida satisface la ecuación en cada punto del intervalo? ¿Qué argumento prueba que no existen otras soluciones con ese dato?

No todas las preguntas se resuelven mediante integración. Los equilibrios se comprueban por sustitución; la unicidad puede reducirse a una ecuación para la diferencia; un extremo de intervalo puede analizarse mediante continuidad y límites. El siguiente problema reúne esas herramientas en una ecuación que cambia de apariencia según la transformación elegida.

EJEMPLO 20.12. **Ejemplo (auditoría completa de una reconstrucción).** Consideremos

$$y' = 2x(1+y), \qquad y(0) = -2,$$

con dominio de ecuación $D = \mathbb{R}^2$.

Reconocer la estructura. Es una ecuación separable, con $a(x) = 2x$ y $b(v) = 1+v$. También es lineal: $y' - 2xy = 2x$. Antes de dividir por $1+y$, registramos que $y = -1$ es una solución constante. No satisface el dato propuesto, pero pertenece a la ecuación y una división la excluiría.

Construir conservando el equilibrio. La sustitución $z = 1+y$ da

$$z' = 2xz, \qquad z(0) = -1.$$

Por [Teorema 21.2],

$$z(x) = -e^{x^2}, \qquad y(x) = -1 - e^{x^2}, \qquad x \in \mathbb{R}.$$

El teorema construye la solución y prueba unicidad para z. Como la sustitución y su inversa son válidas para toda función diferenciable, también prueba unicidad para el problema original.

Verificar. La función está definida y es de clase C^1 en toda la recta. Se cumple

$$y(0) = -2, \qquad y'(x) = -2xe^{x^2} = 2x\big(1+y(x)\big).$$

Su grafo es admisible porque $D = \mathbb{R}^2$. La solución global es maximal: no existe un intervalo real estrictamente mayor que $\mathbb{R}$.

Auditar la separación. Para el dato -2 elegimos $J = (-\infty, -1)$, donde $1+v$ no se anula. Las acumulaciones de [Teorema 21.4] son

$$H(v) = \int_{-2}^{v} \frac{ds}{1+s} = \ln|1+v|, \qquad A(x) = x^2.$$

En J, H es estrictamente decreciente y su imagen es $\mathbb{R}$. La inversa es $H^{-1}(u) = -1 - e^u$. Así, la separación produce la misma solución en todo $\mathbb{R}$. La ecuación $\ln|1+y| = x^2$ por sí sola también permitiría escribir $y = -1 + e^{x^2}$, pero esa rama pertenece a $(-1, \infty)$ y da $y(0) = 0$. El intervalo de estados y el dato inicial seleccionan la rama negativa.

Comprobar exhaustividad para la ecuación. Si dejamos libre el dato inicial, la transformación proporcional da todas las soluciones en cualquier intervalo abierto:

$$y(x) = -1 + Ce^{x^2}, \qquad C \in \mathbb{R}.$$

El valor $C = 0$ recupera el equilibrio. Para el dato indicado, $C = -1$. La separación en una rama no nula describía solo la parte de la familia correspondiente a esa rama; la transformación $z = 1+y$ permite incluir también el equilibrio sin una división.

Leer el comportamiento. La solución satisface $y < -1$. Para $x < 0$ su derivada es positiva, y para $x > 0$ es negativa. Alcanza su máximo -2 en 0 y tiende a $-\infty$ cuando $x \to \pm\infty$. Esta

divergencia ocurre hacia tiempos infinitos; no hay explosión en un extremo finito ni impedimento para la existencia global.

Reunir las dos formulaciones. La tasa compuesta $2t(1+y(t)) = -2te^{t^2}$ es continua. La identidad integral es

$$-2 + \int_0^x 2t(1+y(t))\,dt = -2 - \left(e^{x^2} - 1\right) = -1 - e^{x^2} = y(x).$$

La verificación vale para todo x, incluida la orientación hacia valores negativos. La forma integral y la forma diferencial describen la misma solución.

Una pauta de trabajo que conserva las hipótesis. Los procedimientos del capítulo pueden ordenarse por el paso que reduce la dependencia de la incógnita:

Estructura	Transformación decisiva	Control que no debe omitirse
$y' = g(x)$	Acumular la tasa desde el dato	Continuidad de g e intervalo de integración
$y' = a(x)y$	Multiplicar por la exponencial inversa	Incluir el dato cero; no dividir por y
$y' + p(x)y = q(x)$	Derivar el producto con un factor integrante	Continuidad de coeficientes y factor no nulo
$y' = a(x)b(y)$	Componer con una primitiva de $1/b$	Ceros de b, intervalo de estados e imagen de la inversa
Comparación de soluciones	Buscar una ecuación para su diferencia	Continuidad del coeficiente y punto común
Identidad integral	Aplicar las dos direcciones de FTC	Continuidad de la tasa compuesta; existencia aún por justificar

Una fórmula final debe acompañarse de su intervalo y de la comprobación del dato. Si se reclama unicidad, debemos indicar el argumento y la clase de soluciones a la que se aplica. Si se reclama maximalidad, debemos mostrar por qué no existe una prolongación para la misma ecuación. Si se interpreta un modelo, debemos separar sus consecuencias matemáticas de las hipótesis físicas que permiten utilizarlo.

Qué hemos construido y qué queda abierto. El tomo comenzó con la estructura de los números reales y las funciones, y construyó límites, continuidad y derivación. La integral de Riemann dio una operación de acumulación; el teorema fundamental vinculó esa operación con la derivada. Las funciones elementales y las técnicas de integración ampliaron las posibilidades de cálculo sin reemplazar las condiciones que hacen válidos esos cálculos.

Las ecuaciones diferenciales elementales reúnen ese recorrido. La incógnita es ahora una función cuya tasa depende de ella misma, pero las pruebas siguen descansando en resultados construidos: continuidad sobre intervalos, regla de la cadena, derivación de la inversa, teorema del valor medio, acumulación y Newton–Leibniz. En algunas clases pudimos construir una solución y demostrar que el dato inicial la determina. En otras comprobamos directamente que el mismo dato admite varias trayectorias. Una teoría general de existencia y unicidad, los sistemas y los métodos de aproximación con garantías de error requieren herramientas adicionales. Para abordarlos conservamos las distinciones utilizadas aquí: obtener una candidata, verificarla, demostrar exhaustividad y justificar su intervalo de solución.

El banco siguiente reúne esas distinciones. Cada solución debe precisar el intervalo y justificar que la función obtenida satisface la ecuación y el dato, además de establecer unicidad cuando esa sea la conclusión solicitada.

Ejercicios y soluciones

El banco combina reconocimiento, construcción, comprobación y control de dominios. En cada problema de valor inicial, la fórmula debe acompañarse del intervalo utilizado y de la justificación de unicidad cuando se solicita. Los problemas son originales y se resuelven con las herramientas del tomo.

Nivel A — Reconocimiento y comprensión. Ejercicio A1. Verificar y declarar el intervalo. Comprueba que $y(x) = 1/(x+2)$ resuelve $y' = -y^2$. Determina el mayor intervalo de esta fórmula que contiene 0 y el dato inicial correspondiente.

Ejercicio A2. Ecuación y dato inicial. La familia $y = x^2 + C$ satisface $y' = 2x$. Determina qué miembro cumple $y(2) = 7$ y explica qué información aporta la ecuación y cuál aporta el dato.

Ejercicio A3. Reconocer la estructura. Clasifica como lineales en y, separables, o ambas, las ecuaciones $y' + xy = x$, $y' = xy^2$ y $y' = (1+x)y$. Escribe los coeficientes o factores que justifican tu clasificación.

Ejercicio A4. Equilibrios. Encuentra todas las soluciones constantes de $y' = (1+x^2)(y-2)(y+1)$ sobre $\mathbb{R}$. Justifica que la lista es completa.

Ejercicio A5. División que pierde una solución. En $y' = x(y-3)$ se propone dividir por $y-3$ y afirmar que todas las soluciones cumplen $\ln|y-3| = x^2/2 + C$. Señala qué solución quedó excluida y por qué esa igualdad no la representa.

Ejercicio A6. Orientación de la acumulación. Para $y' = 1 + x$ e $y(1) = 4$, escribe la reconstrucción con punto base y calcula $y(-1)$. Explica el signo de la integral utilizada.

Soluciones del nivel A. **Solución A1.** La derivada es $y' = -1/(x+2)^2 = -y^2$ donde $x \neq -2$. El intervalo que contiene 0 es $(-2, \infty)$ y $y(0) = 1/2$. La fórmula al otro lado del polo corresponde a otro intervalo; no se puede unir a este atravesando -2.

Solución A2. El dato exige $4 + C = 7$, de modo que $C = 3$. La función $y = x^2 + 3$ está definida en $\mathbb{R}$, tiene derivada $2x$ y toma el valor 7 en 2. La ecuación fija la tasa pero deja una constante libre; el dato selecciona esa constante. La exhaustividad de la familia procede de [Teorema 21.1].

Solución A3. La primera es lineal con $p = x, q = x$ y separable como $y' = x(1-y)$. La segunda es separable con $a = x, b(v) = v^2$, pero no lineal en la incógnita: su dependencia cuadrática no tiene la forma $q(x) - p(x)y$. La tercera es separable con $a = 1 + x, b(v) = v$ y lineal homogénea con $p = -(1+x), q = 0$. Todas son de primer orden porque solo aparece la primera derivada.

Solución A4. Una constante c satisface $0 = (1+x^2)(c-2)(c+1)$ para todo x. Como $1 + x^2 > 0$, esto equivale a $(c-2)(c+1) = 0$. Los valores son 2 y -1, y sus funciones constantes verifican la ecuación. La necesidad y suficiencia también son un caso de [Proposición 21.1].

Solución A5. La función $y = 3$ satisface $y' = 0 = x(y-3)$ en toda la recta. Al dividir por $y-3$ se exige $y \neq 3$; además, el logaritmo de cero no está definido. La familia no constante obtenida debe acompañarse de la solución constante. Una comprobación de fórmulas no demuestra exhaustividad si la transformación eliminó casos.

Solución A6. Por [Teorema 21.1], $y(x) = 4 + \int_1^x (1+t)\,dt$. Entonces $\int_1^{-1}(1+t)\,dt = -\int_{-1}^{1}(1+t)\,dt = -2$, y $y(-1) = 2$. Una fórmula equivalente es $y(x) = x + x^2/2 + 5/2$, cuya derivada y valor en 1 confirman el resultado.

Nivel B — Aplicación directa. Ejercicio B1. Reconstrucción con tasa continua. Resuelve en $\mathbb{R}$ $y' = e^{x^2}$, $y(0) = -1$, sin exigir una primitiva elemental. Comprueba el dato y explica por qué la solución es única.

Ejercicio B2. Dato negativo en una tasa proporcional. Resuelve $y' = (1-2x)y$, $y(0) = -4$, y determina dónde crece y dónde decrece la solución.

Ejercicio B3. Factor integrante constante. Resuelve en $\mathbb{R}$ $y' + 3y = 6$, $y(0) = 5$. Verifica la ecuación y calcula el límite futuro.

Ejercicio B4. Factor integrante variable. Resuelve $y' + (2/x)y = x$ con $y(1) = 2$ sobre $(0, \infty)$. Declara por qué no aplicas directamente el teorema lineal en toda $\mathbb{R}$.

Ejercicio B5. Separar en una rama declarada. Resuelve $y' = x/y$, $y(0) = 2$, con valores positivos. Construye una solución sobre el mayor intervalo permitido por esa rama y verifica su unicidad dentro de ella.

Ejercicio B6. Intervalo de una fórmula racional. Resuelve $y' = 3y^2$, $y(0) = 2$, y determina su intervalo maximal que contiene 0.

Soluciones del nivel B. **Solución B1.** La tasa es continua, por lo que [Teorema 21.1] da $y(x) = -1 + \int_0^x e^{t^2}\,dt$. El FTC da $y' = e^{x^2}$; la integral nula en 0 da el dato. El mismo teorema prueba unicidad entre soluciones en $\mathbb{R}$. Cada integral recorre un segmento compacto, aunque la tasa no esté acotada globalmente.

Solución B2. La acumulación del coeficiente es $A = x - x^2$. Por [Teorema 21.2], la solución única en $\mathbb{R}$ es $y = -4e^{x-x^2}$. Su derivada es $(1-2x)y$ y el dato se cumple. Como $y < 0$, la derivada es negativa si $x < 1/2$ y positiva si $x > 1/2$. Decrece en $(-\infty, 1/2)$, crece en $(1/2, \infty)$ y tiene mínimo $-4e^{1/4}$ en $1/2$.

Solución B3. Una solución particular es 2, y la homogénea tiene soluciones Ce^{-3x}. Por [Corolario 21.2], $y = 2 + Ce^{-3x}$. El dato fija $C = 3$, de modo que $y = 2 + 3e^{-3x}$. Se verifica $y' = -9e^{-3x}$ y $y' + 3y = 6$, además de $y(0) = 5$. La unicidad procede de [Teorema 21.3]; el límite cuando $x \to +\infty$ es 2.

Solución B4. En $(0,\infty)$ los coeficientes son continuos. Tomando base 1, $P=2\ln x$ y $\mu=x^2$. La ecuación transformada es $(x^2y)'=x^3$, y $x^2y=2+(x^4-1)/4$. Por tanto, $y=x^2/4+7/(4x^2)$. Su derivada es $x/2-7/(2x^3)$, mientras que $(2/x)y=x/2+7/(2x^3)$; la suma es x y el valor en 1 es 2. [Teorema 21.3] prueba unicidad en ese intervalo. El coeficiente $2/x$ no está definido en 0, por lo que sus hipótesis no se cumplen en toda la recta.

Solución B5. El dominio exige $y\neq 0$; elegimos $J=(0,\infty)$. En [Teorema 21.4], $a(x)=x$, $b(v)=1/v$, $H(v)=(v^2-4)/2$ y $A(x)=x^2/2$. Como $H(J)=(-2,\infty)$ contiene $A(x)$ para todo x, el intervalo es $\mathbb{R}$. La inversa da $y=\sqrt{x^2+4}$. Es positiva, $y(0)=2$ y $y'=x/\sqrt{x^2+4}=x/y$. El teorema prueba unicidad entre soluciones con valores en J.

Solución B6. En la rama positiva, integrar $(1/y)'=-3$ da $1/y=1/2-3x$, y $y=2/(1-6x)$ en $K=(-\infty,1/6)$. Se verifica $y'=12/(1-6x)^2=3y^2$ y $y(0)=2$. Si dos soluciones coinciden, su diferencia satisface $d'=3(y_1+y_2)d$; [Teorema 21.2] prueba unicidad en el intervalo común. Como $y\to+\infty$ al aproximarse a $1/6$ desde la izquierda, no existe una prolongación continua a través de ese punto. Todo intervalo mayor que K lo contendría; así, la solución es maximal.

Nivel C — Combinación estructural. Ejercicio C1. Lineal e integración por partes. Resuelve $y'+y=x^2$, $y(0)=0$, evaluando la integral del factor integrante. Comprueba la fórmula.

Ejercicio C2. Diferencia y orden entre soluciones. Sean y_1,y_2 soluciones de $y'+2xy=e^x$ en $\mathbb{R}$, con $y_1(0)=3$ e $y_2(0)=-1$. Determina y_1-y_2 sin resolver por separado ambas ecuaciones. ¿Pueden cruzarse?

Ejercicio C3. Reconstruir el coeficiente. Se sabe que $y(x)=-2e^{x^2+x}$ es una solución de $y'=a(x)y$ en $\mathbb{R}$. Determina a y, para ese coeficiente, resuelve el problema con dato $z(0)=5$. Explica el papel de la no anulación de la trayectoria dada.

Ejercicio C4. Balance de mezcla. Un tanque tiene inicialmente 10 litros y 1 gramo de soluto. Entran 2 litros por minuto con concentración $1/2$ gramo por litro; sale 1 litro por minuto de mezcla uniforme. Sin reacción química ni límite de capacidad en el modelo, determina cantidad y concentración para $t\geq 0$, y sus límites futuros.

Ejercicio C5. Logística entre equilibrios. Resuelve $y'=2y(1-y)$, $y(0)=1/4$. Prueba existencia futura y unicidad, y determina cuándo alcanza $y=1/2$.

Ejercicio C6. Fórmula y lectura de monotonía. Resuelve $y'=(2x-2)y$, $y(0)=1$, y combina la fórmula con el signo de la tasa para encontrar el mínimo global, los límites en ambos extremos de la recta y el intervalo maximal.

Soluciones del nivel C. **Solución C1.** El factor es e^x. Dos integraciones por partes dan $\int_0^x e^t t^2\,dt=e^x(x^2-2x+2)-2$. Por [Teorema 21.3], $y=x^2-2x+2-2e^{-x}$. La derivada es $2x-2+2e^{-x}$; sumarla a y da x^2, y $y(0)=0$. Los coeficientes son continuos en toda $\mathbb{R}$, de modo que la solución construida allí es única.

Solución C2. La diferencia $d=y_1-y_2$ satisface $d'=-2xd$, $d(0)=4$. Por [Teorema 21.2], $d=4e^{-x^2}>0$ en toda la recta. Por tanto, $y_1>y_2$ en cada punto y no se cruzan. La existencia de ambas soluciones para los datos indicados está garantizada por [Teorema 21.3], pues sus coeficientes son continuos.

Solución C3. La trayectoria nunca se anula, así que la ecuación obliga a $a = y'/y = 2x + 1$. Este coeficiente es continuo y la sustitución verifica la trayectoria propuesta. Para el nuevo dato, [Teorema 21.2] da la solución única $z = 5e^{x^2+x}$. La división aquí reconstruye un coeficiente desde una función conocida no nula; no es una transformación válida para clasificar todas las soluciones sin registrar la solución cero.

Solución C4. El volumen es $V = 10 + t$ litros. La entrada aporta 1 gramo por minuto y la salida aporta $m/(10+t)$ gramos por minuto. Así, $m' + m/(10+t) = 1$, $m(0) = 1$, con coeficientes continuos en $(-10, \infty)$. Una particular es $(10 + t)/2$, y la diferencia respecto de ella resuelve $z' = -z/(10 + t)$ con $z(0) = -4$. Por [Teorema 21.2] y [Corolario 21.2],

$$m(t) = \frac{10+t}{2} - \frac{40}{10+t}, \qquad c(t) = \frac{1}{2} - \frac{40}{(10+t)^2}.$$

Se verifica $m' = 1/2 + 40/(10 + t)^2$ y $m/(10 + t) = 1/2 - 40/(10 + t)^2$, cuya suma es 1; también $m(0) = 1$. Para $t \geq 0$, c crece desde $1/10$ hasta su límite $1/2$ gramo por litro, y $m \to +\infty$ por el volumen creciente. La unicidad lineal y la comprobación aseguran la solución. Un tanque real de capacidad finita limita el tiempo físico de uso; esa capacidad no figura en el modelo.

Solución C5. Separando en $(0, 1)$ obtenemos $\ln(y/(1 - y)) = 2t - \ln 3$, por lo que $y = 1/(1 + 3e^{-2t})$. La función está definida para todo $t \in \mathbb{R}$, toma valores en $(0, 1)$, satisface el dato y su derivada es $6e^{-2t}/(1 + 3e^{-2t})^2 = 2y(1 - y)$. Esto prueba existencia global; [Proposición 21.2] prueba unicidad. La igualdad $y = 1/2$ equivale a $3e^{-2t} = 1$, y ocurre en $t = (\ln 3)/2$.

Solución C6. Por [Teorema 21.2], la solución única es $y = e^{x^2-2x} = e^{(x-1)^2-1}$ en toda $\mathbb{R}$. Es positiva y su derivada es $(2x - 2)y$: decrece para $x < 1$ y crece para $x > 1$. Tiene mínimo global e^{-1} en 1. La fórmula prueba que $y \to +\infty$ cuando $x \to +\infty$ y cuando $x \to -\infty$. El intervalo $\mathbb{R}$ es maximal porque no existe un intervalo real mayor. La divergencia hacia tiempos infinitos no constituye una explosión en tiempo finito.

Nivel D — Inversión y diagnóstico. Ejercicio D1. Una familia sin prueba de exhaustividad. Se verifica que $y = Ce^{x^2}$, con $C > 0$, satisface $y' = 2xy$. ¿Describe todas las soluciones en $\mathbb{R}$? Completa la familia y prueba exhaustividad.

Ejercicio D2. Recuperar un dato anterior. Una solución de $y' + 2y = 4$ satisface $y(1) = 3$. Determina $y(0)$ y justifica que el valor es único.

Ejercicio D3. La condición de equilibrio. Sea $a : I \to \mathbb{R}$ continua en un intervalo abierto no vacío I y supongamos que la función constante $y = c$ resuelve $y' = a(x)(y^2 - 1)$ en I. ¿Cuándo se puede concluir $c = \pm 1$? ¿Qué ocurre si $a \equiv 0$?

Ejercicio D4. Dominio físico y matemático. Un modelo térmico da $T(t) = 20 + 60e^{-t/5}$ en $\mathbb{R}$, con tiempo en minutos y temperatura en grados Celsius. Una experiencia comienza en $t = 0$. Un estudiante elimina los tiempos negativos afirmando que la solución matemática solo existe para $t \geq 0$. Corrige la afirmación y declara la ecuación y el dato.

Ejercicio D5. Representación removible. En $(0, \infty)$ se escribe $y = (x^3 - x^2)/x$ como solución de $y' = 2x - 1$. Construye una prolongación a $\mathbb{R}$. Explica por qué el mismo argumento no prolongaría una solución si el dominio de la ecuación exigiera $x > 0$.

Ejercicio D6. Equivalencia no implica construcción. Se escribe $y(x) = 1 + \int_0^x y(t)^2\,dt$ y se concluye que existe una solución en toda $\mathbb{R}$. Identifica el salto lógico y determina el intervalo maximal del problema correspondiente.

Soluciones del nivel D. **Solución D1.** La verificación solo cubre soluciones positivas. Faltan la nula y las negativas. Para una solución arbitraria y, $(ye^{-x^2})' = e^{-x^2}(y' - 2xy) = 0$, por lo que el producto es constante en el intervalo. Todas las soluciones son Ce^{x^2} con $C \in \mathbb{R}$, como también establece [Teorema 21.2].

Solución D2. La familia completa es $y = 2 + Ce^{-2x}$. El dato posterior exige $Ce^{-2} = 1$, así que $C = e^2$ y $y(0) = 2 + e^2$. La sustitución verifica la ecuación y el valor en 1. [Teorema 21.3], aplicado con base 1, prueba que existe una única solución en $\mathbb{R}$ para ese dato; no se presupone que todo problema inverso de cualquier ecuación sea único.

Solución D3. La sustitución exige $a(x)(c^2 - 1) = 0$ para todo $x \in I$. Si existe un punto con $a(x) \neq 0$, resulta $c^2 = 1$ y $c = \pm 1$. Ambos valores siempre dan soluciones constantes. Si $a \equiv 0$, cualquier constante satisface la ecuación; no se puede concluir que sea un cero del factor de estado. Es exactamente la distinción de [Proposición 21.1].

Solución D4. La fórmula es una solución matemática global de $T' = -(T - 20)/5$, $T(0) = 80$. Su derivada verifica la ley y el coeficiente tiene unidades de minuto inverso. La experiencia utiliza $t \geq 0$, pero eso no elimina la solución diferenciable sobre $\mathbb{R}$. La validez física requiere un ambiente de 20 grados y un coeficiente constante durante el tiempo considerado; su eventual fallo no es una singularidad matemática de la fórmula.

Solución D5. El cociente coincide con $x^2 - x$ para $x > 0$. Definimos $\widetilde{y} = x^2 - x$ en $\mathbb{R}$; es diferenciable, su derivada es $2x - 1$ y conserva los valores anteriores. Es una prolongación para la ecuación definida en toda $\mathbb{R}^2$. Si el dominio exige $x > 0$, ningún grafo con un punto de tiempo 0 es admisible. La simplificación no amplía el dominio de la ecuación.

Solución D6. La incógnita sigue dentro de la integral; no se ha construido una función global. Por [Proposición 21.3], para una candidata continua la identidad equivale a $y' = y^2$, $y(0) = 1$. [Ejemplo 21.9] construye la única solución $y = 1/(1 - x)$ en $(-\infty, 1)$ y prueba que el polo impide prolongarla. Así, la conclusión de existencia global es además falsa en este caso.

Nivel E — Contraejemplos. Ejercicio E1. Verificada pero incompleta. Construye una familia de soluciones de $y' = 1$ en $\mathbb{R}$ que esté correctamente verificada pero no sea exhaustiva. Exhibe una solución omitida y una prueba de la familia completa.

Ejercicio E2. Continuidad sin unicidad. Para $y' = 4\sqrt{|y|}$, $y(0) = 0$, construye dos soluciones globales distintas. Verifica la derivada en cualquier punto donde cambie la expresión.

Ejercicio E3. Una trayectoria que toca un equilibrio. Construye para $y' = 2\sqrt{|y|}$ una solución global no constante que coincida con el equilibrio cero en todo $(-\infty, 2]$. ¿Qué afirmación sobre no cruce o contacto queda refutada?

Ejercicio E4. Dato fuera del dominio. Da un problema inicial con segundo miembro $F(x, y) = 1/y$ y dato no admisible. Explica por qué no puede tener solución, sin intentar integrar.

Ejercicio E5. La suma no conserva una ecuación no homogénea. Exhibe dos soluciones de $y' + y = 1$ cuya suma no sea solución de la misma ecuación. Comprueba las tres afirmaciones.

Ejercicio E6. Una lineal que cruza cero. Construye un problema lineal con coeficientes continuos en $\mathbb{R}$ cuya solución única cambie de signo. Explica por qué una división por la incógnita no puede usarse en todo su intervalo.

Soluciones del nivel E. **Solución E1.** La familia $y = x + C$ con $C > 0$ tiene derivada 1, pero omite $y = x$. Para cualquier solución, $(y - x)' = 0$, por lo que $y - x$ es constante. La familia completa es $x + C$, $C \in \mathbb{R}$. Verificar cada miembro de una subfamilia no demuestra que contenga todas las soluciones.

Solución E2. Una es $y = 0$. Otra es $z(x) = 0$ para $x \leq 0$ y $z(x) = 4x^2$ para $x > 0$. En 0, el cociente incremental es 0 a la izquierda y $4h$ a la derecha, con límite 0; por tanto, $z'(0) = 0$. Para $x > 0$, $z' = 8x = 4\sqrt{4x^2}$; para $x \leq 0$ ambos miembros son cero. Ambas cumplen el dato y son distintas, por ejemplo en 1. El segundo miembro es continuo en todos los estados reales.

Solución E3. Tomamos $y = 0$ si $x \leq 2$ y $y = (x - 2)^2$ si $x > 2$. El cociente incremental en 2 tiene límites cero; la derivada es cero hasta 2 y $2(x - 2)$ después. Esa derivada coincide con $2\sqrt{|y|}$ en todos los puntos. La función no es constante y toca el equilibrio, incluso durante un intervalo entero. Se refuta que toda solución no constante de una ecuación autónoma continua evita sus equilibrios; no se refuta el resultado específico de unicidad logística.

Solución E4. El problema $y' = 1/y$, $y(0) = 0$ tiene dato fuera de $D = \{(x, v) : v \neq 0\}$. Cualquier función que satisfaga el dato tendría en su grafo el punto $(0, 0)$, donde la ecuación no está definida. Por la definición de solución, no existe una solución de ese problema. Una fórmula que resuelva la ecuación lejos del dato no repara la inadmisibilidad.

Solución E5. Podemos elegir $y_1 = 1$ e $y_2 = 1 + e^{-x}$, distintas y globales. La primera satisface $0 + 1 = 1$; la segunda satisface $-e^{-x} + 1 + e^{-x} = 1$. Su suma es $2 + e^{-x}$ y cumple $(y_1 + y_2)' + (y_1 + y_2) = 2$, que no es 1. Una solución particular admite suma con una homogénea, no con cualquier otra particular.

Solución E6. El problema $y' = 1$, $y(0) = 0$ es lineal con $p = 0, q = 1$. Su solución única por [Teorema 21.1] es $y = x$, negativa para $x < 0$ y positiva para $x > 0$. Dividir por y excluye 0, que pertenece al intervalo y al dato. La conservación del signo de la ecuación homogénea proporcional no se extiende a todas las lineales no homogéneas.

Nivel F — Descubrimiento guiado. Ejercicio F1. Descubrir el factor integrante. Sean I un intervalo abierto no vacío, $p, q : I \to \mathbb{R}$ continuas, $x_0 \in I$ e $y_0 \in \mathbb{R}$. (a) Expande $(\mu y)'$ y determina una condición sobre μ para que sea $\mu(y' + py)$. (b) Construye una μ positiva con $\mu(x_0) = 1$. (c) Deduce la fórmula del problema inicial $y(x_0) = y_0$.

Ejercicio F2. Unicidad sin dividir por el estado. Sean u, v soluciones de $y' = y^2$ en un intervalo abierto I, sea $s \in I$ y supongamos $u(s) = v(s)$. (a) Obtén una ecuación para $d = u - v$. (b) Usa el teorema proporcional para demostrar igualdad. (c) Deduce qué sucede con el dato cero.

Ejercicio F3. Construir una inversa con datos nuevos. Resuelve $y' = (1 + x)/(1 + y^2)$, $y(0) = 0$, sin buscar una fórmula por radicales. (a) Construye H. (b) Prueba que su imagen es $\mathbb{R}$. (c) Define y verifica la solución mediante H^{-1}.

Ejercicio F4. Del balance a la concentración. Un tanque ideal tiene $V_0 > 0$, caudales constantes de entrada $u > 0$ y salida $v > 0$, concentración entrante constante $c_{\text{in}} \geq 0$, concentración inicial $c_0 \geq 0$ y mezcla uniforme. Denota por m la cantidad de soluto, con $m(0) = c_0 V_0$. (a) Deduce las ecuaciones para volumen V y cantidad m. (b) Para $c = m/V$, demuestra $c' = (u/V)(c_{\text{in}} - c)$. (c) Obtén una fórmula y demuestra que, para tiempos futuros con $V > 0$, c permanece entre su dato c_0 y c_{in}.

Ejercicio F5. Auditar un empalme. Para $\tau \geq 0$, define $y_\tau = 0$ si $x \leq \tau$ y $y_\tau = 9(x - \tau)^2$ si $x > \tau$. (a) Comprueba diferenciabilidad en τ. (b) Determina $k > 0$ para que todas estas funciones resuelvan $y' = k\sqrt{|y|}$. (c) Comprueba el dato $y(0) = 0$ y explica la falta de unicidad.

Ejercicio F6. Reconstruir la equivalencia integral. Sean I un intervalo abierto no vacío, $x_0 \in I$, $y_0 \in \mathbb{R}$ y $F : D \to \mathbb{R}$. Sea una candidata continua $y : I \to \mathbb{R}$ con grafo contenido en D y $h(t) = F(t, y(t))$ continua. Sin citar [Proposición 21.3], demuestra las dos direcciones entre el problema diferencial con dato (x_0, y_0) y la identidad integral. Indica en qué dirección se construye la derivada.

Soluciones del nivel F. **Solución F1.** (a) El producto es $\mu y' + \mu' y$, así que basta $\mu' = p\mu$. (b) Para $P = \int_{x_0}^{x} p$, la función $\mu = e^P$ satisface $\mu' = p\mu$, es positiva y vale 1 en x_0. (c) La ecuación equivale a $(\mu y)' = \mu q$. Reconstruir desde el dato da $\mu y = y_0 + \int_{x_0}^{x} \mu(t) q(t)\, dt$, y por tanto $y = e^{-P}(y_0 + \int_{x_0}^{x} e^{P(t)} q(t)\, dt)$. El FTC verifica existencia; la unicidad de reconstrucción para μy y la no anulación de μ prueban unicidad para y.

Solución F2. (a) $d' = u^2 - v^2 = (u + v)d$, con coeficiente continuo. (b) Por [Teorema 21.2], $d(x) = d(s) \exp(\int_s^x (u + v)) = 0$, así que $u = v$ en I. (c) La función nula es una solución global; comparar cualquier solución con ella en su intervalo que contiene el dato cero demuestra que también es nula. El argumento incluye el cero y evita la división por y^2.

Solución F3. (a) Con $b(v) = 1/(1 + v^2)$, $H(v) = \int_0^v (1 + s^2)\, ds = v + v^3/3$, y $A = x + x^2/2$. (b) $H' = 1 + v^2 > 0$, así que es estrictamente creciente; sus límites hacia $\pm\infty$ son $\pm\infty$. Continuidad y valor intermedio dan $H(\mathbb{R}) = \mathbb{R}$. (c) Definimos $y = H^{-1}(x + x^2/2)$ en $\mathbb{R}$. La derivada de la inversa da $y' = (1 + x)/(1 + y^2)$ y $H^{-1}(0) = 0$. [Teorema 21.4] prueba unicidad con estados en $\mathbb{R}$. La solución existe sin necesidad de expresar la inversa mediante radicales.

Solución F4. (a) $V' = u - v$, $V = V_0 + (u - v)t$, y $m' = uc_{\text{in}} - vm/V$. (b) Usando $m = cV$, $m' = c'V + c(u - v)$; sustituir en el balance da $c'V = u(c_{\text{in}} - c)$. (c) La diferencia $z = c - c_{\text{in}}$ satisface $z' = -(u/V)z$, de modo que

$$c(t) = c_{\text{in}} + (c_0 - c_{\text{in}}) \exp\left(- \int_0^t \frac{u}{V(s)}\, ds \right).$$

En el intervalo positivo de volumen el coeficiente es continuo, y para $t \geq 0$ la integral es no negativa: el factor pertenece a $(0, 1]$. La fórmula es una combinación entre los dos valores, incluidos los casos de igualdad. u/V tiene unidades de tiempo inverso; la concentración no se obtiene restando directamente los caudales de soluto sin considerar el volumen.

Solución F5. (a) El cociente incremental es cero para $h < 0$ y $9h$ para $h > 0$, así que la derivada en τ es cero. (b) Para $x > \tau$, la derivada es $18(x - \tau)$ y la raíz es $3(x - \tau)$; por tanto, $k = 6$. En $x \leq \tau$ ambos miembros son cero, incluida la unión. (c) Como $\tau \geq 0$, el dato se cumple. Las elecciones $\tau = 0$ y $\tau = 1$ producen valores distintos en $x = 1/2$, respectivamente $9/4$ y 0. La comprobación del empalme permite afirmar que son soluciones globales distintas.

Solución F6. Si y es diferenciable y $y' = h$, entonces y' es continua. Newton–Leibniz sobre el segmento entre x_0 y x, con orientación si es necesaria, da $y(x) - y(x_0) = \int_{x_0}^{x} h(t)\,dt$. El dato produce la identidad. A la inversa, si $y = y_0 + \int_{x_0}^{x} h$, el FTC para integrando continuo muestra que y es diferenciable y $y' = h = F(x, y(x))$; en x_0 la integral cero da el dato. Esta segunda dirección construye la derivada de la candidata. Ninguna dirección construye una candidata para un F arbitrario.

Nivel G — Desafío. Ejercicio G1. Clasificar las soluciones globales. Demuestra que la única solución de $y' = y^2$ definida en toda $\mathbb{R}$ es la nula. Tu prueba debe incluir posibles ceros y explicar por qué las fórmulas no nulas no pueden ser globales.

Ejercicio G2. Problema inverso con datos suficientes. En el modelo $T' = -k(T - T_a)$, se sabe $T_a = 20$, $k > 0$, $T(0) = 80$ y $T(10) = 50$, con tiempo en minutos. Determina k, la trayectoria y el tiempo en que llega a 35. Prueba unicidad del parámetro y de la trayectoria bajo estas hipótesis.

Ejercicio G3. Equilibrio y frontera finita en un modelo. Un tanque ideal tiene volumen inicial 12 litros, entrada 1 litro por minuto con concentración 2 gramos por litro y salida 3 litros por minuto. Su concentración inicial es 5 gramos por litro. (a) Determina V, c, m durante el tiempo físico futuro. (b) Comprueba el balance y los límites al vaciarse. (c) Decide si esos límites justifican continuar la misma ecuación después del vaciado.

Ejercicio G4. Dato, identidad integral y orden global. Sea $g : \mathbb{R} \to \mathbb{R}$ continua y acotada por $|g| \leq M$, con $M \geq 0$, y sea $\lambda > 0$. Para $y' + \lambda y = g$, $y(0) = a$: (a) construye la solución global y verifica su identidad integral; (b) prueba para $t \geq 0$ la cota $|y(t)| \leq |a|e^{-\lambda t} + (M/\lambda)(1 - e^{-\lambda t})$; (c) compara las soluciones con datos $a < b$; (d) explica por qué la cota no prueba que y tenga límite futuro.

Soluciones del nivel G. **Solución G1.** Si una solución global toma valor cero en algún s, el argumento de diferencia lineal de [Ejemplo 21.9], con base s, la compara con la nula y prueba que es cero en toda $\mathbb{R}$. Si nunca se anula, su continuidad en el intervalo implica signo fijo. La función $1/y$ es diferenciable y $(1/y)' = -1$. La función $1/y(x) + x$ tiene derivada nula, así que es constante y $1/y(x) = 1/y(0) - x$ para todo x. Al tomar $x = 1/y(0)$, el miembro derecho es cero, imposible para el recíproco de un valor real no nulo. Por tanto, no hay solución global no nula. La solución cero está directamente verificada.

Solución G2. La fórmula térmica es $T = 20+60e^{-kt}$. El dato en 10 exige $e^{-10k} = 1/2$, de donde $k = (\ln 2)/10 > 0$, único por inyectividad de la exponencial. Entonces $T(t) = 20 + 60e^{-(\ln 2)t/10}$. Alcanzar 35 exige $e^{-kt} = 1/4$, así que $t = (\ln 4)/k = 20$ minutos. Sustituir confirma ambos valores observados. Para ese k, el teorema lineal prueba unicidad de la trayectoria en $\mathbb{R}$. La temperatura ambiental conocida es una hipótesis esencial para esta determinación con los datos dados.

Solución G3. (a) $V = 12 - 2t$, por lo que el intervalo físico es $0 \le t < 6$. De la ecuación de concentración $c' = (1/V)(2 - c)$ obtenemos

$$c(t) = 2 + 3\sqrt{\frac{12 - 2t}{12}}, \qquad m(t) = (12 - 2t)\left(2 + 3\sqrt{\frac{12 - 2t}{12}}\right).$$

La solución matemática está definida en el intervalo abierto $(-\infty, 6)$. (b) Si $R = \sqrt{V/12} > 0$, entonces $R' = -R/V$, y $c' = -3R/V = (2 - c)/V$. Como $m = cV$ y $V' = -2$, $m' = c'V - 2c = 2 - 3c = 2 - 3m/V$, exactamente el balance de entrada menos salida. Se cumplen $c(0) = 5$ y $m(0) = 60$. Al acercarse a 6 desde abajo, $V \to 0$, $c \to 2$ y $m \to 0$. La concentración entrante 2 es el estado constante de equilibrio de la ecuación de concentración mientras $V > 0$. (c) En 6, el cociente m/V y el coeficiente $1/V$ no están definidos. Los límites no convierten un tanque vacío en una mezcla con concentración definida ni permiten una prolongación de la misma ecuación a través de ese punto. Una descripción posterior requiere otro modelo.

Solución G4. (a) El factor integrante da

$$y(t) = e^{-\lambda t}\left(a + \int_0^t e^{\lambda s} g(s)\, ds\right).$$

Las integrales sobre segmentos compactos existen; derivar verifica $y' = g - \lambda y$ y el dato. Los coeficientes continuos y [Teorema 21.3] prueban existencia y unicidad en $\mathbb{R}$. Como la tasa compuesta es continua, Newton–Leibniz da $y(t) = a + \int_0^t (g(s) - \lambda y(s))\, ds$. (b) Para $t \ge 0$, la desigualdad integral da

$$|y(t)| \le |a|e^{-\lambda t} + Me^{-\lambda t}\int_0^t e^{\lambda s}\, ds = |a|e^{-\lambda t} + \frac{M}{\lambda}(1 - e^{-\lambda t}).$$

(c) La diferencia entre las soluciones con datos b y a es $(b - a)e^{-\lambda t} > 0$ para todo tiempo real. Conservan su orden, y su diferencia tiende a cero en el futuro. (d) La acotación no implica convergencia. Para comprobar que incluso esta clase de ecuaciones puede tener una solución acotada sin límite, definamos $h(u) = 3u^2 - 2u^3$ en $[0, 1]$. Para cada entero n, ponemos

$$Y(t) = \begin{cases} h(t - 2n), & 2n \le t \le 2n + 1, \\ h(2n + 2 - t), & 2n + 1 \le t \le 2n + 2. \end{cases}$$

Las expresiones coinciden en las uniones: en los tiempos pares valen cero y en los impares valen uno. Como $h'(0) = h'(1) = 0$, las derivadas también coinciden, y Y es de clase C^1 en $\mathbb{R}$. Además, $0 \le Y \le 1$ y $|Y'| \le 3/2$, pues $h'(u) = 6u(1 - u)$. La función $g = Y' + \lambda Y$ es continua y acotada por $3/2 + \lambda$. Con dato $a = 0$, Y es la solución única de la ecuación correspondiente. Sin embargo, $Y(2n) = 0$ e $Y(2n+1) = 1$ para todos los enteros positivos n, de modo que no tiene límite futuro. Este ejemplo demuestra que las hipótesis de continuidad y acotación sobre la entrada no bastan para deducir convergencia.

Bibliografía

1. Stephen Abbott, *Understanding Analysis*, 2 ed., Springer, 2015, Ejemplar corregido en la segunda impresión, 2016.
2. Tom M. Apostol, *Calculus, Volume I*, 2 ed., Wiley, 1967.
3. Richard Courant and Fritz John, *Introduction to Calculus and Analysis, Volume I*, Springer, 1989.
4. G. H. Hardy, *A Course of Pure Mathematics*, 3 ed., Cambridge University Press, 1921.
5. Niels Jacob and Kristian P. Evans, *A course in analysis, vol. i: Introductory calculus, analysis of functions of one real variable*, World Scientific, 2016.
6. Gustavo Labbé Morales, *Cálculo I*, Apuntes de curso; secciones 1.3–1.5.
7. Zbigniew H. Nitecki, *Calculus Deconstructed: A Second Course in First-Year Calculus*, Mathematical Association of America, 2009.
8. George Pólya and Gábor Szegő, *Problems and Theorems in Analysis I: Series, Integral Calculus, Theory of Functions*, Springer, 1972.
9. Murray H. Protter and Charles B. Morrey, Jr., *A First Course in Real Analysis*, 2 ed., Springer, 1991.
10. Charles C. Pugh, *Real Mathematical Analysis*, 2 ed., Springer, 2015.
11. Tomasz Radożycki, *Solving Problems in Mathematical Analysis, Part I: Sets, Functions, Limits, Derivatives, Integrals, Sequences and Series*, Springer, 2020.
12. ———, *Solving Problems in Mathematical Analysis, Part II: Definite, Improper and Multidimensional Integrals, Functions of Several Variables and Differential Equations*, Springer, 2020, Control de aplicaciones de integrales de una variable, secciones 3.1 y 3.3.
13. Walter Rudin, *Principles of Mathematical Analysis*, 3 ed., McGraw-Hill, 1976.
14. Michael Spivak, *Calculus*, 4 ed., Publish or Perish, 2008.
15. Terence Tao, *Analysis I*, 4 ed., Springer, 2022.
16. Daniel J. Velleman, *How to Prove It: A Structured Approach*, 2 ed., Cambridge University Press, 2006.
17. Vladimir A. Zorich, *Mathematical Analysis I*, Springer, 2004.