Fundamentos

para matemáticos

A. Peña

TEORÍA, EJERCICIOS Y PERSPECTIVA

Matemática²biertA

MATEMATICAABIERTA.CL

Fundamentos para matemáticos

A. Peña

Matemática Abierta, 2026.

Edición completa con soluciones.
A4 · Lectura digital.

Se permite copiar, distribuir y modificar este documento bajo los términos de la GNU Free Documentation License, versión 1.3 o cualquier versión posterior publicada por la Free Software Foundation; sin secciones invariantes, sin textos de cubierta delantera ni textos de cubierta trasera. El texto completo de la licencia se incluye al final del volumen.

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

Índice

Capítulo 1 — ¿Qué significa afirmar algo en matemática? 1
1.1. Los símbolos no hablan por sí solos 1
1.2. Una proposición necesita una interpretación 3
1.3. Variables libres: una frase todavía incompleta 4
1.4. Cuatro preguntas que no debemos confundir 4
1.5. Verdad y demostración no son la misma cosa 6
1.6. Definir no es afirmar un nuevo teorema 7
1.7. Negar una afirmación sin deformarla 8
1.8. Afirmaciones compuestas 8
1.9. Lo explícito y lo implícito 8
1.10. Un primer laboratorio de lectura 9
1.11. Lista de auditoría de una afirmación 9
1.12. Ejercicios 10
1.13. Soluciones desarrolladas 11
Hacia la validez 13

Capítulo 2 — ¿Cuándo una conclusión se sigue realmente de las hipótesis? 15
2.1. La verdad de las frases no basta 15
2.2. Argumento, premisas y conclusión 16
2.3. El condicional no es la inferencia 16
2.4. De un esquema correcto a una prueba correcta 17
2.5. Recíproca, contrapositiva y equivalencia 18
2.6. Condiciones necesarias y suficientes 21
2.7. Ejercicios progresivos 23
2.8. Soluciones desarrolladas 26
2.9. Hacia las variables y los cuantificadores 28

Capítulo 3 — ¿Qué cambia cuando aparecen variables y cuantificadores? 29
3.1. De una fórmula abierta a una afirmación 29
3.2. El cuantificador universal: no dejar ningún caso fuera 30
3.3. El cuantificador existencial: encontrar un testigo 31
3.4. Variables libres, variables ligadas y alcance 32
3.5. El orden de los cuantificadores y la dependencia de los testigos 34
3.6. Negación de enunciados cuantificados y control del alcance 36
3.7. Existencia y unicidad: dos exigencias distintas 39
3.8. Traducción entre prosa matemática y lenguaje cuantificado 42
3.9. Ejercicios graduados: leer, traducir, negar y reparar 46
3.10. Soluciones desarrolladas y cierre del capítulo 48

Capítulo 4 — Conjuntos, relaciones y funciones como lenguaje 54
4.1. Pertenencia, inclusión e igualdad: tres afirmaciones diferentes 54
4.2. Operaciones de conjuntos: leer cada operación como una condición de pertenencia 56
4.3. Pares ordenados y producto cartesiano: por qué importa el orden 59
4.4. Relaciones como subconjuntos de productos cartesianos: ¿qué afirma $R(x, y)$? 62
4.5. ¿Qué convierte una correspondencia en función? Existencia y unicidad por entrada 65

4.6. Imagen y preimagen: valores alcanzados y selección de entradas 68
4.7. Inyectividad, sobreyectividad, biyectividad e igualdad de funciones 72
4.8. Conjuntos solución, fórmulas y dominios de selección 77
4.9. Ejercicios graduados: de la lectura de definiciones a la elección de una prueba 81
4.10. Soluciones desarrolladas: reconstruir el razonamiento, no sólo el resultado 84

Capítulo 5 — ¿Cómo se convierten las definiciones en herramientas de prueba? 90
5.1. Una definición tiene dos direcciones de uso 90
5.2. Ejemplos, no-ejemplos y objetos que quizá no existan 91
5.3. Definiciones, resultados e hipótesis de trabajo 92
5.4. Reformular una definición exige conservar su alcance 92
5.5. Los axiomas de cuerpo: definir una estructura mediante leyes 93
5.6. Unicidad y cancelación: ordenar las dependencias 94
5.7. El cero y los signos se deducen 95
5.8. Un cuerpo no incluye automáticamente un orden 97
5.9. Ejercicios: leer una condición y decidir qué hacer con ella 98
5.10. Soluciones desarrolladas 99

Capítulo 6 — ¿Cómo se demuestra una afirmación? 104
6.1. ¿Qué obliga a hacer la forma del objetivo? 104
6.2. Prueba directa: avanzar desde lo que sabemos 110
6.3. Dos direcciones: bicondicionales y doble inclusión 117
6.4. Contraposición: cambiar el objetivo sin cambiar la afirmación 127
6.5. Prueba por contradicción: mostrar que la alternativa es imposible 134
6.6. Demostración por casos: dividir sin perder nada 143
6.7. Existencia: encontrar, construir o garantizar un objeto 149
6.8. Unicidad: existir no basta 158
6.9. Integrar arquitecturas: elegir, combinar y saber cuándo refutar 167
6.10. Ejercicios y soluciones desarrolladas 175

Capítulo 7 — ¿Cómo se descubre una demostración que todavía no conocemos? 188
7.1. Cuando conocer la arquitectura no basta 188
7.2. Trabajar hacia adelante desde las hipótesis 194
7.3. Trabajar hacia atrás desde el objetivo 202
7.4. Hacer que los dos extremos se encuentren 208
7.5. Definiciones, representaciones y objetos auxiliares 215
7.6. Construir un lema auxiliar 221
7.7. Detectar el cuello de botella y abandonar callejones 228
7.8. Explorar con ejemplos sin confundirlos con pruebas 235
7.9. Del borrador de descubrimiento a la prueba final 243
7.10. Ejercicios y soluciones desarrolladas 251

Capítulo 8 — Naturales, recursión, inducción y buen orden 268
8.1. ¿Qué tiene de especial $\mathbb{N}$? 268
8.2. El principio de inducción 276
8.3. Empezar en otro índice 282
8.4. La hipótesis inductiva: uso y límites 289
8.5. Inducción fuerte 295
8.6. Buen orden y contraejemplo mínimo 301
8.7. Cómo se relacionan los tres principios 310
8.8. Definir por recursión 318
8.9. Probar propiedades de objetos recursivos 325
8.10. Ejercicios y soluciones desarrolladas 333

Capítulo 9 — ¿Cómo sabemos que una afirmación universal es falsa? 351
9.1. ¿Qué significa refutar una afirmación universal? 352
9.2. Cómo buscar un contraejemplo 360
9.3. Recíprocos falsos y condicionales 369
9.4. Cuantificadores intercambiados y testigos 376
9.5. Dominio, operaciones ilegítimas y casos excluidos 383
9.6. Pasos irreversibles y soluciones espurias 392
9.7. Circularidad y "nonproofs" 402
9.8. Evidencia numérica y figuras 414
9.9. Reparar una afirmación o un argumento falso 424
9.10. Ejercicios y soluciones desarrolladas 436

Capítulo 10 — ¿Cómo se escribe una prueba para que otro matemático pueda verificarla? 455
10.1. Una prueba correcta también debe poder auditarse 456
10.2. Introducir objetos, variables, dominios e hipótesis 462
10.3. Notación coherente y nombres estables 469
10.4. Párrafos de prueba y transiciones lógicas 478
10.5. Ecuaciones como parte de la oración 479
10.6. Referencias a definiciones y resultados previos 481
10.7. Pronombres, alcance y ambigüedad 482
10.8. Cuánto detalle escribir y el uso de «claramente» 483
10.9. Auditoría por terceros y edición final 485
10.10. Ejercicios y soluciones desarrolladas 486

Capítulo 11 — Familias, productos, cocientes y tamaño 490
11.1. Familias indexadas: etiquetas y valores 490
11.2. Uniones e intersecciones: alguna condición o todas 491
11.3. Operaciones finitas e infinitas: controlar las condiciones 492
11.4. Productos finitos: conservar las coordenadas 493
11.5. Productos indexados: un elemento es una selección completa 494
11.6. Equivalencias y particiones: agrupar sin solapamientos 495
11.7. Cocientes: comprobar independencia del representante 495
11.8. Igual tamaño, finitud e infinitud 497
11.9. Numerabilidad y convenciones de tamaño 499
11.10. Ejercicios y soluciones desarrolladas 500

Capítulo 12 — ¿Cómo puede una colección convertirse en un número? 505
12.1. Cuando todavía no hemos dado nombre a una cantidad 505
12.2. Primer momento de construcción: los naturales como clases de colecciones finitas 506
12.3. Hacer aritmética con los naturales que acabamos de construir 507
12.4. Definir una operación no demuestra sus leyes 509
12.5. Las leyes de los naturales nacen de las correspondencias 510
12.6. Una equivalencia *entre* naturales: los restos no construyen los naturales 512
12.7. Ejercicios 513
12.8. Soluciones desarrolladas 515

Capítulo 13 — ¿De dónde salen los enteros? 518
13.1. Segundo momento de construcción: los enteros a partir de los naturales 518
13.2. Los enteros heredan unas leyes y adquieren otras 521
13.3. El orden de los enteros y el sentido de la ampliación 523
13.4. Las congruencias en los enteros: construir un cociente no es ampliar el sistema 524
13.5. Ejercicios 525
13.6. Soluciones desarrolladas 526

Capítulo 14 — ¿Por qué las fracciones forman un cuerpo ordenado? 530
14.1. Tercer momento de construcción: los racionales a partir de los enteros 530
14.2. Dividir dentro de los racionales 532
14.3. Las leyes del cuerpo: lo que las fórmulas deben demostrar 534
14.4. Ordenar las clases sin dejarse engañar por los signos 535
14.5. Una recta densamente poblada que todavía tiene una ausencia 537
14.6. Lo que significa haber construido los números 538
14.7. Ejercicios 539
14.8. Soluciones desarrolladas 541

Epílogo — Cuando la aproximación pide nuevos números 544

Apéndice A — Glosario lógico y demostrativo 546

Apéndice B — Formas de objetivo y estrategias iniciales 548

Apéndice C — Errores frecuentes y pruebas defectuosas 549

Apéndice D — Directorio de soluciones desarrolladas 550

GNU Free Documentation License 551

Capítulo 1 — ¿Qué significa afirmar algo en matemática?

Durante muchos años puede estudiarse matemática sin detenerse a preguntar qué clase de cosas son las frases que escribimos. Calculamos

$$2 + 3 = 5,$$

resolvemos ecuaciones como

$$x + 3 = 5,$$

y aprendemos reglas que empiezan con expresiones del tipo «si $x > 0$» o «para todo número real x». Mientras el objetivo principal es calcular, esas diferencias parecen secundarias. Sabemos más o menos qué hay que hacer y seguimos adelante.

Pero en cuanto una demostración entra en escena, esa tolerancia deja de ser suficiente.

Una prueba no consiste simplemente en encadenar símbolos correctos. Consiste en establecer que una afirmación se sigue de ciertas condiciones mediante pasos cuyo significado pueda ser revisado por otra persona. Para eso necesitamos aprender a distinguir cosas que en la práctica escolar suelen aparecer mezcladas: una **expresión**, una **afirmación**, una frase que depende de una variable, una **definición**, una hipótesis temporal y una conclusión.

La primera pregunta de este libro será, por tanto,

¿qué significa afirmar algo en matemática?

No comenzaremos con una lista completa de símbolos lógicos. Antes de aprender a manipularlos necesitamos comprender qué pretenden registrar.

1.1. Los símbolos no hablan por sí solos

Comparemos cuatro cadenas:

$$2 + 3,$$

$$2 + 3 = 5,$$

$$x + 3 = 5,$$

y la frase

Sea n un entero par.

Se parecen porque todas pertenecen al lenguaje matemático. Sin embargo, no cumplen la misma función.

La primera,

$$2 + 3,$$

es una **expresión**. Podemos evaluarla y obtener 5, pero la expresión no afirma por sí sola que algo sea verdadero o falso. Del mismo modo,

$$x^2 + 1$$

es una expresión: una vez fijado el valor de x, determina un valor, pero no formula todavía una afirmación.

La segunda cadena,

$$2 + 3 = 5,$$

sí afirma algo. Bajo la interpretación usual de los números naturales y de los símbolos $+$ y $=$, la afirmación es verdadera.

La tercera,

$$x + 3 = 5,$$

ocupa una posición intermedia. Parece una afirmación, pero su verdad depende de qué sea x. Si $x = 2$, resulta verdadera; si $x = 4$, resulta falsa. Mientras no fijemos, asignemos o cuantifiquemos x, la variable permanece **libre**.

La cuarta frase tampoco pretende describir un hecho consumado. «Sea n un entero par» introduce un objeto del que hablaremos a continuación y fija una condición bajo la cual continuará el argumento.

Estas diferencias son pequeñas en apariencia, pero gobiernan la arquitectura de las pruebas.

Expresiones y afirmaciones. Llamaremos **expresión matemática** a una combinación bien formada de símbolos que, dentro de una interpretación fijada, denota o representa un objeto matemático. Por ejemplo,

$$3, \qquad 2 + 5, \qquad x^2 - 1, \qquad A \cap B$$

son expresiones en contextos adecuados.

Llamaremos **proposición** a un enunciado declarativo interpretado para el que se ha fijado un contexto suficiente para determinar un valor de verdad. Ese contexto incluye los significados de los símbolos y, cuando corresponda, los valores de las variables o su cuantificación. No confundiremos la posibilidad de determinar su valor de verdad con el hecho de conocer ya cuál es.

Esta elección terminológica debe manejarse con cuidado. Distintos libros usan «enunciado», «sentencia», «fórmula» y «proposición» con convenciones ligeramente distintas. En este tratado adoptaremos la siguiente convención de trabajo:

- **enunciado** será el término amplio para una formulación declarativa matemática;
- un enunciado interpretado y **cerrado** —sin variables libres pendientes de asignación en el contexto— que afirma algo susceptible de ser verdadero o falso será una **proposición**;
- un enunciado que todavía contiene variables libres será una **fórmula abierta**; cuando queremos destacar que expresa una propiedad de esas variables, hablaremos también de **predicado**.

Así,

$$x + 3 = 5$$

no será tratado todavía como una proposición mientras x permanezca libre. En cambio,

$$\exists x \in \mathbb{R}\ (x + 3 = 5)$$

sí es una proposición. El discurso «sea x un real» fija el dominio de x, pero no le asigna automáticamente un valor particular: las afirmaciones posteriores se leerán bajo ese contexto.

Al aislar $x + 3 = 5$ del discurso, sigue siendo una fórmula abierta. La diferencia entre una frase aislada y su uso en una demostración merece atención.

Esta convención no pretende imponer una terminología universal. Su función es mantener una distinción estable a lo largo del libro: algunas cadenas **denotan objetos**, otras **expresan propiedades todavía abiertas** y otras **afirman algo ya cerrado**.

Prueba de diagnóstico

Ante una línea matemática, pregúntate primero:

> **¿Esta línea nombra o construye un objeto, o está diciendo que algo es el caso?**

Esa pregunta elemental evita muchos errores posteriores.

1.2. Una proposición necesita una interpretación

Consideremos

$$1 + 1 = 2.$$

Solemos decir que es verdadera sin añadir nada. Y en la práctica ordinaria está bien hacerlo: suponemos la aritmética usual. Pero la verdad no está adherida físicamente a los signos impresos. Necesitamos saber qué representan 1, 2, + y =.

La interpretación no fija sólo qué objetos representan los símbolos individuales. También fija las **operaciones** y **relaciones**. El signo + puede representar suma de enteros, suma de vectores o suma de funciones; el mismo dibujo tipográfico puede desempeñar papeles distintos en estructuras distintas. Por eso leer una fórmula no consiste únicamente en reconocer sus caracteres: consiste en saber qué estructura les da significado.

Esta observación puede parecer filosófica, pero pronto se vuelve técnica. Consideremos ahora:

> Todo elemento no nulo tiene inverso multiplicativo.

¿Es verdadera?

No podemos responder todavía. Si el dominio es $\mathbb{Z}$, la afirmación es falsa: 2 es un entero no nulo y no existe $m \in \mathbb{Z}$ tal que

$$2m = 1.$$

Si el dominio es $\mathbb{Q}$, la afirmación es verdadera: para todo racional $q \neq 0$, el racional $1/q$ satisface

$$q \cdot \frac{1}{q} = 1.$$

La frase no cambió. Cambió el universo en el que la interpretamos.

Dominio. El **dominio** especifica qué objetos pueden ocupar las variables o entrar en una afirmación. A veces se escribe explícitamente:

$$\forall x \in \mathbb{R},$$

pero otras veces queda implícito en el contexto. Una parte importante de aprender a leer matemática consiste en recuperar esos datos implícitos sin inventarlos.

El dominio no es un detalle tipográfico. Puede cambiar el valor de verdad de una proposición, el sentido de una definición e incluso la legitimidad de una expresión.

Por ejemplo,

$$x^2 \geq 0$$

es una **fórmula abierta** natural cuando $x \in \mathbb{R}$ y $\geq$ representa el orden usual. La proposición universal correspondiente es $\forall x \in \mathbb{R},\ x^2 \geq 0$. No debemos trasladar esa escritura mecánicamente a $\mathbb{C}$: allí no estamos trabajando con ese mismo orden.

Una costumbre peligrosa

En matemática elemental muchas hipótesis permanecen invisibles porque todos conocen el escenario. En matemática avanzada los escenarios se multiplican. Una misma palabra —«inverso», «positivo», «continuo», «cerrado»— puede depender de la estructura en la que estamos trabajando.

Leer bien significa aprender a preguntar **¿en qué mundo es verdadera esta frase?**

1.3. Variables libres: una frase todavía incompleta

Volvamos a

$$x^2 = 4.$$

¿Es verdadera o falsa?

La respuesta correcta, tal como está escrita, es que todavía no tenemos una proposición cerrada: x permanece libre.

Podemos convertirla en una proposición de varias maneras.

Asignando un valor:

$$2^2 = 4.$$

Cuantificando existencialmente:

$$\exists x \in \mathbb{R}\ (x^2 = 4).$$

O cuantificando universalmente:

$$\forall x \in \mathbb{R}\ (x^2 = 4).$$

Las tres frases se construyen a partir del mismo núcleo $x^2 = 4$, pero dicen cosas diferentes. La primera es verdadera, la segunda también y la tercera es falsa.

Esto muestra que una variable no es simplemente un espacio vacío donde «podemos poner cualquier número». Su estatus lógico importa. Una variable puede estar libre, puede haber recibido un valor, puede quedar ligada por un cuantificador o puede haber sido introducida bajo una hipótesis.

En el capítulo 3 estudiaremos esta arquitectura con precisión. Por ahora basta una regla de lectura:

si una variable aparece, pregunta quién la controla.

1.4. Cuatro preguntas que no debemos confundir

Hasta aquí ya podemos separar cuatro preguntas que en la práctica suelen mezclarse:

1. **¿Está bien formada esta escritura?**
   Es una pregunta acerca de la forma: si los símbolos están combinados de acuerdo con las reglas del lenguaje que estamos usando.
2. **¿Qué significa?**
   Es una pregunta de interpretación: qué representan los símbolos, cuál es el dominio y qué operaciones o relaciones están en juego.
3. **¿Es verdadera?**
   Es una pregunta acerca del valor de verdad de una proposición ya interpretada.

4. **¿Por qué es verdadera?**
   Es una pregunta de justificación: qué definiciones, hipótesis y resultados permiten establecerla.

Consideremos, por ejemplo,

$$\forall x \in \mathbb{R}, \qquad x^2 + 1 > 0.$$

La fórmula está bien formada; su interpretación usual fija el dominio $\mathbb{R}$, la multiplicación, la suma y el orden; la proposición es verdadera; y todavía queda una cuarta tarea: justificarla. Una posible razón es que $x^2 \geq 0$ para todo real x, de modo que

$$x^2 + 1 \geq 1 > 0.$$

Estas cuatro preguntas pertenecen a niveles diferentes. Una demostración puede estar escrita con símbolos perfectamente bien formados y, sin embargo, apoyarse en una interpretación equivocada. Una proposición puede ser verdadera, pero el argumento ofrecido para justificarla puede ser inválido. Y una expresión puede tener un significado perfectamente claro sin afirmar nada que pueda ser verdadero o falso.

Cuatro niveles de lectura

Cuando una línea matemática produzca confusión, no preguntes solamente «¿la entiendo?». Descompón el problema:

forma → interpretación → verdad → justificación

Gran parte de la madurez matemática consiste en saber en cuál de estos niveles se encuentra la dificultad.

Qué significa que una expresión esté bien definida. Consideremos ahora $1/x$. Los símbolos pueden estar correctamente escritos y podemos entender que se pide el inverso de x. Aun así, falta comprobar que ese inverso exista para los valores permitidos. Sobre los números reales, la expresión tiene un valor para cada $x \neq 0$, pero no para $x = 0$. Por eso necesitamos declarar el dominio o una hipótesis que excluya ese caso.

Diremos que una expresión que pretende designar un objeto está **bien definida en un contexto** cuando, para cada elección admisible de sus parámetros, determina un único objeto del tipo anunciado. El contexto comprende el dominio, la interpretación de las operaciones y las hipótesis disponibles. Si la descripción permite distintas representaciones de un mismo objeto, el resultado debe ser independiente de cuál se use.

La condición se comprueba para cada elección de los parámetros; los resultados pueden variar de una elección a otra. Así, $1/x$ toma valores diferentes en $x = 1$ y $x = 2$, y está bien definida en ambos casos. La unicidad exige que una misma entrada, bajo las mismas condiciones, tenga un valor determinado.

Veamos tres dificultades que conviene distinguir.

Una operación fuera de su dominio. Sobre los reales, $1/(x-1)$ está bien definida bajo la hipótesis $x > 1$, porque entonces $x - 1 > 0$ y el denominador no se anula. La hipótesis $x \geq 1$ permite $x = 1$, para el cual la división no está definida. Leer la fórmula no basta: hay que verificar que las operaciones utilizadas admitan los datos.

Una descripción que no determina un único objeto. La frase «el número real cuyo cuadrado es 4» deja dos posibilidades, 2 y -2. La existencia de soluciones no resuelve la ambigüedad. En cambio, «el número real no negativo cuyo cuadrado es 4» determina 2: de $y^2 = 4$ se sigue $(y-2)(y+2) = 0$, de modo que $y = 2$ o $y = -2$, y la condición $y \geq 0$ excluye la segunda posibilidad. Esta es la convención que da a $\sqrt{4}$ su valor 2; la ecuación $y^2 = 4$ conserva sus dos soluciones.

Una regla que depende de la representación. Supongamos que intentamos asignar a cada número racional el numerador de una fracción que lo representa. Como $1/2 = 2/4$, el

mismo racional recibiría 1 si usamos la primera fracción y 2 si usamos la segunda. La regla propuesta no determina un valor para el racional: determina un valor para la escritura elegida. Para que la asignación esté bien definida como regla sobre racionales, debemos garantizar que dos fracciones que representan el mismo número produzcan el mismo resultado.

Tener un valor determinado y tener un valor de verdad son exigencias distintas. La expresión $1/2$ designa un número; la proposición $1/2 = 3$ es falsa, aunque las expresiones que aparecen en ella estén bien definidas. Asimismo, una definición escrita con claridad puede exigir una prueba de existencia, unicidad o independencia de la representación antes de autorizar el uso del objeto definido.

Comprobar la buena definición

Antes de usar una expresión para designar un objeto, verifica que las operaciones admitan los datos y que la descripción determine un único resultado del tipo anunciado. Si intervienen representaciones o elecciones auxiliares, comprueba que el resultado no dependa de ellas.

Comprobación de lectura. Una descripción propone asignar a cada entero n «el menor entero estrictamente mayor que n». Otra propone asignar a cada real x «el menor real estrictamente mayor que x». ¿Determinan ambas descripciones una regla bien definida en sus respectivos dominios? No basta responder que siempre existen números mayores.

Respuesta razonada. La primera sí: para cada $n \in \mathbb{Z}$, el entero $n + 1$ es mayor que n, y ningún entero se encuentra estrictamente entre n y $n + 1$. Es, por tanto, el menor entero mayor que n, y es único. La segunda falla por **inexistencia del objeto anunciado**, no por una dificultad para calcularlo: si $y > x$ es un real candidato, el número $(x + y)/2$ satisface $x < (x + y)/2 < y$. Siempre existe otro real mayor que x y menor que y. Ningún candidato puede ser el menor. La misma descripción verbal cambia de resultado al cambiar el dominio.

En el capítulo 4 (§4.5) formularemos esta exigencia para funciones; en el capítulo 6 (§6.7), para las fórmulas que construyen testigos; y en el capítulo 11 (§11.7), para reglas sobre clases de equivalencia.

1.5. Verdad y demostración no son la misma cosa

Supongamos que queremos saber si 91 es primo.

Podemos ensayar divisiones y descubrir que

$$91 = 7 \cdot 13.$$

Por tanto, la proposición

91 es primo

es falsa.

Ahora consideremos

97 es primo.

En este caso aritmético elemental, una vez fijada la interpretación usual, la proposición tiene un valor de verdad independientemente de que nosotros hayamos terminado o no la verificación. Nuestra ignorancia momentánea no altera aquello que la frase afirma. Lo que cambia es nuestro **estado de conocimiento** y la justificación de la que disponemos.

Más adelante encontraremos situaciones fundacionales más sutiles en las que convendrá distinguir cuidadosamente **verdad en una estructura**, **demostrabilidad desde un sistema de axiomas** y **conocimiento efectivo de una prueba**. No necesitamos todavía esa teoría; basta con no identificar automáticamente estos conceptos.

Esta distinción es crucial:

- la **verdad** pertenece a aquello que la proposición afirma dentro del contexto matemático fijado;

- una **demostración** es una justificación que establece esa afirmación a partir de definiciones, hipótesis y resultados admitidos.

No definiremos todavía formalmente qué cuenta como demostración. Lo haremos progresivamente. Pero ya podemos descartar dos confusiones frecuentes.

Primera confusión: «lo comprobé en muchos ejemplos, entonces es verdadero». Comprobar ejemplos puede sugerir una conjetura. No establece una afirmación universal.

Por ejemplo, para $n = 1, 2, 3, 4, 5$ podemos verificar que

$$n^2 + n$$

es par. Cinco casos positivos no demuestran todavía que esto ocurra para todo entero positivo.

La afirmación general requiere un argumento que abarque todos los casos. En este ejemplo existe uno muy breve:

$$n^2 + n = n(n + 1).$$

Los enteros n y $n + 1$ son consecutivos, así que uno de ellos es par; por tanto, su producto es par.

El salto conceptual es importante. La prueba no consiste en revisar más valores de n, sino en encontrar una razón que abarque **cualquier** n del dominio.

Segunda confusión: «si es verdadero, debería ser evidente». Tampoco. Una afirmación puede ser verdadera y requerir una idea profunda para demostrarla. La facilidad con la que reconocemos una verdad no forma parte de su contenido matemático.

1.6. Definir no es afirmar un nuevo teorema

Comparemos:

Un entero n es **par** si existe $k \in \mathbb{Z}$ tal que $n = 2k$.

con

6 es par.

La primera frase fija el significado que utilizaremos para una palabra. La segunda usa ese significado para afirmar algo acerca de un objeto particular.

Para verificar la segunda, desplegamos la definición. Necesitamos exhibir un entero k con

$$6 = 2k.$$

Podemos elegir $k = 3$.

Este pequeño ejemplo contiene un mecanismo que aparecerá miles de veces en matemática avanzada:

$$\boxed{\text{definición} \longrightarrow \text{condición verificable.}}$$

Más adelante, cuando una prueba diga «como f es continua», «como G es un grupo» o «como A es compacto», una de las primeras preguntas estratégicas será: **¿qué me permite escribir exactamente esa definición?**

Las definiciones no son adornos introductorios que olvidamos después de leerlas. Son instrumentos de trabajo.

1.7. Negar una afirmación sin deformarla

Si

n es par

es una proposición en un contexto determinado, su negación es simplemente

n no es par.

¿Podemos sustituirla por «n es impar»?

Si sabemos que $n \in \mathbb{Z}$ y usamos las definiciones usuales, sí: todo entero es par o impar, y no ambas cosas a la vez.

Pero observa lo que ocurrió. Para pasar de «no par» a «impar» utilizamos información adicional sobre el dominio y una propiedad de los enteros. La negación lógica por sí sola no inventa esa información.

Esta disciplina será especialmente importante cuando aparezcan cuantificadores. Negar una proposición no consiste en escribir una frase que *suene contraria*; consiste en construir exactamente las condiciones bajo las cuales la proposición original falla.

1.8. Afirmaciones compuestas

Las proposiciones pueden combinarse. Si P y Q representan proposiciones, podemos formar, entre otras,

- «P y Q»;
- «P o Q»;
- «no P»;
- «si P, entonces Q»;
- «P si y sólo si Q».

Estas construcciones se denominan **conectivos lógicos**. En los capítulos siguientes estudiaremos con más cuidado qué significan y cómo gobiernan la forma de una prueba.

Por ahora nos interesa una idea previa: cuando combinamos proposiciones, el valor de verdad de la expresión compuesta depende de los valores de verdad de sus componentes y del significado del conectivo.

Por ejemplo, si

$$P: \quad 8 \text{ es par}$$

y

$$Q: \quad 8 \text{ es primo},$$

entonces P es verdadera y Q es falsa. Por tanto,

P y Q

es falsa, mientras que

P o Q

es verdadera si usamos el «o» inclusivo habitual de la lógica matemática.

Las tablas de verdad sistematizan este tipo de análisis. Serán útiles, pero no deben confundirse con el razonamiento matemático completo. Una tabla de verdad puede analizar una forma proposicional finita; no sustituye una demostración acerca del contenido matemático que aparece dentro de sus componentes.

1.9. Lo explícito y lo implícito

Tomemos la frase

La ecuación $x^2 = 1$ tiene exactamente dos soluciones.

Un lector puede reconocer de inmediato las soluciones 1 y -1. Pero antes de aceptar la afirmación debe preguntar: ¿soluciones **dónde**?

Sobre $\mathbb{R}$ la afirmación es verdadera. Sobre $\mathbb{Z}$ también. Si buscamos soluciones únicamente entre los enteros positivos, sólo encontramos 1, de modo que resulta falsa. El problema no está en la frase como español ordinario, sino en que el discurso matemático exige controlar los objetos sobre los que cuantificamos.

Muchas demostraciones incorrectas nacen de un cambio de escenario no declarado. Se empieza con números reales, se divide por una cantidad cuya no nulidad no se justificó, se extrae una raíz con una convención no fijada o se utiliza una propiedad que pertenece a un dominio más rico.

Por eso una prueba madura mantiene un inventario silencioso:

1. ¿qué objetos están en juego?;
2. ¿en qué conjuntos o estructuras viven?;
3. ¿qué operaciones están permitidas?;
4. ¿qué propiedades ya han sido establecidas?;
5. ¿qué variables están libres y cuáles han sido fijadas?;
6. ¿qué se está suponiendo temporalmente?

No siempre escribiremos ese inventario completo. Pero debemos poder reconstruirlo.

1.10. Un primer laboratorio de lectura

Consideremos el siguiente argumento:

> Sea n un entero. Si n es múltiplo de 4, entonces n es par. Como 6 es par, concluimos que 6 es múltiplo de 4.

Algo ha fallado.

La primera afirmación es verdadera. Si $n = 4k$, entonces

$$n = 2(2k),$$

por lo que n es par.

También es verdadera la segunda afirmación: 6 es par.

Y, sin embargo, la conclusión es falsa.

Este ejemplo será el punto de partida del capítulo 2. Nos obliga a separar dos preguntas:

1. ¿son verdaderas las afirmaciones individuales?;
2. ¿la conclusión se sigue realmente de las premisas?

La segunda pregunta introduce la noción de **validez**. Descubriremos que un argumento no se vuelve correcto por estar formado por frases verdaderas, del mismo modo que una frase falsa dentro de un argumento no basta por sí sola para decir que la forma inferencial es inválida.

1.11. Lista de auditoría de una afirmación

Al terminar este primer capítulo, ante una línea nueva conviene entrenar las siguientes preguntas:

- **¿Qué función cumple?** ¿Es una expresión, una proposición, una definición, una hipótesis, una instrucción o una conclusión?
- **¿Qué significan sus símbolos?**
- **¿Cuál es el dominio?**
- **¿Hay variables libres?**
- **¿Qué datos están implícitos en el contexto?**
- **¿La frase tiene ya un valor de verdad determinado?**
- **¿Estoy confundiendo evidencia experimental con demostración?**

- **¿Estoy usando una definición como herramienta o sólo reconociendo su nombre?**

Estas preguntas no son burocracia lógica. Son una forma de leer más despacio para poder razonar más rápido después.

Qué no hemos hecho todavía

Este capítulo no ha definido formalmente la consecuencia lógica ni la validez; tampoco ha desarrollado la teoría de cuantificadores ni la noción axiomática de estructura. Hemos preparado el vocabulario que permitirá hacerlo sin manipular símbolos a ciegas.

En particular:

- la **validez** será el problema del capítulo 2;
- los **predicados y cuantificadores**, el del capítulo 3;
- las **estructuras y definiciones axiomáticas** aparecerán progresivamente, con los axiomas de cuerpo como primer laboratorio sostenido en el capítulo 5.

1.12. Ejercicios

Ejercicio 1. Clasifica cada elemento como expresión, proposición, predicado con variable libre o instrucción de trabajo:

1. $3^2 + 4^2$;
2. $3^2 + 4^2 = 5^2$;
3. $x^2 + 1 = 0$;
4. «Sea $x \in \mathbb{R}$ con $x > 0$»;
5. «Todo cuadrado de un número real es no negativo».

Ejercicio 2. Explica por qué «todo elemento no nulo tiene inverso multiplicativo» es falsa en $\mathbb{Z}$ y verdadera en $\mathbb{Q}$.

Ejercicio 3. Convierte $x^2 = 4$ en tres proposiciones distintas: una mediante asignación y dos mediante cuantificadores. Determina sus valores de verdad sobre $\mathbb{R}$.

Ejercicio 4. Un estudiante verifica que $n^2 + n$ es par para diez valores consecutivos de n y afirma que ya ha demostrado el caso general. Explica el error y proporciona una demostración válida.

Ejercicio 5. Niega «n es par». Explica qué hipótesis permite reemplazar esa negación por «n es impar».

Ejercicio 6. Analiza qué información falta para interpretar correctamente $x^2 \geq 0$. Formula a continuación una proposición universal verdadera que utilice esa misma desigualdad.

Ejercicio 7. Distingue la función de estas dos frases:

1. «Un entero n es par si existe $k \in \mathbb{Z}$ tal que $n = 2k$».
2. «18 es par».

Después justifica la segunda usando la primera.

Ejercicio 8. Un compañero dice: «Si todas las premisas de un argumento son verdaderas, entonces la conclusión debe ser verdadera». El argumento sobre 6 al final de §1.10 muestra que esta frase necesita una condición adicional. ¿Cuál crees que será esa condición? Formula tu respuesta antes de leer el capítulo 2.

Ejercicio 9. Clasifica cada caso usando la convención terminológica del capítulo: **expresión**, **fórmula abierta/predicado**, **proposición** o **instrucción/definición**. Justifica cada decisión.

1. $x^2 + 1$;
2. $x^2 + 1 = 0$;
3. $\exists x \in \mathbb{R}\ (x^2 + 1 = 0)$;
4. «Definimos $f : \mathbb{R} \to \mathbb{R}$ mediante $f(x) = x^2 + 1$»;
5. «Para todo $x \in \mathbb{R}$, $x^2 + 1 > 0$».

Ejercicio 10. Para la escritura

$$\frac{1}{x} > 0,$$

separa las cuatro preguntas de §1.4: forma, interpretación, verdad y justificación. Explica por qué no puede responderse simplemente «es verdadera» o «es falsa» mientras no se complete el contexto.

Ejercicio 11. Un estudiante observa que

$$2^2, \qquad 4^2, \qquad 6^2$$

son pares y concluye: «el cuadrado de todo entero es par». Encuentra un contraejemplo y formula una afirmación cercana que sí sea verdadera. Demuéstrala desplegando la definición de número par.

Ejercicio 12. Considera la cadena

$$a + b = b + a.$$

Explica qué información debe fijarse antes de tratarla como una proposición. Luego conviértela en una proposición verdadera sobre $\mathbb{Z}$. Para comprobar que no basta especificar el conjunto, define otra operación sobre $\mathbb{Z}$ mediante $a \star b = a$ para cualesquiera $a, b \in \mathbb{Z}$: ¿es verdad que $a \star b = b \star a$ para todos los enteros?

1.13. Soluciones desarrolladas

Solución del ejercicio 1.

1. $3^2 + 4^2$ es una expresión: denota un número.
2. $3^2 + 4^2 = 5^2$ es una proposición y es verdadera bajo la aritmética usual.
3. $x^2 + 1 = 0$ contiene una variable libre. Su verdad depende del valor y del dominio asignados a x.
4. «Sea $x \in \mathbb{R}$ con $x > 0$» funciona como instrucción: introduce un objeto y una hipótesis para el razonamiento posterior; no afirma por sí sola un teorema.
5. Es una proposición universal expresada en lenguaje natural; su dominio está explícitamente fijado en los números reales.

Solución del ejercicio 2. En $\mathbb{Z}$, toma el elemento 2. Si tuviera inverso multiplicativo entero, existiría $m \in \mathbb{Z}$ tal que $2m = 1$, lo que es imposible. Por tanto, la afirmación es falsa en $\mathbb{Z}$.

En $\mathbb{Q}$, sea $q \neq 0$. Entonces $1/q \in \mathbb{Q}$ y

$$q\left(\frac{1}{q}\right) = 1.$$

Así, todo racional no nulo posee inverso multiplicativo racional.

Solución del ejercicio 3. Una asignación posible es

$$2^2 = 4,$$

que es verdadera. Las cuantificaciones pueden ser

$$\exists x \in \mathbb{R}\ (x^2 = 4),$$

verdadera, y

$$\forall x \in \mathbb{R}\ (x^2 = 4),$$

falsa. El mismo núcleo con variable libre produce afirmaciones distintas cuando cambia la forma de cerrar la variable.

Solución del ejercicio 4. Diez casos sólo muestran que la fórmula funciona en esos diez casos. Para demostrar la afirmación universal, sea $n \in \mathbb{Z}$. Entonces

$$n^2 + n = n(n+1).$$

Los enteros consecutivos n y $n+1$ no pueden ser ambos impares; uno de ellos es par. Por consiguiente, el producto $n(n+1)$ es par. Como n era arbitrario, $n^2 + n$ es par para todo entero n.

Solución del ejercicio 5. La negación estricta es «n no es par». Si además sabemos que $n \in \mathbb{Z}$ y usamos la dicotomía usual de los enteros entre pares e impares, podemos concluir equivalentemente que n es impar.

Solución del ejercicio 6. Necesitamos especificar al menos el dominio de x y el significado de la relación $\geq$. Incluso después de fijar $x \in \mathbb{R}$ y el orden usual, la cadena aislada $x^2 \geq 0$ conserva una variable libre: es una fórmula abierta, no una proposición cerrada. Una proposición universal verdadera se obtiene al escribir $\forall x \in \mathbb{R},\ x^2 \geq 0$. En efecto, el cuadrado de un real es no negativo en el orden usual. No debemos suponer que la misma notación tiene automáticamente el mismo sentido en cualquier estructura.

Solución del ejercicio 7. La primera frase es una definición: fija qué significa «par» para enteros. La segunda es una proposición. Para justificarla, basta exhibir un entero k tal que

$$18 = 2k.$$

Tomando $k = 9$, obtenemos $18 = 2 \cdot 9$, por lo que 18 es par.

Solución del ejercicio 8. La condición que falta es que la conclusión **se siga válidamente** de las premisas. No basta con que las premisas sean verdaderas de hecho. El capítulo siguiente convertirá esta intuición en una noción precisa de argumento y validez.

Solución del ejercicio 9.

1. $x^2 + 1$ es una **expresión**: representa un objeto una vez interpretados x, la multiplicación y la suma.
2. $x^2 + 1 = 0$ es una **fórmula abierta** o un **predicado** en la variable x: tiene forma declarativa, pero x sigue libre.
3. $\exists x \in \mathbb{R}\ (x^2 + 1 = 0)$ es una **proposición**: la variable está ligada y el dominio está fijado. Sobre $\mathbb{R}$ es falsa.
4. «Definimos $f : \mathbb{R} \to \mathbb{R}$ mediante $f(x) = x^2 + 1$» funciona como una **instrucción definitoria**: especifica dominio, codominio y regla; no es un nuevo teorema.
5. «Para todo $x \in \mathbb{R}$, $x^2 + 1 > 0$» es una **proposición universal** y es verdadera.

La distinción decisiva entre (2) y (3) no está en la ecuación interior, que es la misma, sino en el estatus de la variable.

Solución del ejercicio 10. La escritura $1/x > 0$ puede analizarse así:

1. **Forma:** la cadena tiene la forma sintáctica de una desigualdad construida con división; escribirla no garantiza por sí solo que todos sus términos estén definidos bajo cualquier asignación.
2. **Interpretación:** debemos fijar qué objetos puede tomar x, qué significa la división y qué relación representa $>$. Además, $x = 0$ debe excluirse para que $1/x$ esté definido en los contextos numéricos usuales.
3. **Verdad:** mientras x permanezca libre, no tenemos todavía una proposición única. Sobre $\mathbb{R}$, por ejemplo, la fórmula es verdadera para $x > 0$ y falsa para $x < 0$.
4. **Justificación:** sólo después de cerrar la variable —por asignación, hipótesis o cuantificación— tiene sentido pedir una demostración de la proposición resultante.

El ejercicio muestra que una pregunta de verdad puede estar mal planteada porque faltan pasos previos de interpretación y cierre de variables.

Solución del ejercicio 11. La afirmación universal propuesta es falsa. Basta el contraejemplo

$$3^2 = 9,$$

que es impar. Una afirmación cercana y verdadera es:

El cuadrado de todo entero par es par.

Sea n un entero par. Por definición existe $k \in \mathbb{Z}$ tal que $n = 2k$. Entonces

$$n^2 = (2k)^2 = 4k^2 = 2(2k^2).$$

Como $2k^2 \in \mathbb{Z}$, la definición de par implica que n^2 es par. Aquí el paso esencial fue sustituir una colección de ejemplos por una razón que utiliza la definición y vale para un entero par arbitrario.

Solución del ejercicio 12. Tal como aparece, $a + b = b + a$ contiene variables libres. Además necesitamos saber en qué dominio viven a y b y qué operación representa $+$. Una forma de cerrarla como proposición es

$$\forall a, b \in \mathbb{Z}, \qquad a + b = b + a,$$

con $+$ interpretado como la suma usual de enteros. Esta proposición es verdadera.

La cadena original ocultaba, al menos, tres datos: el dominio, la interpretación de la operación y la cuantificación universal de las variables. La notación familiar hacía que esos datos parecieran innecesarios, pero estaban trabajando silenciosamente.

Para comprobar que la operación importa, definamos $a \star b = a$ en $\mathbb{Z}$. Si $a = 1$ y $b = 2$, entonces $a \star b = 1$, mientras que $b \star a = 2$. Por tanto, $a \star b = b \star a$ no vale para todos los enteros. Haber conservado el dominio no fue suficiente: cambió la operación.

Hacia la validez

Ya podemos distinguir una expresión de una afirmación, reconocer el papel del dominio, detectar variables libres y separar una definición de una proposición. También hemos visto que la verdad de frases individuales no basta para garantizar que un razonamiento sea correcto.

La pregunta siguiente es inevitable:

¿cuándo una conclusión se sigue realmente de las hipótesis?

Ése será el problema del capítulo 2.

Capítulo 2 — ¿Cuándo una conclusión se sigue realmente de las hipótesis?

En el capítulo anterior aprendimos a distinguir lo que una expresión representa de lo que un enunciado afirma. También vimos que una afirmación puede ser verdadera y, sin embargo, estar acompañada por una justificación incorrecta. Éste es el problema que debemos resolver ahora: no basta examinar una por una las frases de un argumento. Tenemos que estudiar **la relación que permite pasar de unas a otra**.

Supongamos que un estudiante escribe:

> Todo múltiplo de 4 es par. El número 6 es par. Por tanto, 6 es múltiplo de 4.

No ha escrito una frase sin sentido. La primera premisa es verdadera; la segunda también. Ha cometido un error más interesante: las premisas no autorizan la conclusión. Para comprender por qué, necesitamos distinguir el contenido particular de las frases y la estructura del razonamiento.

La pregunta rectora del capítulo será

¿cuándo una conclusión se sigue realmente de las hipótesis?

2.1. La verdad de las frases no basta

Definamos provisionalmente dos enunciados sobre un entero particular n:

$$P: \quad n \text{ es múltiplo de } 4, \qquad Q: \quad n \text{ es par.}$$

Una vez fijado n, podemos examinar sus valores de verdad. La primera premisa del argumento dice «si P, entonces Q»; la segunda afirma Q; la conclusión pretende afirmar P. Su forma es

$$\frac{\begin{array}{l} P \to Q, \\ Q, \end{array}}{P.}$$

El símbolo $\to$ de este esquema representa aquí el condicional que aparece **dentro** de una premisa, y la raya horizontal separa las premisas de la conclusión. Más adelante distinguiremos expresamente ese condicional de la relación «se sigue de».

Tomemos $n = 6$. Entonces P es falsa y Q es verdadera. La premisa condicional sigue siendo verdadera: afirma que ser múltiplo de 4 basta para ser par; no afirma que el 6 sea múltiplo de 4. Por tanto, tenemos premisas verdaderas y conclusión falsa.

Podríamos cambiar el ejemplo y tomar $n = 8$. En ese caso ambas premisas y la conclusión serían verdaderas. ¿Habremos reparado la inferencia? No. Su conclusión acertaría en ese caso, pero el procedimiento usado para obtenerla sería el mismo. Una inferencia no queda justificada por acertar ocasionalmente.

El contraste entre 6 y 8 deja una primera lección: **la verdad efectiva de la conclusión y la corrección del camino que conduce a ella son dos cuestiones distintas**.

2.2. Argumento, premisas y conclusión

Un **argumento** consta de premisas y de una conclusión que se propone obtener a partir de ellas. Por ahora trabajaremos con listas finitas de premisas formuladas mediante proposiciones $P, Q, R, \ldots$ y conectivos, no de la formalización completa de predicados y cuantificadores.

Podemos representar un argumento de dos premisas mediante

$$\begin{array}{l} P_1, \\ P_2, \\ \hline C. \end{array}$$

Las letras no indican, por sí solas, que las premisas sean verdaderas ni que la conclusión sea falsa o verdadera. Representan lugares que pueden ser ocupados por proposiciones. Lo que queremos saber es si existe alguna combinación admisible de valores de verdad en la que las premisas sean todas verdaderas y la conclusión sea falsa.

Validez proposicional

Un argumento proposicional es **válido** si no existe ninguna asignación de valores de verdad a sus proposiciones componentes que haga verdaderas todas las premisas y falsa la conclusión.

Equivalente: **en toda asignación que hace verdaderas las premisas, también es verdadera la conclusión**.

Esta definición no dice que las premisas sean efectivamente verdaderas. Dice que no podemos conservarlas todas como verdaderas y negar la conclusión al mismo tiempo. Por eso una prueba de validez estudia posibilidades, no sólo lo que sucede en el ejemplo elegido.

Consideremos ahora la forma

$$\begin{array}{l} P \to Q, \\ P, \\ \hline Q. \end{array}$$

Si ambas premisas son verdaderas, P es verdadera; como el condicional $P \to Q$ también lo es, Q no puede ser falsa. Ésta es una inferencia válida, tradicionalmente llamada *modus ponens*.

Para contrastarla con el argumento de los múltiplos de 4, basta examinar el siguiente caso:

P	Q	$P \to Q$	Premisas $P \to Q$ y Q	Conclusión P
F	V	V	V y V	F

La fila muestra que la forma del argumento inicial **no es válida**. No necesitamos revisar infinitos ejemplos: una sola asignación con premisas verdaderas y conclusión falsa refuta la supuesta validez.

Qué cuenta como contraejemplo

Cuando queremos refutar una afirmación matemática universal, buscamos un objeto del dominio que no cumpla la propiedad. Cuando queremos refutar la validez de una forma proposicional, buscamos una **asignación de valores** que haga verdaderas todas sus premisas y falsa su conclusión. Ambos procedimientos comparten una idea, pero sus objetos de prueba no son idénticos.

2.3. El condicional no es la inferencia

Conviene detenerse en dos escrituras que suelen confundirse. Una es la proposición condicional

$$P \to Q.$$

La otra es el juicio de consecuencia

$$P \models Q.$$

En el primer caso construimos **una proposición** a partir de P y Q: «si P, entonces Q». En el segundo afirmamos, desde nuestro discurso sobre el argumento, que **toda asignación que hace verdadera P también hace verdadera Q**. El signo $\models$ no es un conectivo dentro de la misma proposición: expresa una relación semántica entre premisas y conclusión.

En lógica proposicional clásica, $P \to Q$ es falsa precisamente cuando P es verdadera y Q es falsa. En las otras tres posibilidades es verdadera. Esta convención recoge una condición muy concreta: para refutar «si P, entonces Q», debemos encontrar un caso en que P se cumpla y Q falle.

P	Q	$P \to Q$
V	V	V
V	F	F
F	V	V
F	F	V

Que el condicional sea verdadero cuando su antecedente es falso **no demuestra el antecedente ni el consecuente**. Solamente significa que no se ha producido la situación que refutaría ese condicional.

Si escribimos, por ejemplo,

> Si un entero es múltiplo de 4, entonces es par,

estamos formulando, en el lenguaje matemático ordinario, una afirmación general sobre enteros. Para analizarla con rigor debemos recordar el dominio y la cuantificación implícitos: para todo entero n, si n es múltiplo de 4, entonces n es par. El capítulo 3 formalizará la cuantificación; aquí basta reconocer que la afirmación matemática concreta requiere algo más que la tabla de verdad de un par de letras.

Justificación matemática. Sea n un entero múltiplo de 4. Por definición, existe un entero k tal que $n = 4k$. Entonces

$$n = 4k = 2(2k).$$

Como $2k$ es entero, n es par. Esta prueba establece el contenido aritmético de la afirmación general; la tabla de verdad sólo explica la forma lógica del condicional utilizado en ella.

2.4. De un esquema correcto a una prueba correcta

Tenemos ahora dos niveles de control. Primero debemos verificar que el paso inferencial utilizado conserve la verdad de las premisas. Después, para demostrar un resultado matemático, debemos justificar las premisas matemáticas de las que partimos o indicar explícitamente que son hipótesis, axiomas o resultados establecidos.

Regresemos a

$$P \to Q, \qquad P \quad \therefore \quad Q.$$

La forma es válida, pero eso no convierte automáticamente a Q en una verdad matemática establecida. Si alguien introduce como premisa un enunciado falso, puede utilizar una regla de inferencia válida y obtener una conclusión que no queda justificada a partir de verdades aceptadas.

En nuestro trabajo ordinario llamaremos **argumento sólido**, para este nivel introductorio, a un argumento válido cuyas premisas son verdaderas en la interpretación considerada. Cuando las premisas son hipótesis provisionales, no afirmamos que sean verdaderas de hecho: demostramos que *si* se cumplen, la conclusión queda garantizada. Esta última distinción permite trabajar rigurosamente con teoremas condicionales sin fingir que sus hipótesis se cumplen en todos los objetos.

Considera la diferencia entre las dos frases siguientes: «6 es múltiplo de 4» y «supongamos que n es múltiplo de 4». La primera es una afirmación concreta falsa. La segunda abre una situación hipotética con un entero n restringido por una condición; podemos extraer consecuencias dentro de ella sin afirmar que todos los enteros satisfagan la condición. Una demostración debe indicar cuándo se abre y cuándo se cierra ese contexto.

Cuando las premisas no pueden ser verdaderas a la vez. La definición de validez de §2.2 presenta un caso que merece examinarse: ¿qué sucede cuando las premisas son **incompatibles**? Consideremos, para dos proposiciones P y R, la forma de argumento

$$\frac{\begin{array}{l}P,\\ \neg P,\end{array}}{R.}$$

Es **válida** según la definición semántica adoptada: no existe ninguna asignación que haga verdaderas simultáneamente P y $\neg P$, y por ello tampoco existe una asignación que haga verdaderas ambas premisas y falsa la conclusión R. La validez no depende en este caso del contenido de R; éste podría reemplazarse por $\neg R$ sin dejar de obtener un argumento válido. A estas premisas se las llama **inconsistentes**.

Este resultado no demuestra R como una verdad matemática independiente de las premisas. Para obtener una conclusión establecida necesitamos premisas justificadas o un razonamiento legítimo bajo hipótesis pertinentes. Un conjunto de hipótesis incompatibles no describe ningún caso en que todas se cumplan: el condicional asociado a ellas no proporciona un ejemplo de su antecedente.

Ejemplo y comprobación. Para un entero n, examinemos «si $4 \mid n$ y $4 \nmid n$, entonces $n = 1$». Es un condicional verdadero para cualquier n: ningún entero cumple a la vez sus dos condiciones iniciales. ¿Hemos demostrado que todos los enteros son iguales a 1?

Respuesta. No. Hemos probado solamente que no hay contraejemplo del tipo exigido por ese condicional, porque sus hipótesis nunca son simultáneamente verdaderas. En particular, $n = 2$ refuta «todos los enteros son iguales a 1» sin refutar el condicional anterior. Una inferencia formalmente válida no sustituye el control de compatibilidad y justificación de las hipótesis. Aquí usamos lógica clásica; la demostración por contradicción se estudiará en el capítulo 6, con el debido control de los supuestos temporales.

La distinción entre **validez**, **verdad de las premisas** y **verdad de la conclusión** será nuestra herramienta para estudiar a continuación las recíprocas, las contrapositivas y las condiciones necesarias y suficientes. Antes de elegir una técnica de prueba, debemos saber exactamente qué afirmación necesita justificarse y qué nos autorizan a concluir las hipótesis.

2.5. Recíproca, contrapositiva y equivalencia

Hemos demostrado que todo múltiplo de 4 es par y hemos visto que la afirmación inversa falla para 6. Este ejemplo permite distinguir operaciones sobre un condicional que se parecen en el lenguaje ordinario, pero tienen comportamientos lógicos muy diferentes. Antes de intentar demostrar «si P, entonces Q», necesitamos saber qué enunciado estamos leyendo y qué cambios conservan realmente su significado.

La recíproca cambia lo que se afirma. Consideremos dos proposiciones P y Q. A partir del condicional

$$P \to Q$$

podemos formar su **recíproca**:

$$Q \to P.$$

La recíproca intercambia antecedente y consecuente. No es una reformulación automática de la proposición original: puede tener otro valor de verdad.

Tomemos un entero arbitrario n y escribamos, para ese entero,

$$P(n): \quad 4 \mid n, \qquad Q(n): \quad 2 \mid n.$$

La notación $d \mid n$ significa que n es múltiplo de d. La afirmación original dice que, para todo entero n, si $4 \mid n$, entonces $2 \mid n$. Su recíproca diría que, para todo entero n, si $2 \mid n$, entonces $4 \mid n$. Esta última es falsa: $n = 6$ satisface el antecedente, pero no el consecuente. Hemos refutado la recíproca **sin refutar la afirmación original**.

Sin embargo, la recíproca no tiene por qué ser falsa: podría ser verdadera por razones adicionales. Lo que debemos evitar es concluirla *únicamente* a partir del condicional original. Una afirmación puede dar origen a una recíproca verdadera, pero su verdad requiere una comprobación o una demostración propia.

La contrapositiva conserva el valor de verdad. Otra transformación consiste en negar ambas proposiciones e intercambiar sus posiciones. Obtenemos la **contrapositiva** de $P \to Q$:

$$\neg Q \to \neg P.$$

Leída en palabras: «si no ocurre Q, entonces no ocurre P». En nuestro ejemplo: «si un entero no es par, entonces no es múltiplo de 4». A diferencia de la recíproca, la contrapositiva expresa una condición lógicamente equivalente al condicional original.

Proposición. En lógica proposicional clásica, los condicionales $P \to Q$ y $\neg Q \to \neg P$ tienen el mismo valor de verdad para toda asignación de valores a P y Q.

Demostración. Basta examinar las cuatro asignaciones posibles. Incluimos la recíproca para que la diferencia sea visible:

P	Q	$P \to Q$	$Q \to P$	$\neg Q \to \neg P$
V	V	V	V	V
V	F	F	V	F
F	V	V	F	V
F	F	V	V	V

La tercera y la quinta columnas coinciden en las cuatro filas. Por definición de equivalencia proposicional, los dos condicionales son equivalentes. En cambio, la tercera y la cuarta columnas difieren, de modo que el condicional y su recíproca no son equivalentes en general. □

Después de la prueba. La tabla establece una propiedad de las *formas lógicas*: cualquier sustitución de P y Q por proposiciones conserva la equivalencia con la contrapositiva. No establece por sí sola que una afirmación aritmética particular sea verdadera. Para ello debemos justificar su contenido matemático, como hicimos con la divisibilidad por 4.

La equivalencia nos proporciona una opción de trabajo. Si necesitamos demostrar $P \to Q$, podemos intentar establecer directamente que P conduce a Q, o demostrar que $\neg Q$ conduce a $\neg P$. El segundo camino no constituye una nueva tesis: es otra manera de demostrar la misma implicación. Volveremos sobre esta estrategia, con su forma completa de escritura, en el capítulo dedicado a los métodos de demostración.

Tres transformaciones que no deben confundirse

Para una implicación $P \to Q$:

- **recíproca:** $Q \to P$;
- **contrapositiva:** $\neg Q \to \neg P$;
- **inversa:** $\neg P \to \neg Q$.

La original equivale a su contrapositiva. La recíproca equivale a la inversa, pero ninguna de estas últimas se deduce en general de la original. En particular, negar el antecedente no autoriza a negar el consecuente.

Un ejemplo en que la contrapositiva ayuda a descubrir la prueba. Vamos a examinar la afirmación siguiente:

Si el cuadrado de un entero es par, entonces el entero es par.

Para un entero arbitrario n, podemos representarla mediante

$$n^2 \text{ es par} \quad \longrightarrow \quad n \text{ es par.}$$

La hipótesis habla de n^2, pero la conclusión habla de n. Desplegar directamente la definición de par en la hipótesis nos daría $n^2 = 2k$ para algún entero k; no resulta inmediato obtener de ahí $n = 2m$. Probemos con la contrapositiva:

$$n \text{ no es par} \quad \longrightarrow \quad n^2 \text{ no es par.}$$

Como estamos trabajando con enteros y utilizamos la clasificación usual en pares e impares, esto equivale a mostrar que **el cuadrado de un entero impar es impar**.

Demostración. Sea n un entero impar. Entonces existe $k \in \mathbb{Z}$ tal que $n = 2k + 1$. Por tanto,

$$\begin{aligned} n^2 &= (2k+1)^2 \\ &= 4k^2 + 4k + 1 \\ &= 2(2k^2 + 2k) + 1. \end{aligned}$$

El número $2k^2 + 2k$ es un entero. La expresión obtenida tiene la forma $2m + 1$, de modo que n^2 es impar. Hemos probado la contrapositiva y, por su equivalencia con la implicación original, queda demostrado que si n^2 es par, entonces n es par. □

Idea de la prueba. La hipótesis original se refería a un cuadrado cuya paridad no sabíamos cómo trasladar al número de partida. Al negar la conclusión, obtuvimos una descripción explícita de ese número, $n = 2k + 1$, que podía elevarse al cuadrado. La decisión de cambiar de forma no fue una maniobra simbólica arbitraria: hizo disponible una definición útil.

Después de la prueba. Hemos demostrado la implicación para un entero arbitrario; por eso el resultado vale para todos los enteros. Su validez no depende de haber probado casos particulares como $n = 2$, $n = 4$ o $n = 6$. Tampoco hemos usado el hecho de que una implicación y su recíproca sean equivalentes: no lo son en general. Usamos solamente la equivalencia con la contrapositiva.

Cuándo dos condiciones son equivalentes. Podemos preguntar ahora si también vale el sentido contrario:

Si un entero es par, entonces su cuadrado es par.

En este caso tenemos una prueba directa breve. Sea $n = 2k$ con $k \in \mathbb{Z}$. Entonces

$$n^2 = (2k)^2 = 2(2k^2),$$

y $2k^2$ es entero. Por tanto, n^2 es par.

Junto con la implicación demostrada antes, hemos establecido **ambas direcciones**. Podemos reunirlas en un único teorema:

Teorema: paridad de un cuadrado

Para todo entero n,

$$n \text{ es par} \quad \Longleftrightarrow \quad n^2 \text{ es par.}$$

El símbolo $\Longleftrightarrow$ expresa aquí una equivalencia entre las dos propiedades de un entero arbitrario. Si P y Q son proposiciones, el **bicondicional** $P \leftrightarrow Q$ significa que se cumplen los dos condicionales:

$$P \leftrightarrow Q \quad \equiv \quad (P \to Q) \wedge (Q \to P).$$

Así, una demostración de «P si y sólo si Q» requiere justificar los dos sentidos. A veces la segunda dirección se obtiene repitiendo un argumento con los papeles intercambiados; otras veces necesita una idea distinta. Aquí una dirección se probó mediante la definición de entero par, y la otra mediante la contrapositiva y la definición de entero impar.

Conviene contrastar este resultado con nuestra afirmación inicial. «Todo múltiplo de 4 es par» no admite el bicondicional correspondiente, porque 6 es par y no es múltiplo de 4. **Una implicación probada es suficiente para afirmar un sentido, no para añadir el otro.**

Auditoría de lectura. Al encontrar un teorema escrito con «si», «sólo si» o «si y sólo si», separaremos las dos propiedades y comprobaremos qué dirección está efectivamente afirmada. Después distinguiremos si el argumento propuesto demuestra la implicación original, una contrapositiva equivalente o una recíproca que exige prueba adicional. Esta disciplina nos conducirá a la siguiente pregunta: ¿qué significa que una condición sea *necesaria*, *suficiente* o ambas cosas?

2.6. Condiciones necesarias y suficientes

Sabemos ya que una implicación y su recíproca son afirmaciones diferentes. Esa distinción tiene una consecuencia para la lectura: en matemática es frecuente expresar las implicaciones sin escribir «si... entonces». Decimos, por ejemplo, que una propiedad es *necesaria* o *suficiente* para otra. Estas palabras no añaden un nuevo conectivo; indican la dirección en la que una condición permite concluir otra.

Regresemos a los enteros. Ser múltiplo de 4 garantiza ser par. Pero ser par no garantiza ser múltiplo de 4: el número 6 vuelve a proporcionarnos un contraejemplo. Podemos describir exactamente la misma relación de dos maneras:

Ser múltiplo de 4 es **suficiente** para ser par.

Ser par es **necesario** para ser múltiplo de 4.

La primera frase dice que basta conocer la divisibilidad por 4 para concluir la paridad. La segunda dice que ningún múltiplo de 4 puede dejar de ser par. Ambas expresan una sola implicación; no son dos resultados distintos.

Definición y dirección de la flecha. Fijemos un dominio D y dos propiedades $P(x)$ y $Q(x)$ de sus elementos. Trabajaremos con las proposiciones que resultan de considerarlas para cada elemento de D; el capítulo 3 precisará el papel de la cuantificación. Diremos que:

- P es una **condición suficiente** para Q si, para todo $x \in D$, de $P(x)$ se sigue $Q(x)$;
- P es una **condición necesaria** para Q si, para todo $x \in D$, de $Q(x)$ se sigue $P(x)$.

En notación abreviada:

$$P \text{ es suficiente para } Q \iff \forall x \in D\ (P(x) \to Q(x)),$$
$$P \text{ es necesaria para } Q \iff \forall x \in D\ (Q(x) \to P(x)).$$

En estas líneas las flechas representan los condicionales correspondientes a un objeto fijado; al hablar de propiedades del dominio entendemos que la implicación vale para **todos** sus elementos. La definición permite leer una misma flecha en ambas direcciones lingüísticas:

$$\boxed{P \to Q: \quad P \text{ es suficiente para } Q; \quad Q \text{ es necesaria para } P.}$$

La asimetría es decisiva. **Suficiente** señala qué dato nos basta para obtener una conclusión; **necesaria** señala qué requisito no puede faltar si se cumple la otra propiedad. Ninguna de las dos palabras, por sí sola, autoriza a invertir la flecha.

Lo necesario no tiene por qué ser suficiente. Sea $n \in \mathbb{Z}$. Consideremos otra vez

$$P(n): \quad 4 \mid n, \qquad Q(n): \quad 2 \mid n.$$

Hemos demostrado que $P(n) \to Q(n)$ para todo entero. Por consiguiente, Q es necesaria para P: si n fuera múltiplo de 4 y no fuera par, contradiríamos el resultado demostrado.

¿Es también suficiente la paridad para la divisibilidad por 4? Para afirmarlo necesitaríamos $Q(n) \to P(n)$ para todo entero. Pero $n = 6$ cumple $Q(6)$ y no cumple $P(6)$. Luego la paridad es **necesaria, pero no suficiente**, para que un entero sea múltiplo de 4.

El mismo contraejemplo permite ver la otra cara: ser múltiplo de 4 es **suficiente, pero no necesario**, para ser par. El entero 6 es par sin satisfacer aquella condición suficiente. Una condición suficiente puede garantizar una propiedad sin ser la única manera de obtenerla.

No confundir necesidad con garantía

Que Q sea necesaria para P significa: «sin Q, no puede darse P». No significa: «si se da Q, entonces se da P». Esa última frase sería la recíproca y requiere otra demostración.

Leer «si», «sólo si» y «si y sólo si». Las formas verbales exigen especial cuidado. Para dos propiedades P y Q:

Escritura	Dirección que afirma
«P **si** Q»	$Q \to P$
«P **sólo si** Q»	$P \to Q$
«P **si y sólo si** Q»	$(P \to Q) \wedge (Q \to P)$

La palabra «sólo» introduce un requisito que debe cumplirse cuando se cumple lo anterior. Así, «un entero es múltiplo de 4 sólo si es par» es verdadero; «un entero es múltiplo de 4 si es par» es falso. Cambiar «sólo si» por «si» cambia la tesis matemática.

Una lectura segura consiste en sustituir siempre la frase por un condicional inequívoco antes de intentar demostrarla: **si se cumple el antecedente, entonces se cumple el consecuente**. En un bicondicional habrá que escribir dos frases y probarlas por separado, salvo que se disponga ya de una equivalencia establecida.

Cuando una condición es necesaria y suficiente. Diremos que P es **necesaria y suficiente** para Q cuando ambas implicaciones son verdaderas:

$$P \to Q \qquad \text{y} \qquad Q \to P.$$

Es exactamente el bicondicional estudiado en §2.5. No estamos introduciendo una tercera relación independiente. En particular, el teorema de paridad del cuadrado permite afirmar, para todo $n \in \mathbb{Z}$:

> Ser par es una condición necesaria y suficiente para que el cuadrado sea par.

La suficiencia se estableció escribiendo $n = 2k$ y comprobando que $n^2 = 2(2k^2)$. La necesidad se estableció demostrando la contrapositiva: si n es impar, entonces n^2 es impar. Cada dirección tiene una justificación matemática identificable.

Este ejemplo señala un límite pedagógico importante: aprender a traducir «necesaria y suficiente» a un bicondicional nos dice **qué tenemos que demostrar**, pero no produce por sí mismo la demostración. Para pasar del lenguaje a una prueba necesitamos además desplegar definiciones y elegir argumentos adecuados.

Encadenar condiciones sin invertirlas. Las condiciones también pueden formar cadenas. Consideremos tres proposiciones P, Q y R.

Proposición. La siguiente forma de argumento es válida:

$$P \to Q, \qquad Q \to R \quad \therefore \quad P \to R.$$

Demostración. Supongamos verdaderas las dos premisas. Para establecer $P \to R$, admitamos provisionalmente P. De $P \to Q$ y P obtenemos Q; de $Q \to R$ y Q obtenemos R. Por tanto, cuando P es verdadera, también lo es R, tal como exige el condicional $P \to R$. □

Idea de la prueba. En lugar de buscar todas las asignaciones posibles, abrimos el supuesto P y aplicamos dos veces el paso válido estudiado en §2.2. No hemos supuesto que P sea verdadera de hecho: la usamos dentro de una demostración condicional.

Ejemplo aritmético. Para todo entero n,

$$8 \mid n \quad \Longrightarrow \quad 4 \mid n \quad \Longrightarrow \quad 2 \mid n.$$

La primera implicación se justifica porque, si $n = 8k$ con $k \in \mathbb{Z}$, entonces $n = 4(2k)$. La segunda ya fue demostrada. Podemos concluir que $8 \mid n$ es una condición suficiente para $2 \mid n$, y que $2 \mid n$ es una condición necesaria para $8 \mid n$. En cambio, nada en esta cadena autoriza a concluir que todo número par sea múltiplo de 8.

Después de la prueba. La transitividad permite reutilizar resultados: si una propiedad garantiza una segunda y ésta garantiza una tercera, la primera garantiza la tercera. El orden de la cadena no es decorativo; invertirla sin nuevas premisas es repetir el error que inició el capítulo.

Una lectura antes de continuar. Ante cualquier afirmación que emplee «necesario», «suficiente» o «sólo si», el lector ya puede reconstruir un procedimiento: fijar el dominio, identificar las dos propiedades, escribir la flecha en el sentido afirmado y preguntar por qué esa implicación es verdadera para todos los elementos del dominio. Si el texto afirma que una condición es necesaria y suficiente, debe encontrar **dos direcciones** justificadas.

En la sección de ejercicios reuniremos estos criterios con los anteriores: distinguir la verdad de una conclusión de la validez de una inferencia, construir contraasignaciones, reconocer recíprocas y contrapositivas, y traducir condiciones sin alterar su sentido.

2.7. Ejercicios progresivos

Los ejercicios recorren tres niveles: identificar una forma de inferencia, traducir correctamente una afirmación y construir o revisar una justificación matemática. Cuando se pida refutar la validez de un esquema, proporciona una **asignación de valores de verdad** con todas las premisas verdaderas y la conclusión falsa. Cuando se pida refutar

una afirmación universal sobre enteros, proporciona un **entero concreto** y comprueba las propiedades implicadas. No son exactamente la misma tarea.

Ejercicio 1. Cuatro formas de inferencia. Determina cuáles de los siguientes esquemas son válidos. Justifica cada respuesta: mediante un razonamiento que descarte las contraasignaciones, si es válido; mediante una contraasignación, si no lo es.

1. $P \to Q,\ P\ \therefore\ Q$.
2. $P \to Q,\ Q\ \therefore\ P$.
3. $P \to Q,\ \neg Q\ \therefore\ \neg P$.
4. $P \to Q,\ \neg P\ \therefore\ \neg Q$.

Ejercicio 2. Una conclusión verdadera obtenida incorrectamente. Para un entero fijado n, escribe P: «$4 \mid n$» y Q: «$2 \mid n$». Un estudiante emplea $P \to Q$ y Q para concluir P.

1. Examina por separado las premisas y la conclusión cuando $n = 8$.
2. Repite el examen cuando $n = 6$.
3. Explica por qué la primera elección no valida el procedimiento y cuál de las dos constituye un contraejemplo a la inferencia.

Ejercicio 3. Lo que permite una disyunción. Recuerda que $P \vee Q$ utiliza el «o» inclusivo. Decide si son válidos los argumentos siguientes y demuestra o refuta cada uno:

1. $P \vee Q,\ \neg P\ \therefore\ Q$.
2. $P \vee Q,\ P\ \therefore\ \neg Q$.

Para el segundo, no basta indicar que la conclusión «podría ser falsa»: exhibe valores concretos de P y Q.

Ejercicio 4. Un conectivo frente a una relación de consecuencia. Considera P y Q como letras proposicionales independientes. Decide si son correctas las afirmaciones siguientes y justifícalas mediante la definición de §2.2 o una asignación de valores:

1. $P \models Q$.
2. $\{P,\ P \to Q\} \models Q$.
3. $P \to Q$ es una tautología.
4. $\{P \to Q,\ \neg Q\} \models \neg P$.

Explica en una oración por qué las expresiones de los apartados 2 y 3 no dicen lo mismo.

Ejercicio 5. Transformar una afirmación sin cambiarla inadvertidamente. Trabaja sobre el dominio $\mathbb{Z}$. Considera la proposición general: «Si un entero es múltiplo de 9, entonces es múltiplo de 3».

1. Escribe su recíproca, su contrapositiva y su inversa, conservando la cuantificación sobre todos los enteros.
2. Demuestra la afirmación original a partir de la definición de divisibilidad.
3. Determina el valor de verdad de la recíproca e inversa. Si son falsas, da un contraejemplo.
4. Justifica la contrapositiva sin volver a efectuar una prueba aritmética desde cero.

Ejercicio 6. Leer «si» y «sólo si». Sean, para $n \in \mathbb{Z}$, $A(n)$: «$12 \mid n$» y $B(n)$: «$6 \mid n$». Traduce cada frase a una implicación cuantificada, decide si es verdadera y justifica tu decisión:

1. «A si B».
2. «A sólo si B».
3. «B si A».
4. «B sólo si A».

Ejercicio 7. La importancia del dominio. Sea $x \in \mathbb{R}$ y consideremos $P(x)$: «$x > 0$» y $Q(x)$: «$x^2 > 0$».

1. ¿Es P suficiente para Q? Demuéstralo.
2. ¿Es P necesaria para Q? Justifica la respuesta.
3. Encuentra una condición necesaria y suficiente para Q entre las propiedades elementales de los números reales y demuestra las dos direcciones.
4. Explica por qué no se puede trasladar sin examen esta formulación a un dominio donde $>$ no esté definido.

Ejercicio 8. Una equivalencia nueva a partir de definiciones. Demuestra que, para todo entero n,

$$n \text{ es par} \quad \iff \quad n^3 \text{ es par.}$$

Separa las dos direcciones. En una de ellas, utiliza la contrapositiva y escribe de manera explícita un entero que certifique la paridad o imparidad correspondiente. No basta verificar varios valores de n.

Ejercicio 9. Encadenar la divisibilidad. Demuestra para todo $n \in \mathbb{Z}$ que

$$12 \mid n \implies 6 \mid n \implies 3 \mid n.$$

Concluye qué condición es suficiente para cuál y cuál es necesaria para cuál. Decide si $3 \mid n \to 12 \mid n$ es verdadera para todo entero; justifica la respuesta.

Ejercicio 10. Diagnóstico y reparación de una prueba. Un estudiante propone la siguiente demostración: «Todo múltiplo de 6 es múltiplo de 3. Como 9 es múltiplo de 3, se sigue que 9 es múltiplo de 6».

1. Identifica las dos propiedades y escribe el esquema de inferencia empleado.
2. Señala el paso no autorizado y demuestra, mediante un entero concreto, que la conclusión propuesta es falsa.
3. Reescribe un resultado verdadero que sí se desprenda de la primera afirmación y proporciona su justificación aritmética.

Ejercicio 11. Validez no significa solidez. Considera el esquema $P \to Q,\ P \ \therefore\ Q$.

1. Demuestra que es válido.
2. Interprétalo con P: «6 es múltiplo de 4» y Q: «6 es par». Calcula el valor de verdad de cada premisa y de la conclusión.
3. ¿Es sólido el argumento así interpretado? Explica por qué el valor de verdad de la conclusión no altera tu respuesta.

Ejercicio 12. Una cadena que parece poder recorrerse hacia atrás. Para un entero arbitrario n, sea

$$P(n):\ 12 \mid n, \qquad Q(n):\ 6 \mid n, \qquad R(n):\ 3 \mid n.$$

Un estudiante razona: «P implica Q y Q implica R. Como R se cumple, entonces P se cumple».

1. Traduce el razonamiento a un esquema proposicional con tres premisas y una conclusión.
2. Refuta su validez mediante una asignación de valores de verdad.
3. Da un entero que realice esas mismas condiciones aritméticas y refute la conclusión universal pretendida.
4. Sustituye la conclusión por una consecuencia que sí se siga de las dos implicaciones, y demuéstrala sin suponer que P sea verdadera de hecho.

2.8. Soluciones desarrolladas

Solución del ejercicio 1.

1. **Válido.** Si P y $P \to Q$ son verdaderas, Q no puede ser falsa, porque ésa es precisamente la única situación que haría falso el condicional. Es *modus ponens*.
2. **Inválido.** Asigna $P = \text{F}$ y $Q = \text{V}$. El condicional es verdadero y la segunda premisa también, pero la conclusión P es falsa. Es la afirmación del consecuente.
3. **Válido.** Si $\neg Q$ es verdadera, Q es falsa. Si además $P \to Q$ es verdadero, P tiene que ser falsa: de otro modo el condicional tendría antecedente verdadero y consecuente falso. Por tanto, $\neg P$ es verdadera. Este esquema es *modus tollens*.
4. **Inválido.** Asigna $P = \text{F}$ y $Q = \text{V}$. Entonces $P \to Q$ y $\neg P$ son verdaderas, pero $\neg Q$ es falsa. Es la negación del antecedente.

La comparación importante no es el nombre de cada forma, sino qué datos permiten descartar una conclusión falsa.

Solución del ejercicio 2. Para $n = 8$, tenemos $P = \text{V}$ y $Q = \text{V}$: 8 es múltiplo de 4 y es par. Por eso son verdaderos $P \to Q$, Q y también la conclusión P.

Para $n = 6$, en cambio, $P = \text{F}$ y $Q = \text{V}$: el 6 es par, pero $4 \nmid 6$. El condicional $P \to Q$ es verdadero, al igual que Q, y la conclusión P es falsa.

La elección de 8 sólo muestra que la conclusión puede ser verdadera en un caso. No establece que esté garantizada por las premisas. La elección de 6 sí muestra, con premisas verdaderas y conclusión falsa, que el esquema no es válido.

Solución del ejercicio 3.

1. **Válido.** Si $\neg P$ es verdadera, P es falsa. Como la disyunción inclusiva $P \vee Q$ es verdadera, su segundo componente Q debe ser verdadero.
2. **Inválido.** Elige $P = \text{V}$ y $Q = \text{V}$. Tanto $P \vee Q$ como P son verdaderas, pero la conclusión $\neg Q$ es falsa. En la disyunción inclusiva, que una alternativa ocurra no excluye que ocurra la otra.

En el primer caso la negación de una alternativa sí permite recuperar la otra; en el segundo, afirmar una alternativa no permite negar la restante.

Solución del ejercicio 4.

1. **Incorrecta.** Con $P = \text{V}$ y $Q = \text{F}$, la premisa P es verdadera y la conclusión Q es falsa. Por tanto, $P \not\models Q$.
2. **Correcta.** Si ambas premisas P y $P \to Q$ son verdaderas, Q debe ser verdadera. No existe contraasignación; así, $\{P, P \to Q\} \models Q$.
3. **Incorrecta.** $P \to Q$ es falsa cuando $P = \text{V}$ y $Q = \text{F}$; por eso no es una tautología.
4. **Correcta.** Si $\neg Q$ es verdadera, Q es falsa. Con $P \to Q$ verdadera, necesariamente P es falsa y $\neg P$ verdadera. No hay contraasignación.

El apartado 2 expresa una **relación de consecuencia entre premisas y conclusión**; el apartado 3 pregunta si **una sola fórmula** es verdadera en todas las asignaciones. No deben confundirse, aunque en lógica proposicional clásica las dos nociones puedan relacionarse mediante transformaciones justificadas.

Solución del ejercicio 5. Para todo $n \in \mathbb{Z}$, la afirmación original es $9 \mid n \to 3 \mid n$. Las transformaciones cuantificadas son:

- **Recíproca:** si $3 \mid n$, entonces $9 \mid n$.
- **Contrapositiva:** si $3 \nmid n$, entonces $9 \nmid n$.
- **Inversa:** si $9 \nmid n$, entonces $3 \nmid n$.

Para probar la original, supongamos $9 \mid n$. Existe $k \in \mathbb{Z}$ con $n = 9k = 3(3k)$; como $3k$ es entero, $3 \mid n$.

La recíproca es falsa: $n = 3$ es múltiplo de 3 y no de 9. La inversa también es falsa: el mismo $n = 3$ no es múltiplo de 9, pero sí es múltiplo de 3. La contrapositiva es verdadera porque §2.5 demostró que toda implicación es lógicamente equivalente a su contrapositiva. Al aplicar esa equivalencia a cada entero obtenemos la afirmación general correspondiente.

Solución del ejercicio 6. Recordemos primero que $A(n)$ significa $12 \mid n$ y $B(n)$ significa $6 \mid n$.

1. «A si B» significa $B(n) \to A(n)$ para todo entero. **Falsa:** $n = 6$ es múltiplo de 6, pero no de 12.
2. «A sólo si B» significa $A(n) \to B(n)$ para todo entero. **Verdadera:** si $n = 12k$, entonces $n = 6(2k)$, y $2k \in \mathbb{Z}$.
3. «B si A» significa $A(n) \to B(n)$ para todo entero. **Verdadera**, por el mismo argumento del apartado 2.
4. «B sólo si A» significa $B(n) \to A(n)$ para todo entero. **Falsa**, de nuevo con $n = 6$.

Los apartados 2 y 3 son dos formulaciones verbales de la misma implicación; los apartados 1 y 4 formulan su recíproca. El ejercicio se resuelve identificando las flechas antes de mirar los ejemplos.

Solución del ejercicio 7.

1. **Sí.** Para todo real x, si $x > 0$, entonces el producto de dos números positivos es positivo; en particular, $x^2 > 0$. Por tanto, P es suficiente para Q.
2. **No.** El real $x = -1$ satisface $x^2 = 1 > 0$, pero no $x > 0$. Como $Q(-1)$ es verdadera y $P(-1)$ falsa, $Q \to P$ no vale universalmente; P no es necesaria para Q.
3. La condición $x \neq 0$ es necesaria y suficiente para $x^2 > 0$ en $\mathbb{R}$. Si $x \neq 0$, o bien $x > 0$ o bien $x < 0$, y en cualquiera de los dos casos $x^2 > 0$. Recíprocamente, si $x^2 > 0$, no puede ocurrir $x = 0$, pues entonces $x^2 = 0$. Hemos justificado ambas direcciones.
4. La escritura usa la relación $>$ y propiedades del orden usual de $\mathbb{R}$. En un dominio sin tal relación especificada, la frase no conserva automáticamente el mismo significado. Antes de discutir su verdad hay que precisar la interpretación.

Solución del ejercicio 8. Primera dirección. Supongamos que n es par. Existe $k \in \mathbb{Z}$ con $n = 2k$; entonces

$$n^3 = (2k)^3 = 2(4k^3).$$

El número $4k^3$ es entero; por definición, n^3 es par.

Segunda dirección. Demostraremos por contraposición que si n es impar, entonces n^3 es impar. Sea $n = 2k + 1$ con $k \in \mathbb{Z}$. Calculamos

$$\begin{aligned} n^3 &= (2k+1)^3 \\ &= 8k^3 + 12k^2 + 6k + 1 \\ &= 2(4k^3 + 6k^2 + 3k) + 1. \end{aligned}$$

Como $4k^3 + 6k^2 + 3k$ es entero, n^3 tiene la forma de un entero impar. Por equivalencia entre un condicional y su contrapositiva, si n^3 es par, entonces n es par. Las dos direcciones establecen el bicondicional. □

Idea de la prueba. La definición de par produce inmediatamente una factorización útil para el cubo. Para la dirección restante, negar la paridad de n permite escribir $n = 2k + 1$ y reconocer la forma de su cubo.

Solución del ejercicio 9. Supongamos $12 \mid n$. Por definición, $n = 12k$ para algún $k \in \mathbb{Z}$. Como $n = 6(2k)$, tenemos $6 \mid n$. Si ahora $6 \mid n$, escribimos $n = 6m = 3(2m)$ para cierto $m \in \mathbb{Z}$, de donde $3 \mid n$. Quedan demostradas ambas implicaciones y, por transitividad, $12 \mid n \to 3 \mid n$.

Ser múltiplo de 12 es suficiente para ser múltiplo de 3, y ser múltiplo de 3 es necesario para ser múltiplo de 12. La implicación recíproca es falsa: $n = 3$ es múltiplo de 3, pero no de 12. Una cadena de implicaciones no puede invertirse sin demostrar las recíprocas.

Solución del ejercicio 10. Denotemos por P la propiedad «ser múltiplo de 6» y por Q la propiedad «ser múltiplo de 3». El estudiante utiliza $P \to Q$ y Q para concluir P, forma inválida de afirmación del consecuente.

La conclusión concreta también es falsa: $9 = 3 \cdot 3$, por lo que $3 \mid 9$; sin embargo, no existe un entero k con $9 = 6k$, pues ello exigiría $k = 3/2 \notin \mathbb{Z}$. La verdad de «9 es múltiplo de 3» no fuerza la divisibilidad por 6.

Un resultado correcto es: **todo múltiplo de** 6 **es múltiplo de** 3. Si $n = 6k$ con $k \in \mathbb{Z}$, entonces $n = 3(2k)$ y $2k \in \mathbb{Z}$. La corrección no consiste en hacer verdadera una conclusión arbitraria, sino en escribir y demostrar la dirección que autorizan las hipótesis.

Solución del ejercicio 11.

1. Si P y $P \to Q$ son verdaderas, la falsedad de Q haría falso el condicional. Por tanto, Q debe ser verdadera: la forma es válida.
2. En la interpretación dada, P es falsa y Q verdadera. El condicional $P \to Q$ es verdadero porque su antecedente es falso; la otra premisa, P, es falsa; la conclusión, Q, es verdadera.
3. El argumento interpretado **no es sólido**, pues la solidez requiere validez y verdad de todas las premisas, y P es falsa. Que Q resulte verdadera no convierte en verdadera la premisa falsa ni repara el argumento. Esta distinción impide confundir acertar con justificar.

Solución del ejercicio 12.

1. El esquema utilizado es

$$P \to Q, \qquad Q \to R, \qquad R \quad \therefore \quad P.$$

2. Asigna $P = \mathrm{F}$, $Q = \mathrm{F}$ y $R = \mathrm{V}$. Los dos condicionales son verdaderos porque sus antecedentes son falsos; la tercera premisa R es verdadera; la conclusión P es falsa. La asignación refuta la validez.
3. La asignación tiene una realización aritmética con $n = 3$: $12 \nmid 3$, $6 \nmid 3$ y $3 \mid 3$. Las dos implicaciones universales de divisibilidad son verdaderas, y también lo es $R(3)$, pero $P(3)$ es falsa.
4. Una consecuencia correcta de las dos implicaciones es $P \to R$. Para demostrarla, suponemos provisionalmente P; de $P \to Q$ obtenemos Q, y de $Q \to R$ obtenemos R. Al cerrar la hipótesis provisional concluimos $P \to R$, **sin afirmar que** P **se cumpla para el entero considerado**. La condición final R no autoriza recorrer la cadena hacia atrás. □

2.9. Hacia las variables y los cuantificadores

En los ejemplos de divisibilidad hemos repetido «para todo entero n», mientras que los esquemas proposicionales se examinaban mediante letras como P y Q. Sabemos que ambos niveles deben coordinarse, pero todavía no hemos descrito con detalle qué hace un cuantificador, qué variables liga y cómo cambia el significado de una negación cuando interviene «para todo» o «existe». Ésa será la pregunta del capítulo siguiente. Con lo aprendido hasta aquí, ya sabemos qué exige una inferencia proposicional; el capítulo 3 precisará cómo formular y negar afirmaciones que dependen de variables.

Capítulo 3 — ¿Qué cambia cuando aparecen variables y cuantificadores?

El capítulo anterior nos enseñó a distinguir una inferencia válida de una conclusión que resulta verdadera por casualidad. Sin embargo, al justificar que todo múltiplo de 4 es par, no razonamos sobre una sola proposición aislada: elegimos un entero n sin fijar su valor, supusimos que $4 \mid n$ y concluimos que $2 \mid n$. Después afirmamos que el razonamiento servía para **todos** los enteros. ¿Qué autorizaba ese último paso?

Consideremos ahora otra afirmación:

> Existe un entero par que no es múltiplo de 4.

El número 6 basta para justificarla. No haríamos lo mismo para establecer que *todos* los enteros pares tienen esa propiedad: esa nueva afirmación es falsa. Las palabras «todos» y «existe» no son adornos de la frase, sino que determinan qué objetos debemos examinar y qué cuenta como una demostración o un contraejemplo.

La pregunta rectora de este capítulo es

¿qué afirmamos al cuantificar una propiedad y qué variables quedan bajo esa afirmación?

Avanzaremos desde las fórmulas abiertas hasta su cuantificación y, después, estudiaremos el alcance, el orden, la negación y la unicidad. El propósito no es sustituir la lectura por una colección de reglas gráficas: queremos poder reconstruir qué dice realmente una fórmula antes de intentar demostrarla.

3.1. De una fórmula abierta a una afirmación

En el capítulo 1 distinguimos una proposición de una fórmula que contiene variables libres. Recuperemos esa distinción en un contexto concreto. Trabajemos en el dominio de los enteros $\mathbb{Z}$ y escribamos

$$P(n): \quad 4 \mid n.$$

La expresión $P(n)$ designa la propiedad «n es múltiplo de 4». Mientras no asignemos un valor a n, ni indiquemos qué hemos afirmado acerca de todos o algunos de sus posibles valores, no disponemos todavía de una proposición cerrada cuya verdad podamos examinar sin información adicional. No obstante, sí sabemos **qué condición** expresa la fórmula y cuál es su dominio.

Si sustituimos n por un entero concreto, obtenemos una proposición. Por ejemplo,

$$P(8): \ 4 \mid 8, \qquad P(6): \ 4 \mid 6.$$

La primera es verdadera; la segunda es falsa. Una fórmula abierta admite distintas evaluaciones según el valor que reciba su variable. Esto no significa que la fórmula esté mal escrita: precisamente nos permite hablar de una misma propiedad para muchos objetos sin reescribir su descripción cada vez.

Contexto y fórmula abierta

Una **fórmula abierta** es una fórmula con al menos una ocurrencia de variable libre. Para evaluarla se necesita una interpretación de sus símbolos y una asignación a las variables libres. Su valor de verdad puede depender de esa asignación.

Una **fórmula cerrada** no contiene ocurrencias libres de variables: las que aparecen en ella están ligadas por cuantificadores. Una fórmula abierta puede convertirse en una proposición evaluable cuando el contexto asigna valores a todas sus variables libres; eso no la convierte, por ese solo hecho, en una fórmula sintácticamente cerrada.

Aquí usamos la palabra «predicado» en el sentido introductorio fijado en el capítulo 1: una propiedad expresada mediante una fórmula abierta, como $P(n)$. No la identificamos todavía con un objeto funcional ni desarrollamos una teoría general de predicados. Conviene separar lo que representa la notación de la afirmación que construiremos con ella.

Dos maneras distintas de cerrar una pregunta. La pregunta «¿es n múltiplo de 4?» puede dar lugar a afirmaciones muy diferentes:

«Todos los enteros son múltiplos de 4»,

«Existe un entero múltiplo de 4».

La primera es falsa —basta tomar $n = 1$—, mientras que la segunda es verdadera —podemos tomar $n = 8$—. La propiedad de partida no cambió: cambió **lo que afirmamos sobre sus posibles argumentos**. Ésta es la función de los cuantificadores.

Lectura guiada. Antes de añadir «todos» o «existe» a una fórmula, identifica tres datos: el dominio de los objetos, la propiedad que han de cumplir y la fuerza de la afirmación que se quiere hacer. En las dos frases anteriores, el dominio es $\mathbb{Z}$ y la propiedad es $4 \mid n$; solamente difiere la fuerza de la afirmación.

3.2. El cuantificador universal: no dejar ningún caso fuera

La notación

$$\forall n \in \mathbb{Z}\ P(n)$$

se lee «para todo entero n, se cumple $P(n)$». El símbolo $\forall$ se llama **cuantificador universal**. La expresión completa afirma que *cada* elemento del dominio satisface la propiedad. Si trabajamos en un universo más amplio U que contiene a $\mathbb{Z}$, la notación restringida $\forall n \in \mathbb{Z}\ P(n)$ abrevia $\forall n\ (n \in \mathbb{Z} \to P(n))$; si el dominio ya está fijado como $\mathbb{Z}$, la escritura anuncia directamente ese dominio. La restricción universal tiene forma condicional: no exige que todos los elementos de U sean enteros. En el ejemplo anterior,

$$\forall n \in \mathbb{Z}\ (4 \mid n)$$

es falsa, porque $4 \nmid 1$. No necesitamos conocer la divisibilidad de todos los demás enteros para rechazar la afirmación: un solo caso contrario basta.

En cambio, consideremos

$$\forall n \in \mathbb{Z}\ (4 \mid n \to 2 \mid n).$$

Ésta es la forma explícita de la afirmación que demostramos en el capítulo 2. Su contenido es condicional: para cada entero, **si** es múltiplo de 4, **entonces** es par. No dice que todos los enteros sean múltiplos de 4; exige que ningún entero que satisfaga el antecedente deje de satisfacer el consecuente.

Proposición. Para todo $n \in \mathbb{Z}$, si $4 \mid n$, entonces $2 \mid n$.

Demostración. Sea n un entero arbitrario. Supongamos que $4 \mid n$. Por definición de divisibilidad, existe $k \in \mathbb{Z}$ tal que $n = 4k$. Por tanto,

$$n = 4k = 2(2k).$$

Puesto que $2k \in \mathbb{Z}$, concluimos que $2 \mid n$. La elección de n no impuso una restricción adicional: el razonamiento se aplica a cualquier entero que cumpla la hipótesis. Por eso queda establecida la afirmación universal. □

Qué permite generalizar. En esta prueba «sea n un entero arbitrario» introduce un objeto del dominio **sin elegir un valor particular** ni suponer una propiedad especial de él. La suposición posterior $4 \mid n$ pertenece al interior de la demostración condicional. Cerrada esa suposición, obtenemos $4 \mid n \to 2 \mid n$ para el entero arbitrario; cerrada también la elección de n, obtenemos el enunciado universal.

Si hubiéramos comenzado por «tomemos $n = 8$», sólo habríamos establecido la implicación en ese caso. Si hubiéramos comenzado por «sea n un entero par», habríamos restringido el objeto desde el principio y necesitaríamos revisar qué enunciado permite justificar el argumento. **Generalizar no consiste en borrar las hipótesis particulares al final de la prueba.**

La condición pertenece dentro del cuantificador

Cuando decimos «todo entero múltiplo de 4 es par», la lectura correcta es

$$\forall n \in \mathbb{Z}\ (4 \mid n \to 2 \mid n).$$

No estamos afirmando conjuntamente que todos los enteros sean múltiplos de 4 y que todos sean pares. El condicional discrimina cuáles objetos activan la hipótesis. Más adelante expresaremos las restricciones de dominio con el mismo cuidado.

La relación entre una fórmula universal y sus casos concretos también funciona en sentido descendente: si hemos demostrado que $\forall n \in \mathbb{Z}\ P(n)$ y sabemos que $a \in \mathbb{Z}$, podemos concluir $P(a)$. Esto se llama **instanciación universal**. El paso requiere que el objeto elegido pertenezca al dominio y que la sustitución sea legítima; la atención al alcance de las variables, que estudiaremos en §3.4, precisará esta última condición.

3.3. El cuantificador existencial: encontrar un testigo

La notación

$$\exists n \in \mathbb{Z}\ P(n)$$

se lee «existe al menos un entero n tal que se cumple $P(n)$». El símbolo $\exists$ se llama **cuantificador existencial**. No exige que todos los enteros tengan la propiedad; exige que alguno la tenga. Si el universo de referencia es más amplio que $\mathbb{Z}$, la notación $\exists n \in \mathbb{Z}\ P(n)$ abrevia $\exists n\ (n \in \mathbb{Z} \wedge P(n))$: el testigo debe ser entero **y** cumplir la propiedad. Compárese esta conjunción con el condicional usado para restringir el universal.

Para probar

$$\exists n \in \mathbb{Z}\ (4 \mid n)$$

podemos presentar $n = 8$ y comprobar que $8 \in \mathbb{Z}$ y $8 = 4 \cdot 2$. El número 8 es un **testigo** de la existencia afirmada. Un testigo no es simplemente un número mencionado: debe pertenecer al dominio y satisfacer exactamente la propiedad propuesta.

Consideremos una afirmación algo más interesante:

Proposición existencial

Existe un entero que es par y no es múltiplo de 4:

$$\exists n \in \mathbb{Z}\ (2 \mid n \wedge 4 \nmid n).$$

Demostración. Tomemos $n = 6$. Es un entero y $6 = 2 \cdot 3$, de modo que $2 \mid 6$. En cambio, no existe $k \in \mathbb{Z}$ tal que $6 = 4k$, pues ello exigiría $k = 3/2 \notin \mathbb{Z}$. Por tanto, $4 \nmid 6$. Hemos exhibido un entero que cumple las dos condiciones y queda probada la existencia. □

Idea de la prueba. El cuantificador existencial nos permitía elegir un objeto conveniente; el conectivo $\wedge$ nos obligaba, sin embargo, a verificar **ambas** condiciones para ese mismo objeto. No habría bastado exhibir un número par y otro distinto que no fuera múltiplo de 4.

Un testigo no demuestra una afirmación universal. Con $n = 6$ demostramos que existe un entero par no divisible por 4. Pero el mismo dato no demuestra

$$\forall n \in \mathbb{Z}\ (2 \mid n \to 4 \nmid n).$$

De hecho, esta última afirmación es falsa: el entero 8 es par y múltiplo de 4. El cambio de cuantificador transforma lo que debemos demostrar. Para una existencia, un testigo verificado basta; para una universal, cada caso debe quedar cubierto por un argumento general.

Además, hay que distinguir las dos escrituras

$$\exists n \in \mathbb{Z}\ (4 \mid n \wedge 2 \mid n) \qquad \text{y} \qquad \exists n \in \mathbb{Z}\ (4 \mid n \to 2 \mid n).$$

La primera exige un entero que cumpla las dos propiedades; $n = 8$ sirve. La segunda sólo exige que cierto condicional sea verdadero. Por ejemplo, $n = 1$ lo satisface porque su antecedente $4 \mid 1$ es falso. Por tanto, **la existencia de una implicación verdadera no garantiza la existencia de un caso que satisfaga su antecedente**. Esta diferencia recupera la lectura del condicional estudiada en §2.3.

Dos tareas de prueba, dos cargas distintas. En términos operativos, el lector puede plantearse una pregunta antes de escribir una solución: «¿debo cubrir todos los objetos del dominio o encontrar al menos uno?». Las palabras «arbitrario» y «testigo» reflejan esas tareas, pero no son fórmulas mágicas. La primera exige controlar qué hipótesis utilizamos; la segunda exige comprobar que el objeto elegido satisface la propiedad completa.

La afirmación existencial no especifica, por sí sola, **un único** objeto ni uno que sepamos calcular mediante un procedimiento. Cuando más adelante estudiemos existencia y unicidad, separaremos ambas exigencias. La distinción entre pruebas constructivas y no constructivas se desarrollará en el capítulo sobre métodos de demostración.

3.4. Variables libres, variables ligadas y alcance

Hemos escrito $\forall n \in \mathbb{Z}\ P(n)$ y $\exists n \in \mathbb{Z}\ P(n)$ como afirmaciones sobre el conjunto completo de enteros. En ambos casos, el cuantificador actúa sobre las apariciones de n que están dentro de la fórmula a la que se aplica. Decimos que esas apariciones quedan **ligadas** por el cuantificador. Para leer expresiones más complejas debemos precisar hasta dónde llega esa acción: su **alcance**.

Variable ligada y alcance

Una ocurrencia de una variable está **ligada** cuando se encuentra dentro del alcance de un cuantificador que declara esa misma variable. Una ocurrencia no ligada por ningún cuantificador es **libre**. El **alcance** es la parte de la fórmula sobre la que actúa el cuantificador; los paréntesis permiten hacerlo explícito.

La condición de libre o ligada se determina para cada **ocurrencia** de la variable, no para la letra considerada aisladamente.

Trabajemos con enteros y examinemos

$$\forall x \in \mathbb{Z}\ (x + y = y + x).$$

El cuantificador universal liga las apariciones de x de la igualdad. En cambio, y permanece libre: el cuantificador $\forall x$ no le asigna un valor ni dice nada directamente sobre todos sus valores. El dominio de y debe fijarse en el contexto si deseamos evaluar la fórmula; por ejemplo, podemos declarar «sea $y \in \mathbb{Z}$». Para obtener una proposición cerrada sin esa asignación previa podemos escribir

$$\forall y \in \mathbb{Z}\ \forall x \in \mathbb{Z}\ (x + y = y + x).$$

Esta última fórmula ya no deja ninguna de esas variables libre. Aquí la igualdad es verdadera en la aritmética usual de los enteros, pero **la verdad del resultado no sustituye el examen sintáctico de las variables**: debemos saber qué estamos afirmando antes de invocar propiedades de la suma.

Los paréntesis fijan una frontera. Comparemos la fórmula cerrada

$$\forall x \in \mathbb{Z}\ (x^2 \geq 0)$$

con la fórmula

$$\big(\forall x \in \mathbb{Z}\ (x^2 \geq 0)\big) \wedge (x < 3).$$

En la segunda, el primer x de $x^2 \geq 0$ está ligado por el cuantificador; el x de $x < 3$ queda **fuera** de su alcance y es libre. El primer componente de la conjunción es una proposición cerrada; el segundo requiere una asignación o un nuevo cuantificador. No podemos declarar que el x final está ligado sólo porque en alguna parte anterior de la línea aparezca $\forall x$.

Escribir bien una fórmula exige que sus fronteras sean legibles. Si queremos afirmar que todo entero tiene cuadrado no negativo y es menor que 3, debemos escribir

$$\forall x \in \mathbb{Z}\ (x^2 \geq 0 \wedge x < 3).$$

Ahora sí, ambas apariciones de x están ligadas, pero la afirmación es falsa: $x = 3$ viola $x < 3$. Cambiar el alcance cambió simultáneamente la condición de la variable y el contenido de la afirmación.

Cambiar una letra sin cambiar la afirmación. Las letras de las variables ligadas son nombres auxiliares. En condiciones de sustitución correcta,

$$\forall x \in \mathbb{Z}\ (x^2 \geq 0) \qquad \text{y} \qquad \forall t \in \mathbb{Z}\ (t^2 \geq 0)$$

expresan la misma afirmación. Renombrar una variable ligada exige cambiar **todas y sólo** las ocurrencias ligadas por ese cuantificador, sin capturar variables que antes eran libres.

Por ejemplo, en

$$\forall x \in \mathbb{Z}\ (x < y)$$

la letra y está libre. Reemplazar sin cuidado x por y produciría $\forall y \in \mathbb{Z}\ (y < y)$, que liga el antiguo y libre y cambia el significado. Una sustitución segura sería $\forall t \in \mathbb{Z}\ (t < y)$, manteniendo y libre y reservando una letra nueva para la variable ligada. Esta precaución recibirá importancia cuando trabajemos con varios cuantificadores.

Laboratorio de lectura: no confundir la fórmula con su cierre. Sea

$$R(x, y): \quad x + y = 0$$

una fórmula abierta en el dominio $\mathbb{Z}$ para ambas variables. La escritura

$$\exists x \in \mathbb{Z}\ (x + y = 0)$$

liga x, pero deja y libre. Si fijamos $y = 5$, la afirmación resultante es verdadera con testigo $x = -5$; si fijamos $y = -2$, sirve $x = 2$. Para decir que **cada** entero y tiene un entero x que satisface la igualdad, escribimos la proposición cerrada

$$\forall y \in \mathbb{Z}\ \exists x \in \mathbb{Z}\ (x + y = 0).$$

En este caso podemos justificarla: sea $y \in \mathbb{Z}$ arbitrario y elijamos $x = -y$, que también es entero; entonces $x + y = -y + y = 0$. La escritura de los cuantificadores explicita una estructura que la demostración ya deja ver: primero se toma un y cualquiera y después se elige un x que puede depender de él.

Después de la prueba. Ésta es una primera advertencia sobre el orden de los cuantificadores, no una licencia para intercambiarlos: afirmar que cada entero tiene su opuesto no equivale a afirmar que exista **un solo entero fijo** que sea el opuesto de todos. Examinaremos rigurosamente esa diferencia en §3.5.

3.5. El orden de los cuantificadores y la dependencia de los testigos

En §3.4 demostramos que cada entero tiene un opuesto: dado $y \in \mathbb{Z}$, elegimos $x = -y$ y obtuvimos $x + y = 0$. El testigo de la existencia se eligió **después** de conocer el entero al que debía corresponder. ¿Podemos cambiar ese orden y elegir primero un único x que sirva para todos los valores de y? La respuesta obliga a leer los cuantificadores de izquierda a derecha y a conservar sus dominios.

Cada objeto puede tener su propio testigo. Comparemos dos proposiciones cerradas sobre los enteros:

$$A: \quad \forall x \in \mathbb{Z}\ \exists y \in \mathbb{Z}\ (x + y = 0),$$
$$B: \quad \exists y \in \mathbb{Z}\ \forall x \in \mathbb{Z}\ (x + y = 0).$$

En A se afirma: «para **cada** entero x, existe **algún** entero y que sumado a él da cero». Podemos comenzar la demostración tomando un x arbitrario y, sólo entonces, elegir $y = -x$. Como $-x \in \mathbb{Z}$, el testigo pertenece al dominio y satisface $x + y = x - x = 0$. Al no haber elegido un valor especial de x, esta construcción demuestra A. Es la misma propiedad de los opuestos estudiada al final de §3.4, con nombres de variables reorganizados.

En B se afirma otra cosa: «existe un entero y que, sumado a **cualquier** entero x, da cero». Ahora el testigo debe ser **uno solo**, elegido antes de recorrer los posibles valores de x. Esta afirmación es falsa.

Demostración de que B es falsa. Supongamos que existe un entero y tal que $x + y = 0$ para todo $x \in \mathbb{Z}$. Al tomar $x = 0$ obtenemos $y = 0$. Al tomar $x = 1$ obtenemos $1 + y = 0$, es decir, $y = -1$. El mismo entero no puede satisfacer $y = 0$ y $y = -1$. Por tanto, no existe el testigo exigido por B. □

La dependencia del testigo

La forma

$$\forall x \in D\ \exists y \in E\ R(x, y)$$

afirma que **cada** x tiene por lo menos un y relacionado con él; el y elegido *puede depender* de x. No exige que ese y sea único ni que debamos disponer de una fórmula para calcularlo.

La forma

$$\exists y \in E\ \forall x \in D\ R(x, y)$$

afirma que **un mismo** y sirve para todos los x. El testigo es uniforme: no se puede cambiar al pasar de un x a otro.

La expresión «puede depender» es deliberada: el primer enunciado no exige que testigos distintos correspondan a valores distintos de x. Si por otras razones existe un testigo uniforme, también podemos utilizarlo para cada x. Lo que no podemos hacer es **suponer** que existe tal testigo cuando sólo hemos demostrado una existencia por cada objeto.

La diferencia persiste aunque la relación sea muy sencilla. Consideremos las proposiciones

$$C: \quad \forall x \in \mathbb{Z}\ \exists y \in \mathbb{Z}\ (y > x),$$
$$D: \quad \exists y \in \mathbb{Z}\ \forall x \in \mathbb{Z}\ (y > x).$$

Proposición. C es verdadera y D es falsa.

Demostración. Para C, sea $x \in \mathbb{Z}$ arbitrario. Elijamos $y = x + 1$. Entonces $y \in \mathbb{Z}$ y $y > x$. Como esta elección funciona para cualquier entero de partida, queda demostrada C.

Para refutar D, supongamos que existe un entero fijo y mayor que todos los enteros. Podemos tomar $x = y + 1$, que también es entero. La propiedad supuesta exigiría $y > y + 1$, lo cual es imposible. Luego D es falsa. □

Idea de la prueba. El procedimiento $y = x + 1$ es una elección legítima para C, pues sólo buscamos un y después de fijar x. Para D, en cambio, la elección de y se realizaría antes: después podríamos poner a prueba ese candidato escogiendo $x = y + 1$. El error de lectura consistiría en utilizar para D la misma respuesta variable que funciona en C.

El dominio importa. Si en D sustituimos el dominio de x por $\{0, 1\}$ y conservamos $y \in \mathbb{Z}$, la proposición resultante **sí** es verdadera: $y = 2$ es mayor que los dos posibles valores de x. No cambió el significado de $\exists$ ni de $\forall$; cambió el conjunto de casos que un testigo uniforme debía cubrir.

Una implicación entre las dos formas, pero no una equivalencia. La diferencia anterior puede precisarse en una proposición general. Sean D y E dominios fijados y $R(x, y)$ una propiedad de pares con $x \in D$ e $y \in E$.

Proposición. Si existe $y \in E$ tal que $R(x, y)$ para todo $x \in D$, entonces, para cada $x \in D$, existe algún $y \in E$ que satisface $R(x, y)$. En símbolos,

$$(\exists y \in E\ \forall x \in D\ R(x, y)) \ \longrightarrow\ (\forall x \in D\ \exists y \in E\ R(x, y)).$$

Demostración. Supongamos verdadera la premisa y fijemos un testigo $y_0 \in E$ que satisfaga $R(x, y_0)$ para todo $x \in D$. Sea ahora $x \in D$ arbitrario. Para ese x podemos utilizar el mismo y_0 como testigo de $\exists y \in E\ R(x, y)$. Como el argumento no depende de una elección especial de x, obtenemos la conclusión universal. □

La recíproca **no** es válida en general: la relación $x + y = 0$ sobre $\mathbb{Z}$ proporciona un contraejemplo, pues A es verdadera y B falsa. Por tanto, el paso de un testigo común a testigos para cada caso está justificado; el paso inverso necesita información adicional.

Conviene observar que la demostración anterior no invoca que los dominios sean finitos ni que los testigos sean únicos. Si el antecedente afirma que existe un testigo común, ya suministra el objeto que necesitamos. Si D está vacío, la conclusión universal también es

verdadera; este caso no altera la implicación demostrada. Más adelante podremos estudiar con detalle cómo interviene el dominio vacío en afirmaciones existenciales y universales.

Cuantificadores iguales y cuantificadores distintos. No todo cambio de orden modifica la afirmación. Si las variables recorren dominios fijados y la fórmula interior se conserva, dos cuantificadores universales consecutivos pueden permutarse:

$$\forall x \in D\ \forall y \in E\ R(x,y) \quad \Longleftrightarrow \quad \forall y \in E\ \forall x \in D\ R(x,y).$$

Ambos enunciados requieren que $R(x,y)$ se cumpla para **todos los pares** permitidos. Del mismo modo,

$$\exists x \in D\ \exists y \in E\ R(x,y) \quad \Longleftrightarrow \quad \exists y \in E\ \exists x \in D\ R(x,y),$$

porque ambos exigen la existencia de **algún par** que satisfaga la propiedad. En el segundo caso no intercambiamos las posiciones de x y y *dentro* de R: sólo modificamos el orden en que anunciamos su existencia.

Con cuantificadores de distinto tipo, en cambio, el orden puede modificar la dependencia y, con ella, el valor de verdad. Los ejemplos A/B y C/D muestran precisamente ese cambio. No basta recordar una regla superficial sobre el intercambio de símbolos: hay que reconstruir quién se elige primero, quién puede depender de quién y qué exige la propiedad.

Auditoría de lectura: reconstruir la afirmación antes de probarla. Examinemos la frase: «Todo entero tiene un entero mayor que él». Una lectura correcta comienza por fijar el primer objeto: **para cada entero** x; después declara la existencia de otro: **hay un entero** y; finalmente especifica la relación: $y > x$. Así obtenemos C, no D. La frase «Existe un entero mayor que todos los enteros» tiene el orden opuesto y expresa D.

En una demostración, los mismos cuantificadores determinan qué elecciones son legítimas. Si el objetivo comienza por $\forall x\,\exists y$, podemos tomar un x arbitrario y buscar después un testigo. Si comienza por $\exists y\,\forall x$, debemos encontrar primero un candidato fijo y demostrar que funciona para todo x. Una solución que cambie de candidato cada vez que cambia x no prueba este segundo objetivo.

Después de la prueba. Ya podemos leer una cadena de cuantificadores como una secuencia de compromisos, no como una decoración de la fórmula. El siguiente problema será negar correctamente esos compromisos: ¿qué debemos afirmar para refutar «todos cumplen» o «existe alguno», especialmente cuando intervienen varios cuantificadores y paréntesis? Ése será el objeto de §3.6.

3.6. Negación de enunciados cuantificados y control del alcance

En §3.5 vimos que una afirmación sobre *cada* objeto no permite elegir automáticamente un testigo común. Ahora preguntaremos algo distinto: ¿qué afirmación necesitamos demostrar para establecer que una proposición cuantificada es falsa? No basta añadir «no» en cualquier lugar de una frase. La negación debe afectar a la **afirmación completa**, incluidos los cuantificadores y los paréntesis que fijan su alcance.

Negar «todos»: encontrar al menos una excepción. Sea D un dominio fijado y $P(x)$ una propiedad de sus elementos. Trabajamos, como hasta aquí, con la lógica clásica y una interpretación dada de los símbolos. Entonces

$$\boxed{\neg(\forall x \in D\ P(x)) \quad \Longleftrightarrow \quad \exists x \in D\ \neg P(x).}$$

La lectura es precisa: «no todos los elementos de D cumplen P» equivale a «hay por lo menos un elemento de D que no cumple P». Para refutar una afirmación universal debemos ofrecer, o justificar que existe, un **contraejemplo** perteneciente al dominio. No afirmamos que todos los elementos fallen.

Ejemplo. Neguemos «todos los enteros pares son múltiplos de 4». Su forma explícita es

$$\forall n \in \mathbb{Z}\ (2 \mid n \to 4 \mid n).$$

Su negación comienza como $\exists n \in \mathbb{Z}\ \neg(2 \mid n \to 4 \mid n)$. Recordando del capítulo 2 que $\neg(A \to B)$ equivale a $A \wedge \neg B$, obtenemos

$$\exists n \in \mathbb{Z}\ (2 \mid n \wedge 4 \nmid n).$$

El entero 6 verifica las dos condiciones: es par y no es múltiplo de 4. Por eso refuta la universal. Elegir $n = 3$ no serviría: aunque $4 \nmid 3$, tampoco se cumple el antecedente «es par», y el condicional no queda falsado.

Lectura de la prueba. Un contraejemplo a $\forall x\,(A(x) \to B(x))$ debe satisfacer *simultáneamente* $A(x)$ y $\neg B(x)$. Negar sólo el consecuente no basta: hay que conservar la hipótesis que hace fallar la implicación.

Negar «existe»: excluir todos los posibles testigos. La segunda equivalencia intercambia los papeles:

$$\boxed{\neg(\exists x \in D\ P(x)) \quad\Longleftrightarrow\quad \forall x \in D\ \neg P(x).}$$

Decir «no existe ningún elemento de D que cumpla P» equivale a afirmar que *cada* elemento del dominio falla esa propiedad. La negación de una existencia es, por tanto, una afirmación universal, no otra afirmación existencial.

Ejemplo. La negación de

$$\exists n \in \mathbb{Z}\ (n^2 = -1)$$

es

$$\forall n \in \mathbb{Z}\ (n^2 \neq -1).$$

Para justificarla no basta ensayar $n = 0, 1, -1$ y no encontrar soluciones. Debemos cubrir todo el dominio. Si $n \in \mathbb{Z}$, su cuadrado es no negativo, por lo que no puede ser -1. Así queda excluido cualquier posible testigo. □

Las dos leyes de negación cuantificada

Para una propiedad $P(x)$ y un dominio fijado D:

$$\neg\forall x \in D\ P(x) \quad\Longleftrightarrow\quad \exists x \in D\ \neg P(x),$$
$$\neg\exists x \in D\ P(x) \quad\Longleftrightarrow\quad \forall x \in D\ \neg P(x).$$

Los paréntesis de las expresiones originales se sobreentienden sólo cuando el alcance resulta inequívoco. Estas equivalencias valen también para el dominio vacío en semántica clásica: allí la universal es verdadera y la existencial falsa.

Por qué las equivalencias expresan exactamente la negación. Consideremos primero $\forall x \in D\ P(x)$. Esta afirmación es verdadera precisamente cuando **ningún** elemento de D constituye una excepción. Por consiguiente, es falsa exactamente cuando existe una excepción: algún $a \in D$ satisface $\neg P(a)$. Hemos recuperado la primera equivalencia sin manipular mecánicamente símbolos.

Para $\exists x \in D\ P(x)$ ocurre lo contrario: su verdad exige al menos un testigo. Negarla significa que ningún candidato puede desempeñar ese papel, es decir, que todo $a \in D$ satisface $\neg P(a)$. En cada paso conservamos el dominio original: una excepción buscada entre enteros no puede reemplazarse por un número de otro conjunto.

Estas leyes se formulan aquí para negación clásica. No estamos equiparando automáticamente una prueba de inexistencia con un algoritmo que inspeccione uno a uno todos los elementos de un dominio infinito: una demostración general puede excluirlos mediante una propiedad común, como $n^2 \geq 0$ en el ejemplo anterior.

Dos cuantificadores: negar desde el exterior hacia el interior. Si una fórmula contiene varios cuantificadores, negamos primero la afirmación exterior y continuamos hasta llegar a la propiedad interna. Para dominios fijos D, E y una relación $R(x, y)$, las leyes anteriores proporcionan

$$\begin{aligned}\neg\big(\forall x \in D\ \exists y \in E\ R(x,y)\big) &\iff \exists x \in D\ \neg\big(\exists y \in E\ R(x,y)\big)\\ &\iff \exists x \in D\ \forall y \in E\ \neg R(x,y).\end{aligned}$$

La negación de «cada x tiene *algún* y relacionado con él» dice «hay *un* x para el cual **ningún** y sirve». El x elegido debe resistir todas las alternativas de y. No significa que «para cada x hay un y que falla», enunciado mucho más débil y generalmente distinto.

De manera análoga,

$$\begin{aligned}\neg\big(\exists y \in E\ \forall x \in D\ R(x,y)\big) &\iff \forall y \in E\ \neg\big(\forall x \in D\ R(x,y)\big)\\ &\iff \forall y \in E\ \exists x \in D\ \neg R(x,y).\end{aligned}$$

Ahora la negación afirma que **cada candidato uniforme** y tiene al menos una excepción x; esta excepción puede depender de y. Vuelve así la dependencia de testigos que distinguimos en §3.5.

Comprobación matemática. Allí demostramos que $C:\ \forall x \in \mathbb{Z}\ \exists y \in \mathbb{Z}\ (y > x)$ es verdadera. Su negación es

$$\exists x \in \mathbb{Z}\ \forall y \in \mathbb{Z}\ (y \leq x),$$

que afirma que existe un entero mayor o igual que todos los enteros. Es falsa: dado cualquier candidato x, $y = x + 1$ lo supera. También demostramos que $D:\ \exists y \in \mathbb{Z}\ \forall x \in \mathbb{Z}\ (y > x)$ es falsa. Su negación es

$$\forall y \in \mathbb{Z}\ \exists x \in \mathbb{Z}\ (y \leq x),$$

que es verdadera: para cualquier y sirve $x = y$. Las dos verificaciones concuerdan con los valores de verdad de C y D. No hemos cambiado sólo los signos: hemos identificado qué clase de testigo exige cada enunciado negado.

Las restricciones de dominio y el alcance de la negación. En §3.2 interpretamos la cuantificación universal restringida como un condicional, mientras que en §3.3 interpretamos la existencial restringida como una conjunción. Esa diferencia importa al negar. Si D está contenido en un universo de referencia U, tenemos

$$\begin{aligned}\neg\big(\forall x \in D\ P(x)\big) &\iff \neg\forall x \in U\ (x \in D \to P(x))\\ &\iff \exists x \in U\ (x \in D \wedge \neg P(x)),\end{aligned}$$

mientras que

$$\begin{aligned}\neg\big(\exists x \in D\ P(x)\big) &\iff \neg\exists x \in U\ (x \in D \wedge P(x))\\ &\iff \forall x \in U\ (x \in D \to \neg P(x)).\end{aligned}$$

En ambos casos la propiedad negada sigue refiriéndose a los elementos de D. No convertimos «ningún entero satisface P» en «todos los objetos del universo, sean enteros o no, dejan de satisfacer P».

También debemos conservar los paréntesis de los conectivos. Sean $A(x)$ y $B(x)$ propiedades sobre un mismo dominio. La negación de «todo objeto satisface ambas propiedades» es

$$\neg\forall x \in D\ (A(x) \wedge B(x)) \quad \Longleftrightarrow \quad \exists x \in D\ (\neg A(x) \vee \neg B(x)).$$

El testigo sólo necesita fallar **alguna** de las dos propiedades, no necesariamente ambas. En cambio,

$$\neg\exists x \in D\ (A(x) \vee B(x)) \quad \Longleftrightarrow \quad \forall x \in D\ (\neg A(x) \wedge \neg B(x)).$$

En este último caso todos los objetos deben fallar **las dos** propiedades. La transformación del conectivo interno obedece a las leyes de De Morgan: después de cambiar el cuantificador, hay que seguir negando la fórmula que queda bajo su alcance.

Una variable libre no queda ligada por negar una fórmula. Volvamos al tipo de expresión de §3.4:

$$F(y): \quad (\forall x \in \mathbb{Z}\ (x^2 \geq 0)) \wedge (y < 3).$$

Aquí y está libre. Su negación, relativa a una asignación o contexto que dé valor a y, es

$$\neg F(y) \quad \Longleftrightarrow \quad (\exists x \in \mathbb{Z}\ (x^2 < 0)) \vee (y \geq 3).$$

El cuantificador sobre x actúa exclusivamente en el primer componente; la variable y sigue libre en el segundo. Como no hay enteros con cuadrado negativo, la negación tiene, para cada valor fijado de y, el mismo valor de verdad que $y \geq 3$. **Negar no cierra una fórmula abierta ni amplía por arte de magia el alcance de un cuantificador.**

Auditoría de una negación. Ante una expresión compleja, primero delimita la fórmula entera que se desea negar y distingue sus variables libres de las ligadas. Después conserva los dominios, transforma cada cuantificador que esté dentro de la negación y lleva ésta hasta las propiedades elementales mediante las equivalencias proposicionales. Por último, traduce el resultado al lenguaje ordinario: debe expresar exactamente una situación incompatible con el enunciado original, no simplemente una frase que suene negativa.

Después de la prueba. Sabemos ya refutar universales mediante excepciones, negar existencias mediante una exclusión general y enfrentar cadenas de cuantificadores sin perder la dependencia de los testigos. En §3.7 estudiaremos una exigencia diferente: no sólo que un objeto exista, sino que sea el **único** que satisface una determinada propiedad.

3.7. Existencia y unicidad: dos exigencias distintas

En §3.3 aprendimos a demostrar una existencia mediante un testigo; en §3.5 advertimos que ese testigo puede depender de una variable elegida antes. Nada de ello garantiza que no haya otros objetos que satisfagan la misma propiedad. Decir «hay alguno» y decir «hay exactamente uno» son compromisos distintos. Para escribir y demostrar el segundo necesitamos reunir una afirmación existencial y otra universal.

De «alguno» a «exactamente uno». Sea D un dominio fijado y $P(x)$ una propiedad de sus elementos. Escribiremos

$$\exists! x \in D\ P(x)$$

para afirmar «existe **un único** x en D que satisface P». El signo $\exists!$ es una abreviatura, no una nueva clase de testigo: su significado se expresa utilizando los cuantificadores y la igualdad que ya conocemos.

Definición: existe un único

La expresión $\exists! x \in D\ P(x)$ significa

$$\boxed{\exists x \in D\ \big(P(x) \land \forall y \in D\ (P(y) \to y = x)\big).}$$

Exige conjuntamente **existencia** —un x del dominio cumple P— y **unicidad** —todo y del mismo dominio que cumpla P coincide con ese x. El nombre de la variable ligada y es auxiliar; debe elegirse sin capturar otras variables libres.

Observemos la posición de los conectivos. El candidato x ha de satisfacer $P(x)$ **y**, para ese mismo candidato, debemos probar la afirmación universal sobre los demás objetos. No bastaría escribir $\exists x \in D\ P(x)$ y añadir informalmente que «es el elegido»: nuestra elección no excluye otras posibilidades.

Existe otra formulación especialmente útil al construir demostraciones. Definamos la condición «a lo sumo uno satisface P» mediante

$$\forall u \in D\ \forall v \in D\ \big((P(u) \land P(v)) \to u = v\big).$$

Esta condición permite **cero o un** objeto: no asegura que haya alguno. Con ella podemos desplegar la unicidad existencial como

$$\begin{aligned} \exists! x \in D\ P(x) \quad \Longleftrightarrow \quad & \big(\exists x \in D\ P(x)\big) \\ & \land \big(\forall u \in D\ \forall v \in D\ ((P(u) \land P(v)) \to u = v)\big). \end{aligned}$$

Justificación de la equivalencia. Si existe un x con la propiedad y todo objeto que la tiene es igual a él, entonces dos candidatos cualesquiera u, v son ambos iguales a x y, por tanto, iguales entre sí. Recíprocamente, si existe un candidato x y dos candidatos cualesquiera han de ser iguales, al tomar uno de ellos igual a x concluimos que todo y con $P(y)$ satisface $y = x$. Así se recupera la definición anterior. □

Tres situaciones que no debemos confundir. Sobre el dominio $\mathbb{Z}$ podemos observar tres problemas diferentes:

Propiedad $P(n)$	Qué ocurre	Consecuencia
$n^2 = -1$	Ningún entero satisface la igualdad, pues $n^2 \geq 0$.	No hay existencia; por tanto, tampoco existe un único testigo.
$n^2 = 1$	Tanto 1 como -1 son enteros distintos que la satisfacen.	Hay existencia, pero falla la unicidad.
$n + 5 = 0$	El entero -5 satisface la igualdad y cualquier solución debe ser igual a -5.	Hay existencia y unicidad.

La tercera fila requiere justificar su última afirmación, no sólo reconocer el candidato por inspección. Si $n + 5 = 0$, sumando -5 a ambos lados obtenemos $n = -5$. Por consiguiente, cualquier entero que satisfaga la propiedad coincide con el testigo presentado.

Lectura guiada. Ante «existe un único», formula dos preguntas separadas: «¿qué objeto satisface la propiedad?» y «¿por qué dos objetos que la satisfacen no pueden ser distintos?». A veces la misma manipulación algebraica ayuda a responder ambas; aun así, las exigencias lógicas siguen siendo dos.

Una demostración completa: el opuesto de cada entero. El ejemplo de los opuestos nos permite combinar esta sección con el orden de cuantificadores de §3.5.

Proposición. Para cada entero a existe un único entero b tal que $a + b = 0$. En símbolos,

$$\forall a \in \mathbb{Z}\ \exists! b \in \mathbb{Z}\ (a + b = 0).$$

Demostración. Sea $a \in \mathbb{Z}$ arbitrario.

Existencia. Tomemos $b = -a$. Como $-a \in \mathbb{Z}$ y $a + (-a) = 0$, tenemos un testigo de la existencia.

Unicidad. Supongamos que $b, c \in \mathbb{Z}$ satisfacen $a + b = 0$ y $a + c = 0$. De estas igualdades se sigue que $a + b = a + c$. Sumando $-a$ a ambos miembros y utilizando las propiedades habituales de la suma de enteros, obtenemos $b = c$. Por tanto, cualesquiera dos candidatos coinciden.

Hemos probado ambas exigencias para un entero a arbitrario; en consecuencia, la afirmación universal queda demostrada. □

Después de la prueba. La estrategia no fue repetir dos veces que $b = -a$. Primero exhibimos un candidato; después supusimos que existían *dos candidatos cualesquiera* y demostramos su igualdad. La segunda operación certifica que no se ha omitido ninguna alternativa. Esta es la estructura reutilizable de una demostración de existencia y unicidad; los métodos técnicos que permiten hallar el candidato pueden variar según el problema.

Unicidad para cada parámetro no significa un testigo común. El cuantificador universal exterior de la proposición anterior importa. La afirmación

$$\forall a \in \mathbb{Z}\ \exists! b \in \mathbb{Z}\ (a + b = 0)$$

permite que el entero único b **dependa de** a: para $a = 2$ es $b = -2$, mientras que para $a = 3$ es $b = -3$. No podemos sustituirla por

$$\exists! b \in \mathbb{Z}\ \forall a \in \mathbb{Z}\ (a + b = 0),$$

que requeriría un solo b capaz de funcionar para todos los valores de a. Esta última es falsa: si existiera tal entero, $a = 0$ exigiría $b = 0$ y $a = 1$ exigiría $b = -1$. El error no está en la unicidad de cada opuesto, sino en intercambiar el orden de las elecciones.

La fórmula $\exists! b \in \mathbb{Z}\ (a + b = 0)$ considerada por sí sola contiene a a libre. Sólo cuando fijamos un valor de a en el contexto o añadimos un cuantificador exterior obtenemos la afirmación correspondiente. La notación de unicidad no elimina la necesidad de controlar las demás variables.

¿Cómo negar «existe exactamente uno»?. Las leyes de §3.6 nos permiten distinguir las dos maneras de fallar. En lógica clásica,

$$\begin{aligned}\neg\big(\exists! x \in D\ P(x)\big) \quad \Longleftrightarrow \quad & \big(\forall x \in D\ \neg P(x)\big) \\ & \vee\ \big(\exists u \in D\ \exists v \in D\ (P(u) \wedge P(v) \wedge u \neq v)\big).\end{aligned}$$

O bien **no hay ningún** objeto con la propiedad, o bien **hay al menos dos distintos**. Esta disyunción es exhaustiva: si algún objeto cumple P pero no existe uno único, tiene que haber otro candidato distinto. No debemos confundirla con la mera negación de la existencia, que contempla únicamente el primer escenario.

El dominio vacío y la unicidad

Si $D = \varnothing$, la condición «a lo sumo uno» es verdadera: no hay dos elementos del dominio que puedan contradecirla. Sin embargo, $\exists! x \in D\ P(x)$ es falsa, porque su componente existencial carece de testigo. Este caso confirma que «a lo sumo uno» y «exactamente uno» no son sinónimos.

Auditoría de un argumento incompleto. Leamos el razonamiento: «6 es par y no es múltiplo de 4; por tanto, existe un único entero par que no es múltiplo de 4». La premisa sí justifica una existencia, como vimos en §3.3, pero no contiene ninguna prueba de unicidad. De hecho, 2 es otro entero par no divisible por 4, distinto de 6. La conclusión es falsa. La reparación correcta consiste en conservar únicamente lo demostrado —«existe al menos uno»—, no en añadir una frase solemne sobre la elección del testigo.

Después de la prueba. Ya disponemos de una pauta precisa para leer el signo $\exists!$: desplegarlo, comprobar por separado el testigo y la igualdad de candidatos, conservar el dominio y vigilar los parámetros exteriores. El siguiente tramo del capítulo llevará estas herramientas a un trabajo de traducción: pasar de la prosa matemática a los símbolos y volver de ellos a una frase cuyo significado podamos defender.

3.8. Traducción entre prosa matemática y lenguaje cuantificado

Hasta ahora hemos aprendido a reconocer qué hace cada cuantificador, a vigilar su alcance y a separar existencia de unicidad. Falta una tarea que reúne todas las anteriores: **escribir una afirmación matemática en símbolos sin alterar lo que dice y recuperar, desde esos símbolos, una frase inequívoca**. Una fórmula sintácticamente correcta puede ser una mala traducción. La verificación no consiste en comprobar si «se parece» a la frase, sino en reconstruir, en ambas direcciones, las condiciones bajo las cuales sería verdadera o falsa.

Antes de simbolizar: fijar el contrato de lectura. Trabajaremos principalmente sobre los enteros $\mathbb{Z}$. Cuando aparezcan $P(x)$, $Q(x)$ o $R(x, y)$, serán propiedades o relaciones cuyo significado se declarará antes de utilizarlas. Una frase matemática bien traducida debe permitir identificar cinco elementos:

1. **Dominio.** ¿De qué objetos se habla? ¿De todos los enteros o sólo de los pares? ¿Qué conjuntos recorren las otras variables?
2. **Fuerza.** ¿Se afirma algo de cada objeto, se pide al menos uno, se excluye todo testigo o se exige exactamente uno?
3. **Dependencia.** Si hay varios cuantificadores, ¿puede elegirse un testigo después de conocer otro objeto o debe servir uno mismo para todos?
4. **Conectivos.** ¿La frase afirma una condición, una conjunción, una alternativa o una equivalencia? ¿Cuál es el antecedente de cada «si»?
5. **Alcance.** ¿Hasta dónde llega cada cuantificador y cada negación? ¿Quedó alguna variable libre que el contexto deba fijar?

Esta lista funciona como auditoría, no como sustituto del razonamiento. La pregunta final es siempre la misma: **¿podría una situación hacer verdadera la frase y falsa la fórmula, o viceversa?** Si la respuesta es sí, no hemos conservado el significado.

«Todos los que. . . » y «alguno que. . . »: el conectivo cambia. Comencemos por una afirmación ya conocida: «todo entero divisible por 6 es divisible por 3». El objeto variable es un entero n; no afirmamos que todos sean divisibles por 6, sino que esa condición garantiza la divisibilidad por 3. Por tanto,

$$\boxed{\forall n \in \mathbb{Z}\ (6 \mid n \to 3 \mid n).}$$

Comprobación de la traducción. El cuantificador recorre todos los enteros, pero la hipótesis $6 \mid n$ aparece dentro del condicional. Cuando $n = 1$, el condicional es verdadero porque su antecedente es falso; la frase tampoco exige que 1 sea múltiplo de 6. Si hubiéramos escrito $\forall n \in \mathbb{Z}\ (6 \mid n \wedge 3 \mid n)$, habríamos afirmado falsamente que *todos* los enteros cumplen ambas condiciones. Para demostrar la afirmación correcta, dado $6 \mid n$ escribimos $n = 6k = 3(2k)$ con $k \in \mathbb{Z}$.

En cambio, «existe un entero par que no es divisible por 4» pide que **un mismo testigo** satisfaga dos condiciones:

$$\boxed{\exists n \in \mathbb{Z}\ (2 \mid n \wedge 4 \nmid n).}$$

El entero 6 es un testigo. Reemplazar $\wedge$ por $\to$ daría otra afirmación: un entero impar, por ejemplo 1, satisface $2 \mid n \to 4 \nmid n$ sin ser el entero par requerido. Las formas de cuantificación restringida vistas en §§3.2–3.3 explican la diferencia: una restricción universal se expresa mediante un condicional y una existencial mediante una conjunción.

«Si», «sólo si» y «si y sólo si»: identificar la dirección. Introduzcamos dos propiedades de un entero: $P(n)$ significa $12 \mid n$ y $Q(n)$ significa $6 \mid n$. La oración «un entero es divisible por 12 **sólo si** es divisible por 6» declara que ser divisible por 6 es una **condición necesaria** para ser divisible por 12:

$$\forall n \in \mathbb{Z}\ (P(n) \to Q(n)).$$

La oración «un entero es divisible por 12 **si** es divisible por 6» invierte las funciones gramaticales: ahora la condición anunciada después de «si» sería suficiente para P:

$$\forall n \in \mathbb{Z}\ (Q(n) \to P(n)).$$

Esta segunda afirmación es falsa: $n = 6$ verifica $Q(6)$ y no $P(6)$. La primera es verdadera, pues $n = 12k = 6(2k)$ siempre que $12 \mid n$. Un pequeño contraejemplo distingue dos fórmulas que, a primera vista, contienen las mismas letras.

La locución «si y sólo si» exige las dos direcciones. Por ejemplo, la proposición «un entero es par si y sólo si su cuadrado es par» se expresa como

$$\forall n \in \mathbb{Z}\ \big(2 \mid n \leftrightarrow 2 \mid n^2\big).$$

Esta traducción fija el *enunciado*; demostrar su verdad requiere justificar ambas implicaciones, como hicimos con los bicondicionales en el capítulo 2. **Traducir correctamente no es todavía demostrar.**

Tres lecturas de una condición

Con el mismo dominio y las mismas propiedades P, Q:

$$\begin{aligned} \text{«}P \text{ si } Q\text{»}: &\quad Q \to P, \\ \text{«}P \text{ sólo si } Q\text{»}: &\quad P \to Q, \\ \text{«}P \text{ si y sólo si } Q\text{»}: &\quad P \leftrightarrow Q. \end{aligned}$$

Cuando P, Q contienen una variable libre, estas expresiones siguen siendo fórmulas abiertas hasta que se fije su valor o se cuantifique esa variable. Por ejemplo, para afirmar cualquiera de ellas *sobre todos los enteros* debemos anteponer $\forall n \in \mathbb{Z}$ con el alcance adecuado.

El orden no se recupera contando símbolos. «Cada entero tiene un entero mayor que él» se lee en dos elecciones consecutivas: tomamos primero un entero cualquiera x y después buscamos un entero y mayor que ese x:

$$\forall x \in \mathbb{Z}\ \exists y \in \mathbb{Z}\ (y > x).$$

La frase «hay un entero que es mayor que todos los enteros» cambia la elección: primero exige un entero fijo y y luego afirma que supera a cualquier x:

$$\exists y \in \mathbb{Z}\ \forall x \in \mathbb{Z}\ (y > x).$$

En §3.5 probamos que la primera es verdadera y la segunda falsa. Este contraste muestra por qué el mero recuento de un $\forall$ y un $\exists$ no basta: **su orden codifica la dependencia del testigo**.

La misma cautela vale cuando interviene la unicidad. «Cada entero tiene un único opuesto aditivo entero» significa

$$\forall a \in \mathbb{Z}\ \exists! b \in \mathbb{Z}\ (a + b = 0).$$

El entero b es único *para el valor de a que se ha fijado*, no un candidato que deba servir a todos los enteros. La fórmula $\exists! b \in \mathbb{Z}\ \forall a \in \mathbb{Z}\ (a + b = 0)$ dice otra cosa y es falsa, como vimos en §3.7. Si escribiéramos solamente $\exists! b \in \mathbb{Z}\ (a + b = 0)$ sin declarar a, dejaríamos una variable libre y, por tanto, una tarea de interpretación pendiente.

«No todos», «ninguno» y «no existe uno único». En el lenguaje corriente, la negación puede quedar lejos de aquello que niega. Conviene identificar primero **la afirmación completa** y sólo entonces simbolizarla. «No todos los enteros pares son divisibles por 4» se expresa como

$$\neg\forall n \in \mathbb{Z}\ (2 \mid n \to 4 \mid n),$$

y, por §3.6, equivale a

$$\exists n \in \mathbb{Z}\ (2 \mid n \wedge 4 \nmid n).$$

La frase afirma que existe una excepción, no que *ningún* par sea divisible por 4. El entero 6 verifica la negación; el entero 8 muestra por qué sería incorrecto afirmar que todos los pares incumplen la propiedad.

«No existe ningún entero mayor que todos los enteros» comienza por negar **toda** una existencia uniforme:

$$\neg\exists y \in \mathbb{Z}\ \forall x \in \mathbb{Z}\ (y > x) \quad\Longleftrightarrow\quad \forall y \in \mathbb{Z}\ \exists x \in \mathbb{Z}\ (y \leq x).$$

Leída de regreso, la fórmula de la derecha dice: «para cada posible candidato y, hay al menos un entero x que no es menor que él». Sirve $x = y$, aunque también podría elegirse $x = y + 1$. La transformación conserva el orden **después** de aplicar las leyes de negación; no introduce un candidato uniforme nuevo.

Por último, «no existe un único entero cuyo cuadrado sea 1» significa

$$\neg\exists! n \in \mathbb{Z}\ (n^2 = 1).$$

La frase es verdadera porque 1 y -1 son dos soluciones distintas. No podemos traducirla por $\forall n \in \mathbb{Z}\ (n^2 \neq 1)$, que afirmaría que no hay soluciones y es falsa. §3.7 explica las dos formas de fracasar una existencia única: ausencia de todo testigo o presencia de dos distintos.

«Al menos dos» y «a lo sumo uno» sin nuevos cuantificadores. La prosa a veces expresa cantidades sin utilizar las palabras «todos» o «alguno». Sean $P(x)$ una propiedad y D su dominio. «Hay al menos dos elementos distintos de D que cumplen P» significa

$$\exists u \in D\ \exists v \in D\ \big(P(u) \wedge P(v) \wedge u \neq v\big).$$

La desigualdad $u \neq v$ es indispensable: sin ella podríamos repetir el mismo testigo dos veces y demostrar solamente una existencia. En cambio, «a lo sumo uno cumple P» se traduce como

$$\forall u \in D\ \forall v \in D\ \big((P(u) \wedge P(v)) \to u = v\big).$$

Esta última fórmula puede ser verdadera aunque no exista ningún objeto con la propiedad. Para afirmar «exactamente uno», añadimos una existencia como hicimos en §3.7. La elección entre «al menos», «a lo sumo» y «exactamente» forma parte del significado, no de una corrección estilística de la frase.

Lectura inversa: recuperar una oración completa. Tomemos, con $n, m \in \mathbb{Z}$, la fórmula

$$\forall n \in \mathbb{Z}\,\big(2 \mid n \to \exists m \in \mathbb{Z}\,(2 \nmid m \wedge m > n)\big).$$

No conviene leerla acumulando símbolos sueltos. El cuantificador exterior ordena considerar **cada entero** n. El condicional acota la exigencia a los casos en los que n es par. La existencia interior pide entonces **un entero** m que cumpla conjuntamente dos condiciones: ser impar y superar a ese n. Una frase fiel es: «Todo entero par tiene algún entero impar mayor que él». La afirmación es verdadera: dado n par, $m = n + 1$ es impar y mayor que n.

A continuación, consideremos

$$\exists n \in \mathbb{Z}\,\forall m \in \mathbb{Z}\,\big(2 \mid n \wedge m > n\big).$$

Una lectura literal dice: «existe un entero par n tal que **todo** entero m es mayor que n». Ésta es falsa, porque para cualquier candidato n podemos elegir $m = n$. Aunque aparecen las palabras «par», «existe» y «mayor», la fórmula no traduce la oración anterior: ha cambiado la posición de la conjunción, ha desaparecido la condición «impar» y se ha invertido el orden de dependencia. El contraste impide confundir semejanza superficial con equivalencia semántica.

Laboratorio de reparación: tres versiones de una frase. Queremos formalizar «todo entero divisible por 4 es par». Un borrador podría proponer:

$$\begin{aligned} F_1 &: \quad \forall n \in \mathbb{Z}\,(4 \mid n \wedge 2 \mid n), \\ F_2 &: \quad \exists n \in \mathbb{Z}\,(4 \mid n \to 2 \mid n), \\ F_3 &: \quad \forall n \in \mathbb{Z}\,(4 \mid n \to 2 \mid n). \end{aligned}$$

F_1 impone ambas propiedades a todos los enteros y falla, por ejemplo, en $n = 1$. F_2 sólo pide que un caso haga verdadero el condicional y también lo satisface $n = 1$; no cubre todos los múltiplos de 4. Únicamente F_3 conserva a la vez el dominio, la universalidad y el antecedente condicionado. No seleccionamos F_3 porque sepamos de antemano que es verdadera: la seleccionamos porque **afirma exactamente lo que dice la oración**. La comprobación de verdad y la demostración vienen después.

Auditoría bidireccional de una traducción

Prosa → símbolos: subraya dominio, cuantificadores, dependencias, conectivos y negaciones; construye la fórmula con paréntesis explícitos.

Símbolos → prosa: lee los cuantificadores en orden, declara en qué casos se activa cada condicional y nombra todas las condiciones que debe satisfacer un testigo.

Control final: ensaya un caso límite o un contraejemplo que distinga la fórmula elegida de una traducción vecina. No sustituyas la equivalencia semántica por el hecho de que ambas afirmaciones tengan casualmente el mismo valor de verdad en un ejemplo.

Después de la traducción. Hemos pasado de identificar cuantificadores individuales a reconstruir afirmaciones completas. Ya podemos comparar una frase con su fórmula, reconocer una traducción demasiado fuerte o demasiado débil y diagnosticar qué hipótesis, variables o elecciones se perdieron. Los ejercicios que siguen reúnen las ocho primeras secciones del capítulo: lectura, traducción, negación, prueba, unicidad y reparación de argumentos. Sus soluciones desarrolladas se encuentran en §3.10.

3.9. Ejercicios graduados: leer, traducir, negar y reparar

Los ejercicios reúnen lo aprendido en §§3.1–3.8. No basta reconocer símbolos: en cada respuesta debe quedar claro **qué afirma la fórmula**, cuáles son sus dominios y qué justifica la respuesta. Cuando se pida una refutación, proporciona un contraejemplo del dominio y verifica las condiciones pertinentes; cuando se pida existencia, comprueba el testigo; cuando se pida unicidad, excluye explícitamente otros candidatos. Las soluciones desarrolladas se presentan en §3.10.

Salvo indicación contraria, los números y las variables aritméticas pertenecen a $\mathbb{Z}$; $a \mid b$ significa que existe $k \in \mathbb{Z}$ tal que $b = ak$. Las letras P, Q, R designarán las propiedades especificadas en cada ejercicio, no proposiciones cerradas por defecto.

Nivel I. Reconocer lo que efectivamente se afirma.

Ejercicio 1. Ocurrencias libres y ligadas. Examina la fórmula

$$(\forall x \in \mathbb{Z}\ (x + y = y + x)) \wedge (x < 3).$$

1. Identifica **cada ocurrencia** de x e y como libre o ligada; indica el alcance del cuantificador.
2. ¿Es la fórmula completa sintácticamente cerrada? ¿Qué datos adicionales permitirían evaluarla?
3. Escribe una fórmula cerrada que afirme «para todos los enteros x e y, $x+y = y+x$». Explica por qué tu modificación no consiste simplemente en declarar que el x final ya estaba ligado.

Ejercicio 2. Universal, existencia y clase de justificación. Considera

$$A:\ \forall n \in \mathbb{Z}\ (6 \mid n \to 3 \mid n), \qquad B:\ \exists n \in \mathbb{Z}\ (3 \mid n \wedge 6 \nmid n).$$

1. Traduce A y B a oraciones completas, conservando conectivos y cuantificadores.
2. Determina el valor de verdad de cada una. Para A, presenta un argumento general a partir de la divisibilidad; para B, presenta y verifica un testigo.
3. Explica por qué tu testigo para B no serviría, por sí solo, para demostrar A.

Ejercicio 3. El dominio vacío como prueba de lectura. Sea $D = \varnothing$ y sea $P(x)$ una propiedad definida para los elementos de D. Determina el valor de verdad de cada afirmación bajo la semántica clásica y justifica tu respuesta sin inventar un elemento del dominio:

1. $\forall x \in D\ P(x)$.
2. $\exists x \in D\ P(x)$.
3. $\forall u \in D\ \forall v \in D\ ((P(u) \wedge P(v)) \to u = v)$.
4. $\exists! x \in D\ P(x)$.

Indica cuáles expresan «a lo sumo uno» y «exactamente uno», y por qué no son intercambiables.

Ejercicio 4. «Si» y «sólo si»: una dirección puede cambiarlo todo. Para cada $n \in \mathbb{Z}$, sean $P(n)$: «$4 \mid n$» y $Q(n)$: «$2 \mid n$». Formaliza mediante una fórmula **cuantificada** cada oración; determina su valor de verdad y justifícalo:

1. «$P(n)$ sólo si $Q(n)$, para todo entero n».
2. «$P(n)$ si $Q(n)$, para todo entero n».
3. «$P(n)$ si y sólo si $Q(n)$, para todo entero n».

Una respuesta correcta debe explicar cuál propiedad es antecedente en cada caso. No decidas la dirección de la flecha por el valor de verdad que esperas obtener.

Nivel II. Conservar el alcance y las dependencias.

Ejercicio 5. Un opuesto para cada entero no es un opuesto común. Sean

$$A:\ \forall a \in \mathbb{Z}\ \exists b \in \mathbb{Z}\ (a+b=0), \qquad B:\ \exists b \in \mathbb{Z}\ \forall a \in \mathbb{Z}\ (a+b=0).$$

1. Traduce ambos enunciados a la prosa y señala cuándo se elige b en cada uno.
2. Demuestra A exhibiendo un testigo que pueda depender de a.
3. Refuta B sin limitarte a afirmar que «los cuantificadores no conmutan»: obtén dos condiciones incompatibles para un supuesto testigo fijo.

Ejercicio 6. Negaciones que alcanzan hasta la propiedad interior. Niega las dos proposiciones siguientes. Escribe primero la negación de **toda** la fórmula, después una fórmula equivalente sin negaciones antepuestas a cuantificadores ni a conectivos, y finalmente una oración inequívoca en español.

$$A:\ \forall n \in \mathbb{Z}\ (3 \mid n \to 9 \mid n),$$

$$B:\ \exists y \in \mathbb{Z}\ \forall x \in \mathbb{Z}\ (y > x).$$

Determina si A y B son verdaderas o falsas y justifica el valor de sus negaciones. En A, un contraejemplo debe cumplir el antecedente y fallar en el consecuente.

Ejercicio 7. Traducir una restricción y reparar otra traducción. Formaliza: «Todo entero mayor que 3 tiene algún entero estrictamente mayor que él». Conserva el dominio de ambas variables y el alcance de la existencia. Después examina el borrador

$$\forall n \in \mathbb{Z}\ \big(n > 3 \land \exists m \in \mathbb{Z}\ (m > n)\big).$$

1. Lee literalmente el borrador y explica por qué no expresa la oración inicial.
2. Reemplaza únicamente lo necesario para producir una traducción fiel.
3. Justifica la verdad de la oración inicial y aporta un entero que refute el borrador. Distingue el defecto semántico de la evaluación aritmética.

Ejercicio 8. Separar las dos obligaciones de $\exists!$. Estudia las afirmaciones

$$A:\ \exists! n \in \mathbb{Z}\ (n+5=0), \qquad B:\ \exists! n \in \mathbb{Z}\ (n^2=1).$$

1. Despliega A mediante cuantificadores ordinarios e igualdad. Demuestra por separado existencia y unicidad.
2. Determina el valor de verdad de B. Explica si falla la existencia o la unicidad y presenta los datos que lo prueban.
3. Escribe en símbolos «a lo sumo un entero tiene cuadrado 1» y evalúa esa afirmación. Explica por qué no equivale a B.

Nivel III. Auditar fórmulas y argumentos.

Ejercicio 9. Renombrar sin capturar una variable libre. Fija $D = \{0, 1\}$, considera $x \in D$ y deja $y \in \mathbb{Z}$ libre en

$$F(y):\quad \forall x \in D\ (x < y).$$

Una persona intenta renombrar la variable ligada y propone $G:\ \forall y \in D\ (y < y)$.

1. Identifica exactamente qué ocurrencias cambiaron de condición libre o ligada. Propón un renombrado correcto usando t.
2. Asigna $y = 2$ a la fórmula original y compara sus valores de verdad con el de G.
3. Explica por qué el error es una **captura de variable** y no una elección inocente de letras. ¿Qué diferencia sintáctica permanece incluso si dos fórmulas coinciden en valor de verdad bajo alguna asignación?

Ejercicio 10. La generalización ilegítima. Se ofrece la siguiente «demostración»: «Tomemos $n = 20$. Como $10 \mid 20$ y $5 \mid 20$, hemos probado que para todo entero n, si $10 \mid n$, entonces $5 \mid n$».

1. Formula la conclusión pretendida con cuantificador, dominio y paréntesis.
2. Señala el paso que no está justificado; distingue la comprobación de un caso del uso de un entero arbitrario.
3. Reescribe la demostración completa a partir de la definición de divisibilidad. Señala qué hipótesis introduces y cuándo puedes cerrar la generalización.

Ejercicio 11. ¿Qué dirección permite el testigo uniforme?. Sean D, E dominios fijados, posiblemente vacíos, y $R(x, y)$ una propiedad de pares. Considera

$$U:\ \exists y \in E\ \forall x \in D\ R(x,y), \qquad V:\ \forall x \in D\ \exists y \in E\ R(x,y).$$

1. Demuestra $U \to V$ sin suponer que los dominios sean finitos ni no vacíos.
2. Para mostrar que la recíproca no es válida en general, toma $D = E = \{0, 1\}$ y $R(x, y)$: «$x = y$». Comprueba por separado el valor de U y el de V.
3. Compara además las parejas $\forall x \in D\ \forall y \in E\ R(x,y)$ y $\forall y \in E\ \forall x \in D\ R(x,y)$, así como $\exists x \in D\ \exists y \in E\ R(x,y)$ y $\exists y \in E\ \exists x \in D\ R(x,y)$. Explica qué permite intercambiar esos cuantificadores iguales **con dominios fijos y sin alterar** R, y contrástalo con la permutación de $\forall$ y $\exists$ del ejercicio.

Ejercicio 12. Negación completa de una unicidad parametrizada. Sea

$$S:\quad \forall a \in \mathbb{Z}\ \exists! b \in \mathbb{Z}\ (a + b = 0).$$

1. Expande $\exists!$ en una fórmula con $\exists$, $\forall$, igualdad y conectivos ordinarios, prestando atención a qué variable sigue dependiendo de a.
2. Obtén una fórmula equivalente a $\neg S$ **sin usar** $\exists!$. Debe mostrar los dos modos de fracaso posibles para algún entero a: que no haya solución o que existan dos soluciones distintas. Especifica los dominios de todos los cuantificadores.
3. Traduce la negación a prosa y determina su valor de verdad mediante un argumento válido para un a arbitrario. Explica por qué comprobar sólo $a = 0$ y $a = 1$ no bastaría para justificar el valor de verdad de S.

Comprobación final de aprendizaje. Antes de consultar las soluciones, revisa para cada ejercicio si has conservado todos los dominios, distinguido variables libres y ligadas, ubicado correctamente los condicionales y justificado las afirmaciones universales o existenciales con el tipo de prueba que requieren. Una fórmula bien formada no es, por ese solo hecho, una traducción fiel; una afirmación verdadera no queda demostrada por cualquier argumento.

3.10. Soluciones desarrolladas y cierre del capítulo

Estas soluciones explicitan tanto la respuesta como la **obligación lógica** que la justifica. Conviene compararlas con un intento propio: acertar el valor de verdad no compensa una traducción equivocada, y disponer de un ejemplo no sustituye una demostración universal.

Solución 1. Ocurrencias libres y ligadas. 1. Alcance y ocurrencias. En

$$(\forall x \in \mathbb{Z}\ (x + y = y + x)) \wedge (x < 3)$$

el alcance de $\forall x \in \mathbb{Z}$ es sólo la igualdad $x + y = y + x$ situada entre los primeros paréntesis. Las **dos ocurrencias de** x **en esa igualdad** están ligadas; las **dos ocurrencias de** y están libres. El x que aparece en $x < 3$ también está libre: se encuentra fuera del alcance del cuantificador, aunque tenga el mismo nombre.

2. Cierre y evaluación. La fórmula completa es abierta, porque contiene las ocurrencias libres de x y y. Para evaluarla en la aritmética usual basta asignarles valores enteros, por

ejemplo $x = a$ e $y = b$ con $a, b \in \mathbb{Z}$. El primer miembro de la conjunción será verdadero por conmutatividad de la suma; el valor de la fórmula entera dependerá de si $a < 3$. Esa asignación permite evaluar la fórmula, pero **no la cierra sintácticamente**.

3. La afirmación solicitada. Una escritura cerrada es

$$\forall x \in \mathbb{Z}\, \forall y \in \mathbb{Z}\, (x + y = y + x).$$

También podemos intercambiar estos dos universales, conservando la igualdad. No hemos «extendido» retrospectivamente el primer cuantificador hasta el último x de la fórmula de partida: hemos escrito **una afirmación nueva**, cuyo único contenido es la conmutatividad para todos los pares. La conjunción $x < 3$ no pertenece a la oración que se pidió formalizar.

Solución 2. Universal, existencia y clase de justificación. 1. Lectura. A dice: «Para todo entero n, si n es divisible por 6, entonces es divisible por 3». B dice: «Existe un entero n divisible por 3 y no divisible por 6». El condicional de A no debe convertirse en una conjunción, ni la conjunción de B en un condicional.

2. Valores de verdad y pruebas. Ambas afirmaciones son verdaderas. Para A, sea $n \in \mathbb{Z}$ arbitrario y supongamos $6 \mid n$. Existe $k \in \mathbb{Z}$ tal que $n = 6k = 3(2k)$; como $2k \in \mathbb{Z}$, se sigue $3 \mid n$. Cerrada la hipótesis condicional y al ser arbitrario n, queda demostrado A.

Para B sirve el testigo $n = 3$: es entero y $3 = 3 \cdot 1$, así que $3 \mid 3$; no existe $k \in \mathbb{Z}$ con $3 = 6k$, pues exigiría $k = 1/2$. Por ello $6 \nmid 3$ y las dos condiciones se verifican para el **mismo** testigo.

3. Diferencia de prueba. El ejemplo $n = 3$ demuestra B, pero no demuestra A: la universal necesita un argumento aplicable a cada entero que satisfaga su antecedente. De hecho, $n = 3$ ni siquiera cumple $6 \mid n$, por lo que no puede mostrar por sí solo cómo se deriva $3 \mid n$ de esa hipótesis.

Solución 3. El dominio vacío como prueba de lectura. Con $D = \varnothing$ obtenemos, por la semántica clásica:

Fórmula	Valor	Razón
$\forall x \in D\ P(x)$	Verdadera	No existe un elemento de D que constituya una excepción.
$\exists x \in D\ P(x)$	Falsa	No hay ningún elemento disponible como testigo.
$\forall u \in D\, \forall v \in D\ ((P(u) \wedge P(v)) \to u = v)$	Verdadera	No existen pares $u, v \in D$ que puedan contradecir la condición.
$\exists! x \in D\ P(x)$	Falsa	Exige existencia, imposible en un dominio vacío.

La tercera fórmula expresa **a lo sumo uno**: impide dos testigos diferentes y resulta compatible con que no haya ninguno. La cuarta expresa **exactamente uno**: exige, además, un testigo. Su distinción no depende de inventar objetos para evaluar el dominio vacío.

Solución 4. «Si» y «sólo si»: una dirección puede cambiarlo todo. Recordemos que $P(n)$ significa $4 \mid n$ y $Q(n)$ significa $2 \mid n$.

1. «P sólo si Q». La fórmula correcta es

$$\forall n \in \mathbb{Z}\, (P(n) \to Q(n)),$$

es decir, $\forall n \in \mathbb{Z}\ (4 \mid n \to 2 \mid n)$. Es verdadera: si $n = 4k$, entonces $n = 2(2k)$. En esta oración P **es el antecedente** y Q la condición necesaria.

2. «P si Q». Se formaliza

$$\forall n \in \mathbb{Z}\ (Q(n) \to P(n)).$$

Ahora Q **es el antecedente**. La universal es falsa: $n = 2$ satisface $2 \mid 2$, pero no $4 \mid 2$.

3. «P si y sólo si Q». Corresponde a

$$\forall n \in \mathbb{Z}\ (P(n) \leftrightarrow Q(n)),$$

que equivale a exigir conjuntamente ambas direcciones. Es falsa por $n = 2$: una de las implicaciones falla. La dirección de cada flecha se determina por la **gramática de la frase**, no por el deseo de que la proposición resulte verdadera.

Solución 5. Un opuesto para cada entero no es un opuesto común. 1. Lectura y orden. A afirma: «Cada entero a tiene algún entero b que sumado a él da cero». Primero se fija a y después puede escogerse b en función de ese valor. B afirma: «Existe un entero b que sumado a todos los enteros a da cero». Aquí se elige b **una sola vez**, antes de considerar cada a.

2. Demostración de A. Sea $a \in \mathbb{Z}$ arbitrario. Tomemos $b = -a \in \mathbb{Z}$. Entonces $a + b = a - a = 0$. La construcción proporciona un testigo para cualquier a sin imponerle una condición especial; por eso A es verdadera.

3. Refutación de B. Supongamos que un entero fijo b satisface $a + b = 0$ para todo $a \in \mathbb{Z}$. La instancia $a = 0$ obliga a $b = 0$; la instancia $a = 1$ obliga a $b = -1$. Como $0 \neq -1$, la suposición es imposible. Por tanto, B es falsa. Lo decisivo no es mencionar una regla de intercambio de cuantificadores, sino demostrar que ningún testigo uniforme puede cumplir el enunciado.

Solución 6. Negaciones que alcanzan hasta la propiedad interior. 1. Primera proposición. Partimos de la negación completa:

$$\neg\big[\forall n \in \mathbb{Z}\ (3 \mid n \to 9 \mid n)\big].$$

La dualidad de cuantificadores y $\neg(H \to C) \leftrightarrow (H \wedge \neg C)$ dan

$$\exists n \in \mathbb{Z}\ (3 \mid n \wedge 9 \nmid n).$$

En prosa: «Existe un entero divisible por 3 que no es divisible por 9». Es verdadera con $n = 3$: $3 = 3 \cdot 1$ y $9 \nmid 3$. En consecuencia, la proposición original A es falsa y su negación verdadera. No serviría elegir un entero que sólo incumpla el consecuente sin cumplir el antecedente.

2. Segunda proposición. Su negación inicial es

$$\neg\big[\exists y \in \mathbb{Z}\ \forall x \in \mathbb{Z}\ (y > x)\big].$$

Al transformar sucesivamente ambos cuantificadores y negar $y > x$, obtenemos

$$\forall y \in \mathbb{Z}\ \exists x \in \mathbb{Z}\ (y \leq x).$$

En prosa: «Para cada entero y existe un entero x mayor o igual que él». Dado y, podemos elegir $x = y$, por lo que la negación es verdadera y la afirmación original B es falsa. El testigo x se selecciona *después* de fijar y: no se está afirmando la existencia de un x común para todos.

Solución 7. Traducir una restricción y reparar otra traducción. La oración original se escribe

$$\forall n \in \mathbb{Z}\,\big(n > 3 \to \exists m \in \mathbb{Z}\,(m > n)\big).$$

1. Lectura del borrador. La conjunción propuesta dice literalmente: «Todo entero n es mayor que 3 **y**, además, posee algún entero mayor que él». Exige $n > 3$ incluso para los enteros que no cumplen la restricción de la oración inicial. El error es semántico: se colocó la condición del caso dentro de una conjunción en vez de usarla como antecedente.

2. Reparación mínima. Sustituimos $\wedge$ por $\to$, dejando intactos los dominios, la propiedad $m > n$ y el alcance de $\exists m$. Esto expresa la obligación **sólo para los** n **mayores que** 3.

3. Verdad y contraejemplo. Sea $n \in \mathbb{Z}$ arbitrario. Si $n > 3$, el entero $m = n + 1$ verifica $m > n$, así que la proposición corregida es verdadera. El borrador es falso: al tomar $n = 3$, la primera parte $3 > 3$ es falsa y toda la conjunción falla, aunque existan enteros mayores que 3. Distinguir el cambio de significado de la comprobación numérica evita reparar una fórmula sólo para obtener un valor de verdad deseado.

Solución 8. Separar las dos obligaciones de $\exists!$. 1. Despliegue y prueba de A. Según la definición de existencia única,

$$\exists n \in \mathbb{Z}\,\big[(n + 5 = 0) \wedge \forall m \in \mathbb{Z}\,((m + 5 = 0) \to m = n)\big].$$

Existencia: $n = -5$ es entero y $-5 + 5 = 0$. *Unicidad:* si otro entero m satisface $m + 5 = 0$, al sumar -5 a ambos lados obtenemos $m = -5 = n$. Luego A es verdadera; hemos probado algo más que hallar por inspección un candidato.

2. Evaluación de B. B es falsa **por falta de unicidad**, no por falta de existencia: $1 \in \mathbb{Z}$ y $-1 \in \mathbb{Z}$ satisfacen $n^2 = 1$, pero $1 \neq -1$.

3. «A lo sumo uno». Su fórmula es

$$\forall u \in \mathbb{Z}\,\forall v \in \mathbb{Z}\,\big((u^2 = 1 \wedge v^2 = 1) \to u = v\big).$$

Es falsa: $u = 1$ y $v = -1$ hacen verdadero el antecedente y falso el consecuente. Esta fórmula expresa **sólo** la exigencia de no tener dos soluciones diferentes; B exige además la existencia de una solución. No son fórmulas equivalentes en general, aunque en este ejemplo ambas sean falsas. Por ejemplo, con la propiedad $n^2 = -1$ la afirmación «a lo sumo uno» es verdadera, pero «exactamente uno» es falsa.

Solución 9. Renombrar sin capturar una variable libre. 1. Diagnóstico y corrección. En $F(y) = \forall x \in D\,(x < y)$, la ocurrencia de x en la desigualdad está ligada por $\forall x$; la ocurrencia de y está libre y se interpreta con un valor entero asignado desde fuera. En $G = \forall y \in D\,(y < y)$, ambas ocurrencias de y están ligadas por el nuevo cuantificador: la que antes era libre ha sido **capturada**. El renombrado correcto es

$$\forall t \in D\,(t < y),$$

que conserva libre a y y cambia únicamente las ocurrencias ligadas de x.

2. Comprobación de verdad. Con $D = \{0, 1\}$ y $y = 2$, $F(2)$ exige $0 < 2$ y $1 < 2$; es verdadera. G exige en particular $0 < 0$, que es falso, de modo que G es falsa. La transformación errónea ni siquiera preservó la evaluación elegida.

3. Error sintáctico independiente del valor. Una variable libre no puede pasar inadvertidamente a estar ligada cuando cambiamos el nombre de otra variable. $F(y)$ es una fórmula abierta; G es una fórmula cerrada. Incluso si asignamos $y = 0$ y ambas evaluaciones resultan falsas, esa coincidencia accidental no repara la captura ni convierte las dos fórmulas en equivalentes.

Solución 10. La generalización ilegítima. 1. Enunciado pretendido. La conclusión que se desea establecer es

$$\forall n \in \mathbb{Z}\ (10 \mid n \to 5 \mid n).$$

2. Paso injustificado. La comprobación $10 \mid 20$ y $5 \mid 20$ sólo informa sobre $n = 20$. No justifica que la implicación se cumpla para **todo** entero: el argumento había fijado un valor particular, no un n arbitrario. La conclusión puede ser verdadera aunque el razonamiento presentado no la haya probado.

3. Prueba reparada. Sea $n \in \mathbb{Z}$ arbitrario. Supongamos que $10 \mid n$. Por definición, existe $k \in \mathbb{Z}$ tal que $n = 10k$. Entonces

$$n = 10k = 5(2k).$$

Como $2k \in \mathbb{Z}$, la definición de divisibilidad implica $5 \mid n$. Hemos demostrado la conclusión bajo la hipótesis $10 \mid n$ y podemos cerrar esa suposición para obtener $10 \mid n \to 5 \mid n$. Finalmente, como n se introdujo sin fijar su valor ni añadir condiciones especiales, la generalización universal es legítima. □

Solución 11. ¿Qué dirección permite el testigo uniforme?. 1. Demostración de $U \to V$. Supongamos U: existe $y_0 \in E$ tal que $R(x, y_0)$ para **todo** $x \in D$. Sea $x \in D$ arbitrario. El propio $y_0 \in E$ sirve como testigo de $\exists y \in E\ R(x, y)$. Por tanto V es verdadera. Si D es vacío, V es verdadera sin instanciaciones; si E es vacío, U no puede ser verdadera y la implicación también lo es. En ningún caso hemos necesitado finitud ni no vaciedad de los dominios.

2. Contramodelo de la recíproca. Sean $D = E = \{0, 1\}$ y $R(x, y)$ la igualdad $x = y$. V es verdadera: para cada $x \in D$ elegimos $y = x \in E$. En cambio, U es falsa: si $y = 0$, falla la igualdad para $x = 1$; si $y = 1$, falla para $x = 0$. Como ésos son todos los candidatos, no hay testigo uniforme. Así, $V \to U$ no es válida en general.

3. Permutaciones que sí están justificadas. Las fórmulas $\forall x \in D\ \forall y \in E\ R(x, y)$ y $\forall y \in E\ \forall x \in D\ R(x, y)$ afirman que se cumple R para **cada par permitido**: son equivalentes. Las fórmulas $\exists x \in D\ \exists y \in E\ R(x, y)$ y $\exists y \in E\ \exists x \in D\ R(x, y)$ exigen la existencia de **algún par permitido**: también son equivalentes. Estas equivalencias conservan dominios fijos y el orden de los argumentos de R; subsisten incluso si alguno de los dominios es vacío. Intercambiar $\forall$ y $\exists$ altera, en cambio, cuándo puede elegirse el testigo, como muestra el contramodelo anterior.

Solución 12. Negación completa de una unicidad parametrizada. 1. Expansión. La afirmación S puede desplegarse como

$$\forall a \in \mathbb{Z}\ \exists b \in \mathbb{Z}\ \big[(a + b = 0) \land \forall c \in \mathbb{Z}\ ((a + c = 0) \to c = b)\big].$$

La variable a se fija antes de buscar b; el posible testigo b puede depender de a. La variable auxiliar c compara con ese mismo candidato todas las demás soluciones, sin capturar a a ni a b.

2. Negación sin $\exists!$. Para refutar S debe existir **algún** entero a para el cual no haya exactamente una solución. En lógica clásica esto ocurre de dos maneras excluyentes: no hay ninguna solución, o existen dos soluciones distintas. Una fórmula que muestra ambos casos es

$$\boxed{\begin{aligned}\neg S \quad \Longleftrightarrow \quad & \exists a \in \mathbb{Z}\ \big[(\forall b \in \mathbb{Z}\ (a + b \neq 0)) \\ & \qquad \lor \big(\exists b \in \mathbb{Z}\ \exists c \in \mathbb{Z}\ (a + b = 0 \land a + c = 0 \land b \neq c)\big)\big].\end{aligned}}$$

Justificación: la expansión de $\exists!$ exige existencia y que cualquier segundo candidato coincida con el primero. Si falla la existencia obtenemos el primer disyunto; si hay soluciones, pero

falla la unicidad, hay dos testigos b, c con la propiedad y $b \neq c$, que satisfacen el segundo. Recíprocamente, cualquiera de esos dos casos contradice la existencia única. En toda la fórmula se conservan los dominios $\mathbb{Z}$.

3. Traducción y valor de verdad. La negación dice: «Hay un entero a que no tiene ningún opuesto aditivo entero, o tiene por lo menos dos opuestos aditivos enteros distintos». Esta afirmación es falsa, porque S es verdadera para **todo** entero a. En efecto, sea $a \in \mathbb{Z}$ arbitrario. El entero $b = -a$ satisface $a + b = 0$. Si otro entero c cumple $a + c = 0$, al sumar $-a$ a ambos lados concluimos $c = -a = b$. Hay existencia y unicidad para ese a arbitrario, y por tanto para todos. Verificar solamente $a = 0$ y $a = 1$ no demuestra una proposición universal: no examina los restantes enteros ni proporciona una razón que cubra todos los casos.

Cierre: del lenguaje cuantificado al lenguaje de conjuntos, relaciones y funciones. A lo largo de este capítulo hemos pasado de una propiedad abierta a una afirmación cerrada, identificado el dominio y el alcance de cada cuantificador, distinguido testigos individuales de testigos uniformes y separado existencia de unicidad. Los ejercicios nos obligaron a recorrer el camino en ambos sentidos: interpretar fórmulas y construirlas, justificar proposiciones y detectar argumentos que parecían suficientes sin serlo.

El paso siguiente consiste en dar nombres y estructura a algunos objetos que ya hemos utilizado informalmente. Al escribir $R(x, y)$ hablamos de una relación entre elementos de dominios fijados; al afirmar $\forall x \in D\ \exists! y \in E\ R(x, y)$ expresamos que a cada x le corresponde exactamente un y. En el **capítulo 4, «Conjuntos, relaciones y funciones como lenguaje»**, precisaremos las nociones de conjunto, relación, dominio, codominio y función necesarias para interpretar y emplear estas escrituras. No estamos dando aquí por construida una función ni anticipando la teoría axiomática de conjuntos: llevamos al nuevo capítulo una arquitectura lógica que ahora sabemos leer.

La transición al capítulo siguiente no requiere memorizar recetas de transformación: exige conservar el dominio y el alcance de las afirmaciones, reconocer qué testigos se necesitan y justificar cada paso de una prueba. Ésta es la disciplina de lectura que aplicaremos al estudiar conjuntos, relaciones y funciones.

Capítulo 4 — Conjuntos, relaciones y funciones como lenguaje

En el capítulo anterior aprendimos a distinguir afirmaciones como «todos los enteros cumplen cierta condición», «algún entero la cumple» y «cada entero determina un único objeto con determinada propiedad». El lenguaje de los cuantificadores hace explícitas las obligaciones de esas frases. Sin embargo, la práctica matemática requiere algo más: necesitamos hablar de **los objetos que satisfacen una condición**, comparar colecciones de objetos y describir correspondencias entre dominios.

Por ejemplo, la frase «n es un entero par» expresa una propiedad de n. Podemos reunir bajo un mismo nombre a los enteros que la satisfacen y preguntar si todo número de esa colección pertenece también a otra. Podemos considerar pares ordenados que cumplen una relación y, cuando cada objeto del primer dominio tiene una única pareja en el segundo, describir una función. En cada caso, los cuantificadores conservan su significado: el nuevo vocabulario permite organizar las afirmaciones, no eludir su justificación.

Nuestra pregunta será: **¿qué debemos saber sobre un conjunto, una relación o una función para leer correctamente las afirmaciones en las que intervienen?** Comenzaremos por pertenencia, inclusión e igualdad de conjuntos; después estudiaremos operaciones, pares ordenados, relaciones y funciones, hasta llegar a imágenes, preimágenes, propiedades de las correspondencias y conjuntos solución.

Este capítulo presenta una capa de lenguaje compartida por los demás libros. Usaremos conjuntos de números y otros conjuntos previamente especificados; no justificaremos aquí los axiomas de existencia de conjuntos ni admitiremos que cualquier condición produzca automáticamente un conjunto en un universo irrestricto. Tampoco supondremos que una relación sea una función antes de verificar las condiciones que la definen. Nuestro objetivo es aprender a **formular, interpretar y demostrar** las afirmaciones elementales de este lenguaje.

4.1. Pertenencia, inclusión e igualdad: tres afirmaciones diferentes

Pertenecer a un conjunto no es estar incluido en él. Consideremos los conjuntos finitos

$$A = \{2, 4\}, \qquad B = \{2, 4, 6\}.$$

La escritura $2 \in A$ afirma que el número 2 es un **elemento** de A; el símbolo $\in$ expresa pertenencia. También tenemos $6 \notin A$ y $6 \in B$. Para leer cualquiera de estas afirmaciones debemos identificar tanto el objeto situado a la izquierda como el conjunto situado a la derecha.

Pertenencia e inclusión

Para un objeto x y un conjunto A, la expresión $x \in A$ afirma que x es un elemento de A.

Para conjuntos A y B, la expresión $A \subseteq B$ afirma que **todo elemento de A es también elemento de B**. Su significado lógico es

$$A \subseteq B \quad \Longleftrightarrow \quad \forall x\, (x \in A \to x \in B).$$

El universo del cuantificador se entiende fijado por el contexto; equivalentemente, podemos escribir $\forall x \in A\ (x \in B)$. La inclusión permite la igualdad: $A \subseteq A$.

Así, $A \subseteq B$ es verdadera: los dos elementos de A, a saber, 2 y 4, están en B. En cambio, $B \subseteq A$ es falsa; 6 es un contraejemplo, pues $6 \in B$ pero $6 \notin A$.

La diferencia no es tipográfica. Consideremos el conjunto $C = \{2, \{2\}\}$. Tiene dos elementos distintos: el número 2 y el conjunto $\{2\}$. Por tanto,

$$2 \in C, \qquad \{2\} \in C, \qquad \{2\} \subseteq C.$$

La última inclusión se justifica porque el **único elemento de** $\{2\}$, que es 2, pertenece a C. La pertenencia $\{2\} \in C$ tiene otra justificación: ese conjunto figura como elemento de C. No debe inferirse, en general, $A \in B$ a partir de $A \subseteq B$, ni la implicación inversa. Para nuestros primeros conjuntos, de hecho, $A \subseteq B$, pero $A \notin B$: los elementos de B son números, no el conjunto A.

Igualdad: los mismos elementos, no la misma presentación. El orden de enumeración y la repetición de elementos no alteran un conjunto. En particular,

$$\{2, 4\} = \{4, 2, 2\}.$$

¿Qué criterio permite demostrar una igualdad aunque los conjuntos no se presenten mediante listas? Adoptaremos el criterio extensional habitual: dos conjuntos son iguales exactamente cuando tienen los mismos elementos.

$$\boxed{A = B \iff \forall x\ (x \in A \leftrightarrow x \in B).}$$

El bicondicional puede desplegarse en dos obligaciones:

$$A = B \iff (A \subseteq B) \wedge (B \subseteq A).$$

Justificación. Si $A = B$, cualquier elemento de uno pertenece al otro, luego valen ambas inclusiones. Recíprocamente, si $A \subseteq B$ y $B \subseteq A$, para cualquier x se cumplen las dos implicaciones $x \in A \to x \in B$ y $x \in B \to x \in A$. Por consiguiente, $x \in A \leftrightarrow x \in B$ para todo x; el criterio extensional da $A = B$. □

Lectura de la demostración. La igualdad de conjuntos no se prueba comparando su apariencia. Se transforma una igualdad en dos afirmaciones universales sobre pertenencia. Más adelante estudiaremos la doble inclusión como estrategia general de demostración; por ahora nos importa comprender exactamente por qué funciona.

El conjunto vacío no es una excepción a la inclusión. Denotaremos por $\varnothing$ el conjunto que no tiene elementos. Entonces, para cualquier conjunto A,

$$\varnothing \subseteq A.$$

Demostración. Para negar esta inclusión tendría que existir un objeto x con $x \in \varnothing$ y $x \notin A$. Pero no existe ningún $x \in \varnothing$. Por ello, para todo objeto x, el condicional $x \in \varnothing \to x \in A$ es verdadero y se cumple la inclusión. □

Esto **no** significa que $\varnothing \in A$ para todo conjunto A. Por ejemplo, $\varnothing \notin \{2, 4\}$, mientras que $\varnothing \in \{\varnothing, 2\}$. La inclusión es universal y puede ser verdadera sin elementos en el antecedente; la pertenencia exige que un objeto concreto sea uno de los elementos del conjunto.

De una propiedad a un conjunto solución. Regresemos a las fórmulas del capítulo 3. Dentro de un dominio ya especificado, podemos escribir el conjunto de los objetos que satisfacen cierta propiedad. Por ejemplo, con dominio $\mathbb{Z}$ y propiedad $P(n):\ n+5=0$, escribimos

$$S = \{n \in \mathbb{Z} : n + 5 = 0\}.$$

El signo «:» se lee «tal que». Para cualquier entero a,

$$a \in S \quad \Longleftrightarrow \quad a + 5 = 0.$$

Como $-5+5=0$, tenemos $-5 \in S$; y si $a \in S$, entonces $a+5=0$, de donde $a=-5$. Estas dos observaciones justifican $S=\{-5\}$: ningún otro entero pertenece a S y -5 sí pertenece. La propiedad $P(n)$, la proposición $P(-5)$ y el conjunto S son objetos de lenguaje diferentes, aunque estrechamente relacionados.

Una precaución de alcance

La expresión $\{n \in \mathbb{Z} : P(n)\}$ selecciona elementos **dentro del conjunto ya fijado** $\mathbb{Z}$. No adoptaremos la regla irrestricta «toda propiedad define un conjunto de todos los objetos que la satisfacen»; una regla así requiere precisiones fundacionales que exceden este capítulo. El conjunto solución relativo a un dominio reaparecerá en §4.8.

Comprobación de lectura. Si alguien afirma que $A \subseteq B$ porque ha verificado únicamente que $2 \in A$ y $2 \in B$, ¿qué falta? Falta examinar **todos** los elementos de A. La coincidencia de un elemento acredita dos pertenencias, pero no una inclusión universal. Si, en cambio, encuentra un $a \in A$ con $a \notin B$, ese solo testigo refuta la inclusión. La lógica de los cuantificadores vuelve a determinar el trabajo de prueba.

Después de la prueba. Pertenencia, inclusión e igualdad permiten traducir preguntas sobre colecciones en afirmaciones controlables elemento por elemento. En §4.2 estudiaremos cómo se forman nuevas colecciones a partir de conjuntos dados y cómo leer sus operaciones mediante condiciones de pertenencia.

4.2. Operaciones de conjuntos: leer cada operación como una condición de pertenencia

La sección anterior mostró que $A = B$ puede establecerse comprobando $x \in A \leftrightarrow x \in B$ para cada objeto del universo pertinente. Adoptaremos ahora esa misma perspectiva para formar conjuntos a partir de otros ya dados. La pregunta «¿qué elementos tiene el nuevo conjunto?» debe responderse mediante una condición precisa; su nombre o su dibujo no reemplazan esa respuesta.

Unión e intersección: «o» e «y» aplicados a la pertenencia. Sean A y B conjuntos. Su **unión**, denotada por $A \cup B$, reúne los objetos que pertenecen a A, a B o a ambos. Su **intersección**, denotada por $A \cap B$, reúne los que pertenecen simultáneamente a los dos conjuntos. Para cualquier objeto x del universo de referencia, las definiciones se traducen como

$$\boxed{\begin{aligned} x \in A \cup B \quad &\Longleftrightarrow \quad (x \in A) \vee (x \in B), \\ x \in A \cap B \quad &\Longleftrightarrow \quad (x \in A) \wedge (x \in B). \end{aligned}}$$

El «o» es **inclusivo**: un elemento que está en ambos conjuntos también pertenece a su unión. No se excluye de $A \cup B$ precisamente por aparecer en los dos. Por el contrario, para pertenecer a $A \cap B$ no basta cumplir una sola de las condiciones.

Tomemos el universo finito $U = \{1, 2, 3, 4, 5, 6\}$ y dos subconjuntos

$$A = \{1, 2, 4\}, \qquad B = \{2, 3, 4\}.$$

Entonces

$$A \cup B = \{1, 2, 3, 4\}, \qquad A \cap B = \{2, 4\}.$$

El 2 pertenece a la unión **y** a la intersección; el 1 pertenece a la unión, pero no a la intersección; el 5 no pertenece a ninguna. El ejemplo sirve para comprobar el significado de los conectivos, no para demostrar las identidades generales que formularemos más abajo.

Cuatro operaciones, cuatro pruebas de pertenencia

Para conjuntos A, B y para un conjunto de referencia U con $A \subseteq U$:

Operación	Condición que caracteriza a sus elementos
Unión $A \cup B$	$x \in A \lor x \in B$.
Intersección $A \cap B$	$x \in A \land x \in B$.
Diferencia $A \setminus B$	$x \in A \land x \notin B$.
Complemento de A en U, $U \setminus A$	$x \in U \land x \notin A$.

Estas condiciones son bicondicionales de pertenencia. En particular, **definir** una operación no equivale a demostrar que cumple una identidad propuesta: para esa segunda tarea debemos justificar una equivalencia válida para cualquier objeto pertinente.

Diferencia: conservar una pertenencia y excluir otra. La **diferencia** de A y B es el conjunto de los elementos de A que no están en B:

$$A \setminus B = \{x \in A : x \notin B\}.$$

Su condición de pertenencia contiene una conjunción y una negación:

$$x \in A \setminus B \quad \Longleftrightarrow \quad (x \in A) \land (x \notin B).$$

En nuestro ejemplo, $A \setminus B = \{1\}$, mientras que $B \setminus A = \{3\}$. La operación **no es conmutativa en general**: cambiar el orden intercambia el conjunto del cual partimos con el conjunto cuyos elementos excluimos. El 1 distingue las dos diferencias, pues pertenece a $A \setminus B$ pero no a $B \setminus A$.

Una precisión lógica: $x \in A \setminus B$ exige *ambas* condiciones sobre el **mismo** x. No bastaría presentar un elemento de A y otro, distinto, que no esté en B; esa elección no aporta un testigo de la conjunción.

Complemento: «lo que falta» depende de un universo. Para hablar de complemento debemos fijar un conjunto de referencia U y suponer $A \subseteq U$. El **complemento de A relativo a U** es

$$A^{\mathrm{c}}_U := U \setminus A = \{x \in U : x \notin A\}.$$

Si declaramos desde el principio que todos los elementos considerados pertenecen a U, podemos abreviar su condición de pertenencia como $x \in A^{\mathrm{c}}_U \leftrightarrow x \notin A$ **para** $x \in U$. Sin esa restricción, el bicondicional omitiría la exigencia $x \in U$: no estamos reuniendo objetos ajenos al conjunto de referencia.

Con $U = \{1, 2, 3, 4, 5, 6\}$ y $A = \{1, 2, 4\}$, obtenemos

$$A^{\mathrm{c}}_U = \{3, 5, 6\}.$$

Este complemento no coincide con $B \setminus A = \{3\}$: los números 5 y 6 también pertenecen a U y están fuera de A, aunque no pertenezcan a B. Además, si ampliamos el universo a $V = U \cup \{7\}$, entonces $V \setminus A = \{3, 5, 6, 7\}$. Por tanto, la expresión informal «todos los elementos que no están en A» es incompleta hasta especificar **dónde** se buscan esos elementos.

El universo no desaparece al negar

La negación $x \notin A$ expresa una condición; el complemento $U \setminus A$ designa un conjunto cuyos elementos se seleccionan **dentro de** U. El capítulo 3 enseñó a mantener el dominio de una negación cuantificada; aquí aplicamos la misma disciplina a una operación de conjuntos. No usamos una comprensión irrestricta sobre «todos los objetos».

Una prueba completa: distribuir la intersección sobre la unión. La semejanza entre $\cap, \cup$ y los conectivos $\wedge, \vee$ sugiere una identidad. Para conjuntos arbitrarios A, B, C, afirmamos que

$$\boxed{A \cap (B \cup C) = (A \cap B) \cup (A \cap C).}$$

Demostración por pertenencia. Sea x un objeto arbitrario. Si $x \in A \cap (B \cup C)$, entonces $x \in A$ y $x \in B \cup C$. Por definición de unión, $x \in B$ o $x \in C$. En el primer caso, $x \in A \cap B$; en el segundo, $x \in A \cap C$. En cualquiera de los casos, $x \in (A \cap B) \cup (A \cap C)$. Hemos probado la primera inclusión.

Recíprocamente, supongamos $x \in (A \cap B) \cup (A \cap C)$. Entonces $x \in A \cap B$ o $x \in A \cap C$. En ambos casos $x \in A$; además, en el primero $x \in B$ y en el segundo $x \in C$. Por ello $x \in B \cup C$ y, junto con $x \in A$, obtenemos $x \in A \cap (B \cup C)$. Hemos probado la inclusión inversa.

Como ambas inclusiones valen para un x arbitrario, los dos conjuntos tienen exactamente los mismos elementos y son iguales. □

Estructura de la prueba. El trabajo no consistió en verificar algunos conjuntos particulares ni en apelar a un diagrama. Partimos de una pertenencia, desplegamos definiciones, usamos una alternativa para separar casos y volvimos a reunir los resultados en una pertenencia. La igualdad exigió las **dos direcciones**, como anticipaba §4.1. Podemos condensar el núcleo lógico —sin reemplazar la explicación precedente— en

$$\begin{aligned} x \in A \cap (B \cup C) &\iff (x \in A) \wedge \big((x \in B) \vee (x \in C)\big) \\ &\iff \big((x \in A) \wedge (x \in B)\big) \vee \big((x \in A) \wedge (x \in C)\big) \\ &\iff x \in (A \cap B) \cup (A \cap C). \end{aligned}$$

La equivalencia central es la distributividad proposicional. Leerla en términos de objetos muestra exactamente por qué autoriza la identidad de conjuntos.

Diagnóstico: al negar una unión no permanece una alternativa. Alguien propone la igualdad

$$A \setminus (B \cup C) \stackrel{?}{=} (A \setminus B) \cup (A \setminus C).$$

Su argumento es que «si un elemento queda fuera de la unión, basta con que quede fuera de uno de sus dos conjuntos». Ahí está el error: para no pertenecer a $B \cup C$ debe estar fuera de **ambos**. En el universo finito $U = \{1\}$, tomemos $A = B = \{1\}$ y $C = \varnothing$. El miembro izquierdo es $\varnothing$, pero el derecho es $\varnothing \cup \{1\} = \{1\}$; los conjuntos son diferentes.

La reparación correcta reemplaza la unión final por una intersección:

$$\boxed{A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C).}$$

Justificación. Para cualquier objeto x,

$$\begin{aligned} x \in A \setminus (B \cup C) &\iff (x \in A) \wedge \neg\big((x \in B) \vee (x \in C)\big) \\ &\iff (x \in A) \wedge (x \notin B) \wedge (x \notin C) \\ &\iff x \in (A \setminus B) \cap (A \setminus C). \end{aligned}$$

La segunda equivalencia usa De Morgan, y la última exige que x pertenezca a **ambas** diferencias. Esta cadena es válida aunque alguno de los conjuntos sea vacío; el contraejemplo anterior solamente refutó la versión incorrecta, no fue la prueba de la versión correcta. □

Comprobación de aprendizaje. Sin enumerar elementos, podemos afirmar para todo conjunto A que $A \cup \varnothing = A$, $A \cap \varnothing = \varnothing$ y $A \setminus A = \varnothing$: en cada caso se despliega la condición de pertenencia y se comprueba que coincide con la del conjunto propuesto. Estas identidades simples ofrecen una prueba de control para cualquier definición que utilicemos.

Respuesta razonada. Para cualquier x, pertenecer a $A \cup \varnothing$ equivale a $x \in A$ o a una condición imposible; por eso equivale a $x \in A$. Pertenecer a $A \cap \varnothing$ exigiría $x \in \varnothing$, imposible. Pertenecer a $A \setminus A$ exigiría simultáneamente $x \in A$ y $x \notin A$, también imposible. La extensionalidad da las tres igualdades.

Después de la prueba. Las operaciones de conjuntos han convertido conectivos sobre pertenencia en nuevas colecciones, y la extensionalidad nos permitió probar identidades sin depender de cómo se dibujen. El siguiente paso introduce objetos de otra naturaleza: un **par ordenado**, cuyo orden importa, y el producto cartesiano de dos conjuntos.

4.3. Pares ordenados y producto cartesiano: por qué importa el orden

En §4.1 observamos que $\{1, 2\} = \{2, 1\}$: cambiar el orden de los elementos enumerados no produce otro conjunto. Pero hay situaciones en las que importa distinguir el primer lugar del segundo. Una dirección de partida y una de llegada, o una entrada y un resultado, no desempeñan el mismo papel. Para expresar esa diferencia necesitamos un objeto que registre **dos componentes con posiciones determinadas**.

Un par ordenado no es una enumeración de dos elementos. Escribimos (a, b) para el **par ordenado** cuya primera componente es a y cuya segunda componente es b. El criterio que caracteriza su igualdad es

$$\boxed{(a, b) = (c, d) \quad \Longleftrightarrow \quad (a = c) \wedge (b = d).}$$

La igualdad exige comparar las componentes *en la misma posición*. De aquí se sigue que $(a, b) = (b, a)$ exactamente cuando $a = b$: si los pares son iguales, la igualdad de primeras componentes da $a = b$; recíprocamente, si $a = b$, las dos escrituras designan el mismo par. En particular, $(1, 2) \neq (2, 1)$, aunque $\{1, 2\} = \{2, 1\}$. Tampoco confundiremos (a, b) con $\{a, b\}$: el primero registra posiciones; el segundo sólo registra pertenencias.

El criterio de igualdad es parte del significado

Para verificar que dos pares ordenados son iguales debemos comprobar **dos igualdades**: primera componente con primera y segunda con segunda. Para refutar esa igualdad basta encontrar una posición en la que las componentes sean distintas. La expresión (a, a) sigue siendo un par ordenado con dos posiciones, aunque ambas estén ocupadas por el mismo objeto.

Una precisión fundacional. No estamos postulando que una lista entre paréntesis constituya, sin más, un nuevo objeto. Los pares pueden representarse mediante conjuntos ya dados. Una codificación habitual es

$$(a, b) := \{\{a\}, \{a, b\}\}.$$

Veamos por qué esta codificación conserva las posiciones. El único objeto que pertenece a **todos** los conjuntos que son elementos de $\{\{a\}, \{a, b\}\}$ es a; los objetos que pertenecen a **alguno** de esos conjuntos son precisamente los de $\{a, b\}$. Por tanto, si dos pares codificados son iguales, sus primeras componentes coinciden: $a = c$, y además $\{a, b\} = \{a, d\}$. Si $b = a$, esta última igualdad obliga a $d = a = b$. Si $b \neq a$, la pertenencia $b \in \{a, d\}$ obliga a $b = d$. En ambos casos coinciden las segundas componentes. El recíproco es inmediato: si $a = c$ y $b = d$,

las dos codificaciones son idénticas. Ésta es una **justificación del criterio de igualdad**; no emprenderemos aquí una axiomatización de la existencia de los conjuntos utilizados en la codificación.

Lectura guiada. ¿Por qué no bastaría definir (a, b) como $\{a, b\}$? Porque entonces $(1, 2)$ y $(2, 1)$ serían iguales por extensionalidad, y perderíamos precisamente la información que pretendíamos conservar. No se trata de una diferencia entre llaves y paréntesis: una definición propuesta debe satisfacer la propiedad que necesitamos de ella.

Dos conjuntos dados determinan un producto cartesiano. Sean A y B conjuntos fijados. El **producto cartesiano** de A por B se define como el conjunto de todos los pares ordenados cuya primera componente pertenece a A y cuya segunda componente pertenece a B:

$$\boxed{A \times B := \{(a, b) : a \in A \text{ y } b \in B\}.}$$

Se trata de una construcción relativa a los conjuntos A y B previamente especificados: no reunimos pares procedentes de un universo irrestricto. La definición proporciona una prueba de pertenencia que utilizaremos repetidamente:

$$\boxed{(x, y) \in A \times B \quad \Longleftrightarrow \quad (x \in A) \wedge (y \in B).}$$

La conjunción obliga a comprobar **ambas** condiciones y a respetar las posiciones. Si escribimos $\forall x \in A\ \forall y \in B\ ((x, y) \in A \times B)$, la afirmación resultante es verdadera por definición; escribir solamente $(x, y) \in A \times B$ deja pendientes los valores de x, y o su declaración en el contexto. Recuperamos así la diferencia entre fórmula abierta y cuantificación trabajada en el capítulo 3.

Ejemplo finito. Sean $A = \{1, 2\}$ y $B = \{2, 3\}$. Al combinar cada primera componente permitida con cada segunda componente permitida, obtenemos

$$A \times B = \{(1, 2), (1, 3), (2, 2), (2, 3)\}.$$

El par $(1, 3)$ pertenece al producto porque $1 \in A$ y $3 \in B$. El par $(3, 1)$ no pertenece: falla $3 \in A$ y también $1 \in B$. El par $(2, 2)$ sí pertenece; la definición no exige componentes diferentes. Obsérvese que $\{1, 3\} = \{3, 1\}$ no convierte a $(1, 3)$ y $(3, 1)$ en el mismo par. Asimismo, $A \times B$ es un **conjunto de pares**; no es el par ordenado (A, B), cuyas componentes son los propios conjuntos.

Cambiar los factores cambia las posiciones disponibles. El producto $B \times A$ exige que la primera componente pertenezca a B y la segunda a A. Con los conjuntos del ejemplo,

$$B \times A = \{(2, 1), (2, 2), (3, 1), (3, 2)\}.$$

Como $(1, 3) \in A \times B$ pero $(1, 3) \notin B \times A$, se sigue que $A \times B \neq B \times A$ en este ejemplo. Concluir que **nunca** son iguales sería igualmente incorrecto: si $A = B$, ambas expresiones son $A \times A$, y si uno de los factores es vacío ambos productos son vacíos, como probaremos enseguida. La conclusión general justificada es que el producto cartesiano **no es conmutativo en general**, no que la igualdad sea imposible en todo caso.

El cambio de posiciones, sin embargo, sí permite relacionar los productos mediante una regla precisa: si $(a, b) \in A \times B$, entonces $(b, a) \in B \times A$. En efecto, de la primera pertenencia obtenemos $a \in A$ y $b \in B$; para el par invertido, b ocupa la primera posición permitida por B y a la segunda permitida por A. Esta correspondencia entre pares invertidos **no** establece la igualdad de los dos productos, porque generalmente $(a, b) \neq (b, a)$.

Producto cartesiano y conjunto vacío

Para cualesquiera conjuntos A, B,

$$\boxed{A \times B = \varnothing \quad \Longleftrightarrow \quad (A = \varnothing) \vee (B = \varnothing).}$$

El vacío en cualquiera de los factores. Demostración. Si $A = \varnothing$, no existe ninguna primera componente $a \in A$; por tanto, no existe ningún par (a, b) en $A \times B$. Lo mismo ocurre si $B = \varnothing$, pues no hay segunda componente disponible. En cualquiera de esos casos, $A \times B = \varnothing$.

Recíprocamente, supongamos $A \times B = \varnothing$ y que ninguno de los factores es vacío. Entonces existen $a \in A$ y $b \in B$. La definición del producto garantiza $(a, b) \in A \times B$, contradiciendo que el producto sea vacío. Luego, si el producto es vacío, al menos uno de los factores debe serlo. □

Lectura de la prueba. Para probar que un producto no es vacío necesitamos **dos testigos**, uno de cada factor, y formar con ellos un par. No basta saber que A contiene un elemento si desconocemos si B contiene alguno. A la inversa, demostrar que uno de los factores carece de elementos descarta todos los pares sin necesidad de enumerarlos. En particular, $\varnothing \times A = A \times \varnothing = \varnothing$, incluso si A contiene muchos elementos.

Una inclusión que se demuestra componente por componente. Proposición. Si $A \subseteq C$ y $B \subseteq D$, entonces

$$A \times B \subseteq C \times D.$$

Demostración. Sea $(x, y) \in A \times B$ un par arbitrario. Por la condición de pertenencia, $x \in A$ y $y \in B$. Las hipótesis de inclusión dan $x \in C$ y $y \in D$, respectivamente. Por definición, $(x, y) \in C \times D$. Hemos cubierto todos los pares del primer producto y queda probada la inclusión. □

No hemos demostrado la recíproca. Por ejemplo, si $A = \varnothing$, se tiene $A \times B = \varnothing \subseteq C \times D$ para cualesquiera B, C, D, incluso aunque $B \nsubseteq D$. Elegir $B = \{1\}$ y $D = \varnothing$ hace explícito el fallo. Este caso límite muestra por qué una igualdad o inclusión entre productos no debe transformarse en una afirmación sobre sus factores sin examinar las hipótesis de no vaciedad.

Auditoría: tres confusiones en una sola justificación. Se propone lo siguiente: «Como $1 \in A$ y $3 \in B$, el par $(3, 1)$ pertenece a $A \times B$; además, $A \times B = B \times A$ porque $\{1, 3\} = \{3, 1\}$». Para los conjuntos $A = \{1, 2\}$ y $B = \{2, 3\}$, el argumento incurre en tres errores distintos. Primero, ubica 3 en la primera posición sin verificar $3 \in A$. Segundo, invierte la segunda componente sin verificar $1 \in B$. Tercero, usa la igualdad de dos **conjuntos no ordenados** para concluir la igualdad de pares o de productos, aunque sus criterios de igualdad son diferentes.

La reparación comienza por conservar las posiciones: de $1 \in A$ y $3 \in B$ se deduce $(1, 3) \in A \times B$, y también $(3, 1) \in B \times A$. Pero $(1, 3) \notin B \times A$, pues $1 \notin B$; ese solo par refuta la igualdad de los productos en el ejemplo. **Una afirmación sobre pertenencias individuales sólo autoriza el par cuyas componentes ocupan las posiciones indicadas por la definición.**

Comprobación de aprendizaje. Si $A = \{2\}$ y $B = \{2\}$, calcula ambos productos y explica por qué coinciden. Después toma $B = \varnothing$ y determina el producto sin enumerar pares. Compara por último $(2, 2)$ con $\{2\}$: la coincidencia del valor de sus componentes no identifica los tipos de objetos.

Respuesta razonada. Con $A = B = \{2\}$, ambos productos son $\{(2, 2)\}$: sólo hay una elección para cada posición. Si $B = \varnothing$, ambos productos son vacíos porque falta una componente. El par $(2, 2)$ registra dos posiciones; $\{2\}$ es un conjunto con un elemento. Con la codificación adoptada, $(2, 2) = \{\{2\}\}$, que tampoco coincide con $\{2\}$.

Después de la prueba. Sabemos reconocer pares por sus componentes y productos por una condición conjuntiva de pertenencia. Esto proporciona el universo de pares dentro del cual podremos seleccionar aquellos que cumplan una propiedad adicional. En §4.4 llamaremos **relación** a un subconjunto apropiado de un producto cartesiano; su definición exigirá conservar los dominios y no confundir una condición entre dos objetos con una función.

4.4. Relaciones como subconjuntos de productos cartesianos: ¿qué afirma $R(x, y)$?

El producto cartesiano reúne todos los pares que podemos formar a partir de dos conjuntos. Una relación selecciona **algunos** de esos pares, de acuerdo con una condición. El paso no consiste en abandonar el lenguaje lógico: consiste en registrar mediante un conjunto qué instancias de una condición entre dos objetos son verdaderas. Conservaremos los dos conjuntos de referencia, pues de ellos depende qué pares pueden considerarse y qué significan los cuantificadores de nuestras afirmaciones.

Definición: una relación selecciona pares de un producto fijado. Sean A y B conjuntos. Una **relación binaria de A en B** es un conjunto R tal que

$$\boxed{R \subseteq A \times B.}$$

Sus elementos son pares ordenados. Para $x \in A$ e $y \in B$, escribiremos indistintamente $x\,R\,y$ o $R(x, y)$ como abreviaturas de la proposición $(x, y) \in R$:

$$\boxed{x\,R\,y \quad \Longleftrightarrow \quad (x, y) \in R.}$$

La letra R designa aquí un **conjunto de pares**; la escritura $R(x, y)$ expresa una **afirmación** acerca de dos objetos en las posiciones indicadas. No denota, por sí sola, un valor obtenido al aplicar una función llamada R al argumento x. Si R se describe mediante una condición $P(x, y)$, debemos indicar los conjuntos de referencia y entender

$$R = \{(x, y) \in A \times B : P(x, y)\}, \qquad (x, y) \in R \iff P(x, y) \quad (x \in A,\ y \in B).$$

La comprensión está **restringida** al producto ya fijado. La condición no autoriza a reunir en un conjunto todos los pares imaginables, con independencia de sus dominios. Una relación de A en A recibe también el nombre de *relación sobre A*.

Tres niveles que no deben confundirse

- $A \times B$ es el conjunto de **pares permitidos** por las posiciones.
- $R \subseteq A \times B$ es el conjunto de **pares seleccionados**.
- $R(x, y)$ es la **proposición** que afirma que el par concreto (x, y) fue seleccionado.

De $x \in A$ e $y \in B$ se deduce $(x, y) \in A \times B$, **no** necesariamente $(x, y) \in R$. La última pertenencia requiere además verificar la condición propia de R.

Leer una relación finita sin perder las posiciones. Tomemos

$$A = \{1, 2, 4\}, \qquad B = \{2, 3\}, \qquad R = \{(x, y) \in A \times B : x < y\}.$$

Enumerar el producto produce seis pares; imponer la condición $x < y$ conserva exactamente tres:

$$\boxed{R = \{(1, 2), (1, 3), (2, 3)\}.}$$

Por tanto, $1\,R\,2$ es verdadera y $2\,R\,2$ es falsa: el par $(2, 2)$ pertenece al producto, pero no satisface la desigualdad estricta. También $4\,R\,3$ es falsa; ninguna segunda componente permitida supera a 4. En cambio, invertir $(1, 3)$ produce $(3, 1)$, que ni siquiera pertenece

al producto $A \times B$ de referencia. Recordemos que cambiar las posiciones no conserva automáticamente la pertenencia.

Podemos leer la relación como una tabla de comprobación. Cada celda pregunta si el par determinado por su fila y su columna pertenece a R:

Primera componente	$y = 2$	$y = 3$
$x = 1$	Sí	Sí
$x = 2$	No	Sí
$x = 4$	No	No

La tabla hace visibles tres situaciones: 1 está relacionado con **dos** elementos de B; 2, con **uno**; 4, con **ninguno**. Una relación no exige que todos los elementos del conjunto inicial participen, ni que cada uno tenga una única segunda componente. Estos requisitos adicionales serán precisamente el objeto de §4.5.

Lectura lógica. La proposición «cada elemento de A se relaciona con algún elemento de B» se escribe

$$\forall x \in A\ \exists y \in B\ R(x, y).$$

Es **falsa** para nuestra relación: $x = 4$ es un contraejemplo, porque no existe $y \in B$ con $4 < y$. La proposición «existe un elemento de A relacionado con dos elementos distintos de B» es, en cambio, verdadera: $x = 1$ y los testigos $y = 2$, $z = 3$ la verifican. Los cuantificadores no son un adorno añadido a la lista de pares: determinan qué propiedades generales estamos afirmando acerca de ella.

Conjuntos de referencia y elementos efectivamente relacionados. El conjunto A fija las primeras componentes **admisibles**, pero algunas quizá no aparezcan en ningún par de R. Llamamos **dominio efectivo de la relación** al conjunto

$$\operatorname{dom}(R) := \{x \in A : \exists y \in B\ (x, y) \in R\}.$$

Por definición, $\operatorname{dom}(R) \subseteq A$. Para el ejemplo anterior,

$$\operatorname{dom}(R) = \{1, 2\} \neq A :$$

el número 4 pertenece a A, pero no es primera componente de ningún par de R. Análogamente, las segundas componentes que aparecen efectivamente forman el conjunto $\{y \in B : \exists x \in A\ (x, y) \in R\}$, que en el ejemplo es $\{2, 3\}$. Dejaremos el estudio sistemático de las imágenes para §4.6; aquí basta distinguir el conjunto de referencia B de los valores realmente presentes.

La palabra «dominio» exige contexto

En el caso de relaciones, distinguiremos el **conjunto inicial declarado** A del **dominio efectivo** $\operatorname{dom}(R)$. Para funciones, en §4.5, exigiremos que todas las entradas del dominio declarado tengan exactamente una salida, por lo que allí ambos conjuntos coincidirán. No intercambiemos esas convenciones antes de enunciar las hipótesis.

El mismo conjunto de pares puede contemplarse con distintos conjuntos de referencia. Por ejemplo, $T = \{(1, 2)\}$ es una relación de $\{1\}$ en $\{2\}$, pero también de $\{1, 3\}$ en $\{2\}$. En el primer contexto es verdadera la fórmula $\forall x \in A\ \exists y \in B\ T(x, y)$; en el segundo es falsa, porque 3 no aparece como primera componente. **El conjunto de pares por sí solo no determina qué afirmación cuantificada estamos evaluando si omitimos los conjuntos declarados.** Cuando hagan falta ambos datos, especificaremos A, B y T conjuntamente.

Los casos extremos también son relaciones legítimas: $\varnothing \subseteq A \times B$ y $A \times B \subseteq A \times B$. La **relación vacía** no selecciona par alguno; la **relación total** selecciona todos los pares disponibles. Si $A = \varnothing$ o $B = \varnothing$, el producto es vacío y la única relación posible como subconjunto de él es la vacía. Las afirmaciones universales con dominio inicial vacío son vacuamente verdaderas; eso no crea un par ni un testigo.

Comparar relaciones mediante implicaciones lógicas. Sean $R, S \subseteq A \times B$. La inclusión entre relaciones no pide comparar dibujos ni nombres: pide comprobar que **todo par que satisface la primera condición satisface también la segunda**.

Proposición. Para dos relaciones con los mismos conjuntos de referencia,

$$\boxed{R \subseteq S \quad \Longleftrightarrow \quad \forall x \in A\, \forall y \in B\, \big(R(x,y) \to S(x,y)\big).}$$

Demostración. Supongamos primero $R \subseteq S$. Dados $x \in A$ e $y \in B$ arbitrarios, si $R(x,y)$ es verdadera, entonces $(x,y) \in R$; la inclusión garantiza $(x,y) \in S$, es decir, $S(x,y)$. Queda probado el condicional para todo par del producto.

Recíprocamente, supongamos válida la fórmula universal y tomemos un par cualquiera $(x,y) \in R$. Como $R \subseteq A \times B$, sabemos que $x \in A$ e $y \in B$. La hipótesis universal, aplicada a esas componentes, convierte $R(x,y)$ en $S(x,y)$; por tanto, $(x,y) \in S$. Hemos probado $R \subseteq S$. □

Después de la prueba. Para establecer $R = S$ podemos aplicar la proposición en ambas direcciones, igual que hicimos con la igualdad de conjuntos en §4.1. El hecho de que las relaciones sean conjuntos permite reutilizar las técnicas de demostración ya conocidas. En particular, sobre el producto fijado, $R \cap S$ selecciona exactamente los pares que verifican **ambas** condiciones, y $R \cup S$, los que verifican **al menos una**.

Cuando ambos conjuntos son el mismo: propiedades de una relación. Si $R \subseteq A \times A$, las dos componentes pertenecen a un mismo conjunto y podemos formular ciertas propiedades sin cambiar los tipos de las posiciones. Introduciremos tres como ejemplos de traducción de lenguaje a cuantificadores, sin desarrollar aquí una teoría de clasificación de relaciones:

Propiedad de R sobre A	Condición lógica
Reflexiva	$\forall x \in A\, R(x,x)$.
Simétrica	$\forall x, y \in A\, (R(x,y) \to R(y,x))$.
Transitiva	$\forall x, y, z \in A\, ((R(x,y) \land R(y,z)) \to R(x,z))$.

La relación $\leq$ restringida a $A = \{1,2,3\}$ es reflexiva: todo $x \in A$ satisface $x \leq x$; es transitiva porque de $x \leq y$ e $y \leq z$ se sigue $x \leq z$. No es simétrica: $1 \leq 2$, pero $2 \not\leq 1$. Por otra parte, la relación $Q = \{(1,2),(2,3)\}$ sobre el mismo A **no** es transitiva: están los dos pares del antecedente para $x = 1, y = 2, z = 3$, pero falta $(1,3)$. Un solo triple adecuado refuta la afirmación universal de transitividad.

Estas propiedades se aplican a una relación *sobre* un conjunto fijado; no debemos trasladar mecánicamente la fórmula $R(y,x)$ a una relación de A en B cuando quizá $y \notin A$ o $x \notin B$. La inversión de posiciones exige comprobar primero los dominios pertinentes.

Auditoría de un argumento: confundir producto, relación y función. Leemos el razonamiento siguiente: «Si $x \in A$ e $y \in B$, entonces $(x,y) \in A \times B$; por tanto $R(x,y)$. Además, puesto que R es una relación, a cada x le corresponde un único y». Las dos conclusiones son injustificadas. La primera confunde **pertenecer al producto** con **pertenecer al subconjunto** R. La segunda agrega condiciones de existencia y unicidad que no aparecen en la definición $R \subseteq A \times B$.

El ejemplo $x < y$ anterior permite diagnosticar ambas fallas sin ambigüedad: $(2,2) \in A \times B$ pero $(2,2) \notin R$; el elemento 1 tiene dos parejas y el 4 no tiene ninguna. La reparación consiste en probar por separado que un par cumple la condición de R y, si queremos obtener una función, verificar para **cada entrada** la existencia de un resultado y la imposibilidad de dos resultados distintos. La definición exacta de función y sus obligaciones de prueba se establecerán en la sección siguiente.

Comprobación de aprendizaje. Sea $A = \{1,2\}$ y $B = \{2\}$, y definamos $S = \{(x,y) \in A \times B : x < y\}$. Determina S y $\mathrm{dom}(S)$, decide si $\forall x \in A\ \exists y \in B\ S(x,y)$ es verdadera y explica por qué el producto contiene un par que S no contiene. La respuesta correcta exige **enumerar pares con posiciones**, **aplicar la condición** y **evaluar el cuantificador**, en ese orden.

Respuesta razonada. Los pares disponibles son $(1,2)$ y $(2,2)$; sólo el primero cumple $x < y$. Así, $S = \{(1,2)\}$ y $\mathrm{dom}(S) = \{1\}$. La universal es falsa: $x = 2$ no tiene pareja en S. El par $(2,2)$ pertenece al producto pero no a S, pues $2 < 2$ es falso.

Después de la prueba. Una relación es un conjunto de pares y sus afirmaciones pueden leerse mediante pertenencia y cuantificadores. Lo que todavía falta para hablar de función no es una nueva notación, sino una exigencia estructural: para cada entrada del dominio declarado debe existir **exactamente una** salida del codominio. Ésa será la pregunta de §4.5.

4.5. ¿Qué convierte una correspondencia en función? Existencia y unicidad por entrada

En §4.4 vimos que una relación puede dejar entradas sin pareja o asociar a una entrada varias segundas componentes. Una **función** excluye ambas posibilidades, pero no prohíbe que *entradas diferentes* compartan una salida. Esta distinción será decisiva: la unicidad se exige **después de fijar cada entrada**, no entre todas las entradas del dominio.

Definición: una relación total y univaluada respecto de dos conjuntos. Sean A y B conjuntos declarados y $G \subseteq A \times B$. Diremos que G es el **grafo de una función de** A **en** B cuando

$$\boxed{\forall x \in A\ \exists! y \in B\ ((x,y) \in G).}$$

Si usamos la abreviatura $G(x,y)$ de §4.4, esta fórmula se escribe $\forall x \in A\ \exists! y \in B\ G(x,y)$. No afirma que *exista un único par en todo el grafo*: permite un par por cada entrada, incluso cuando hay muchas entradas.

Para hacer explícito qué hemos de probar, desplegamos el cuantificador de existencia única, con $x \in A$ fijado:

$$\exists y \in B\ \big[(x,y) \in G\ \wedge \forall z \in B\ ((x,z) \in G \to z = y)\big].$$

La definición equivale a exigir simultáneamente estas dos condiciones:

$$\begin{aligned} &\text{Existencia para cada entrada:} && \forall x \in A\ \exists y \in B\ ((x,y) \in G), \\ &\text{Unicidad para cada entrada:} && \forall x \in A\ \forall y, z \in B\ \big(((x,y) \in G \wedge (x,z) \in G) \to y = z\big). \end{aligned}$$

Demostración de la equivalencia. Si cada x tiene exactamente una pareja, la primera condición se obtiene tomando esa pareja y la segunda comparando cualesquiera dos parejas de la *misma* entrada: ambas deben coincidir con la única. Recíprocamente, fijemos $x \in A$. La primera condición proporciona un $y \in B$ con $(x,y) \in G$. Si otro $z \in B$ también cumple $(x,z) \in G$, la segunda condición, aplicada a ese x, y, z, da $z = y$. Existe, pues, exactamente una pareja para el x arbitrario. $\square$

Prueba de función: dos obligaciones independientes

Existencia: para un $x \in A$ arbitrario, producir algún $y \in B$ y comprobar $(x, y) \in G$.

Unicidad: dados $y, z \in B$ tales que $(x, y), (x, z) \in G$, demostrar $y = z$.

Un ejemplo particular no prueba la existencia universal; encontrar una salida para cada entrada tampoco demuestra la unicidad si puede haber otras.

La primera condición equivale a $\mathrm{dom}(G) = A$, donde $\mathrm{dom}(G)$ es el dominio efectivo definido en §4.4: por construcción ya sabemos que $\mathrm{dom}(G) \subseteq A$, y la existencia aporta la inclusión inversa. La segunda condición suele llamarse *univaluación* o propiedad de **un solo valor por entrada**. Una relación puede ser univaluada sin estar definida en todo A; aún no sería una función $A \to B$ según nuestra convención.

Tres relaciones para separar los dos requisitos. Fijemos $A = \{1, 2, 4\}$ y $B = \{2, 3\}$ y comparemos

$$G = \{(1, 2), (2, 3), (4, 2)\},$$
$$H = \{(1, 2), (2, 3)\},$$
$$K = \{(1, 2), (1, 3), (2, 3), (4, 2)\}.$$

El conjunto G cumple los dos requisitos: cada elemento de A aparece como primera componente y ninguna primera componente tiene dos valores diferentes. Por tanto, determina una función de A en B. Que 1 y 4 tengan ambos salida 2 **no** viola la definición: las entradas son distintas.

La relación H satisface la unicidad, pero falla la existencia para $x = 4$: no hay par que comience por 4. Su dominio efectivo es $\{1, 2\}$; puede servir de grafo de una función definida sobre ese conjunto más pequeño, pero **no** de una función cuyo dominio declarado sea A.

La relación K satisface la existencia para las tres entradas, pero falla la unicidad en $x = 1$, puesto que $(1, 2), (1, 3) \in K$ y $2 \neq 3$. Por tanto, tampoco es el grafo de una función $A \to B$. En el ejemplo de la relación $x < y$ de §4.4 fallaban *ambas* condiciones: 4 no tenía salida y 1 tenía dos. Tenemos así contraejemplos que permiten distinguir las dos obligaciones, en lugar de confundirlas bajo la frase imprecisa «la regla no funciona».

Dominio, codominio, grafo y valor: datos distintos. En este libro una función $f : A \to B$ tiene **dominio declarado** A, **codominio declarado** B y un grafo $G_f \subseteq A \times B$ que verifica la condición de existencia y unicidad. Si se prefiere explicitar estos datos, podemos representar la función mediante la terna (A, B, G_f): ésta es una convención adoptada para controlar los tipos de entrada y salida, no una afirmación de que todos los textos identifiquen funciones de la misma manera.

Una vez demostrada la condición, para cada $x \in A$ denotamos por $f(x)$ el **único** $y \in B$ tal que $(x, y) \in G_f$. Ésta es la lectura precisa de la notación funcional:

$$\boxed{f(x) = y \quad \Longleftrightarrow \quad (x, y) \in G_f \qquad (x \in A,\ y \in B).}$$

Podemos escribir $x \mapsto f(x)$ para indicar la asignación y reconstruir el grafo a partir de los valores:

$$G_f = \{(x, f(x)) : x \in A\}.$$

Justificación. Cada par de la derecha pertenece a G_f por la definición de $f(x)$. Recíprocamente, si $(x, y) \in G_f$, entonces $x \in A$ y la unicidad obliga a $y = f(x)$, de modo que el par aparece en el conjunto de la derecha. Las dos inclusiones prueban la igualdad. □

En particular, $f(x)$ es un **elemento de** B, mientras que G_f es un **conjunto de pares** y f es la función con sus datos declarados. Escribir una fórmula o dibujar flechas puede ser

una forma de especificar la asignación; ninguna de esas presentaciones exime de comprobar que está definida y es univaluada sobre el dominio escogido.

El codominio no se deduce sólo de los pares. Sea $A_0 = \{1\}$ y $G_0 = \{(1, 2)\}$. El mismo grafo permite declarar una función $A_0 \to \{2\}$ o una función $A_0 \to \{2, 3\}$. Ambas asignan 2 a la entrada 1, pero sus codominios declarados son diferentes; por nuestra convención, no son la misma función como dato completo. La igualdad de funciones se estudiará sistemáticamente en §4.7. Por ahora, no confundamos codominio con el conjunto de valores efectivamente alcanzados.

Una prueba con dominio infinito: $n \mapsto n + 1$. Consideremos la relación, construida dentro de un producto previamente fijado,

$$G = \{(n, m) \in \mathbb{Z} \times \mathbb{Z} : m = n + 1\}.$$

Proposición. G es el grafo de una función $f : \mathbb{Z} \to \mathbb{Z}$ y, para todo $n \in \mathbb{Z}$, $f(n) = n + 1$.

Demostración. Fijemos un entero n arbitrario. Para la **existencia**, elegimos $m = n + 1$. Como la suma de enteros es un entero, $m \in \mathbb{Z}$ y la igualdad que define G se satisface; luego $(n, m) \in G$. Para la **unicidad**, supongamos $(n, m) \in G$ y $(n, m') \in G$ con $m, m' \in \mathbb{Z}$. La definición da $m = n + 1$ y $m' = n + 1$; por transitividad de la igualdad, $m = m'$. Las dos condiciones valen para todo $n \in \mathbb{Z}$, así que G determina una función con dominio y codominio $\mathbb{Z}$. Su valor en n es precisamente el entero construido, $f(n) = n + 1$. □

Lectura de la prueba. La fórmula $m = n + 1$ ofrece un candidato, pero el argumento verifica también que pertenece al codominio y que otro candidato no puede diferir de él. Si sustituyéramos el codominio por un conjunto que excluyera algunos valores $n + 1$, la misma fórmula podría dejar de definir una función con los datos anunciados.

Casos vacíos y una distinción lógica sutil. Si $A = \varnothing$, el producto $A \times B$ es vacío y su única relación es $G = \varnothing$. La afirmación $\forall x \in A\ \exists! y \in B\ G(x, y)$ resulta verdadera **vacuamente**, cualquiera que sea B, incluso si también es vacío. Así, existe la función vacía $\varnothing \to B$, con grafo vacío. No hemos deducido que exista un elemento de B: no hubo ninguna entrada que exigiera testigo.

Si, en cambio, $A \neq \varnothing$ y $B = \varnothing$, no puede existir una función $A \to B$. En efecto, tomando $a \in A$, la existencia exigiría algún $y \in \varnothing$, lo cual es imposible. Estos casos muestran por qué la definición exige un cuantificador **para cada entrada** y por qué no debe sustituirse por la afirmación incondicional $\exists y \in B$.

Auditoría: tres deducciones que no autoriza una fórmula. Alguien escribe: «Como $y = x^2$ da una salida, es una función de $\mathbb{Z}$ en $\mathbb{Z}$; además, al ser función, cada y procede de un único x; por último, el codominio es necesariamente el conjunto de los cuadrados». El primer paso podría justificarse, pero no **sólo** con la frase «da una salida»: hay que verificar para cada entero x que $x^2 \in \mathbb{Z}$ y que, si $y = x^2$ y $z = x^2$, entonces $y = z$. Esas comprobaciones sí demuestran que el grafo de $y = x^2$ define una función $\mathbb{Z} \to \mathbb{Z}$.

La segunda conclusión es falsa: 1 y -1 son entradas diferentes con el mismo valor 1. La unicidad de la definición compara **salidas para una entrada fija**, no entradas que comparten una salida. La tercera conclusión también es falsa: el codominio declarado puede ser $\mathbb{Z}$ aunque sólo se alcancen enteros cuadrados. Las propiedades que discriminan entradas y valores alcanzados pertenecen a §4.7; aquí basta no atribuirlas automáticamente a toda función.

Comprobación de aprendizaje. Para $A = \{0, 1\}$, $B = \{0, 1, 2\}$ y $T = \{(0, 1), (1, 1)\}$, comprueba existencia y unicidad para cada entrada, escribe $f(0)$ y $f(1)$, e identifica su dominio efectivo y su codominio declarado. Después agrega $(0, 2)$ al grafo: ¿qué obligación falla, y con qué dos testigos? Si, en vez de agregarlo, eliminas $(1, 1)$, ¿qué obligación falla? La respuesta debe señalar una **entrada concreta** en cada caso.

Respuesta razonada. Las entradas 0 y 1 tienen, cada una, la única salida 1; por tanto $f(0) = f(1) = 1$, el dominio efectivo es $\{0, 1\} = A$ y el codominio declarado es $B = \{0, 1, 2\}$. Agregar $(0, 2)$ viola la unicidad en la entrada 0, con salidas distintas 1 y 2. Eliminar $(1, 1)$ del grafo original viola la existencia en la entrada 1.

Después de la prueba. Ahora podemos distinguir una relación arbitraria de un grafo funcional y leer $f(x)$ como el único valor asociado a x. La siguiente pregunta no cambia la definición de función: estudia qué **subconjuntos del codominio se alcanzan** y qué **entradas corresponden a un conjunto de valores**. En §4.6 introduciremos imagen y preimagen sin confundir una preimagen con una función inversa.

4.6. Imagen y preimagen: valores alcanzados y selección de entradas

La sección anterior definió una función $f : A \to B$ mediante un dominio A, un codominio B y un grafo que asigna a cada $x \in A$ un único valor $f(x) \in B$. Esto todavía no responde dos preguntas diferentes: si escogemos **algunas entradas**, ¿qué valores producen? Si escogemos **algunos valores del codominio**, ¿qué entradas los producen? La primera pregunta conduce a la *imagen*; la segunda, a la *preimagen*. En ambas fijaremos los conjuntos antes de aplicar las definiciones.

La imagen: partir de entradas y reunir los valores obtenidos. Sea $S \subseteq A$. Definimos la **imagen de S por f** como el conjunto

$$\boxed{f(S) := \{f(x) : x \in S\} \subseteq B.}$$

No se está aplicando f a un conjunto como si éste fuera necesariamente un elemento de A. La notación $f(S)$ designa el **conjunto de los valores** de $f(x)$ cuando x recorre S; la extensión de una función a subconjuntos se define expresamente aquí. En particular, para $y \in B$,

$$\boxed{y \in f(S) \quad \Longleftrightarrow \quad \exists x \in S\ (f(x) = y).}$$

Justificación. Si $y \in f(S)$, por definición es el valor $f(x)$ de alguna entrada $x \in S$; tal entrada es el testigo del existencial. Recíprocamente, si existe $x \in S$ con $f(x) = y$, el valor y figura entre los reunidos por la definición de $f(S)$. Las dos direcciones justifican el bicondicional. □

Esta fórmula también explica por qué **no** se exige que cada valor del codominio sea alcanzado: sólo sabemos que $f(S) \subseteq B$. Llamamos **imagen de la función** o conjunto de valores efectivamente alcanzados a

$$\mathrm{Im}(f) := f(A) = \{y \in B : \exists x \in A\ (f(x) = y)\}.$$

Usaremos $\mathrm{Im}(f)$ cuando interese distinguirla de la imagen de un subconjunto particular. El codominio B forma parte de los datos declarados de la función, mientras que $\mathrm{Im}(f)$ se determina mediante sus valores. Son iguales en algunas funciones, pero no por definición.

La preimagen: fijar un conjunto de valores y seleccionar entradas. Sea ahora $T \subseteq B$. Definimos su **preimagen por f**, también llamada *imagen inversa del conjunto T*, por

$$\boxed{f^{-1}(T) := \{x \in A : f(x) \in T\} \subseteq A.}$$

La prueba de pertenencia es directa: para todo $x \in A$,

$$\boxed{x \in f^{-1}(T) \quad \Longleftrightarrow \quad f(x) \in T.}$$

La imagen pregunta por un **testigo de entrada** para cada valor candidato; la preimagen pregunta por una **condición sobre el valor de una entrada ya fijada**. La definición de preimagen no presupone que exista una entrada para cada $y \in T$, ni que tal entrada sea única. Puede haber cero, una o varias entradas que se seleccionen mediante la misma condición. La comprensión se efectúa dentro del dominio previamente declarado A, sin construir conjuntos a partir de un universo irrestricto.

Leer los símbolos sin intercambiar los dominios

Para $f : A \to B$, $S \subseteq A$ y $T \subseteq B$:

Construcción	Conjunto resultante	Criterio para pertenecer
$f(S)$	Subconjunto de B.	$y \in f(S) \iff \exists x \in S\ (f(x) = y)$.
$f^{-1}(T)$	Subconjunto de A.	$x \in f^{-1}(T) \iff f(x) \in T$, para $x \in A$.

Las letras S y T representan **conjuntos**, no entradas y salidas individuales. En general, $f(S)$ no es un elemento de B, ni $f^{-1}(T)$ es un elemento de A.

Un ejemplo que impide confundir imagen, codominio y función inversa. Trabajemos con los conjuntos finitos

$$A = \{-2, -1, 0, 1, 2\}, \qquad B = \{0, 1, 4, 9\},$$

y con la función $f : A \to B$ definida por $f(n) = n^2$. Cada entero de A tiene exactamente un cuadrado y cada uno de esos cuadrados pertenece a B; la justificación de que la regla define una función usa las obligaciones del §4.5. Su grafo es

$$G_f = \{(-2, 4), (-1, 1), (0, 0), (1, 1), (2, 4)\}.$$

No aparece ningún par con segunda componente 9. Por tanto,

$$\mathrm{Im}(f) = f(A) = \{0, 1, 4\} \quad \text{y} \quad \mathrm{Im}(f) \neq B.$$

Si seleccionamos $S = \{-2, 0, 1\}$, sus valores son 4, 0 y 1, de modo que $f(S) = \{0, 1, 4\}$. No debemos incorporar -1 ni 2 al **conjunto imagen**: son entradas de A, aunque sus valores sean los mismos que los de algunas entradas de S. Si elegimos un subconjunto menor, $S_0 = \{-2, 2\}$, obtenemos $f(S_0) = \{4\}$; dos entradas pueden contribuir a un solo elemento de la imagen.

Por otro lado, para $T = \{1, 9\} \subseteq B$, las entradas seleccionadas son únicamente -1 y 1: ambos cuadrados valen 1, y ninguna entrada tiene cuadrado 9. Así,

$$f^{-1}(\{1, 9\}) = \{-1, 1\}, \qquad f^{-1}(\{9\}) = \varnothing.$$

También $f^{-1}(\{0, 1\}) = \{-1, 0, 1\}$. En el primer cálculo, el valor 9 pertenece a T y al codominio, pero no produce ninguna entrada; en el segundo, 1 produce **dos entradas distintas**. Precisamente por eso la preimagen está bien definida como *conjunto*, aunque no pueda asignarse a cada $y \in B$ una única entrada $x \in A$ con $f(x) = y$.

Las fibras: fijar una salida y examinar sus entradas. La preimagen de un conjunto puede reunir entradas que producen varios valores. Si fijamos un solo valor $b \in B$, obtenemos un conjunto especialmente útil, llamado **fibra de f sobre b**:

$$F_b := f^{-1}(\{b\}) = \{x \in A : f(x) = b\}.$$

La fibra puede ser vacía, contener una entrada o contener varias. Su existencia como *conjunto* está garantizada por la definición de preimagen; no estamos afirmando que exista

una entrada con valor b. Esta diferencia será importante cuando estudiemos inyectividad y sobreyectividad y, después, al resolver los ejercicios sobre fibras.

Para pasar de un ejemplo finito a una función con infinitas entradas, consideremos $p : \mathbb{Z} \to \{0, 1, 2\}$ definida por

$$p(n) = \begin{cases} 0, & \text{si } n \text{ es par,} \\ 1, & \text{si } n \text{ es impar.} \end{cases}$$

La clasificación usual de los enteros en pares e impares asegura que cada entrada pertenece a exactamente uno de los casos. Por tanto, la regla asigna un único valor admisible a cada entero. Si escribimos $P_b := p^{-1}(\{b\})$ para las fibras de esta nueva función, P_0 consta de los enteros pares, P_1 de los impares y $P_2 = \varnothing$, pues ninguna entrada recibe valor 2. El codominio anunciado contiene, así, un elemento cuya fibra es vacía. A la inversa, una sola fibra puede contener infinitas entradas.

Comprobación. ¿Cuál es $p^{-1}(\{1, 2\})$? ¿Permite calcular este conjunto afirmar que existe una función inversa $p^{-1} : \{0, 1, 2\} \to \mathbb{Z}$?

Respuesta razonada. La preimagen es el conjunto de los enteros impares: todas esas entradas producen el valor 1, mientras que ninguna produce 2. No existe la función inversa indicada: al valor 2 no le correspondería ninguna entrada y a 0 o 1 les corresponderían varias. La operación de tomar preimágenes no exige invertir la función.

¿Por qué $f^{-1}(T)$ no presupone una función inversa?. El exponente -1 en $f^{-1}(T)$ es una **notación para la preimagen de un conjunto**, no una afirmación de que ya exista una función $f^{-1} : B \to A$. Si quisiéramos definir una función inversa mediante la regla «a y le corresponde el x tal que $f(x) = y$», deberíamos probar, para cada $y \in B$, la existencia y la unicidad de esa entrada. Nuestro ejemplo refuta ambos requisitos: 9 no tiene ninguna preimagen individual y 1 tiene dos, -1 y 1.

Podemos invertir formalmente los pares del grafo y formar una relación de B en A, pero en este ejemplo esa relación no es el grafo de una función $B \to A$. Esto no impide calcular $f^{-1}(T)$ para **todo** subconjunto $T \subseteq B$: basta evaluar la condición $f(x) \in T$ para cada $x \in A$. Las hipótesis bajo las cuales sí puede invertirse una función se identificarán en §4.7, sin darlas por supuestas aquí.

Lectura guiada. La frase «el valor 1 tiene una preimagen» es ambigua si no se aclara si se habla de una entrada individual o del *conjunto* de entradas. Es más preciso escribir $f^{-1}(\{1\}) = \{-1, 1\}$. La imagen inversa de un singleton puede contener varios elementos, uno o ninguno; no se obtiene por ello una función inversa.

Una demostración por pertenencia: la preimagen conserva intersecciones. Proposición. Para una función $f : A \to B$ y subconjuntos $T, U \subseteq B$,

$$\boxed{f^{-1}(T \cap U) = f^{-1}(T) \cap f^{-1}(U).}$$

Demostración. Tomemos una entrada arbitraria $x \in A$. Por la definición de preimagen,

$$\begin{aligned} x \in f^{-1}(T \cap U) &\iff f(x) \in T \cap U \\ &\iff (f(x) \in T) \wedge (f(x) \in U) \\ &\iff (x \in f^{-1}(T)) \wedge (x \in f^{-1}(U)) \\ &\iff x \in f^{-1}(T) \cap f^{-1}(U). \end{aligned}$$

Cada paso aplica una definición de pertenencia; el segundo despliega la intersección y el tercero vuelve a traducir mediante preimágenes. Como el bicondicional vale para cualquier $x \in A$ y ambos miembros son subconjuntos de A, la extensionalidad establece la igualdad. □

Después de la prueba. La afirmación no depende de que f sea inyectiva ni de que alcance todo B: sólo utiliza que cada entrada tiene un valor definido. El mismo método y el conectivo «o» prueban

$$f^{-1}(T \cup U) = f^{-1}(T) \cup f^{-1}(U).$$

Para esta segunda igualdad, el paso central sería $f(x) \in T \cup U \iff f(x) \in T \vee f(x) \in U$. No se concluye una identidad por semejanza gráfica: se identifica la equivalencia de pertenencia que la prueba necesitaría.

La imagen de una intersección: una inclusión que no siempre es igualdad. Para subconjuntos $S, V \subseteq A$ se cumple

$$\boxed{f(S \cap V) \subseteq f(S) \cap f(V).}$$

Demostración. Sea $y \in f(S \cap V)$. Existe $x \in S \cap V$ tal que $f(x) = y$. Ese mismo testigo pertenece a S y a V; por consiguiente, $y \in f(S)$ e $y \in f(V)$, y obtenemos la inclusión. □

¿Es válida la inclusión inversa? Regresemos a la función cuadrado de nuestro ejemplo y elijamos $S = \{-1\}$ y $V = \{1\}$. Entonces $S \cap V = \varnothing$, por lo que $f(S \cap V) = \varnothing$; sin embargo,

$$f(S) \cap f(V) = \{1\} \cap \{1\} = \{1\}.$$

El error de la supuesta recíproca está en los testigos. Saber que $y \in f(S)$ y $y \in f(V)$ proporciona una entrada en S y otra en V que alcanzan el mismo valor y; no garantiza que sea una sola entrada que pertenezca a ambos conjuntos. En cambio, sí vale siempre

$$f(S \cup V) = f(S) \cup f(V),$$

pues un testigo de pertenencia a $S \cup V$ pertenece a al menos uno de los conjuntos y, recíprocamente, cualquier testigo en uno de ellos también pertenece a la unión. De nuevo, el alcance del cuantificador existencial decide qué operaciones pueden intercambiarse con la imagen.

Dos relaciones útiles y los casos vacíos. La imagen y la preimagen permiten leer de manera exacta qué sucede al ir de las entradas a los valores y regresar. Para $S \subseteq A$ y $T \subseteq B$ tenemos

$$\boxed{S \subseteq f^{-1}(f(S)), \qquad f(f^{-1}(T)) = T \cap \operatorname{Im}(f).}$$

Primera relación. Si $x \in S$, su valor $f(x)$ pertenece a $f(S)$ por definición; por tanto $x \in f^{-1}(f(S))$. Esto demuestra la inclusión. Puede ser estricta: con la función cuadrado y $S = \{1\}$, resulta $f(S) = \{1\}$ y $f^{-1}(f(S)) = \{-1, 1\}$; reaparece una entrada externa a S que comparte el mismo valor.

Segunda relación. Para demostrar la igualdad, tomemos $y \in B$. La condición $y \in f(f^{-1}(T))$ significa que existe $x \in f^{-1}(T)$ con $f(x) = y$. Como $x \in f^{-1}(T)$ equivale a $f(x) \in T$, esto significa que $y \in T$ y que existe $x \in A$ con $f(x) = y$, es decir, $y \in T \cap \operatorname{Im}(f)$. Recíprocamente, si y pertenece a esa intersección, su pertenencia a $\operatorname{Im}(f)$ aporta un $x \in A$ con $f(x) = y$ y su pertenencia a T implica $x \in f^{-1}(T)$. Ese testigo da $y \in f(f^{-1}(T))$. Ambas inclusiones quedan probadas. □

La intersección con $\operatorname{Im}(f)$ no puede omitirse: para $T = \{9\}$ en nuestro ejemplo, $f(f^{-1}(T)) = \varnothing \neq T$, puesto que 9 nunca se alcanza. Finalmente, las definiciones, sin excepciones adicionales, proporcionan

$$f(\varnothing) = \varnothing, \qquad f^{-1}(\varnothing) = \varnothing, \qquad f^{-1}(B) = A.$$

La primera igualdad no tiene testigos $x \in \varnothing$; la segunda no admite entradas cuyo valor pertenezca al vacío; la tercera reúne todas las entradas porque la función toma sus valores en B. Si $A = \varnothing$, la función vacía $A \to B$ tiene imagen vacía y la preimagen de *cualquier* $T \subseteq B$ es vacía. Ninguna de estas afirmaciones obliga a que B sea no vacío.

Auditoría: dos errores de dirección y un supuesto oculto. Consideremos la afirmación defectuosa: «Sea $f : A \to B$ y $T \subseteq B$. Como $f^{-1}(T)$ está definido, cada $y \in T$ debe proceder de exactamente un x; por tanto $f(f^{-1}(T)) = T$. Además, si $S, V \subseteq A$, es evidente que $f(S \cap V) = f(S) \cap f(V)$». El argumento introduce una función inversa no demostrada y confunde testigos que pueden ser diferentes.

La reparación exige tres pasos separados. Primero, $f^{-1}(T)$ es un **conjunto de entradas**, cuya definición no afirma existencia ni unicidad de preimágenes individuales. Segundo, $f(f^{-1}(T))$ sólo recupera los elementos de T que efectivamente pertenecen a $\mathrm{Im}(f)$; su igualdad correcta es $T \cap \mathrm{Im}(f)$. Tercero, de una entrada común en $S \cap V$ se deducen las dos pertenencias de la imagen, pero de dos entradas distintas con valor común no se obtiene una entrada común. Por ello la relación general para las imágenes de intersecciones es una **inclusión**, y la función cuadrado proporciona el contraejemplo a la igualdad universal.

Comprobación de aprendizaje. Para la función cuadrado finita de esta sección, calcula $f(\{-1, 2\})$, $f^{-1}(\{4, 9\})$ y $f^{-1}(f(\{-2\}))$. En cada caso indica si el resultado es subconjunto de A o de B. Después explica por qué $f^{-1}(\{9\})$ existe como conjunto, aunque ningún elemento de A tenga valor 9. Justifica tus respuestas a partir de las definiciones, no sólo enumerando pares.

Respuesta razonada. La imagen $f(\{-1, 2\}) = \{1, 4\} \subseteq B$ reúne los cuadrados de esas dos entradas. La preimagen $f^{-1}(\{4, 9\}) = \{-2, 2\} \subseteq A$ selecciona las entradas con cuadrado 4 o 9; ninguna tiene cuadrado 9. Finalmente, $f(\{-2\}) = \{4\}$, de modo que $f^{-1}(f(\{-2\})) = \{-2, 2\} \subseteq A$. El conjunto $f^{-1}(\{9\}) = \varnothing$ está definido por una condición que ninguna entrada satisface; no necesita un testigo para existir como conjunto vacío.

Después de la prueba. Ya distinguimos el codominio declarado, los valores efectivamente alcanzados y los conjuntos de entradas seleccionados mediante condiciones sobre valores. Las fallas de ciertas igualdades han revelado dos preguntas estructurales: ¿cuándo pueden dos entradas diferentes compartir un valor? y ¿cuándo se alcanza cada elemento del codominio? En §4.7 responderemos mediante inyectividad, sobreyectividad y biyectividad, y fijaremos además la igualdad de funciones con todos sus datos explícitos.

4.7. Inyectividad, sobreyectividad, biyectividad e igualdad de funciones

En §4.5 exigimos que cada entrada de una función tenga exactamente una salida; en §4.6 distinguimos el codominio del conjunto de valores alcanzados y encontramos distintas entradas con una misma imagen. Ninguno de esos hechos modifica la definición de función. Ahora formularemos dos **propiedades adicionales**, que responden a preguntas independientes: ¿pueden dos entradas tener el mismo valor?, ¿se alcanza cada elemento del codominio? Su combinación permitirá decidir cuándo existe una función inversa. Por último precisaremos qué significa que dos funciones sean iguales bajo la convención de datos declarados adoptada en este libro.

Inyectividad: una salida común obliga a identificar las entradas. Sea $f : A \to B$ una función. Decimos que f es **inyectiva** si

$$\boxed{\forall x, x' \in A \; (f(x) = f(x') \to x = x').}$$

La definición de **función** no exige que entradas distintas produzcan valores distintos: ésa es precisamente la condición adicional de inyectividad. También puede escribirse de manera equivalente

$$\forall x, x' \in A\ (x \neq x' \to f(x) \neq f(x')).$$

Justificación de la equivalencia. Si la primera condición vale y $x \neq x'$, una igualdad $f(x) = f(x')$ obligaría a $x = x'$, contradicción; por tanto los valores son diferentes. Recíprocamente, si vale la segunda y $f(x) = f(x')$, no pueden ser diferentes las entradas, pues entonces los valores tendrían que ser diferentes. Hemos utilizado la contraposición en ambas direcciones. □

¿Qué hay que hacer para *demostrar* inyectividad? Tomar $x, x' \in A$ arbitrarios, suponer $f(x) = f(x')$ y deducir $x = x'$. ¿Qué basta para *refutarla*? Encontrar dos entradas concretas $x \neq x'$ para las cuales $f(x) = f(x')$:

$$\boxed{\neg\,\mathrm{Iny}(f) \quad\Longleftrightarrow\quad \exists x, x' \in A\ (x \neq x' \wedge f(x) = f(x')).}$$

La unicidad de la definición de función (§4.5) fija una entrada y compara sus posibles **salidas**; la inyectividad fija un valor común y compara las **entradas** que lo producen. Intercambiar esas obligaciones llevaría a concluir erróneamente que toda función es inyectiva.

Ejemplo y contraejemplo. La función $u : \mathbb{Z} \to \mathbb{Z}$ dada por $u(n) = n + 1$ es inyectiva. En efecto, si $u(m) = u(n)$, entonces $m + 1 = n + 1$ y, restando 1, resulta $m = n$. En cambio, la función cuadrado $f : A \to B$ de §4.6, con $A = \{-2, -1, 0, 1, 2\}$, $B = \{0, 1, 4, 9\}$ y $f(n) = n^2$, no es inyectiva: $-1 \neq 1$ y $f(-1) = f(1) = 1$. Que cada entrada tenga un cuadrado **único** no impide esta coincidencia.

Sobreyectividad: ningún elemento del codominio queda sin alcanzar. Decimos que una función $f : A \to B$ es **sobreyectiva sobre el codominio declarado** B si

$$\boxed{\forall y \in B\ \exists x \in A\ (f(x) = y).}$$

Aquí el cuantificador universal recorre B, no únicamente $\mathrm{Im}(f)$. Ésa es la diferencia entre la propiedad de sobreyectividad y la afirmación, siempre verdadera, de que cada valor *alcanzado* procede de alguna entrada. La definición admite una reformulación conjuntista:

$$\boxed{f \text{ es sobreyectiva} \quad\Longleftrightarrow\quad \mathrm{Im}(f) = B.}$$

Demostración. Ya sabemos por §4.6 que $\mathrm{Im}(f) \subseteq B$. Si f es sobreyectiva, cada $y \in B$ tiene un testigo $x \in A$ con $f(x) = y$, luego $y \in \mathrm{Im}(f)$; esto prueba $B \subseteq \mathrm{Im}(f)$. Recíprocamente, si los dos conjuntos son iguales, un $y \in B$ pertenece a la imagen y la definición de imagen aporta una entrada $x \in A$ con $f(x) = y$. □

Para probar sobreyectividad debemos fijar un $y \in B$ **arbitrario**, construir una entrada $x \in A$ y comprobar su valor. No basta escribir una expresión candidata para x si no se verifica que pertenece al dominio. Para refutarla, en cambio, basta hallar un elemento del codominio que no sea alcanzado:

$$\boxed{\neg\,\mathrm{Sob}(f) \quad\Longleftrightarrow\quad \exists y \in B\ \forall x \in A\ (f(x) \neq y).}$$

Nuestra función cuadrado finita no es sobreyectiva hacia $B = \{0, 1, 4, 9\}$: el valor $9 \in B$ no es cuadrado de ninguna entrada de A. Sin cambiar ni una entrada ni una asignación, podemos declarar otra función $\widetilde{f} : A \to \{0, 1, 4\}$ mediante $\widetilde{f}(n) = n^2$. Esta nueva función **sí** es sobreyectiva: $0 = \widetilde{f}(0)$, $1 = \widetilde{f}(1)$ y $4 = \widetilde{f}(2)$. No son iguales como funciones en la convención de este libro, pues sus codominios difieren; y ninguna es inyectiva, porque las entradas -1 y 1 siguen compartiendo el valor 1.

Dos propiedades, dos obligaciones de prueba

Propiedad de $f : A \to B$	Para demostrarla	Para refutarla
Inyectiva	De $f(x) = f(x')$, con $x, x' \in A$ arbitrarios, deducir $x = x'$.	Exhibir $x \neq x'$ en A con igual valor.
Sobreyectiva	Para cada $y \in B$, exhibir $x \in A$ tal que $f(x) = y$.	Exhibir $y \in B$ sin ninguna entrada que lo alcance.

La primera compara **entradas**; la segunda exige **testigos para los elementos del codominio**. No se deduce ninguna de las dos solamente de que f sea una función.

Cuatro posibilidades que debemos distinguir. Conjuntos finitos permiten separar las propiedades sin apoyarnos en un dibujo ambiguo. Sean a, b, c tres objetos distintos. Los siguientes ejemplos especifican sus dominios, codominios y asignaciones:

Función y datos declarados	Inyectiva	Sobreyectiva	Justificación decisiva
$p : \{0,1\} \to \{a,b,c\}$; $p(0) = a$, $p(1) = b$	Sí	No	Dos valores distintos; c no se alcanza.
$q : \{0,1\} \to \{a,b,c\}$; $q(0) = q(1) = a$	No	No	Dos entradas comparten a; faltan b, c.
$r : \{0,1,2\} \to \{a,b\}$; $r(0) = a$, $r(1) = r(2) = b$	No	Sí	$1 \neq 2$ comparten b; a y b se alcanzan.
$s : \{0,1\} \to \{a,b\}$; $s(0) = a$, $s(1) = b$	Sí	Sí	No hay colisión y cada valor tiene una entrada.

En cada fila se comprueba previamente que hay exactamente una salida permitida para cada entrada: estamos comparando **funciones**, no relaciones que puedan incumplir su propia definición. La tabla demuestra mediante ejemplos que la inyectividad y la sobreyectividad son condiciones independientes. Una función que satisface ambas se llama **biyectiva**:

$$\boxed{f \text{ es biyectiva} \iff (f \text{ es inyectiva}) \wedge (f \text{ es sobreyectiva}).}$$

Por tanto, para cada $y \in B$ existe **exactamente un** $x \in A$ tal que $f(x) = y$. En símbolos,

$$\boxed{f \text{ es biyectiva} \iff \forall y \in B\ \exists! x \in A\ (f(x) = y).}$$

Justificación de la segunda equivalencia. Si f es sobreyectiva, existe al menos una entrada para cada y; si es además inyectiva, cualesquiera dos entradas con ese valor coinciden. Recíprocamente, la existencia única para cada y da inmediatamente la sobreyectividad. Si $f(x) = f(x') = y$ para dos entradas, ambas verifican la propiedad que tiene un único testigo para y, de modo que $x = x'$ y f es inyectiva. □

La función inversa existe precisamente bajo ambas condiciones. En §4.6 vimos que $f^{-1}(T)$ designa la **preimagen de un conjunto** $T \subseteq B$ sin necesitar que exista una función inversa. Ahora sí podemos formular el problema de invertir una función $f : A \to B$: queremos una función $g : B \to A$ que asigne a cada $y \in B$ la entrada de la que procede. La relación candidata está determinada por el grafo de f al invertir sus pares:

$$G_f^{\mathrm{op}} := \{(y,x) \in B \times A : (x,y) \in G_f\}.$$

Proposición. Esta relación es el grafo de una función $g : B \to A$ si y sólo si f es biyectiva. En ese caso, g se denomina **función inversa** de f, y podemos escribir $g = f^{-1}$.

Demostración. Supongamos primero que f es biyectiva. Para cada $y \in B$, la sobreyectividad proporciona un $x \in A$ con $f(x) = y$; por tanto $(y,x) \in G_f^{\mathrm{op}}$. Si también (y,x') está en la relación, entonces $f(x) = f(x') = y$ y la inyectividad obliga a $x = x'$. Hay exactamente una salida x para cada entrada y: la relación invertida es un grafo funcional de B en A.

Recíprocamente, supongamos que G_f^{op} es el grafo de una función $g : B \to A$. Su condición de existencia dice que para cada $y \in B$ hay un $x \in A$ con $(x,y) \in G_f$, es decir, $f(x) = y$: f es sobreyectiva. Si $f(x) = f(x') = y$, la relación invertida contiene (y,x) y (y,x'); la unicidad de g implica $x = x'$, así que f es inyectiva. Por tanto, f es biyectiva. □

La prueba también determina las **dos identidades de ida y vuelta**:

$$\boxed{g(f(x)) = x \quad (x \in A), \qquad f(g(y)) = y \quad (y \in B).}$$

En efecto, g asigna a $f(x)$ la única entrada que lo produce; ésta es x. De manera recíproca, $g(y)$ fue escogida precisamente porque su imagen es y. Ambas igualdades necesitan los dominios indicados. No son propiedades de la mera preimagen de conjuntos, sino de la función inversa cuya existencia acaba de demostrarse.

Ejemplo infinito con demostración. La función $u : \mathbb{Z} \to \mathbb{Z}$ dada por $u(n) = n + 1$ ya era inyectiva. Para cualquier $m \in \mathbb{Z}$, la entrada $n = m - 1$ también es un entero y satisface $u(n) = (m-1)+1 = m$. Luego es sobreyectiva y, por tanto, biyectiva. Su inversa es $g : \mathbb{Z} \to \mathbb{Z}$, $g(m) = m - 1$: se comprueban directamente $g(u(n)) = n$ y $u(g(m)) = m$. Una fórmula para la entrada candidata se convierte en prueba sólo después de verificar su pertenencia al dominio y las igualdades exigidas.

Advertencia de notación. Para una función biyectiva, la escritura f^{-1} puede designar su función inversa $B \to A$. La expresión $f^{-1}(T)$, cuando T es un **subconjunto** de B, designa la preimagen de T tal como la definimos en §4.6. Son construcciones relacionadas, pero no debe confundirse el valor individual $g(y) \in A$ con el conjunto $f^{-1}(\{y\}) = \{g(y)\} \subseteq A$.

Recuperar las propiedades mediante imágenes y preimágenes. Las relaciones de §4.6 permiten formular criterios que no mencionan directamente las palabras «inyectiva» ni «sobreyectiva».

Proposición. Para cualquier función $f : A \to B$,

$$\boxed{f \text{ es inyectiva} \quad \Longleftrightarrow \quad \forall S \subseteq A \ (f^{-1}(f(S)) = S).}$$

Demostración. Supongamos que f es inyectiva. Ya probamos $S \subseteq f^{-1}(f(S))$. Para la otra inclusión, si $x \in f^{-1}(f(S))$, entonces $f(x) \in f(S)$; existe, pues, $s \in S$ con $f(s) = f(x)$. La inyectividad da $s = x$, de donde $x \in S$. Recíprocamente, supongamos la igualdad para todos los subconjuntos de A. Si $f(x) = f(x')$, tomemos $S = \{x\}$. Entonces $x' \in f^{-1}(f(S)) = S$, y por pertenencia al conjunto unitario concluimos $x' = x$. □

Para la sobreyectividad, recordemos la identidad demostrada en §4.6: $f(f^{-1}(T)) = T \cap \mathrm{Im}(f)$. De ella se obtiene

$$\boxed{f \text{ es sobreyectiva} \quad \Longleftrightarrow \quad \forall T \subseteq B \ (f(f^{-1}(T)) = T).}$$

Demostración. Si f es sobreyectiva, $\mathrm{Im}(f) = B$ y la identidad anterior se reduce a $T \cap B = T$. Recíprocamente, aplicando la igualdad supuesta a $T = B$ obtenemos $f(f^{-1}(B)) = B$. Como $f^{-1}(B) = A$, resulta $f(A) = B$, es decir, sobreyectividad. □

Lectura guiada. En el primer criterio, la inyectividad impide que, al regresar de la imagen, reaparezcan entradas externas a S con el mismo valor. En el segundo, la sobreyectividad impide que falten valores de T al regresar de sus entradas. La diferencia entre los criterios se corresponde exactamente con las dos preguntas planteadas al comenzar la sección.

Igualdad de funciones: deben coincidir los datos, no sólo la fórmula. Nuestra convención de §4.5 representa una función por sus datos (A, B, G_f). Sean $f : A \to B$ y $h : C \to D$ dos funciones. **En este libro**, afirmaremos $f = h$ exactamente cuando coincidan sus dominios declarados, codominios declarados y valores en cada entrada:

$$\boxed{f = h \iff A = C \ \wedge \ B = D \ \wedge \ G_f = G_h.}$$

Esta formulación compara conjuntos y no evalúa h fuera de su dominio. Una vez comprobados $A = C$ y $B = D$, podemos escribir el último requisito como $\forall x \in A \ (f(x) = h(x))$: con dominio y codominio comunes, la igualdad punto a punto equivale a $G_f = G_h$. **Justificación.** Si todos los valores coinciden, cada par $(x, f(x))$ del primer grafo es el mismo par $(x, h(x))$ del segundo, y recíprocamente; la reconstrucción del grafo de §4.5 y la extensionalidad dan la igualdad. En la otra dirección, si los grafos coinciden, para cualquier x el par $(x, f(x))$ pertenece también a G_h; la unicidad de la salida de h obliga a $h(x) = f(x)$. □

No basta observar que dos funciones están descritas por la misma expresión. Consideremos $f : \{0, 1\} \to \{0, 1\}$ y $h : \{0, 1\} \to \{0, 1, 2\}$, ambas dadas por $n \mapsto n$. Sus valores y sus grafos coinciden, pero $f \neq h$ según nuestra convención, porque sus codominios declarados son distintos. Tampoco es suficiente coincidir en algunas entradas cuando el dominio y el codominio sí son comunes: basta un solo $a \in A$ con $f(a) \neq h(a)$ para refutar la igualdad.

Algunos textos identifican funciones por su grafo y no consideran el codominio como parte de su identidad. Nuestra definición no pretende invalidar esa otra convención: fija qué deberemos verificar cuando una afirmación de **igualdad de funciones** aparezca en esta obra. Es indispensable declarar la convención antes de argumentar a partir de ella.

Los conjuntos vacíos ponen a prueba las definiciones. La única función de $\varnothing$ en un conjunto B tiene grafo vacío. Es **inyectiva**: no existen dos entradas distintas que puedan violar la condición universal. Es **sobreyectiva** exactamente cuando $B = \varnothing$, porque su imagen es $\varnothing$. En consecuencia, $\varnothing \to \varnothing$ es biyectiva y su función inversa también es la función vacía. Si $B \neq \varnothing$, la función vacía $\varnothing \to B$ no es sobreyectiva: cualquier $y \in B$ carece de entrada.

Como ya vimos en §4.5, no existe función $A \to \varnothing$ cuando $A \neq \varnothing$. Por otra parte, las funciones vacías $\varnothing \to B$ y $\varnothing \to D$ tienen el mismo grafo, pero no son iguales bajo nuestra convención si $B \neq D$. Los casos vacíos verifican que hemos mantenido separadas la lógica de los cuantificadores, la existencia de funciones y los datos que forman su identidad.

Auditoría de un razonamiento: tres conclusiones injustificadas. Se propone este argumento: «Toda función es inyectiva porque cada entrada tiene una sola salida. Además es sobreyectiva porque su imagen está incluida en el codominio. Finalmente, dos funciones dadas por $x \mapsto x^2$ son iguales, aunque se declaren con codominios diferentes». Las tres afirmaciones confunden propiedades distintas. La **funcionalidad** no impide que -1 y 1 compartan imagen; la inclusión $\mathrm{Im}(f) \subseteq B$ no aporta la inclusión inversa; y, en nuestra convención, la misma fórmula no borra las diferencias entre los codominios declarados.

Para reparar el argumento hay que plantear las obligaciones pertinentes: ante la inyectividad, comenzar con $f(x) = f(x')$; ante la sobreyectividad, tomar $y \in B$ arbitrario y buscar un testigo; ante la igualdad, verificar dominio, codominio y coincidencia punto a punto. La función cuadrado de §4.6 refuta las dos primeras inferencias; sus posibles codominios ilustran

la tercera. Ninguna reparación puede basarse únicamente en cambiar las palabras «entrada» y «salida».

Comprobación de aprendizaje. Determina si la función $u : \mathbb{Z} \to \mathbb{Z}$, $u(n) = n + 1$, es inyectiva, sobreyectiva y biyectiva; para la sobreyectividad comienza con un entero m **arbitrario**. Después, para la función cuadrado finita de §4.6, identifica una colisión y un valor del codominio no alcanzado. Por último, explica por qué $f : \{0\} \to \{0\}$ y $h : \{0\} \to \{0, 1\}$, ambas con valor 0 en la única entrada, tienen grafos iguales pero no son iguales como funciones en nuestra convención.

Respuesta razonada. Si $u(n) = u(k)$, cancelar 1 da $n = k$, así que u es inyectiva. Para cada entero m, el entero $m - 1$ satisface $u(m - 1) = m$: es sobreyectiva y, por ambas propiedades, biyectiva. En el cuadrado finito, $f(-1) = f(1) = 1$ es una colisión y $9 \in B$ no se alcanza. Las dos funciones del último ejemplo tienen grafo $\{(0, 0)\}$, pero sus codominios $\{0\}$ y $\{0, 1\}$ difieren; por la convención declarada, son distintas.

Después de la prueba. Las propiedades de una función pueden leerse como afirmaciones cuantificadas sobre sus entradas y su codominio; la biyectividad permite invertir el grafo porque asegura existencia y unicidad en la dirección contraria. La igualdad requiere además controlar los datos declarados. En §4.8 volveremos al lenguaje de las fórmulas abiertas para estudiar cómo una propiedad selecciona un **conjunto solución dentro de un dominio fijado**, y cómo cambian sus elementos cuando cambia ese dominio.

4.8. Conjuntos solución, fórmulas y dominios de selección

En §4.1 reunimos a los enteros que satisfacían una ecuación; en §§4.5–4.7 usamos condiciones sobre entradas y valores para definir funciones y estudiar sus propiedades. Podemos ahora hacer explícito el principio que conecta ambas operaciones: **una fórmula abierta describe una condición; un conjunto solución reúne los elementos de un dominio previamente fijado que la satisfacen**. El dominio no es una información secundaria que podamos suprimir después de escribir la condición.

Definición: seleccionar dentro de un conjunto, no de un universo irrestricto. Sea D un conjunto especificado y sea $P(x)$ una fórmula con x como variable libre, interpretable para cada $x \in D$. Si hay parámetros en P, sus valores y sus dominios se consideran fijados en el contexto. Denotamos por **conjunto solución de P en D** al conjunto

$$\boxed{S_D(P) := \{x \in D : P(x)\}.}$$

La escritura de comprensión es una selección **dentro de** D. Su criterio exacto de pertenencia, para cualquier objeto a al que sean aplicables las expresiones, es

$$\boxed{a \in S_D(P) \quad \Longleftrightarrow \quad (a \in D) \wedge P(a).}$$

En particular, si ya hemos declarado $a \in D$, podemos abreviar el criterio como $a \in S_D(P) \iff P(a)$. **La abreviatura sólo vale bajo esa hipótesis**. Si P está definida exclusivamente sobre D, ni siquiera debemos evaluar $P(a)$ para un objeto externo sin extender primero su interpretación.

Demostración de las dos direcciones. Supongamos $a \in S_D(P)$. La definición del conjunto por comprensión dice simultáneamente que $a \in D$ y que la sustitución $P(a)$ es verdadera. Recíprocamente, si $a \in D$ y $P(a)$ es verdadera, a cumple la condición que selecciona los elementos de D y, por definición, pertenece a $S_D(P)$. La equivalencia no requiere resolver previamente todos los casos del problema. □

Tres objetos de lenguaje distintos

- $P(x)$ es una **fórmula abierta**, cuyo valor de verdad depende de una asignación para x y, cuando existan, de los parámetros.
- $P(a)$ es una **afirmación evaluada** para un objeto admisible a.
- $S_D(P)$ es un **conjunto**, construido seleccionando elementos del dominio declarado D.

No confundimos el conjunto con la condición que lo define ni escribimos $S_D(P) = P(x)$: la igualdad carecería del sentido que pretendemos expresar.

El mismo predicado puede seleccionar conjuntos diferentes. Consideremos la fórmula $P(x) :\ x^2 = 1$, interpretable sobre los enteros, y dos dominios fijados:

$$D = \{-2, -1, 0, 1, 2\}, \qquad E = \{0, 1, 2\}.$$

Evaluando la misma condición en cada dominio obtenemos

$$S_D(P) = \{-1, 1\}, \qquad S_E(P) = \{1\}.$$

La fórmula no ha cambiado: $(-1)^2 = 1$ es verdadera en ambos contextos como igualdad aritmética. Lo que cambia es **si -1 está disponible para la selección**. Al trabajar dentro de E, la verdad de la igualdad no basta para introducir -1 en $S_E(P)$ porque $-1 \notin E$.

El ejemplo sugiere una relación general cuya prueba exige conservar explícitamente el dominio de cada conjunto.

Proposición (restricción del dominio). Sean $D \subseteq E$ dos conjuntos y sea $P(x)$ una misma condición interpretable en E, con los mismos parámetros e interpretación cuando restringimos x a D. Entonces

$$\boxed{S_D(P) = D \cap S_E(P).}$$

Demostración. Ambos miembros son subconjuntos de E. Basta tomar $a \in E$ arbitrario, donde $P(a)$ está definida. Por definición de conjunto solución,

$$\begin{aligned} a \in S_D(P) &\iff (a \in D) \wedge P(a) \\ &\iff (a \in D) \wedge \big(a \in S_E(P)\big) \\ &\iff a \in D \cap S_E(P). \end{aligned}$$

En el segundo paso hemos usado $D \subseteq E$: cuando $a \in D$, ya sabemos que $a \in E$, por lo que $P(a)$ equivale, en ese contexto, a pertenecer a $S_E(P)$. El bicondicional de pertenencia prueba la igualdad por extensionalidad. □

Lectura guiada. La conclusión no afirma que los conjuntos solución sean siempre iguales cuando $D \subseteq E$: afirma que el conjunto relativo a D se obtiene intersectando D con el conjunto relativo a E. En particular, $S_D(P) \subseteq S_E(P)$; la inclusión puede ser estricta, como en el ejemplo anterior. Tampoco podemos aplicar la proposición si al cambiar de dominio cambiamos silenciosamente el significado de P o de alguno de sus parámetros.

Cuantificar equivale a preguntar por la extensión de un conjunto solución. Fijados D y P, las afirmaciones existencial, universal y de existencia única pueden leerse mediante el conjunto solución:

$$\boxed{\begin{aligned} &\exists x \in D\ P(x) \iff S_D(P) \neq \varnothing, \\ &\forall x \in D\ P(x) \iff S_D(P) = D, \\ &\exists! x \in D\ P(x) \iff \exists a \in D\ \big(S_D(P) = \{a\}\big). \end{aligned}}$$

Justificación. En la primera línea, un testigo $a \in D$ que cumple $P(a)$ es exactamente un elemento de $S_D(P)$; que el conjunto no sea vacío proporciona, a la inversa, un testigo. Para la segunda línea, ya sabemos que $S_D(P) \subseteq D$; la afirmación universal aporta la inclusión $D \subseteq S_D(P)$, y la igualdad de los conjuntos recupera $P(a)$ para cada $a \in D$. Para la tercera línea, un testigo único a pertenece al conjunto solución y cualquier otro elemento de éste debe coincidir con a, de modo que $S_D(P) = \{a\}$. Recíprocamente, si el conjunto solución es ese singleton, a existe en D, satisface P y cualquier otro testigo pertenece a $\{a\}$, luego es igual a a. □

Si $D = \varnothing$, necesariamente $S_D(P) = \varnothing$: la afirmación universal es verdadera, pero ni la existencial ni la de existencia única lo son. El conjunto vacío no inventa testigos; revela la diferencia entre exigir algo de **cada** elemento y exigir que **exista** al menos uno.

La negación de una afirmación universal también queda representada sin cambiar el dominio:

$$\neg\big(\forall x \in D\ P(x)\big) \iff \exists x \in D\ \neg P(x) \iff D \setminus S_D(P) \neq \varnothing.$$

El testigo de la negación debe pertenecer a D. Buscar un objeto externo que no cumpla P no refuta una proposición cuantificada **solamente sobre** D.

Los conectivos construyen operaciones de conjuntos solución. Si $P(x)$ y $Q(x)$ son fórmulas interpretables sobre el **mismo dominio** D, sus conjuntos solución obedecen a las reglas

$$\boxed{\begin{aligned} S_D(P \wedge Q) &= S_D(P) \cap S_D(Q), \\ S_D(P \vee Q) &= S_D(P) \cup S_D(Q), \\ S_D(\neg P) &= D \setminus S_D(P). \end{aligned}}$$

La disyunción es inclusiva, tal como la definimos al estudiar la unión en §4.2. La tercera identidad expresa un complemento **relativo a** D, no el pretendido conjunto de todos los objetos que no verifican P.

Demostración de la primera identidad. Los dos conjuntos son subconjuntos de D; por ello basta tomar $a \in D$ arbitrario, donde $P(a)$ y $Q(a)$ están definidas:

$$\begin{aligned} a \in S_D(P \wedge Q) &\iff (a \in D) \wedge \big(P(a) \wedge Q(a)\big) \\ &\iff \big((a \in D) \wedge P(a)\big) \wedge \big((a \in D) \wedge Q(a)\big) \\ &\iff (a \in S_D(P)) \wedge (a \in S_D(Q)) \\ &\iff a \in S_D(P) \cap S_D(Q). \end{aligned}$$

La segunda equivalencia usa que repetir la misma condición $a \in D$ en una conjunción no cambia su valor de verdad. Por extensionalidad, los conjuntos son iguales. Para la unión se emplea la equivalencia lógica entre $D(a) \wedge (P(a) \vee Q(a))$ y $(D(a) \wedge P(a)) \vee (D(a) \wedge Q(a))$, donde $D(a)$ abrevia **únicamente en esta explicación** la afirmación $a \in D$; para el complemento se utiliza $a \in D \wedge \neg P(a)$. □

Una consecuencia útil. Si $\forall x \in D\ (P(x) \to Q(x))$, entonces $S_D(P) \subseteq S_D(Q)$. En efecto, cualquier a del primer conjunto pertenece a D y cumple $P(a)$; la implicación da $Q(a)$ y por ello pertenece al segundo. Recíprocamente, esa inclusión permite recuperar la implicación para todo $x \in D$. Así,

$$\boxed{\forall x \in D\ (P(x) \to Q(x)) \quad \iff \quad S_D(P) \subseteq S_D(Q).}$$

Del mismo modo, $S_D(P) = S_D(Q)$ equivale a $\forall x \in D\ (P(x) \leftrightarrow Q(x))$. **Dos fórmulas pueden seleccionar exactamente los mismos objetos en D sin ser expresiones idénticas**. Lo que demostramos es su equivalencia sobre un dominio y bajo una interpretación,

no su identidad como cadenas de símbolos ni su equivalencia automática en cualquier otro universo.

Ecuaciones, funciones y relaciones como conjuntos solución. La notación de §4.6 es un caso de este mismo mecanismo. Sea $f : A \to B$ una función y sea $T \subseteq B$. La fórmula $P_T(x) :\ f(x) \in T$ tiene sentido para $x \in A$ y selecciona

$$\boxed{S_A(P_T) = \{x \in A : f(x) \in T\} = f^{-1}(T).}$$

Si elegimos un $b \in B$, la ecuación $f(x) = b$ define el conjunto de soluciones

$$\boxed{\{x \in A : f(x) = b\} = f^{-1}(\{b\}).}$$

Justificación. Para $x \in A$, pertenecer al conjunto de la izquierda significa $f(x) = b$; pertenecer a la preimagen de $\{b\}$ significa $f(x) \in \{b\}$, que es la misma igualdad. La extensionalidad prueba la identidad. □

Volvamos a la función cuadrado de §4.6, cuyo dominio es $A = \{-2, -1, 0, 1, 2\}$. La ecuación $f(x) = 1$ tiene conjunto solución $\{-1, 1\}$; la ecuación $f(x) = 9$ tiene conjunto solución vacío **en ese dominio**, aunque 9 pertenezca al codominio declarado. Que una ecuación tenga soluciones depende tanto de su condición como del conjunto de entradas disponibles.

Para una relación $R \subseteq A \times B$, la fórmula $P_R(x) :\ \exists y \in B\ R(x, y)$ selecciona las entradas que están relacionadas con algún elemento de B:

$$\boxed{S_A(P_R) = \mathrm{dom}(R).}$$

Esta igualdad se obtiene directamente de la definición de dominio efectivo en §4.4. No afirma que R sea funcional: la existencia de alguna segunda componente para ciertas entradas no garantiza la unicidad de ninguna de ellas. Asimismo, si una fórmula tiene **dos variables libres** $P(x, y)$, su conjunto de soluciones es naturalmente un subconjunto del producto previamente fijado:

$$\{(x, y) \in A \times B : P(x, y)\}.$$

La cuantificación $\exists y \in B$ liga la variable y y deja libre a x; el conjunto solución resultante vuelve a estar formado por **elementos de** A, no por pares. Antes de escribir llaves conviene, por tanto, identificar qué variables permanecen libres y de qué tipo serán los elementos seleccionados.

Parámetros: fijarlos antes de resolver. Consideremos la familia de fórmulas $P_t(n) : n^2 = t$ con $n \in \mathbb{Z}$ y con un parámetro $t \in \mathbb{Z}$ fijado. Cada elección del parámetro determina un conjunto solución distinto, que podemos denotar por

$$S_t := \{n \in \mathbb{Z} : n^2 = t\}.$$

Por ejemplo, $S_0 = \{0\}$ y $S_1 = \{-1, 1\}$. Es importante no escribir $\{n \in \mathbb{Z} : n^2 = t\} = \{-1, 1\}$ sin haber declarado $t = 1$: la primera expresión todavía depende del parámetro. Si dejamos t sin fijar, no hemos identificado un único conjunto de enteros, sino una familia dependiente de una asignación.

En cambio, el enunciado $\forall t \in \mathbb{Z}\ \exists n \in \mathbb{Z}\ (n^2 = t)$ es una **proposición cerrada** y resulta falso: $t = -1$ no tiene testigo entero, pues todo cuadrado entero es no negativo. La mera presencia de una letra t en una fórmula abierta, la fijación de su valor y la cuantificación sobre todos sus valores constituyen tres situaciones diferentes. Esta distinción recupera el tratamiento de variables, alcance y dependencia estudiado en el capítulo 3.

Auditoría de un argumento: borrar el dominio después de escribir llaves. Un razonamiento afirma: «Sea $D = \{1\}$ y sea $P(x)$ la propiedad $x = x$ para enteros. Como $P(2)$ es verdadera, $2 \in \{x \in D : P(x)\}$. Además, el conjunto de soluciones de $\neg P$ contiene todos los objetos que no pertenecen a D». Ambas conclusiones son falsas. El primer paso elimina la condición $2 \in D$, que no se cumple; de hecho, $S_D(P) = \{1\}$. El segundo sustituye la condición $x \in D \wedge \neg P(x)$ por $x \notin D$, que expresa otra propiedad: en este caso $S_D(\neg P) = \varnothing$.

Reparación. Para probar $a \in S_D(P)$ debemos justificar tanto $a \in D$ como $P(a)$. Para negar la propiedad sin cambiar de dominio, formamos $D \setminus S_D(P)$. Si cambiamos deliberadamente el dominio de D a un conjunto mayor E, recalculamos la selección y relacionamos ambos resultados mediante $S_D(P) = D \cap S_E(P)$, siempre que se conserve la interpretación de la condición.

Comprobación de aprendizaje. Con $D = \{-2, -1, 0, 1, 2\}$, $E = \{0, 1, 2\}$ y $P(x) : x^2 = 1$, determina $S_D(P)$, $S_E(P)$ y $S_D(\neg P)$. Demuestra por pertenencia que $S_E(P) = E \cap S_D(P)$ y explica por qué, pese a la verdad de $P(-1)$ como igualdad aritmética, $-1 \notin S_E(P)$. Por último, con $f : A \to B$ y $b \in B$, expresa «la ecuación $f(x) = b$ tiene exactamente una solución en A» mediante un conjunto solución y relaciona tu respuesta con la definición de biyectividad de §4.7 sin olvidar que la condición debe valer **para cada** $b \in B$.

Respuesta razonada. Se obtiene $S_D(P) = \{-1, 1\}$, $S_E(P) = \{1\}$ y $S_D(\neg P) = \{-2, 0, 2\}$. Para $x \in D$, pertenecer a $E \cap S_D(P)$ equivale a $x \in E$ y $x^2 = 1$, es decir, a pertenecer a $S_E(P)$; fuera de D ninguno de los dos conjuntos tiene elementos. Aunque $P(-1)$ sea verdadera, $-1 \notin E$, por lo que no se selecciona. La ecuación $f(x) = b$ tiene solución única en A si y sólo si existe $a \in A$ tal que $f^{-1}(\{b\}) = \{a\}$. Exigir esto para cada $b \in B$ equivale a que f sea biyectiva; el testigo a puede depender de b.

Después de la prueba. Hemos conectado fórmulas, cuantificadores, operaciones de conjuntos, relaciones y funciones mediante un mismo criterio: la pertenencia equivale a satisfacer una condición **dentro de un dominio declarado**. En §4.9 ejercitaremos esa lectura en ambas direcciones: formular condiciones a partir de conjuntos, construir conjuntos a partir de condiciones y detectar errores de alcance, testigos y tipos antes de intentar una demostración.

4.9. Ejercicios graduados: de la lectura de definiciones a la elección de una prueba

Los ejercicios de este capítulo se agrupan según la **actividad intelectual que exigen**, no sólo según el símbolo que aparece en su enunciado. Primero traduciremos una afirmación; después distinguiremos una conclusión legítima de otra que no se sigue de las hipótesis; construiremos ejemplos y contraejemplos; completaremos pruebas cuyos pasos están orientados; por último, elegiremos nosotros mismos el método. La dificultad aumenta, aunque cada familia utiliza también conceptos de las anteriores.

En todos los problemas debe conservarse el dominio de cada variable. Al demostrar una igualdad de conjuntos, justifica sus dos inclusiones o un bicondicional de pertenencia; al refutar una afirmación universal, identifica un contraejemplo y verifica que cumple las hipótesis. Para funciones rige la convención de §§4.5 y 4.7: el dominio y el codominio declarados forman parte de sus datos. Las soluciones desarrolladas se presentan, con la misma numeración, en §4.10.

Familia I. Traducción y control de tipos — lectura de definiciones. Ejercicio 1. Un conjunto descrito de dos maneras. Sean U un conjunto y $A, B, C \subseteq U$. Considera $H = (A \cup B) \setminus C$.

(a) Escribe una condición lógica necesaria y suficiente para $x \in H$ en términos de las tres pertenencias $x \in A$, $x \in B$ y $x \in C$. (b) Construye una fórmula $P(x)$, interpretable en U, para la cual $H = S_U(P)$. (c) Explica por qué escribir $x \notin C$ sin conservar la pertenencia a $A \cup B$ no caracteriza a H. No se pide enumerar elementos.

Ejercicio 2. ¿Cuándo representa una relación una función? Sean $R \subseteq A \times B$ y $R(x, y)$ la abreviatura de $(x, y) \in R$. Traduce «cada elemento de A se relaciona con exactamente un elemento de B» a una fórmula cuantificada. Después sepárala en dos enunciados: existencia para cada entrada y unicidad para cada entrada. Escribe finalmente, mediante cuantificadores restringidos, la negación de la condición de existencia. Indica en cada fórmula cuáles son las variables ligadas.

Ejercicio 3. Imágenes, preimágenes y sobreyectividad. Sea $f : A \to B$, con $S \subseteq A$ y $T \subseteq B$.

(a) Traduce $y \in f(S)$ mediante un cuantificador sobre S y una igualdad entre valores. (b) Traduce $x \in f^{-1}(T)$ mediante una condición sobre $f(x)$, declarando el dominio de x. (c) Expresa la sobreyectividad de f de dos modos equivalentes: mediante cuantificadores y mediante el conjunto $\mathrm{Im}(f)$. Explica por qué ninguna de estas expresiones convierte $f^{-1}(T)$ en una función inversa.

Familia II. Reparación de argumentos — detectar la hipótesis que falta. Ejercicio 4. Pertenencia e inclusión no son intercambiables. Se argumenta: «Si $A \subseteq B$, entonces $A \in B$, porque todos los elementos de A están en B». Proporciona conjuntos finitos concretos que refuten esa implicación. Construye también un ejemplo que refute la implicación inversa $A \in B \Rightarrow A \subseteq B$. En cada caso señala cuál es el *objeto* cuyo estatuto de pertenencia se está confundiendo con una condición universal.

Ejercicio 5. Un conjunto de pares no es automáticamente el grafo de una función. Sean $A = \{1, 2\}$, $B = \{0, 1\}$ y $R = \{(1, 0), (1, 1)\} \subseteq A \times B$. Una demostración declara: «Cada par de R tiene una primera y una segunda componente; por eso R define una función $A \to B$». Identifica por separado el fallo de existencia y el fallo de unicidad. Modifica el conjunto de pares para construir un grafo $G \subseteq A \times B$ que sí determine una función $A \to B$, y justifica ambos requisitos para tu elección.

Ejercicio 6. Lo que una imagen no permite recuperar sin hipótesis. Se propone: «Como $f(S) \subseteq f(T)$, necesariamente $S \subseteq T$». Construye un contraejemplo con una función $f : A \to B$ y subconjuntos $S, T \subseteq A$ que no satisfagan la conclusión. Luego añade una hipótesis adecuada sobre f y demuestra que, con ella, la implicación sí es válida. Explica por qué la inclusión $S \subseteq T \Rightarrow f(S) \subseteq f(T)$ no requiere esa hipótesis adicional.

Familia III. Ejemplos y contraejemplos — separar propiedades. Ejercicio 7. Existencia frente a unicidad. Sean $A = \{p, q\}$, con $p \neq q$, y $B = \{0, 1\}$. Construye dos relaciones distintas de A en B: una que satisfaga la existencia de alguna pareja para cada entrada pero viole la unicidad, y otra que satisfaga la univaluación pero deje alguna entrada sin pareja. Para cada relación escribe sus pares y verifica la condición que cumple y la que incumple mediante entradas y salidas concretas.

Ejercicio 8. Una misma lista de pares y dos funciones diferentes. Sean $A = \{p, q\}$, $B = \{u, v\}$, $C = \{u, v, w\}$, donde $p \neq q$ y u, v, w son tres objetos distintos. Usa el grafo $G = \{(p, u), (q, v)\}$ para declarar $f : A \to B$ y $g : A \to C$. Comprueba que ambas son funciones y que sus grafos son iguales. Determina, justificándolo, cuáles son inyectivas y cuáles sobreyectivas. Decide si $f = g$ según la convención de este libro e identifica exactamente qué dato produce la diferencia.

Ejercicio 9. La misma fórmula en dos dominios. Elige conjuntos finitos de enteros $D \subsetneq E$ y dos fórmulas $P(x)$ y $Q(x)$ interpretables en E, con los parámetros fijados, tales que $S_D(P) = S_E(P)$ pero $S_D(Q) \subsetneq S_E(Q)$. Calcula los cuatro conjuntos solución y verifica las dos relaciones por pertenencia. Explica por qué la inclusión $D \subseteq E$ garantiza $S_D(R) \subseteq S_E(R)$ para cualquier condición R que conserve su interpretación, pero no garantiza igualdad.

Familia IV. Demostraciones guiadas — identificar la obligación de cada paso. Ejercicio 10. Negación de una intersección. Para conjuntos arbitrarios A, B, C, demuestra

$$A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C).$$

Sigue esta guía: fija x arbitrario; traduce la pertenencia al miembro izquierdo; aplica la ley de De Morgan; distribuye la condición $x \in A$ sobre la disyunción; reconoce las dos diferencias; concluye mediante extensionalidad. Explica cuál de los pasos garantiza las dos inclusiones y por qué un diagrama o una comprobación finita no sería, por sí solo, una demostración general.

Ejercicio 11. De una condición a una función. Sea

$$G = \{(n, m) \in \mathbb{Z} \times \mathbb{Z} : m = 2n + 1\}.$$

(a) Prueba que G es el grafo de una función $f : \mathbb{Z} \to \mathbb{Z}$ mediante existencia y unicidad para una entrada arbitraria. (b) Demuestra que f es inyectiva. (c) Da un entero del codominio que no sea alcanzado y justifica que no existe una entrada con ese valor. (d) Calcula $f^{-1}(\{1, 3\})$ y explica por qué tu resultado es un conjunto de entradas, no la función inversa de f.

Ejercicio 12. Recuperar sólo los valores alcanzados. Sea $f : A \to B$ y $T \subseteq B$. En §4.6 se demostró $f(f^{-1}(T)) = T \cap \mathrm{Im}(f)$. Utiliza esa identidad o demuestra directamente las dos inclusiones para establecer el criterio

$$f(f^{-1}(T)) = T \quad \Longleftrightarrow \quad T \subseteq \mathrm{Im}(f).$$

Comprueba por qué las dos direcciones siguen siendo válidas cuando $T = \varnothing$. Después aplica el criterio a la función cuadrado de §4.6 con $T = \{9\}$ e identifica qué falla.

Familia V. Elección autónoma de método — decidir cómo comenzar. Ejercicio 13. Fibras y propiedades de una función. Sea $f : A \to B$ y, para cada $b \in B$, escribe $F_b = f^{-1}(\{b\})$. Demuestra, eligiendo tu estrategia, las dos equivalencias siguientes:

$$f \text{ es inyectiva} \Longleftrightarrow \forall b \in B\ \forall x, x' \in F_b\ (x = x'),$$
$$f \text{ es biyectiva} \Longleftrightarrow \forall b \in B\ \exists! x \in A\ (f(x) = b).$$

No supongas que todas las fibras son no vacías al tratar la primera línea. Examina explícitamente el caso $A = \varnothing$ y explica qué cambia si B también es vacío.

Ejercicio 14. El dominio efectivo de una intersección de relaciones. Sean $R, S \subseteq A \times B$. Estudia la afirmación

$$\mathrm{dom}(R \cap S) \subseteq \mathrm{dom}(R) \cap \mathrm{dom}(S).$$

Decide si es verdadera y demuéstrala o refútala. Averigua después si la inclusión inversa vale siempre. Si no vale, construye relaciones sobre los mismos conjuntos de referencia para las cuales un elemento pertenezca a ambos dominios efectivos, pero no al dominio efectivo de la intersección. Identifica el punto preciso donde no es legítimo reutilizar el mismo testigo.

Ejercicio 15. ¿Determinan las fibras a la función? Sean $f, g : A \to B$ funciones con **el mismo dominio y el mismo codominio declarados**. Decide si vale la equivalencia

$$f = g \quad \Longleftrightarrow \quad \forall b \in B\ (f^{-1}(\{b\}) = g^{-1}(\{b\})).$$

Elige una estrategia para cada dirección. Si utilizas la extensionalidad de las preimágenes, explica cómo pasar de información sobre conjuntos de entradas a la igualdad $f(x) = g(x)$ para un $x \in A$ arbitrario. Revisa el caso $A = \varnothing$ y declara qué parte de la conclusión dependería de nuestra convención si los codominios fueran diferentes.

Cierre de la práctica. Las quince preguntas no se resuelven todas mediante el mismo gesto: unas se traducen, otras se refutan, otras requieren producir testigos y otras exigen las dos direcciones de una equivalencia. Antes de abordar §4.10 conviene anotar, para cada ejercicio, qué se da, qué se pide, qué definiciones hacen falta y qué propiedad del dominio o del codominio podría perderse si se omitiera. Las soluciones explicitan tanto la respuesta como la decisión de método.

4.10. Soluciones desarrolladas: reconstruir el razonamiento, no sólo el resultado

Conservamos los números y los datos de §4.9. Cada solución identifica qué definición debe desplegarse, qué objeto o testigo exige la prueba y qué conclusión queda efectivamente justificada. Un ejemplo resuelve un problema existencial o refuta una afirmación universal cuando corresponde; no sustituye una demostración sobre conjuntos arbitrarios. Como siempre, las letras que representan dominios y codominios están fijadas antes de cuantificar.

Solución 1. Un conjunto descrito de dos maneras. Método elegido: desplegar pertenencia y volver a formar un conjunto solución. De la definición de diferencia, para cualquier objeto x,

$$\begin{aligned} x \in H &\iff x \in A \cup B \ \wedge \ x \notin C \\ &\iff ((x \in A) \vee (x \in B)) \wedge (x \notin C). \end{aligned}$$

Esto responde a (a). Para (b), como $A, B, C \subseteq U$, fijamos en U la fórmula abierta

$$P(x): \quad ((x \in A) \vee (x \in B)) \wedge (x \notin C).$$

Entonces $S_U(P) = \{x \in U : P(x)\} = H$: ambos conjuntos tienen los mismos elementos según el bicondicional anterior; la condición de pertenecer a U se cumple automáticamente para un elemento de $A \cup B$. No hemos reunido objetos de un universo irrestricto.

Para (c), $x \notin C$ sólo excluye a C; no garantiza que x proceda de la unión. Por ejemplo, con $U = \{1, 2\}$, $A = \{1\}$ y $B = C = \varnothing$, el objeto 2 no pertenece a C, pero tampoco a $H = \{1\}$. El error consiste en eliminar un miembro de la conjunción que define la diferencia.

Solución 2. ¿Cuándo representa una relación una función?. Método elegido: desplegar el cuantificador de existencia única. La traducción solicitada es

$$\forall x \in A \, \exists! y \in B \; R(x, y).$$

Equivale a la conjunción de dos enunciados:

$$\begin{aligned} &\forall x \in A \, \exists y \in B \; R(x, y), \\ &\forall x \in A \, \forall y, z \in B \; ((R(x, y) \wedge R(x, z)) \to y = z). \end{aligned}$$

La primera línea exige **una salida para cada entrada**; la segunda compara dos salidas de **una misma entrada** y obliga a que coincidan. Para ver la equivalencia, la existencia única implica ambas condiciones; a la inversa, para un x arbitrario elegimos un y por la primera condición y usamos la segunda para descartar cualquier z diferente.

La negación de la condición de existencia es

$$\exists x \in A \, \forall y \in B \; \neg R(x, y).$$

Afirma que **alguna** entrada no tiene pareja. En la primera fórmula están ligadas x e y; en la descomposición también lo está z; en la negación vuelven a estar ligadas x e y. Los conjuntos A, B y la relación R son datos fijados, no variables libres que haya que cuantificar en estas fórmulas. El cambio $\neg\forall x \, \exists y \rightsquigarrow \exists x \, \forall y \, \neg$ proviene de negar ambos cuantificadores, no de intercambiar arbitrariamente su orden.

Solución 3. Imágenes, preimágenes y sobreyectividad. Método elegido: traducir las dos nociones de conjunto antes de tratar la propiedad de la función. Para $y \in B$,

$$y \in f(S) \quad \Longleftrightarrow \quad \exists x \in S\,(f(x) = y).$$

La variable x es una entrada testigo; el objeto y es un valor. Para $x \in A$,

$$x \in f^{-1}(T) \quad \Longleftrightarrow \quad f(x) \in T.$$

Aquí el elemento del conjunto resultante es **una entrada**, y no necesitamos resolver la ecuación $f(x) = y$ para todos los posibles y.

La sobreyectividad hacia el codominio declarado B se expresa equivalentemente por

$$\forall y \in B\ \exists x \in A\ (f(x) = y), \qquad \mathrm{Im}(f) = B.$$

La equivalencia se debe a que, por definición, $y \in \mathrm{Im}(f)$ significa que existe una entrada cuyo valor es y, y siempre $\mathrm{Im}(f) \subseteq B$. Finalmente, $f^{-1}(T)$ nombra un **subconjunto de** A para cualquier $T \subseteq B$; una función inversa $B \to A$ requeriría que **cada** $y \in B$ tuviera exactamente una entrada que lo produjera. Nada de lo escrito en (a) o (b) presupone esa condición.

Solución 4. Pertenencia e inclusión no son intercambiables. Método elegido: refutar cada implicación con un contraejemplo que distinga objeto y elementos. Para la primera, tomemos $A = \{1\}$ y $B = \{1, 2\}$. Se cumple $A \subseteq B$, porque el único elemento de A es $1 \in B$. Sin embargo, $A \notin B$: los elementos enumerados de B son 1 y 2, no el conjunto $\{1\}$. La inclusión habla de los **elementos de** A; la conclusión errónea intenta convertir al **propio conjunto** A en elemento de B.

Para la implicación inversa, tomemos $A = \{1\}$ y $B = \{\{1\}\}$. Ahora $A \in B$, pues el único elemento de B es precisamente $\{1\}$. No obstante, $A \nsubseteq B$: su único elemento, 1, no es un elemento de B, cuyo único elemento es $\{1\}$. La información sobre la pertenencia del objeto A no afirma nada, por sí sola, sobre la pertenencia de **cada elemento suyo**.

Solución 5. Un conjunto de pares no es automáticamente el grafo de una función. Método elegido: inspeccionar por separado las dos obligaciones para cada entrada. El conjunto inicial es $A = \{1, 2\}$; en $R = \{(1, 0), (1, 1)\}$ no aparece ningún par con primera componente 2. Por eso falla la existencia universal: para $x = 2$ no existe $y \in B$ tal que $(2, y) \in R$. También falla la unicidad: la entrada 1 está emparejada con 0 y 1, dos elementos distintos de B.

Una reparación posible es

$$G = \{(1, 0), (2, 1)\} \subseteq A \times B.$$

Para $x = 1$ existe la salida 0 y es la única en G con primera componente 1; para $x = 2$ existe la salida 1 y es igualmente única. Como A contiene exactamente esas dos entradas, queda probado $\forall x \in A\, \exists! y \in B\, ((x, y) \in G)$. No bastaba comprobar que *cada par* tiene dos componentes: lo exigido se cuantifica sobre **todas las entradas de** A.

Solución 6. Lo que una imagen no permite recuperar sin hipótesis. Método elegido: buscar una colisión para refutar y usar la inyectividad para reparar. Sea $A = \{0, 1\}$, $B = \{a\}$ y $f : A \to B$ la función constante $f(0) = f(1) = a$. Tomemos $S = \{0\}$ y $T = \{1\}$. Entonces $f(S) = f(T) = \{a\}$ y, en particular, $f(S) \subseteq f(T)$, pero $S \nsubseteq T$ porque $0 \in S$ y $0 \notin T$.

La hipótesis adicional suficiente es que f sea **inyectiva**. Supongamos ahora $f(S) \subseteq f(T)$ y tomemos $x \in S$ arbitrario. Su valor $f(x)$ pertenece a $f(S)$ y, por la inclusión, a $f(T)$. Existe

entonces $t \in T$ con $f(t) = f(x)$. La inyectividad obliga a $t = x$; como $t \in T$, tenemos $x \in T$. Por arbitrariedad de x, $S \subseteq T$. El paso decisivo es **identificar dos testigos mediante la inyectividad**.

La implicación directa $S \subseteq T \Rightarrow f(S) \subseteq f(T)$ no necesita esa hipótesis: si $y \in f(S)$, un testigo $x \in S$ satisface $f(x) = y$; la inclusión $S \subseteq T$ garantiza que **ese mismo** x pertenece a T, por lo que $y \in f(T)$.

Solución 7. Existencia frente a unicidad. Método elegido: construir cada ejemplo para que falle exactamente el requisito solicitado. Consideremos

$$R = \{(p, 0), (p, 1), (q, 0)\}, \qquad S = \{(p, 0)\}.$$

En R, tanto p como q tienen al menos una pareja: p se relaciona con 0 y q con 0. Pero p también se relaciona con 1 y $0 \neq 1$; falla la univaluación. En S, la única entrada que aparece, p, tiene una sola salida, 0; por tanto, si (x, y) y (x, z) pertenecen a S, forzosamente $y = z = 0$. La univaluación se cumple. En cambio, $q \in A$ carece de pareja, por lo que falla la existencia para cada entrada.

Ambos conjuntos están incluidos en $A \times B$, pero **ninguno** determina una función $A \to B$: uno incumple unicidad y el otro existencia. Esta comprobación evita confundir «a lo sumo una salida» con «exactamente una salida».

Solución 8. Una misma lista de pares y dos funciones diferentes. Método elegido: fijar todos los datos declarados antes de comparar propiedades. Para $f : A \to B$ y $g : A \to C$ usamos el mismo grafo $G = \{(p, u), (q, v)\}$. Ambas son funciones: cada una de las dos entradas p, q tiene una salida única y los valores u, v pertenecen tanto a B como a C. Sus grafos coinciden literalmente.

Ambas son **inyectivas**: las dos entradas distintas tienen valores distintos, $u \neq v$. La función f es **sobreyectiva**, pues $f(p) = u$ y $f(q) = v$ alcanzan todos los elementos de $B = \{u, v\}$. La función g **no** es sobreyectiva hacia $C = \{u, v, w\}$, porque $w \in C$ y no existe una entrada cuyo valor sea w. Por consiguiente, f es biyectiva y g no.

Según la convención de §4.7, $f \neq g$: aunque los dominios y los grafos sean idénticos, sus **codominios declarados** son diferentes ($B \neq C$, pues $w \in C \setminus B$). La igualdad de la lista de pares no basta, en este libro, para igualar dos funciones con distinto codominio.

Solución 9. La misma fórmula en dos dominios. Método elegido: introducir un objeto nuevo que satisfaga sólo una de las dos condiciones. Elegimos

$$D = \{0, 1\} \subsetneq E = \{0, 1, 2\}, \qquad P(x) : x = 1, \qquad Q(x) : x \geq 1.$$

Las fórmulas son las mismas en ambos dominios y no poseen parámetros indeterminados. Al evaluar pertenencia obtenemos

$$S_D(P) = \{1\}, \quad S_E(P) = \{1\},$$
$$S_D(Q) = \{1\}, \quad S_E(Q) = \{1, 2\}.$$

La igualdad de los dos conjuntos asociados a P se prueba porque 1 es el único entero de ambos dominios que satisface $x = 1$. Para Q, la inclusión $\{1\} \subseteq \{1, 2\}$ es propia: 2 está en $S_E(Q)$, pero no en D y por tanto tampoco en $S_D(Q)$.

En general, con $D \subseteq E$ y una condición R interpretada de la misma manera, si $x \in S_D(R)$ entonces $x \in D$ y $R(x)$. Como $x \in E$, también $x \in S_E(R)$; esto prueba $S_D(R) \subseteq S_E(R)$. La inclusión inversa no está garantizada: los elementos de $E \setminus D$ pueden satisfacer la fórmula, como ocurre con 2 y Q. Un cambio de dominio es, por ello, un dato matemático y no una modificación inocua de notación.

Solución 10. Negación de una intersección. Método elegido: un bicondicional de pertenencia que prueba simultáneamente ambas inclusiones. Para un objeto arbitrario x,

$$\begin{aligned} x \in A \setminus (B \cap C) &\iff (x \in A) \wedge \neg((x \in B) \wedge (x \in C)) \\ &\iff (x \in A) \wedge ((x \notin B) \vee (x \notin C)) \\ &\iff ((x \in A) \wedge (x \notin B)) \vee ((x \in A) \wedge (x \notin C)) \\ &\iff x \in (A \setminus B) \cup (A \setminus C). \end{aligned}$$

El segundo paso es la ley de De Morgan; el tercero es la distributividad de la conjunción sobre la disyunción; el último reconoce diferencia y unión. Cada flecha es un **bicondicional**: leída de izquierda a derecha da $A \setminus (B \cap C) \subseteq (A \setminus B) \cup (A \setminus C)$, y de derecha a izquierda da la inclusión contraria. Como x era arbitrario, la extensionalidad concluye la igualdad de los conjuntos.

Un dibujo puede orientar la conjetura y una enumeración finita verificar un caso, pero ninguno de esos procedimientos, sin un argumento general, establece el resultado para **todos** los conjuntos A, B, C.

Solución 11. De una condición a una función. Método elegido: demostrar primero que la relación es funcional; sólo entonces estudiar sus propiedades. (a) Sea $n \in \mathbb{Z}$ arbitrario. El entero $m = 2n + 1$ pertenece a $\mathbb{Z}$ y satisface $(n, m) \in G$, lo que prueba existencia. Si $(n, m), (n, m') \in G$, entonces $m = 2n + 1 = m'$, lo que prueba unicidad. Por tanto, G es el grafo de la función $f : \mathbb{Z} \to \mathbb{Z}$ dada por $f(n) = 2n + 1$.

(b) Si $f(n) = f(k)$, se cumple $2n+1 = 2k+1$. Restando 1 y dividiendo entre 2 obtenemos $n = k$. Las entradas eran arbitrarias, luego f es inyectiva.

(c) El entero 0 pertenece al codominio, pero no es alcanzado. Si existiera $n \in \mathbb{Z}$ con $2n + 1 = 0$, tendríamos $n = -\frac{1}{2}$, que no es entero. Así, f no es sobreyectiva hacia $\mathbb{Z}$; en realidad sus valores son enteros impares.

(d) Por definición,

$$\begin{aligned} f^{-1}(\{1, 3\}) &= \{n \in \mathbb{Z} : 2n + 1 = 1 \text{ o } 2n + 1 = 3\} \\ &= \{0, 1\}. \end{aligned}$$

La respuesta es un **conjunto de entradas**, no una función $\mathbb{Z} \to \mathbb{Z}$ inversa. De hecho, tal función inversa no existe para los dominios declarados, pues 0 no tiene antecedente.

Solución 12. Recuperar sólo los valores alcanzados. Método elegido: reducir la igualdad a una inclusión usando una identidad ya demostrada. Por §4.6,

$$f(f^{-1}(T)) = T \cap \operatorname{Im}(f).$$

Por tanto,

$$\begin{aligned} f(f^{-1}(T)) = T &\iff T \cap \operatorname{Im}(f) = T \\ &\iff T \subseteq \operatorname{Im}(f). \end{aligned}$$

Justifiquemos también la segunda equivalencia: si la intersección es T, todo $t \in T$ pertenece a la imagen; si T está incluido en la imagen, intersectarlo con ella no elimina ningún elemento. Para $T = \varnothing$, ambas condiciones siguen siendo verdaderas: $f^{-1}(\varnothing) = \varnothing$, su imagen es vacía y $\varnothing \subseteq \operatorname{Im}(f)$, aunque el dominio de f sea vacío.

En la función cuadrado de §4.6, $A = \{-2, -1, 0, 1, 2\}$, $B = \{0, 1, 4, 9\}$ y $\operatorname{Im}(f) = \{0, 1, 4\}$. Para $T = \{9\}$, ninguna entrada tiene cuadrado 9; entonces

$$f^{-1}(T) = \varnothing, \qquad f(f^{-1}(T)) = \varnothing \neq \{9\} = T.$$

Lo que falla es exactamente $T \subseteq \mathrm{Im}(f)$. Pertenecer al **codominio** no equivale a ser un valor **alcanzado**.

Solución 13. Fibras y propiedades de una función. Método elegido: en la primera equivalencia comparar entradas de una misma fibra; en la segunda separar existencia y unicidad para cada valor. Recordemos que

$$F_b = \{x \in A : f(x) = b\}.$$

Si f es inyectiva y $x, x' \in F_b$, entonces $f(x) = b = f(x')$ y por inyectividad $x = x'$. Recíprocamente, supongamos que para todo $b \in B$ cualesquiera $x, x' \in F_b$ son iguales. Dados $x, x' \in A$ con $f(x) = f(x')$, fijamos $b = f(x) \in B$. Ambas entradas pertenecen a F_b, por lo que $x = x'$. Ésta es la inyectividad. **No** exigimos que cada F_b tenga un elemento: la condición sólo dice que contiene *a lo sumo uno*.

Para la segunda equivalencia, supongamos primero que f es biyectiva. Por sobreyectividad, para cada $b \in B$ existe $x \in A$ con $f(x) = b$; si x' tiene el mismo valor, la inyectividad da $x' = x$. Existe exactamente una entrada. Recíprocamente, si cada $b \in B$ posee exactamente una entrada, la parte de existencia demuestra sobreyectividad. Para probar inyectividad, dadas $x, x' \in A$ con igual valor b, la unicidad de la entrada que produce b obliga a $x = x'$. Se verifican las dos propiedades, luego f es biyectiva.

Caso vacío. Si $A = \varnothing$ y $B \neq \varnothing$, existe la función vacía $f : A \to B$: es inyectiva porque no hay dos entradas que la contradigan, y cada fibra es vacía, de modo que la primera equivalencia se cumple. No es sobreyectiva ni biyectiva: para un $b \in B$ no existe entrada alguna, y el lado derecho de la segunda equivalencia también es falso. Si $A = B = \varnothing$, la única función vacía es inyectiva y sobreyectiva (ambas condiciones son universales sobre vacíos), por lo que es biyectiva. La segunda condición, universal sobre $B = \varnothing$, también es verdadera. Si $B = \varnothing$ pero $A \neq \varnothing$, **no existe** una función $A \to B$; no hay un caso adicional de las equivalencias que evaluar.

Solución 14. El dominio efectivo de una intersección de relaciones. Método elegido: probar la inclusión con un testigo común y refutar su recíproca separando testigos. Sea $x \in \mathrm{dom}(R \cap S)$. Existe un $y \in B$ tal que $(x, y) \in R \cap S$. En particular, $(x, y) \in R$ y $(x, y) \in S$. El mismo y demuestra $x \in \mathrm{dom}(R)$ y $x \in \mathrm{dom}(S)$, de donde

$$\mathrm{dom}(R \cap S) \subseteq \mathrm{dom}(R) \cap \mathrm{dom}(S).$$

La inclusión inversa puede fallar. Tomemos $A = \{a\}$, $B = \{0, 1\}$ y

$$R = \{(a, 0)\}, \qquad S = \{(a, 1)\}.$$

Ambas son relaciones de A en B y sus dominios efectivos son $\{a\}$; sin embargo, $R \cap S = \varnothing$, por lo que $\mathrm{dom}(R \cap S) = \varnothing$. Así, a pertenece a la intersección de los dominios, pero no al dominio de la intersección.

El paso ilegítimo sería concluir, de $\exists y \in B\ R(x, y)$ y $\exists z \in B\ S(x, z)$, que existe **un mismo** y para el que se cumplen ambas relaciones. En nuestro ejemplo, los testigos son 0 y 1; no hay un par compartido.

Solución 15. ¿Determinan las fibras a la función?. Método elegido: usar la definición de preimagen en la ida y elegir la fibra del valor efectivo en la vuelta. La equivalencia es verdadera bajo la hipótesis declarada de que $f, g : A \to B$ comparten dominio y codominio.

Si $f = g$, para cualquier $b \in B$ y $x \in A$ tenemos

$$\begin{aligned} x \in f^{-1}(\{b\}) &\iff f(x) = b \\ &\iff g(x) = b \\ &\iff x \in g^{-1}(\{b\}). \end{aligned}$$

La extensionalidad da la igualdad de las fibras para cada b, lo que prueba la dirección directa.

Recíprocamente, supongamos iguales las fibras para todo $b \in B$ y fijemos $x \in A$ arbitrario. Como $f : A \to B$ es una función, $b := f(x)$ es un elemento de B. Por construcción, $x \in f^{-1}(\{b\})$; por igualdad de esa fibra con $g^{-1}(\{b\})$, resulta $x \in g^{-1}(\{b\})$. La definición de preimagen implica $g(x) = b = f(x)$. Como esto vale para toda entrada y los datos declarados de dominio y codominio coinciden, concluimos $f = g$.

Si $A = \varnothing$, ambas funciones tienen grafo vacío y, dado que comparten codominio, son iguales según nuestra convención. Para todo $b \in B$, sus fibras son igualmente vacías; el razonamiento no requiere elegir una entrada inexistente. Si los codominios declarados fueran diferentes, la conclusión **«las funciones son iguales»** ya no seguiría de la sola coincidencia de los valores o de los grafos: el libro exige también igualdad de codominio. Además, la fórmula propuesta cuantifica sobre un **mismo** B y no puede trasladarse sin especificar previamente un conjunto de valores común.

Después de las soluciones. Los quince problemas vuelven a una misma disciplina: no pasar de una pertenencia particular a una afirmación universal, distinguir testigos que pueden ser diferentes, conservar las posiciones de los pares y declarar los dominios de las condiciones. En el capítulo 5 estudiaremos cómo convertir las definiciones en herramientas de prueba. El trabajo realizado aquí proporciona el punto de partida: desplegar una definición, reconocer las obligaciones que impone y justificar cada paso sin perder las hipótesis.

Capítulo 5 — ¿Cómo se convierten las definiciones en herramientas de prueba?

En el capítulo anterior, afirmar que una relación era una función nos obligaba a comprobar existencia y unicidad de una salida para cada entrada. No bastaba reconocer un dibujo familiar. La definición decía exactamente qué información debíamos tener y qué podía faltar.

Ahora estudiaremos ese mecanismo como una herramienta general. Ante una palabra nueva —«estable», «invertible», «cuerpo»— necesitamos aprender dos movimientos: extraer consecuencias de que un objeto satisfaga una definición y averiguar qué habría que demostrar para que la satisfaga. Una misma definición interviene de modos distintos según esté entre los datos o en la conclusión.

La pregunta de este capítulo será **¿cómo pasar de un nombre matemático a condiciones con las que podamos razonar?** Comenzaremos con conjuntos y funciones ya conocidos. Después leeremos una definición axiomática completa: la de cuerpo. Ese laboratorio mostrará por qué algunas reglas algebraicas se aceptan como punto de partida y otras necesitan prueba, aunque todas resulten familiares.

Usaremos como conocimientos de entrada la aritmética escolar de los enteros y racionales y las leyes usuales de los números reales. Cuando trabajemos con un cuerpo arbitrario, cada paso se justificará mediante sus axiomas o resultados previamente demostrados. No construiremos aquí los sistemas numéricos desde primeros principios.

5.1. Una definición tiene dos direcciones de uso

Sea A un conjunto y sea $f : A \to A$ una función. Diremos que un subconjunto $S \subseteq A$ es **estable bajo** f cuando

$$\forall x \in S \ (f(x) \in S).$$

Intuitivamente, aplicar f a una entrada de S no nos saca de S. Hemos fijado un término para una condición precisa. No hemos dicho todavía qué subconjuntos la cumplen.

Usar la definición como hipótesis. Si sabemos que S es estable y disponemos de $a \in S$, podemos concluir $f(a) \in S$. El dato general proporciona información sobre una entrada particular. Pero no permite concluir $f(a) = a$: permanecer en un conjunto no significa permanecer en el mismo punto.

Usarla como objetivo. Si queremos demostrar que S es estable, debemos comenzar con un $x \in S$ arbitrario y justificar $f(x) \in S$. Comprobarlo únicamente para un elemento especial de S no prueba la condición universal.

Por ejemplo, con $A = \mathbb{Z}$, $f(n) = n + 2$ y

$$S = \{n \in \mathbb{Z} : n \geq 3\},$$

tomamos $n \in S$. Entonces $n \geq 3$, de modo que $f(n) = n + 2 \geq 5 \geq 3$; además, $n + 2$ es entero. Por definición de S, $f(n) \in S$. Como no fijamos ningún valor especial de n, queda probada la estabilidad.

La estructura del argumento fue:

$$\begin{aligned} n \in S &\Longrightarrow n \geq 3 \\ &\Longrightarrow f(n) \geq 3 \\ &\Longrightarrow f(n) \in S. \end{aligned}$$

Desplegamos la definición de S al entrar y la recuperamos al salir. La definición de estabilidad organiza el argumento completo.

Hipótesis y objetivo

Una definición del tipo «un objeto admisible x tiene la propiedad P cuando cumple $Q(x)$» fija una equivalencia entre $P(x)$ y $Q(x)$ en ese dominio. Desde la hipótesis $P(x)$ podemos obtener $Q(x)$; para alcanzar el objetivo $P(x)$ basta probar $Q(x)$.

Esta lectura de una **definición** no convierte en bicondicional cualquier «si» de un teorema. En un resultado, cada dirección requiere la prueba correspondiente.

Comprobación. Si S es estable bajo f y $a \in S$, ¿podemos afirmar que $f(f(a)) \in S$? ¿Debemos suponer que $f(a) = a$?

Respuesta razonada. La primera aplicación de la definición da $f(a) \in S$. Aplicándola otra vez a esa entrada, obtenemos $f(f(a)) \in S$. No hemos identificado las entradas: sólo hemos utilizado dos veces una propiedad universal. La expresión compuesta tiene sentido porque $f : A \to A$.

5.2. Ejemplos, no-ejemplos y objetos que quizá no existan

Una definición que reúne varias condiciones exige cumplirlas **todas**. Para producir un no-ejemplo basta incumplir una, pero debemos señalar cuál. Ante una condición cuantificada, conviene recuperar el dominio y la forma de negarla.

Para la estabilidad bajo $f : A \to A$, los conjuntos $\varnothing$ y A siempre cumplen la definición. En el primer caso no hay entradas que puedan violarla; en el segundo, que $f(x) \in A$ forma parte de declarar $f : A \to A$. Son ejemplos útiles porque impiden añadir silenciosamente la exigencia de que S sea no vacío o propio.

Con $f(n) = n+2$ sobre $\mathbb{Z}$, el subconjunto $T = \{3, 4\}$ no es estable: $3 \in T$, pero $f(3) = 5 \notin T$. Este solo testigo refuta la universal. Por eso, para un $S \subseteq A$ ya fijado,

$$\begin{gathered} S \text{ no es estable bajo } f \\ \Longleftrightarrow \\ \exists x \in S \ (f(x) \notin S). \end{gathered}$$

No sería una negación correcta exigir que **todas** las entradas salieran de S. Una universal falla en cuanto aparece una excepción.

Hay otra precaución: **dar un nombre no demuestra existencia**. Podemos llamar «entero dominante» a un entero d que cumpla $n \leq d$ para todo $n \in \mathbb{Z}$. La condición es precisa, pero ningún entero la satisface: si d fuera candidato, el entero $d+1$ sería mayor que él. Definir una clase de objetos puede dar una clase vacía.

El problema es aún más delicado si intentamos introducir **un objeto determinado** mediante una descripción. La frase «sea r la raíz real de $r^2 = 1$» no selecciona un único real: tanto 1 como -1 cumplen la condición. La frase «sea r la raíz real positiva de $r^2 = 1$» sí lo hace. La existencia la proporciona 1; si $r > 0$ y $r^2 = 1$, la factorización $(r-1)(r+1) = 0$ en los reales, con $r + 1 > 0$, obliga a $r = 1$. Aquí utilizamos aritmética real conocida; más adelante derivaremos el principio del producto nulo desde axiomas abstractos.

Comprobación. Para $f : A \to A$ y $a \in A$, ¿cuándo es estable el singleton $\{a\}$?

Respuesta razonada. Su única entrada es a. La definición exige $f(a) \in \{a\}$, que equivale a $f(a) = a$. Así, un singleton es estable exactamente cuando su elemento es un punto fijo, es

decir, una entrada que la función deja sin cambio. Esto no significa que todos los elementos de un conjunto estable deban ser puntos fijos.

5.3. Definiciones, resultados e hipótesis de trabajo

Una **definición** fija significado. Un **teorema** afirma un resultado y exige justificación desde los puntos de partida disponibles. Un **lema** es un resultado utilizado como apoyo para otro; un **corolario**, una consecuencia obtenida a partir de un resultado anterior. Esas etiquetas describen el papel de una afirmación en la exposición, no grados diferentes de verdad o de rigor.

La palabra **proposición** tiene además dos usos que conviene distinguir. En los capítulos iniciales nombraba un enunciado interpretado con valor de verdad. Como encabezado de una exposición —«Proposición»— suele anunciar un resultado demostrado, igual que «Teorema», con una diferencia editorial de énfasis. El contexto permite saber cuál de los dos usos está en juego.

Veamos cómo una definición se convierte en una pequeña cadena de resultados.

Lema. Si $S, T \subseteq A$ son estables bajo $f : A \to A$, entonces $S \cap T$ es estable bajo f.

Estrategia. El objetivo es universal: partiremos de una entrada arbitraria de la intersección. Su pertenencia nos permitirá utilizar las dos hipótesis de estabilidad.

Demostración. Sea $x \in S \cap T$. Entonces $x \in S$ y $x \in T$. Como cada conjunto es estable, $f(x) \in S$ y $f(x) \in T$. Por definición de intersección, $f(x) \in S \cap T$. Esto vale para cualquier entrada de la intersección, por lo que ésta es estable. Si la intersección es vacía, la condición universal también se cumple. □

Corolario. Si S, T y U son estables bajo la misma función, entonces $(S \cap T) \cap U$ es estable.

Demostración. El lema da primero la estabilidad de $S \cap T$. Podemos aplicarlo otra vez a ese conjunto y a U. □

La segunda prueba no necesita volver a abrir toda la definición: usa un resultado cuya demostración ya está disponible. Una prueba puede apoyarse en definiciones, hipótesis y teoremas, pero debe conservar el origen de cada paso.

Las hipótesis «misma función» y «subconjuntos de A» no son decoración. Si cambiamos de función durante el argumento, la segunda condición de estabilidad podría no informar sobre el valor $f(x)$ que necesitamos. Y si una expresión no está definida para la entrada elegida, ni siquiera hemos construido una afirmación interpretable.

Comprobación. En la prueba del lema, ¿por qué no bastaría saber únicamente que S es estable?

Respuesta razonada. De $x \in S \cap T$ y la estabilidad de S sólo obtenemos $f(x) \in S$. Todavía falta $f(x) \in T$. Por ejemplo, con $A = S = \mathbb{Z}$, $f(n) = n + 1$ y $T = \{0\}$, la intersección es $\{0\}$ y no es estable. Se satisface la única hipótesis de estabilidad que conservamos, pero falla la conclusión.

5.4. Reformular una definición exige conservar su alcance

La condición de estabilidad puede escribirse de otra manera:

$$S \text{ es estable bajo } f \quad \Longleftrightarrow \quad f(S) \subseteq S.$$

Demostración. Supongamos primero que S es estable. Si $y \in f(S)$, por definición de imagen existe $x \in S$ con $y = f(x)$. La estabilidad da $y \in S$; por tanto, $f(S) \subseteq S$. Recíprocamente, supongamos esa inclusión y tomemos $x \in S$. Su valor pertenece a $f(S)$ y, por la inclusión, a S. Ésta es la condición de estabilidad. □

La equivalencia conecta dos escrituras ya definidas. Podemos elegir la que haga más transparente una prueba, porque conocemos cómo pasar en ambas direcciones.

En cambio, reemplazar la inclusión por $f(S) = S$ cambia la condición. La igualdad exige además que cada elemento de S sea alcanzado desde una entrada de S. Con $f(n) = n + 1$ sobre los enteros y

$$S = \{n \in \mathbb{Z} : n \geq 0\},$$

tenemos $f(S) = \{n \in \mathbb{Z} : n \geq 1\} \subsetneq S$. El conjunto es estable, pero $0 \in S$ no se alcanza desde una entrada no negativa. La nueva condición sería más fuerte.

Para comprobar una reformulación debemos mirar tanto los símbolos como los cuantificadores. La condición $\forall x \in S\ f(x) \in S$ permite que distintos elementos tengan imágenes distintas; no equivale a $\exists y \in S\, \forall x \in S\ f(x) = y$, que pide un valor común y además exige que S sea no vacío. El conjunto vacío separa inmediatamente ambas frases.

Comprobación. Si $f(S) = S$, ¿es estable S? ¿Queda justificada la recíproca?

Respuesta razonada. La igualdad implica $f(S) \subseteq S$, así que sí hay estabilidad. La recíproca es falsa por el ejemplo anterior. Una dirección válida no basta para intercambiar definiciones; se necesitan ambas.

5.5. Los axiomas de cuerpo: definir una estructura mediante leyes

Una operación binaria sobre F es una función $F \times F \to F$. Declarar su dominio y codominio significa que a cada par de elementos de F le corresponde un único elemento de F. Como vimos en C04, eso es más información que escribir un símbolo entre dos letras.

Un control previo: el cierre de las operaciones.

No basta que una fórmula pueda calcularse en un conjunto más grande para que describa una **operación binaria sobre** F. Primero hay que comprobar que el resultado permanezca en F para *todos* los pares de entradas de $F \times F$. Ésta es la condición de **cierre**. Sólo después corresponde examinar leyes como la asociatividad o la existencia de neutros e inversos.

Consideremos $F = \mathbb{Q} \setminus \{0\}$ y tratemos de usar la suma racional ordinaria como operación $F \times F \to F$. Aunque $1, -1 \in F$, se cumple

$$1 + (-1) = 0 \notin F.$$

La suma es una operación legítima sobre $\mathbb{Q}$, pero su restricción a estas entradas **no** es una operación binaria *sobre* F, porque hay parejas cuyo resultado sale de ese conjunto. La familiaridad de la fórmula no sustituye la verificación de su codominio. De hecho, aquí ni siquiera hemos llegado al problema de si existirían inversos: ha fallado un dato anterior de la estructura propuesta.

Comprobación. Si $E = \{0, 1\}$ y usamos la suma y el producto habituales de enteros, ¿tenemos dos operaciones binarias $E \times E \to E$?

Respuesta razonada. No. El producto sí tiene sus valores en E: los únicos productos posibles son 0 y 1. En cambio, $1 + 1 = 2 \notin E$. La suma ordinaria no está cerrada en E; modificarla, por ejemplo definiendo otra suma, sería estudiar una estructura distinta y obligaría a declarar la nueva regla antes de comprobar sus axiomas.

Un **cuerpo** será un conjunto F con operaciones $+$ y $\cdot$ de $F \times F$ en F y elementos distinguidos $0, 1 \in F$, que satisfacen las leyes siguientes. Escribiremos ab para $a \cdot b$.

Definición de cuerpo

Para todos $a, b, c \in F$:

Leyes aditivas

$$(a + b) + c = a + (b + c), \quad a + b = b + a,$$
$$a + 0 = a.$$

Para cada $a \in F$ existe algún $u \in F$ tal que $a + u = 0$.

Leyes multiplicativas

$$(ab)c = a(bc), \quad ab = ba,$$
$$a1 = a.$$

Para cada $a \in F$ **distinto de** 0 existe algún $v \in F$ tal que $av = 1$.

Compatibilidad y no trivialidad

$$a(b + c) = ab + ac, \qquad 0 \neq 1.$$

La conmutatividad proporciona también $0 + a = a$ y $1a = a$. Permite obtener la distributividad por el otro lado:

$$(b + c)a = a(b + c) = ab + ac = ba + ca.$$

En esta presentación esa segunda distributividad es una consecuencia breve; podríamos haber elegido una lista equivalente que la incluyera explícitamente. Lo esencial es declarar qué lista utilizamos y justificar las equivalencias cuando importan.

Los inversos se postulan mediante **existencia**, no mediante existencia única. Todavía no escribiremos $-a$ ni a^{-1} como si la unicidad estuviera probada: compararemos candidatos en la sección siguiente. La hipótesis $a \neq 0$ pertenece sólo al axioma de inverso multiplicativo; todo elemento, incluido 0, tiene inverso aditivo.

Llamamos **axiomas de cuerpo** a estas condiciones porque son puntos de partida cuando trabajamos con un cuerpo arbitrario. Pero, si nos dan operaciones concretas y queremos demostrar que forman un cuerpo, esos mismos axiomas se convierten en obligaciones que hay que verificar. El papel lógico depende de la tarea: podemos suponer que estamos en una estructura de cierta clase o demostrar que un candidato pertenece a ella.

No aparecen en la lista ni un orden $<$ ni una propiedad de completitud. Tampoco se han postulado $a0 = 0$, cancelación, reglas de signos o ausencia de productos nulos con factores no nulos. Aunque esas expresiones recuerden hechos escolares, dentro de este laboratorio deben derivarse.

Comprobación. Escribe el axioma de inverso multiplicativo con cuantificadores e identifica quién puede depender de quién.

Respuesta razonada. Su forma es

$$\forall a \in F \, \big(a \neq 0 \Longrightarrow \exists v \in F \, (av = 1)\big).$$

Primero se fija a; cuando es no nulo, el testigo v puede depender de él. No se exige un único v común para todos los a, ni se permite aplicar la existencia al caso $a = 0$.

5.6. Unicidad y cancelación: ordenar las dependencias

Antes de calcular con inversos como objetos determinados, tenemos que justificar que dos candidatos coinciden. La estrategia general será comparar los candidatos usando las propiedades que ambos cumplen. No usaremos una cancelación multiplicativa cuya prueba aún no hemos dado.

Los neutros son únicos. Supongamos que $e \in F$ también es neutro aditivo, es decir, que $a + e = a$ para todo $a \in F$. Al elegir $a = 0$ obtenemos $0 + e = 0$. Por otro lado, el neutro 0 y la conmutatividad dan $0 + e = e$. Luego $e = 0$.

Si $j \in F$ también cumple $aj = a$ para todo a, al elegir $a = 1$ obtenemos $1j = 1$. Como 1 es el neutro multiplicativo y el producto es conmutativo, $1j = j$. Por tanto, $j = 1$.

La comparación sólo necesitó las propiedades de los candidatos y la posibilidad de usarlos como entradas. No necesitó dividir ni restar.

Los inversos aditivos son únicos. Fijemos $a \in F$ y supongamos $a+u=0$ y $a+v=0$. Entonces

$$\begin{aligned} u &= u+0 \\ &= u+(a+v) \\ &= (u+a)+v \\ &= 0+v=v. \end{aligned}$$

Usamos neutro, la ecuación de v, asociatividad y la ecuación de u junto con conmutatividad. Existencia y unicidad justifican escribir $-a$ para ese elemento. Definimos también $b-a := b+(-a)$.

Ahora podemos demostrar **cancelación aditiva**. Si $a+b=a+c$, sumamos $-a$ a ambos lados y reagrupamos:

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

y del otro lado obtenemos c. Así, $b=c$. La igualdad se preserva al aplicar la misma operación a ambos miembros; la asociatividad explica por qué podemos eliminar el término común. La recíproca, $b=c \Rightarrow a+b=a+c$, se sigue de sustituir iguales.

Los inversos multiplicativos son únicos. Sea $a \neq 0$. Si $au=1$ y $av=1$, tenemos

$$u = u1 = u(av) = (ua)v = 1v = v.$$

La demostración usa asociatividad, conmutatividad, neutro y las ecuaciones de los candidatos. Como el axioma garantiza su existencia para $a \neq 0$, podemos denotar el único inverso por a^{-1}. En este libro no asignamos por esta notación un inverso multiplicativo a 0.

Podemos probar ahora **cancelación multiplicativa con factor no nulo**. Si $a \neq 0$ y $ab=ac$, multiplicamos ambos lados por a^{-1}:

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

y análogamente obtenemos c en el otro miembro. Luego $b=c$. La hipótesis $a \neq 0$ se utilizó al disponer de a^{-1}; no puede borrarse al final.

Lectura de la prueba. No hemos justificado la unicidad de los inversos mediante cancelación multiplicativa: hicimos primero la comparación directa de candidatos y después dedujimos la cancelación. Éste es el orden de dependencias que evita un argumento circular.

Comprobación. ¿Por qué sería insuficiente demostrar $au=av=1$ y escribir a continuación «cancelamos a, luego $u=v$» si estuviéramos usando esa unicidad para introducir y justificar la propia regla de cancelación?

Respuesta razonada. Estaríamos utilizando como herramienta el resultado cuya justificación depende de lo que intentamos probar. La cadena $u=u(av)=(ua)v=v$ evita ese círculo y sólo usa axiomas ya disponibles. Si la cancelación se hubiera demostrado antes por otra vía independiente, sí podría usarse; hay que conocer sus dependencias, no prohibirla por su nombre.

5.7. El cero y los signos se deducen

Absorción del cero y prohibición de invertirlo. Fijemos $a \in F$ y escribamos $t=a0$. Como $0+0=0$, la distributividad da

$$t = a0 = a(0+0) = a0+a0 = t+t.$$

Podemos escribir el miembro izquierdo como $t+0$. La cancelación **aditiva**, ya demostrada, aplicada a $t+0=t+t$, da $0=t$. Por tanto, $a0=0$; por conmutatividad, también $0a=0$.

Esto prueba que 0 no posee inverso multiplicativo. Si algún v satisficiera $0v=1$, la absorción del cero daría simultáneamente $0v=0$, de donde $0=1$, contra la no trivialidad. No hemos dividido por cero para concluir que no se puede hacerlo.

Opuesto del opuesto y producto con un signo. La igualdad $a+(-a)=0$ y la conmutatividad muestran que a es un inverso aditivo de $-a$. Por unicidad,

$$-(-a)=a.$$

De igual manera, $-0=0$, porque $0+0=0$.

Para identificar $(-a)b$, buscaremos una propiedad que lo caracterice. La distributividad y la absorción del cero dan

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

Así, $(-a)b$ es un inverso aditivo de ab. Por la unicidad establecida,

$$(-a)b=-(ab).$$

La conmutatividad permite obtener también $a(-b)=-(ab)$. En particular, con $a=1$ en la primera identidad,

$$(-1)b=-(1b)=-b.$$

No dijimos que $(-1)b=-b$ fuera cierto «por la regla de los signos»: acabamos de justificar esa regla desde leyes anteriores.

Dos signos y una conclusión general. Aplicando las identidades ya probadas en el orden indicado,

$$(-a)(-b)=-(a(-b))=-(-(ab))=ab.$$

La primera igualdad usa la regla para un signo en el primer factor; la segunda, la regla para el segundo; la tercera, el opuesto del opuesto. Como ahora podemos localizar cada herramienta, la cadena no oculta una dependencia circular.

También obtenemos el **principio del producto nulo**:

$$ab=0 \quad\Longleftrightarrow\quad a=0 \text{ o } b=0.$$

Si $ab=0$, distinguimos dos casos. Si $a=0$, ya tenemos una alternativa. Si $a\neq 0$, la igualdad $ab=0=a0$ permite cancelar a y concluir $b=0$. Recíprocamente, cualquiera de las dos igualdades $a=0$ o $b=0$ hace nulo el producto por absorción. La distinción de casos permite usar un resultado condicionado sin suponer gratuitamente su hipótesis.

Comprobación. ¿Se puede cancelar el factor común 0 en $0\cdot 0=0\cdot 1$?

Respuesta razonada. Ambos miembros son 0, pero cancelar produciría $0=1$, que contradice un axioma. La regla sólo vale con factor común no nulo. Este ejemplo satisface la igualdad inicial y muestra por qué conservar la condición es indispensable.

5.8. Un cuerpo no incluye automáticamente un orden

Modelos conocidos y un no-ejemplo. Con las operaciones usuales, $\mathbb{Q}$ y $\mathbb{R}$ son cuerpos. En esta etapa tomamos sus leyes aritméticas elementales como conocimiento de entrada, no como resultados construidos desde teoría de conjuntos. Para los racionales, el inverso de un elemento no nulo p/q, con $p, q \in \mathbb{Z}$, $q \neq 0$ y $p \neq 0$, es q/p; pertenece de nuevo a $\mathbb{Q}$ y su producto con p/q es 1. Para los reales se dispone análogamente del recíproco de todo elemento no nulo.

Los enteros, con sus operaciones usuales, no forman un cuerpo. El elemento $2 \neq 0$ no tiene inverso multiplicativo entero. Si $n \leq 0$, entonces $2n \leq 0$; si $n \geq 1$, entonces $2n \geq 2$. En ninguno de los dos casos se obtiene $2n = 1$. Un solo incumplimiento de un axioma basta para descartar el candidato; no hace falta que todos fallen.

Los complejos, presentados mediante pares. Podemos reconocer otro modelo sin suponer un orden para sus elementos. Consideremos $\mathbb{R}^2$ con

$$(a, b) + (c, d) := (a + c, b + d),$$

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

y constantes $0 = (0, 0)$ y $1 = (1, 0)$. A este sistema lo llamamos $\mathbb{C}$.

Las operaciones producen pares de reales. La suma es asociativa y conmutativa componente a componente, su neutro es $(0, 0)$ y el opuesto de (a, b) es $(-a, -b)$. El producto es conmutativo porque las expresiones $ac - bd$ y $ad + bc$ no cambian al intercambiar los pares, usando conmutatividad real; su neutro es $(1, 0)$.

Para la asociatividad multiplicativa, si $z = (a, b)$, $w = (c, d)$ y $u = (e, h)$, expandir ambos productos $(zw)u$ y $z(wu)$ da el mismo par (P, Q), donde

$$P = ace - bde - adh - bch,$$
$$Q = ach - bdh + ade + bce.$$

La distributividad se comprueba igualando componentes: tanto $z(w + u)$ como $zw + zu$ tienen primera componente $ac + ae - bd - bh$ y segunda $ad + ah + bc + be$. Finalmente, $(0, 0) \neq (1, 0)$ porque $0 \neq 1$ en los reales.

Si $(a, b) \neq (0, 0)$, el real $a^2 + b^2$ es positivo: los cuadrados reales son no negativos y al menos uno es positivo. En particular, es no nulo y tiene recíproco real. El par

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

es un inverso multiplicativo: la primera componente del producto es $(a^2 + b^2)/(a^2 + b^2) = 1$ y la segunda es $(-ab + ba)/(a^2 + b^2) = 0$. Hemos verificado la última condición que faltaba. Por tanto, estas operaciones forman un cuerpo.

Aquí el orden de los **reales** ayudó a verificar que un denominador real era no nulo. No hemos introducido por eso un orden sobre el cuerpo de pares.

Qué habría que añadir para hablar de cuerpo ordenado. Un **cuerpo ordenado** es un cuerpo F provisto además de una relación $<$ que sea un orden estricto total y compatible con las operaciones. Por orden estricto total entendemos transitividad y que, para cualesquiera a, b, se cumpla exactamente una de las alternativas $a < b$, $a = b$ o $b < a$. La compatibilidad exige:

$$a < b \Longrightarrow a + c < b + c,$$

para todos a, b, c, y

$$0 < a \text{ y } 0 < b \Longrightarrow 0 < ab.$$

Estas condiciones son información adicional. En un cuerpo ordenado, el cuadrado de un elemento no nulo es positivo. Si $a > 0$, lo da la compatibilidad multiplicativa. Si $a < 0$, sumando $-a$ obtenemos $0 < -a$, por lo que $0 < (-a)^2 = a^2$. La tricotomía cubre ambos casos cuando $a \neq 0$. En particular, $0 < 1^2 = 1$ y, sumando -1, resulta $-1 < 0$.

En $\mathbb{C}$, el elemento $i = (0, 1)$ cumple $i^2 = (-1, 0) = -1$ y es no nulo. Si existiera un orden compatible con la estructura de cuerpo, su cuadrado sería positivo. Pero acabamos de demostrar que -1 sería negativo. La tricotomía impide ambas cosas. Por tanto, $\mathbb{C}$ **es un cuerpo que no admite un orden de cuerpo**. Esto no prohíbe ordenar sus pares de alguna manera; prohíbe que ese orden satisfaga simultáneamente las compatibilidades indicadas.

Comprobación. ¿De la hipótesis «F es un cuerpo» podemos concluir $a^2 > 0$ para todo $a \neq 0$?

Respuesta razonada. No: esa hipótesis ni siquiera declara una relación $>$ sobre F. La conclusión es válida en un cuerpo **ordenado**, con el orden y las compatibilidades anteriores. Los complejos muestran además que no basta intentar añadir cualquier orden a todo cuerpo. Debemos verificar qué estructura está disponible antes de formular una desigualdad.

5.9. Ejercicios: leer una condición y decidir qué hacer con ella

Antes de comenzar, escribe qué datos están fijados y qué definición permite interpretar la conclusión. En las pruebas sobre un cuerpo arbitrario puedes usar los resultados de §§5.6–5.7, indicando dónde intervienen.

Familia I. Traducción y control de definiciones. Ejercicio 1. Negar sin exigir demasiado. Sean $f : A \to A$ y $S \subseteq A$. Escribe con cuantificadores la estabilidad de S y su negación. Describe qué debe proporcionarse para probar cada una. Explica qué sucede cuando $S = \varnothing$.

Ejercicio 2. Dos órdenes de cuantificadores. Compara las afirmaciones

$$\forall a \in F \, \big(a \neq 0 \Longrightarrow \exists v \in F \, (av = 1)\big)$$

y

$$\exists v \in F \, \forall a \in F \, (a \neq 0 \Longrightarrow av = 1).$$

Explica cuál aparece en la definición de cuerpo. Decide si la segunda vale en $\mathbb{Q}$, comprobando las entradas $a = 1$ y $a = 2$.

Ejercicio 3. ¿Dato o resultado? Con la lista exacta de §5.5, clasifica las siguientes afirmaciones: (a) $a(b + c) = ab + ac$; (b) todo a tiene un único inverso aditivo; (c) $a0 = 0$; (d) $0 \neq 1$; (e) $ab = ac \Rightarrow b = c$. Para las que no sean axiomas, señala un resultado previo que las justifique o una hipótesis que falte.

Familia II. Reparación de argumentos. Ejercicio 4. Una cancelación indiscriminada. Se escribe: «En un cuerpo, si $ax = ay$, multiplicamos por a^{-1} y obtenemos $x = y$». Localiza la omisión, produce un contraejemplo a la afirmación sin condiciones y escribe una versión correcta con su prueba.

Ejercicio 5. Unicidad supuesta. Para probar que el inverso aditivo es único, alguien escribe: «Si $a+u = 0$ y $a+v = 0$, entonces $u = -a = v$, porque $-a$ es el inverso de a». Explica por qué esta justificación no sirve para establecer por primera vez la unicidad. Reemplázala por una prueba que compare u y v sin usar la notación $-a$.

Ejercicio 6. Una desigualdad sin estructura. Se propone: «En cualquier cuerpo, el cuadrado de un elemento no nulo es mayor que cero; por eso $x^2 = -1$ nunca tiene solución». Identifica la hipótesis omitida y repara el enunciado. Explica qué demuestra el ejemplo de $i \in \mathbb{C}$ y qué no demuestra acerca de posibles órdenes sobre un conjunto de pares.

Familia III. Ejemplos, no-ejemplos y existencia. Ejercicio 7. Tres comportamientos de una definición. Para $A = \mathbb{Z}$ y $f(n) = 0$, encuentra: (a) un subconjunto no vacío S con $f(S) = S$; (b) un subconjunto estable T con $f(T) \subsetneq T$; (c) un subconjunto no estable U. Justifica las tres decisiones. ¿El conjunto vacío es estable? ¿Cumple $f(\varnothing) = \varnothing$?

Ejercicio 8. Nombrar no proporciona un testigo. Se llama «real excepcional» a un real r que cumpla $r^2 + 1 = 0$. ¿Existe alguno? Justifica la respuesta y distingue entre definir una propiedad y afirmar que hay un objeto que la satisface. ¿Qué cambia si se usa la misma ecuación sobre $\mathbb{C}$?

Ejercicio 9. ¿Por qué pedir $0 \neq 1$? En el conjunto $E = \{e\}$ define $e + e = e$ y $ee = e$, y toma como constantes $0_E = 1_E = e$. Comprueba todas las condiciones de §5.5 salvo $0 \neq 1$. Examina expresamente el axioma de inverso multiplicativo, en vez de decir que «es obvio». ¿Es un cuerpo bajo nuestra definición?

Familia IV. Demostraciones guiadas. Ejercicio 10. Unir conjuntos estables. Demuestra que si $S, T \subseteq A$ son estables bajo $f : A \to A$, entonces $S \cup T$ es estable. Guía: toma una entrada arbitraria de la unión, considera las alternativas de su pertenencia y aplica en cada caso la hipótesis adecuada. ¿Necesitas que los conjuntos sean disjuntos?

Ejercicio 11. Resolver y verificar. Sean F un cuerpo, $a, b \in F$ y $a \neq 0$. Demuestra que la ecuación $ax = b$ tiene una única solución en F. Guía: construye un candidato mediante el inverso, comprueba la ecuación y compara dos soluciones. Después clasifica las soluciones cuando $a = 0$, distinguiendo $b = 0$ y $b \neq 0$.

Ejercicio 12. Invertir de nuevo. Sea $a \neq 0$ en un cuerpo. Prueba primero que $a^{-1} \neq 0$, usando absorción del cero y $0 \neq 1$. Después demuestra $(a^{-1})^{-1} = a$ mediante la unicidad del inverso. Señala por qué la primera parte es necesaria para escribir la segunda expresión.

Familia V. Elección autónoma de método. Ejercicio 13. Una función sobre un cuerpo. Sean $a, b \in F$ y $a \neq 0$. Define $g : F \to F$ mediante $g(x) = ax + b$. Decide si g es biyectiva. Si lo es, construye una función inversa y verifica que recupera las entradas en ambas direcciones. No presupongas que F sea ordenado ni que sus elementos sean reales.

Ejercicio 14. El símbolo cambia, la obligación permanece. Sobre $\mathbb{Q}$ define una operación $x \star y = x + y + 1$, donde el lado derecho usa la suma racional usual. Determina un neutro bilateral para $\star$ y, para cada x, un elemento u con $x \star u = u \star x = e$, donde e es ese neutro. Comprueba también asociatividad. Explica por qué imponer $x \star u = 0$ sería una pregunta distinta.

Ejercicio 15. Cuando toda parte es estable. Sea $f : A \to A$. Demuestra la equivalencia entre: (a) todo subconjunto $S \subseteq A$ es estable bajo f; (b) $f(x) = x$ para todo $x \in A$; (c) $f(S) = S$ para todo $S \subseteq A$. Trata el caso $A = \varnothing$ y explica por qué, aunque la estabilidad de un solo conjunto no exige igualdad con su imagen, la condición universal de (a) sí fuerza (c).

5.10. Soluciones desarrolladas

Solución 1. Negar sin exigir demasiado. La estabilidad se escribe $\forall x \in S\ (f(x) \in S)$; su negación es $\exists x \in S\ (f(x) \notin S)$. Para la primera hay que justificar la pertenencia de la imagen de una entrada **arbitraria**; para la segunda basta exhibir una entrada de S cuya imagen quede fuera y comprobar ambas pertenencias.

Si $S = \varnothing$, la universal es verdadera y la existencial falsa: no existe una entrada que pueda actuar como contraejemplo. No necesitamos construir $f(x)$ para un supuesto elemento del vacío. El método lo determina la forma lógica de la condición.

Solución 2. Dos órdenes de cuantificadores. La primera afirmación es el axioma: permite elegir el inverso después de conocer a. La segunda exige un mismo v que sirva para todos los elementos no nulos. Si valiera en $\mathbb{Q}$, al sustituir $a = 1$ tendríamos $v = 1$; al sustituir $a = 2$, tendríamos $2v = 1$, incompatible con $v = 1$ porque $2 \neq 1$ en $\mathbb{Q}$. Por tanto, allí es falsa.

La estrategia fue confrontar dos instancias del supuesto testigo uniforme. No afirmamos que la segunda fórmula sea falsa en absolutamente toda estructura; el ejemplo racional basta para mostrar que no es una exigencia de la definición general de cuerpo.

Solución 3. ¿Dato o resultado?.

(a) es el axioma de distributividad y (d) el de no trivialidad. (b) es un resultado: la existencia está en la lista, pero la unicidad se demostró comparando candidatos en §5.6. (c) se dedujo en §5.7 mediante distributividad y cancelación aditiva.

(e), tal como está, no es válida para todos los factores: falta $a \neq 0$. La versión con esa hipótesis es la cancelación multiplicativa de §5.6. Con $a = 0$, $b = 0$ y $c = 1$, la igualdad inicial es verdadera y la conclusión falsa. Clasificar una frase exige atender a su formulación exacta, no sólo reconocer su nombre.

Solución 4. Una cancelación indiscriminada. La existencia de a^{-1} exige $a \neq 0$. Si $a = 0$, $x = 0$ y $y = 1$, entonces $ax = ay = 0$, pero $x \neq y$.

La reparación es: «Si $a \neq 0$ y $ax = ay$, entonces $x = y$». Bajo esa hipótesis, el axioma proporciona el inverso, y

$$\begin{aligned} x &= 1x = (a^{-1}a)x = a^{-1}(ax) \\ &= a^{-1}(ay) = (a^{-1}a)y \\ &= 1y = y. \end{aligned}$$

Cada reagrupación usa asociatividad. El contraejemplo y la prueba cumplen funciones distintas: el primero detecta que la condición no puede omitirse; la segunda demuestra que la condición añadida basta.

Solución 5. Unicidad supuesta. El argumento identifica sin justificación cada candidato con un objeto ya seleccionado como «el» inverso. Eso presupone precisamente que todos los candidatos coinciden. Para establecer la unicidad por primera vez, partimos sólo de $a + u = 0$ y $a + v = 0$ y escribimos

$$\begin{aligned} u &= u + 0 = u + (a + v) \\ &= (u + a) + v = 0 + v = v. \end{aligned}$$

La conmutatividad permite usar $u + a = 0$; la asociatividad permite reagrupar. Ningún paso emplea notación que requiera la unicidad pendiente. Después de esta comparación y de la existencia axiomática sí podemos introducir $-a$.

Solución 6. Una desigualdad sin estructura. Se omitió que el cuerpo esté provisto de un orden compatible. La reparación es: «En un cuerpo ordenado, todo cuadrado no nulo es positivo; en consecuencia, $x^2 = -1$ no tiene solución».

En efecto, $0 < 1$ y por traslación $-1 < 0$. Si $x = 0$, su cuadrado es $0 \neq -1$. Si $x \neq 0$, su cuadrado es positivo por la prueba de §5.8 y no puede ser -1. Los casos cubren todos los elementos.

El complejo i satisface $i^2 = -1$, por lo que refuta la pretendida prohibición en todo cuerpo. También impide un orden **compatible con sus operaciones**. No demuestra que el conjunto de pares no pueda recibir otra relación de orden; lo que falla es la exigencia conjunta de orden y compatibilidad algebraica.

Solución 7. Tres comportamientos de una definición. Podemos elegir $S = \{0\}$, $T = \{0, 2\}$ y $U = \{2\}$. Como la función es constante cero, $f(S) = \{0\} = S$ y $f(T) = \{0\} \subsetneq T$. Esto hace estable a T, aunque no recupere todos sus elementos como valores. En cambio, $2 \in U$ y $f(2) = 0 \notin U$, de modo que U no es estable.

El vacío sí es estable y su imagen es vacía. La imagen reúne valores procedentes de entradas del conjunto; sin entradas no reúne ningún valor, aunque la fórmula global de la función sea constante. Es importante no escribir erróneamente $f(\varnothing) = \{0\}$.

Solución 8. Nombrar no proporciona un testigo. No existe tal real. Para cualquier $r \in \mathbb{R}$, tenemos $r^2 \geq 0$, luego $r^2 + 1 \geq 1 > 0$. La definición describe una propiedad que ningún real cumple; no produce un testigo por haber introducido un nombre.

En $\mathbb{C}$ sí existen ejemplos: $i^2 + 1 = 0$ y también $(-i)^2 + 1 = i^2 + 1 = 0$, usando la regla de dos signos. No necesitamos afirmar aquí que sean los únicos para responder a la pregunta de existencia. El cambio de dominio altera qué objetos pueden satisfacer la ecuación.

Solución 9. ¿Por qué pedir $0 \neq 1$?. Toda operación escrita con elementos de E produce e, así que las operaciones están bien definidas y son internas. Las dos asociaciones de una suma o producto, y ambos órdenes de sus factores, dan e. También las dos expresiones de la distributividad dan e.

Las constantes elegidas actúan como neutros porque $e + e = e$ y $ee = e$. El propio e es un inverso aditivo de la única entrada, pues $e + e = e = 0_E$. El axioma de inverso multiplicativo tiene forma condicional para las entradas distintas de 0_E: no hay ninguna, de modo que la condición universal se cumple vacíamente.

Falla exactamente la no trivialidad: $0_E = 1_E$. Por ello no es un cuerpo según nuestra definición. La conclusión se obtiene comprobando condiciones, no descartando el ejemplo por parecer demasiado pequeño.

Solución 10. Unir conjuntos estables. Sea $x \in S \cup T$. Si $x \in S$, su imagen pertenece a S por estabilidad y, por tanto, a $S \cup T$. Si $x \notin S$, la pertenencia a la unión obliga a $x \in T$; entonces $f(x) \in T \subseteq S \cup T$. En todos los casos, la imagen permanece en la unión, que es estable.

No hace falta disjunción de conjuntos. Cuando x pertenece a ambos, cualquiera de las dos hipótesis sirve; no hay conflicto. La elección de casos viene de la disyunción en la definición de unión, a diferencia de la conjunción que organizó la prueba para la intersección.

Solución 11. Resolver y verificar. Como $a \neq 0$, existe $a^{-1} \in F$. Tomamos $x = a^{-1}b$, que pertenece a F por cierre del producto. Entonces

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

El candidato es una solución. Si x e y fueran dos soluciones, tendríamos $ax = b = ay$ y la cancelación con $a \neq 0$ daría $x = y$. Esto prueba unicidad, además de existencia.

Si $a = 0$ y $b = 0$, todo $x \in F$ es solución por absorción del cero. No hay unicidad, pues al menos 0 y 1 son distintos. Si $a = 0$ y $b \neq 0$, ninguna entrada satisface $0x = b$, ya que el miembro izquierdo siempre vale 0. Clasificar primero los casos evita aplicar una fórmula que no está definida en todos ellos.

Solución 12. Invertir de nuevo. Si $a^{-1} = 0$, entonces $aa^{-1} = a0 = 0$. Pero la propiedad del inverso dice $aa^{-1} = 1$, de donde $0 = 1$, contradicción. Por tanto, $a^{-1} \neq 0$ y existe su inverso multiplicativo.

La conmutatividad da $a^{-1}a = aa^{-1} = 1$. Así, a es un candidato a inverso de a^{-1}. Por unicidad, $(a^{-1})^{-1} = a$. El primer paso legitima la expresión que aparece en el segundo; encontrar una fórmula familiar no autoriza omitir su condición de existencia.

Solución 13. Una función sobre un cuerpo. La función está bien definida porque $ax + b \in F$. Si $g(x) = g(y)$, entonces $ax + b = ay + b$; cancelando aditivamente b obtenemos $ax = ay$, y cancelando el factor no nulo a concluimos $x = y$. Por tanto, g es inyectiva.

Para $t \in F$ arbitrario, elegimos $x = a^{-1}(t - b)$. La distributividad, la asociatividad y las identidades de inversos dan

$$\begin{aligned} g(x) &= a(a^{-1}(t - b)) + b \\ &= (t - b) + b = t. \end{aligned}$$

Así, es sobreyectiva y, por ambas propiedades, biyectiva. Definimos $h : F \to F$ por $h(t) = a^{-1}(t - b)$. La igualdad anterior verifica $g(h(t)) = t$. En la otra dirección,

$$\begin{aligned} h(g(x)) &= a^{-1}((ax + b) - b) \\ &= a^{-1}(ax) = x. \end{aligned}$$

Hemos comprobado ambas recuperaciones sin usar un orden. La idea de elegir x proviene de leer la ecuación $ax + b = t$ como una obligación y deshacer las operaciones con resultados ya justificados.

Solución 14. El símbolo cambia, la obligación permanece. El neutro debe satisfacer $x \star e = x$ y $e \star x = x$. La primera ecuación se traduce en $x + e + 1 = x$, que obliga a $e = -1$ en los racionales. Comprobamos las dos direcciones: $x \star (-1) = x - 1 + 1 = x$ y $(-1) \star x = -1 + x + 1 = x$.

Para hallar u, usamos ahora el neutro correcto:

$$x \star u = -1 \quad \Longleftrightarrow \quad x + u + 1 = -1.$$

El candidato es $u = -x - 2$. Sustituirlo da $x \star u = -1$ y $u \star x = -1$. La asociatividad se comprueba expandiendo:

$$(x \star y) \star z = x + y + z + 2 = x \star (y \star z).$$

La operación es interna porque suma racionales. Pedir $x \star u = 0$ sería pedir otro valor final: daría $u = -x - 1$, que no es un inverso respecto del neutro -1. El nombre «inverso» depende de la operación y de su neutro, no del aspecto del símbolo 0 de la aritmética usual.

Solución 15. Cuando toda parte es estable. De (a) a (b). Fijemos $x \in A$. El singleton $\{x\}$ es un subconjunto de A y, por (a), es estable. Por tanto, $f(x) \in \{x\}$, lo que significa $f(x) = x$.

De (b) a (c). Para cualquier $S \subseteq A$, si $y \in f(S)$ existe $x \in S$ con $y = f(x) = x$; luego $y \in S$. Recíprocamente, si $y \in S$, entonces $f(y) = y$ y esa entrada muestra que $y \in f(S)$. Por doble inclusión, $f(S) = S$.

De (c) a (a). Para cualquier S, la igualdad $f(S) = S$ implica $f(S) \subseteq S$; por la equivalencia de §5.4, S es estable. Las tres implicaciones cierran el ciclo y prueban equivalencia.

Si $A = \varnothing$, la única función $A \to A$ es la vacía y el único subconjunto es $\varnothing$. (a) se cumple por estabilidad del vacío; (b), por vacuidad de su universal; (c), por $f(\varnothing) = \varnothing$. La prueba general sigue siendo válida.

La clave fue poder escoger **cada singleton**. La estabilidad de un conjunto aislado no ofrece esa información. Al cuantificar sobre todos los subconjuntos, la hipótesis se vuelve mucho más fuerte y fuerza que cada entrada sea un punto fijo.

De las definiciones a la arquitectura de las demostraciones. En este capítulo hemos usado definiciones para extraer datos, fijar objetivos, construir ejemplos y localizar errores. El laboratorio de cuerpos mostró que incluso reglas conocidas necesitan un lugar preciso en la cadena de razones: una existencia no implica por sí sola unicidad; una cancelación tiene hipótesis; una desigualdad exige un orden.

En el capítulo 6 organizaremos las formas de demostración que ya hemos utilizado: prueba directa, dos direcciones de una equivalencia, casos, contraposición, contradicción, existencia y unicidad. Su estudio partirá de una disciplina que ahora podemos aplicar: antes de elegir una estrategia, saber exactamente qué dicen los datos y qué obliga a demostrar la conclusión.

Capítulo 6 — ¿Cómo se demuestra una afirmación?

En los capítulos anteriores ya hemos demostrado muchas cosas. Probamos que ciertas fórmulas eran equivalentes, que algunas relaciones definían funciones, que dos conjuntos coincidían, que ciertos objetos existían y que otros eran únicos. En el capítulo 5, además, seguimos cadenas de consecuencias dentro de una estructura axiomática y vimos que una regla familiar no puede utilizarse antes de haberla justificado.

Sin embargo, todavía no hemos detenido la marcha para estudiar las **formas de demostración** que estaban actuando en esos argumentos. A veces partíamos de las hipótesis y avanzábamos hasta la conclusión. Otras veces separábamos dos direcciones. En algunos problemas dividíamos la situación en casos; en otros suponíamos que la conclusión fallaba y observábamos qué se volvía imposible. Cuando el objetivo era existencial, buscábamos un objeto; cuando pedía unicidad, comparar un candidato con cualquier otro era parte esencial de la tarea.

Este capítulo hará explícita esa arquitectura.

La pregunta no será simplemente «¿qué métodos de prueba existen?». Una lista de nombres —prueba directa, contraposición, contradicción, casos— puede aprenderse con rapidez y, aun así, servir de muy poco. Antes de escoger una técnica debemos comprender **qué afirmación tenemos delante y qué tendría que quedar establecido para poder decir que está demostrada**.

Ése será el principio rector:

> **La forma del objetivo orienta la arquitectura de la prueba, pero no proporciona por sí sola el contenido matemático que la completa.**

Saber que una afirmación tiene la forma $P \Rightarrow Q$ nos dice que debemos justificar el paso desde P hasta Q. No nos dice qué definición, identidad, estimación o idea intermedia hará posible ese paso. Reconocer la arquitectura evita empezar a ciegas; descubrir el contenido que la llena será una cuestión distinta, que retomaremos con mayor profundidad en el capítulo 7.

Comenzaremos, por tanto, antes de cualquier técnica particular: aprenderemos a leer una afirmación como una **lista de obligaciones**.

6.1. ¿Qué obliga a hacer la forma del objetivo?

Consideremos una afirmación sencilla:

Para todo $x \in A$, si $P(x)$, entonces $Q(x)$.

Podemos escribirla como

$$\forall x \in A \, \big(P(x) \Rightarrow Q(x)\big).$$

Antes de intentar demostrarla, conviene preguntarnos qué tendría que ocurrir para que el argumento fuese completo.

La cuantificación universal nos impide escoger un elemento especialmente favorable. Debemos trabajar con un $x \in A$ arbitrario. La implicación nos permite suponer $P(x)$ y nos obliga a llegar a $Q(x)$. Por tanto, una arquitectura natural es:

$$\boxed{\begin{aligned} x \in A \text{ arbitrario} &\longrightarrow P(x) \text{ como hipótesis} \\ &\longrightarrow Q(x) \text{ como objetivo.} \end{aligned}}$$

Esto todavía no es una prueba. Es el **esqueleto** de una prueba posible.

Entre $P(x)$ y $Q(x)$ puede haber un paso inmediato o diez ideas difíciles. Tal vez necesitemos desplegar una definición, introducir un objeto auxiliar, usar un teorema anterior o separar casos. La forma lógica nos ha dicho dónde empezar y qué final necesitamos alcanzar; no ha construido el puente.

Leer la conclusión antes de calcular. Una dificultad frecuente al comenzar a demostrar es empezar a manipular expresiones antes de decidir qué se está tratando de establecer. Supongamos que tenemos conjuntos A, B, C y queremos demostrar

$$A \cap (B \cup C) \subseteq (A \cap B) \cup C.$$

Podríamos comenzar a dibujar diagramas o a aplicar leyes de conjuntos. Pero la forma del objetivo ya contiene una instrucción más elemental.

Demostrar

$$X \subseteq Y$$

significa demostrar

$$\forall x \ (x \in X \Rightarrow x \in Y).$$

Por ello, nuestro trabajo comienza así:

1. tomar un elemento arbitrario $x \in A \cap (B \cup C)$;
2. utilizar esa pertenencia como dato;
3. demostrar $x \in (A \cap B) \cup C$.

Ahora sí podemos desplegar las definiciones:

$$x \in A \cap (B \cup C)$$

nos dice simultáneamente que

$$x \in A$$

y que

$$x \in B \cup C.$$

La segunda pertenencia significa que ocurre al menos una de estas alternativas:

$$x \in B \qquad \text{o} \qquad x \in C.$$

Si $x \in B$, entonces junto con $x \in A$ obtenemos $x \in A \cap B$, y por tanto

$$x \in (A \cap B) \cup C.$$

Si $x \in C$, la misma conclusión se obtiene directamente. En ambos casos llegamos al conjunto de destino.

La demostración completa podría escribirse sin toda esta conversación preparatoria:

Demostración. Sea $x \in A \cap (B \cup C)$. Entonces $x \in A$ y $x \in B \cup C$. Si $x \in B$, entonces $x \in A \cap B$ y, por tanto, $x \in (A \cap B) \cup C$. Si $x \in C$, también $x \in (A \cap B) \cup C$. Luego

$$A \cap (B \cup C) \subseteq (A \cap B) \cup C.$$

□

Lo importante aquí no es la identidad de conjuntos. Es la disciplina previa: **la inclusión nos dijo qué forma debía tener la prueba**. La disyunción apareció después, al desplegar la pertenencia a una unión, y por eso surgió naturalmente una prueba por casos.

Una forma puede contener otras formas. Las obligaciones de prueba suelen estar anidadas.

En el ejemplo anterior, el objetivo exterior era una inclusión. Eso produjo una implicación universal. Después, una de las hipótesis se convirtió en una disyunción:

$$x \in B \cup C \quad \Longleftrightarrow \quad x \in B \text{ o } x \in C.$$

El argumento tuvo entonces una estructura dentro de otra:

$$\text{universal } \supset \text{ implicación } \supset \text{ casos.}$$

Esta observación será importante durante todo el capítulo. Una demostración real rara vez consiste en aplicar una única etiqueta de principio a fin. Podemos comenzar con una prueba directa y encontrar dentro de ella una existencia que requiera construir un testigo; o iniciar una demostración de unicidad y necesitar una contradicción local; o probar una equivalencia y usar técnicas distintas en cada dirección.

Por eso conviene hablar de **arquitectura** y no sólo de «método».

Si el objetivo es una igualdad. Supongamos ahora que queremos demostrar

$$A = B.$$

¿Qué exige exactamente esta conclusión?

Para conjuntos, la igualdad significa que tienen los mismos elementos. Una manera operativa de probarlo es establecer las dos inclusiones:

$$A \subseteq B \qquad \text{y} \qquad B \subseteq A.$$

Cada inclusión vuelve a convertirse en una implicación sobre un elemento arbitrario. La arquitectura completa es, por tanto,

$$\boxed{A = B \quad \Longleftarrow \quad (A \subseteq B) \text{ y } (B \subseteq A).}$$

Obsérvese que no basta demostrar sólo una dirección. De

$$A \subseteq B$$

no se sigue que $A = B$. El conjunto A podría ser una parte propia de B.

La misma idea aparece en un bicondicional:

$$P \Longleftrightarrow Q.$$

Para demostrarlo debemos justificar

$$P \Rightarrow Q$$

y también

$$Q \Rightarrow P.$$

Las dos direcciones pueden ser muy distintas. Que una resulte inmediata no vuelve innecesaria la otra.

Más adelante estudiaremos esta estructura con detalle. Por ahora basta retener la regla de lectura: **cuando la conclusión afirma equivalencia, debemos identificar todas las direcciones que forman parte de ella**.

Si el objetivo es existencial. Consideremos ahora

$$\exists x \in A\ P(x).$$

Aquí la obligación cambia por completo. Ya no tenemos que tratar a todos los elementos de A. Basta encontrar uno que funcione.

Pero hay dos tareas, no una:

1. producir o justificar un candidato $a \in A$;
2. verificar que cumple $P(a)$.

Esquemáticamente,

$$\boxed{\text{elegir } a \quad + \quad a \in A \quad + \quad P(a).}$$

Un candidato sin verificación no es una prueba.

Por ejemplo, sea $f : \mathbb{R} \to \mathbb{R}$ dada por

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

Queremos probar que existe un real cuya imagen es 5:

$$\exists x \in \mathbb{R}\ (f(x) = 5).$$

Podemos elegir $x = 2$. Pero la frase «tomemos $x = 2$» no completa el argumento. Debemos comprobar:

$$2 \in \mathbb{R}$$

y

$$f(2) = 2^2 + 1 = 5.$$

Entonces sí tenemos un testigo válido.

En este ejemplo encontramos el testigo explícitamente. No todas las pruebas de existencia serán así. En §6.7 distinguiremos cuidadosamente entre una construcción, una elección apoyada en resultados previos y una existencia que se obtiene sin exhibir de manera efectiva el objeto.

Si el objetivo exige unicidad. La notación

$$\exists! x \in A\ P(x)$$

contiene dos obligaciones diferentes:

existe al menos uno

y

existe a lo sumo uno.

En símbolos, podemos organizarla como

$$\exists x \in A\ P(x)$$

junto con

$$\forall x, y \in A\ \big(P(x) \land P(y) \Rightarrow x = y\big).$$

Encontrar un candidato sólo resuelve la primera mitad.

Esta separación ya apareció en el capítulo 5 con neutros, inversos y soluciones de ecuaciones. Ahora la convertimos en arquitectura general. Una demostración de unicidad suele comenzar suponiendo que hay dos candidatos x e y que cumplen la misma propiedad y tratando de demostrar

$$x = y.$$

De nuevo, la forma del objetivo no nos dice **cómo** obtendremos esa igualdad. Tal vez usemos cancelación, una desigualdad, una propiedad de orden, una inyectividad o una contradicción. Pero sí nos dice qué debemos comparar.

Si queremos refutar una universal. La palabra «demostrar» no siempre significa confirmar la afirmación que alguien propone. A veces la tarea correcta es mostrar que es falsa.

Supongamos que se afirma

$$\forall x \in A\ P(x).$$

Su negación es

$$\exists x \in A\ \neg P(x).$$

Por tanto, para refutar la universal basta encontrar un solo elemento del dominio que viole la propiedad y verificar la violación.

Consideremos la afirmación:

Para toda función $f : \mathbb{R} \to \mathbb{R}$, si $f(0) = 0$, entonces f es inyectiva.

Para refutarla necesitamos una función con dos características:

1. que satisfaga la hipótesis $f(0) = 0$;
2. que no sea inyectiva.

Tomemos

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

Se cumple

$$f(0) = 0.$$

Pero

$$f(1) = 1 = f(-1),$$

aunque

$$1 \neq -1.$$

Por tanto, f no es inyectiva. Hemos construido un contraejemplo completo.

No bastaría escribir simplemente «$f(x) = x^2$». Un contraejemplo también tiene obligaciones: debe pertenecer al dominio de la afirmación, satisfacer sus hipótesis y violar exactamente su conclusión.

Hipótesis disponibles y objetivo pendiente. Una forma útil de leer cualquier problema de demostración es separar dos columnas conceptuales:

qué puedo usar	qué debo obtener.

En la primera columna entran:

- hipótesis explícitas;
- definiciones ya activadas por esas hipótesis;
- resultados anteriores que realmente sean aplicables;
- propiedades del dominio que hayan sido establecidas.

En la segunda columna entra la forma precisa de la conclusión.

Esta separación parece elemental, pero previene errores profundos. Si la conclusión es «f es inyectiva», no sirve demostrar que f es sobreyectiva. Si el objetivo es una igualdad de conjuntos, una sola inclusión no basta. Si pedimos existencia, una descripción sin testigo o sin argumento de existencia queda incompleta. Si pedimos unicidad, verificar un candidato no excluye otros.

También ayuda a detectar razonamientos circulares. Si colocamos silenciosamente en la columna de «lo que puedo usar» aquello mismo que estamos intentando demostrar, el argumento puede parecer fluido aunque no haya avanzado.

Una prueba no es la historia de cómo encontramos la idea. Hay una última distinción importante.

El camino mediante el cual descubrimos una demostración puede incluir tanteos, ejemplos, cálculos fallidos, dibujos, conjeturas y cambios de estrategia. La prueba final no necesita reproducir toda esa historia. Debe presentar una cadena verificable que vaya desde los datos aceptados hasta la conclusión.

Por ejemplo, quizá descubramos una inclusión dibujando varios diagramas. El diagrama puede sugerir la afirmación y ayudarnos a encontrar una ruta. Pero si el objetivo es demostrar una inclusión entre conjuntos abstractos, la prueba final debe justificar por qué **todo** elemento del primer conjunto pertenece al segundo.

De manera análoga, comprobar veinte valores puede sugerir una proposición universal, pero no sustituye el argumento para un elemento arbitrario.

En el capítulo 7 volveremos a la pregunta del descubrimiento. Aquí trabajaremos principalmente con la arquitectura verificable de la prueba terminada.

Antes de demostrar

Antes de escribir una prueba, identifica:

1. el dominio de los objetos;
2. las hipótesis que puedes usar;
3. la forma exacta del objetivo;
4. las obligaciones que esa forma genera.

Sólo después tiene sentido preguntar qué estrategia matemática permitirá cumplirlas.

Un pequeño mapa de obligaciones. Sin convertirlo en un recetario, podemos resumir las formas que aparecerán en este capítulo:

Objetivo	Obligación inicial
$\forall x \in A\, P(x)$	tomar $x \in A$ arbitrario y demostrar $P(x)$
$P \Rightarrow Q$	suponer P y demostrar Q
$P \Leftrightarrow Q$	demostrar $P \Rightarrow Q$ y $Q \Rightarrow P$
$A \subseteq B$	tomar $x \in A$ arbitrario y demostrar $x \in B$
$A = B$	demostrar $A \subseteq B$ y $B \subseteq A$

Objetivo	Obligación inicial
$\exists x \in A\, P(x)$	justificar un testigo $a \in A$ y verificar $P(a)$
$\exists! x \in A\, P(x)$	probar existencia y además unicidad
refutar $\forall x \in A\, P(x)$	producir $a \in A$ con $\neg P(a)$

La tabla no decide por nosotros. Sólo transforma una frase matemática en una tarea precisa.

Comprobación. Supongamos que queremos demostrar:

> Si $f : A \to B$ y $g : B \to C$ son inyectivas, entonces $g \circ f : A \to C$ es inyectiva.

¿Cuál es la obligación inmediata? ¿Qué datos podemos usar?

Respuesta razonada. El objetivo «$g \circ f$ es inyectiva» debe desplegarse antes de hacer cálculos. Tomamos $x, y \in A$ y suponemos

$$(g \circ f)(x) = (g \circ f)(y).$$

Debemos demostrar $x = y$.

La hipótesis de que g es inyectiva permite pasar de

$$g(f(x)) = g(f(y))$$

a

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

Después, la inyectividad de f permite concluir

$$x = y.$$

La prueba completa es breve porque, una vez desplegado el objetivo, las hipótesis encajan exactamente con las dos etapas necesarias. La arquitectura no inventó las propiedades de f y g, pero permitió ver cómo debían utilizarse.

Lo que esta lectura todavía no resuelve. Algunos problemas se vuelven casi automáticos una vez que entendemos su forma. Otros no.

Saber que debemos demostrar Q a partir de P no nos dice siempre qué hacer con P. Saber que necesitamos un testigo no nos dice dónde encontrarlo. Saber que dos objetos deben ser iguales no nos dice qué comparación será útil.

Ahí comienza la matemática que no cabe en una tabla.

En las próximas secciones estudiaremos varias arquitecturas fundamentales: prueba directa, dos direcciones, contraposición, contradicción, casos, existencia y unicidad. Las aprenderemos como maneras de organizar obligaciones, no como comandos que se aplican mecánicamente.

Comenzaremos con la más transparente: **partir de las hipótesis y avanzar hasta la conclusión**.

6.2. Prueba directa: avanzar desde lo que sabemos

La arquitectura más inmediata para una implicación

$$P \Rightarrow Q$$

consiste en aceptar P como hipótesis y construir, a partir de ella, una cadena de razones que termine en Q. A esta forma de argumento la llamaremos **prueba directa**.

El esquema mínimo es

$$\boxed{\begin{aligned} &P \longrightarrow \text{consecuencias justificadas} \\ &\quad \longrightarrow Q. \end{aligned}}$$

La palabra «directa» puede inducir a error si la entendemos como «fácil», «corta» o «sin ideas auxiliares». Una prueba directa puede ser larga y puede necesitar definiciones, resultados anteriores, construcciones intermedias o varios pasos ingeniosos. Lo que la caracteriza es otra cosa: **trabajamos bajo las hipótesis de la afirmación y avanzamos hacia la conclusión, sin reemplazar globalmente la implicación por una forma lógicamente equivalente**.

Esto enlaza con una distinción del capítulo anterior. Una definición puede proporcionarnos información apenas sabemos que un objeto la satisface. En una prueba directa, buena parte del trabajo consiste precisamente en preguntar:

> ¿Qué puedo extraer de las hipótesis y qué tendría que obtener para reconocer la conclusión?

No hay que demostrar la hipótesis. Supongamos que queremos probar

> Si una función $f : A \to B$ es inyectiva y $f(x) = f(y)$, entonces $x = y$.

La afirmación parece casi tautológica porque la conclusión forma parte de la definición de inyectividad. Pero conviene observar la arquitectura.

No debemos demostrar que f es inyectiva. **Eso es una hipótesis.** Tampoco debemos demostrar que $f(x) = f(y)$. También se nos concede dentro del antecedente. Nuestro trabajo comienza después de aceptar ambos datos.

Desplegamos la definición:

$$\begin{aligned} f \text{ es inyectiva} \quad \Longrightarrow \quad & \forall u, v \in A, \\ & f(u) = f(v) \Rightarrow u = v. \end{aligned}$$

Aplicándola a $u = x$ y $v = y$, la igualdad $f(x) = f(y)$ produce directamente $x = y$.

La demostración puede escribirse en una línea, pero la lección es más general: **en una implicación no tenemos que justificar el antecedente; tenemos que justificar que, siempre que el antecedente se cumple, la conclusión también se cumple**.

Esto evita un error frecuente. Si queremos demostrar

$$P \Rightarrow Q,$$

no sirve comenzar demostrando P para un caso especial y después demostrar Q en ese mismo caso. Eso podría establecer que P y Q son verdaderas en una situación concreta, pero todavía no justificaría la dependencia lógica entre ellas.

Del dato a la meta mediante definiciones. Consideremos ahora un ejemplo donde la conclusión no aparece de inmediato.

Sean A, B, C conjuntos y supongamos que

$$A \subseteq B.$$

Queremos probar

$$A \cap C \subseteq B \cap C.$$

Leemos primero el objetivo. Para demostrar la inclusión, tomamos un elemento arbitrario

$$x \in A \cap C$$

y debemos obtener

$$x \in B \cap C.$$

La hipótesis de pertenencia nos da dos datos:

$$x \in A \qquad \text{y} \qquad x \in C.$$

La hipótesis global $A \subseteq B$ transforma el primero en

$$x \in B.$$

Ahora tenemos exactamente las dos condiciones necesarias para concluir

$$x \in B \cap C.$$

Así queda probado el resultado.

Proposición. Si $A \subseteq B$, entonces, para todo conjunto C,

$$A \cap C \subseteq B \cap C.$$

Demostración. Sea $x \in A \cap C$. Entonces $x \in A$ y $x \in C$. Como $A \subseteq B$, se sigue que $x \in B$. Por tanto, $x \in B \cap C$. Puesto que x era arbitrario,

$$A \cap C \subseteq B \cap C.$$

□

La prueba es directa, pero contiene varias capas: una cuantificación universal implícita sobre los elementos, la definición de inclusión y la definición de intersección. Ninguna capa fue añadida por intuición ornamental; cada una apareció porque alguna hipótesis o algún objetivo la exigía.

Las flechas de una prueba no tienen que ser equivalencias. En una prueba directa solemos encadenar consecuencias:

$$P \Rightarrow R_1 \Rightarrow R_2 \Rightarrow \cdots \Rightarrow Q.$$

No necesitamos que cada paso pueda recorrerse en sentido contrario.

Ésta es una diferencia importante entre una **cadena de consecuencias** y una **cadena de equivalencias**. Si desde P obtenemos R, y desde R obtenemos Q, entonces hemos demostrado que P implica Q, aunque Q no implique R ni R implique P.

Por ejemplo, de

$$x > 3$$

podemos concluir

$$x > 0$$

y de ahí

$$x^2 > 0.$$

La cadena

$$x > 3 \Rightarrow x > 0 \Rightarrow x^2 > 0$$

es perfectamente útil para avanzar. Pero no podemos invertirla: $x^2 > 0$ no implica $x > 0$, pues también vale para números negativos no nulos.

Una prueba directa puede, por supuesto, contener equivalencias locales. Lo que no debemos hacer es tratarlas como requisito universal. Para probar una implicación sólo necesitamos preservar la verdad **hacia adelante** desde las hipótesis hasta la conclusión.

Un ejemplo con funciones: usar una propiedad dentro de otra. Sean

$$f : A \to B \qquad \text{y} \qquad g : B \to C.$$

Supongamos que la composición

$$g \circ f : A \to C$$

es inyectiva. Queremos demostrar que f es inyectiva.

Antes de comenzar, desplegamos el objetivo. Debemos tomar $x, y \in A$, suponer

$$f(x) = f(y)$$

y demostrar

$$x = y.$$

A partir de la igualdad de imágenes podemos aplicar g a ambos lados:

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

Esto equivale a

$$(g \circ f)(x) = (g \circ f)(y).$$

Como $g \circ f$ es inyectiva, concluimos

$$x = y.$$

Ya hemos satisfecho la definición de inyectividad de f.

Proposición. Si $g \circ f$ es inyectiva, entonces f es inyectiva.

Demostración. Sean $x, y \in A$ y supongamos $f(x) = f(y)$. Aplicando g a ambos miembros obtenemos

$$g(f(x)) = g(f(y)),$$

es decir,

$$(g \circ f)(x) = (g \circ f)(y).$$

La composición es inyectiva, luego $x = y$. Por tanto, f es inyectiva. □

Aquí aparece una característica importante de la prueba directa: una hipótesis puede no ser utilizable en su forma original hasta que construyamos la situación a la que se aplica. La inyectividad de $g \circ f$ habla de igualdades entre **valores de la composición**. Partíamos, en cambio, de una igualdad entre valores de f. Aplicar g fue el paso que hizo coincidir nuestros datos con la forma de la hipótesis disponible.

No toda hipótesis tiene que usarse inmediatamente. En una demostración algo más larga, algunas hipótesis permanecen inactivas durante varios pasos. Esto no significa que sean irrelevantes.

Supongamos que R es una relación sobre un conjunto A y que sabemos que R es simétrica y transitiva. Queremos demostrar:

Si xRy, entonces xRx.

La hipótesis particular es

$$xRy.$$

La simetría nos da

$$yRx.$$

Ahora disponemos simultáneamente de

$$xRy \qquad \text{y} \qquad yRx.$$

La transitividad puede aplicarse con el elemento intermedio y, y obtenemos

$$xRx.$$

La prueba completa es:

Demostración. Supongamos xRy. Como R es simétrica, yRx. Por transitividad, de xRy y yRx se sigue xRx. □

La hipótesis de transitividad estaba disponible desde el comienzo, pero no era aplicable hasta producir una segunda relación mediante simetría. Esto ilustra una regla útil: **una hipótesis puede necesitar que fabriquemos primero las condiciones que permiten activarla**.

Una afirmación universal añade una capa. Muchas proposiciones matemáticas tienen la forma

$$\forall x \in A \, \big(P(x) \Rightarrow Q(x)\big).$$

Una prueba directa debe respetar ambas capas:

1. elegir $x \in A$ arbitrario;
2. suponer $P(x)$;
3. demostrar $Q(x)$.

La palabra **arbitrario** importa. No significa que elijamos un objeto al azar, sino que no utilizaremos ninguna propiedad especial que no esté garantizada para todo elemento del dominio considerado.

Por ejemplo, queremos demostrar:

Para todo real x, si $x \geq 2$, entonces $x^2 \geq 2x$.

Sea $x \in \mathbb{R}$ y supongamos

$$x \geq 2.$$

Entonces

$$x - 2 \geq 0.$$

Además, de $x \geq 2$ se sigue $x \geq 0$. El producto de dos reales no negativos es no negativo, de modo que

$$x(x-2) \geq 0.$$

Al desarrollar,

$$x^2 - 2x \geq 0,$$

y por tanto

$$x^2 \geq 2x.$$

Hemos usado propiedades del orden de los reales. No podríamos trasladar esta prueba mecánicamente a un cuerpo arbitrario: allí la relación $\geq$ puede no estar siquiera definida. El dominio no es un dato decorativo; determina qué reglas forman parte de la caja de herramientas disponible.

Elegir una representación que haga visible la conclusión. En algunos argumentos directos, el paso decisivo consiste en reescribir una hipótesis de una forma que se conecte con la definición del objetivo.

Sea n un entero. Supongamos que existen enteros a y b tales que

$$n = 6a + 10b.$$

Queremos demostrar que n es par.

La conclusión «n es par» significa que debemos exhibir un entero k con

$$n = 2k.$$

La hipótesis ya contiene casi esa forma:

$$n = 6a + 10b = 2(3a + 5b).$$

Como $a, b \in \mathbb{Z}$, también

$$3a + 5b \in \mathbb{Z}.$$

Tomando

$$k = 3a + 5b,$$

obtenemos

$$n = 2k.$$

Por definición, n es par.

El argumento no consistió simplemente en «hacer álgebra». La definición de paridad fijó la forma que queríamos alcanzar, y la factorización por 2 reorganizó la hipótesis para que esa forma apareciera.

Podemos resumir este tipo de movimiento como

$$\begin{aligned} \text{hipótesis} &\longrightarrow \text{representación útil} \\ &\longrightarrow \text{definición del objetivo.} \end{aligned}$$

¿Cuándo se termina una prueba directa?. Una demostración termina cuando hemos satisfecho **exactamente** la obligación que generaba el objetivo.

Si el objetivo era

$$f \text{ es inyectiva},$$

terminamos cuando, para entradas arbitrarias, hemos demostrado

$$f(x) = f(y) \Rightarrow x = y.$$

Si el objetivo era

$$A \subseteq B,$$

terminamos cuando un elemento arbitrario de A ha quedado demostrado como elemento de B.

Si el objetivo era

$$n \text{ es par},$$

terminamos cuando hemos producido la representación exigida por la definición.

Seguir calculando después de haber alcanzado la conclusión puede ocultar dónde terminó realmente el argumento. Detenerse en el punto correcto también forma parte de escribir una prueba legible.

Qué no es una prueba directa completa. Comparemos tres intentos de demostrar:

Si $g \circ f$ es inyectiva, entonces f es inyectiva.

Intento A.

Como $g \circ f$ es inyectiva, entonces f es inyectiva.

La conclusión puede ser correcta, pero falta el vínculo. Se ha repetido esencialmente el enunciado sin demostrarlo.

Intento B.

Sean $x, y \in A$. Como $g \circ f$ es inyectiva, tenemos $x = y$.

Aquí la hipótesis se usa fuera de sus condiciones de aplicación. La inyectividad no permite concluir $x = y$ para dos entradas arbitrarias: requiere primero saber que sus imágenes bajo $g \circ f$ son iguales.

Intento C.

Sean $x, y \in A$ y supongamos $f(x) = f(y)$. Entonces $g(f(x)) = g(f(y))$, así que $(g \circ f)(x) = (g \circ f)(y)$. Como $g \circ f$ es inyectiva, $x = y$.

Éste sí cumple la obligación exacta de demostrar que f es inyectiva.

La diferencia entre B y C es pequeña en longitud y enorme en lógica. C explicita la condición que permite activar la hipótesis decisiva.

Prueba directa

Para demostrar directamente

$$P \Rightarrow Q,$$

trabajamos suponiendo P y construimos una cadena de inferencias justificadas hasta obtener Q.

La cadena puede usar definiciones, resultados anteriores y objetos auxiliares. Sus pasos no tienen que ser equivalencias; basta que preserven correctamente la consecuencia hacia la conclusión.

Comprobación. Sean $f : A \to B$ y $g : B \to C$. Decide si la siguiente afirmación puede demostrarse directamente con los datos disponibles:

Si f es sobreyectiva y g es sobreyectiva, entonces $g \circ f$ es sobreyectiva.

¿Qué obligación genera el objetivo y cómo se activan las dos hipótesis?

Respuesta razonada. Para probar que $g \circ f$ es sobreyectiva debemos tomar un elemento arbitrario

$$c \in C$$

y encontrar $a \in A$ tal que

$$(g \circ f)(a) = c.$$

Como g es sobreyectiva, para ese c existe algún $b \in B$ con

$$g(b) = c.$$

Ahora usamos la sobreyectividad de f: para ese b existe algún $a \in A$ tal que

$$f(a) = b.$$

Entonces

$$(g \circ f)(a) = g(f(a)) = g(b) = c.$$

Como c era arbitrario, $g \circ f$ es sobreyectiva.

La prueba avanza de manera directa, pero contiene dos elecciones existenciales anidadas. La forma global es una implicación; dentro de ella, la definición del objetivo nos obliga a construir una preimagen para un elemento arbitrario del codominio.

Lo que la prueba directa deja abierto. Cuando las hipótesis se despliegan y conducen de manera natural a la conclusión, la prueba directa suele ser la arquitectura más transparente. Pero no siempre ocurre así.

Podemos encontrarnos con una implicación en la que trabajar desde P no revela ninguna ruta razonable hacia Q, mientras que negar Q produce información mucho más concreta. En ese momento no cambiaremos la afirmación que queremos demostrar: cambiaremos la **forma equivalente desde la que la atacamos**.

Antes de llegar allí, estudiaremos otra obligación fundamental que ya apareció varias veces: demostrar una afirmación que exige **dos direcciones**.

6.3. Dos direcciones: bicondicionales y doble inclusión

Algunos objetivos no piden llegar a una conclusión en una sola dirección. Piden demostrar que dos condiciones contienen exactamente la misma información o que dos conjuntos tienen exactamente los mismos elementos.

En esos casos aparece una obligación doble.

Para un bicondicional,

$$P \iff Q,$$

debemos demostrar

$$P \Rightarrow Q$$

y

$$Q \Rightarrow P.$$

Para una igualdad de conjuntos,

$$A = B,$$

una estrategia fundamental es demostrar

$$A \subseteq B$$

y

$$B \subseteq A.$$

Las dos estructuras son cercanas porque la igualdad de conjuntos puede leerse elemento a elemento:

$$A = B \quad \iff \quad \forall x \, (x \in A \iff x \in B).$$

Pero conviene no confundirlas. El símbolo $=$ expresa igualdad de objetos; el símbolo $\iff$ conecta proposiciones. La relación entre ambos aparece porque podemos traducir la igualdad de conjuntos en equivalencia de condiciones de pertenencia.

Probar una dirección no prueba la equivalencia. Supongamos que queremos demostrar:

> Para una función $f : A \to B$, las siguientes condiciones son equivalentes:
>
> 1. f es inyectiva;
> 2. para todos $x, y \in A$, si $x \neq y$, entonces $f(x) \neq f(y)$.

La primera condición se expresa como

$$f(x) = f(y) \Rightarrow x = y.$$

La segunda, como

$$x \neq y \Rightarrow f(x) \neq f(y).$$

Estas dos implicaciones son contrapositivas entre sí. Por tanto, sabemos por lógica que son equivalentes.

Pero si el ejercicio pide **demostrar la equivalencia**, escribir sólo

> Si f es inyectiva y $x \neq y$, entonces $f(x) \neq f(y)$

no basta. Eso establece una dirección:

$$(1) \Rightarrow (2).$$

Todavía falta justificar

$$(2) \Rightarrow (1).$$

Veamos ambas.

De (1) a (2). Supongamos que f es inyectiva y tomemos $x, y \in A$ con $x \neq y$. Si ocurriera $f(x) = f(y)$, la inyectividad implicaría $x = y$, contradiciendo la hipótesis. Por tanto,

$$f(x) \neq f(y).$$

De (2) a (1). Supongamos ahora que

$$x \neq y \Rightarrow f(x) \neq f(y)$$

para todos $x, y \in A$. Queremos demostrar que f es inyectiva. Tomemos $x, y \in A$ y supongamos

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

Si $x \neq y$, la hipótesis daría

$$f(x) \neq f(y),$$

lo cual contradice la igualdad que acabamos de suponer. Luego

$$x = y.$$

Así, f es inyectiva.

La equivalencia ha quedado probada porque hemos cerrado las dos obligaciones.

Las direcciones pueden necesitar ideas distintas. Nada obliga a que

$$P \Rightarrow Q$$

y

$$Q \Rightarrow P$$

se demuestren del mismo modo.

Consideremos la siguiente caracterización de los subconjuntos estables introducidos en el capítulo 5. Sea $f : A \to A$ y $S \subseteq A$. Recordemos que S es estable bajo f cuando

$$\forall x \in S \ (f(x) \in S).$$

Queremos demostrar:

S es estable bajo f si y sólo si

$$f(S) \subseteq S.$$

La primera dirección comienza con estabilidad. Sea

$$y \in f(S).$$

Por definición de imagen, existe $x \in S$ tal que

$$y = f(x).$$

Como S es estable,

$$f(x) \in S.$$

Luego $y \in S$. Así,

$$f(S) \subseteq S.$$

La dirección inversa parte de una hipótesis de conjuntos. Supongamos

$$f(S) \subseteq S$$

y tomemos $x \in S$. Por definición de imagen,

$$f(x) \in f(S).$$

La inclusión da entonces

$$f(x) \in S.$$

Como esto vale para todo $x \in S$, el conjunto es estable.

Las dos pruebas utilizan definiciones distintas en órdenes distintos. La primera comienza con un elemento de la **imagen** y recupera una preimagen. La segunda comienza con un elemento de S y observa dónde cae su imagen. La equivalencia es simétrica como afirmación; su demostración no tiene por qué ser simétrica como recorrido.

Doble inclusión. Pasemos ahora a la igualdad de conjuntos.

Para demostrar

$$A = B,$$

es suficiente probar

$$A \subseteq B$$

y

$$B \subseteq A.$$

Esta técnica se llama **doble inclusión**.

No se trata de una convención arbitraria. Proviene de la definición extensional de igualdad: dos conjuntos son iguales cuando tienen los mismos elementos. La primera inclusión impide que A tenga elementos que falten en B; la segunda impide que B tenga elementos que falten en A.

Veamos un ejemplo que no depende de un diagrama particular.

Sean A, B, C conjuntos. Demostraremos

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

Primera inclusión. Tomemos

$$x \in A \cap (B \cup C).$$

Entonces

$$x \in A$$

y

$$x \in B \cup C.$$

La segunda condición significa que

$$x \in B \qquad \text{o} \qquad x \in C.$$

Si $x \in B$, entonces

$$x \in A \cap B.$$

Si $x \in C$, entonces

$$x \in A \cap C.$$

En cualquiera de los dos casos,

$$x \in (A \cap B) \cup (A \cap C).$$

Por tanto,

$$A \cap (B \cup C) \subseteq (A \cap B) \cup (A \cap C).$$

Segunda inclusión. Ahora tomemos

$$x \in (A \cap B) \cup (A \cap C).$$

Entonces ocurre una de dos cosas:

$$x \in A \cap B$$

o

$$x \in A \cap C.$$

En el primer caso, $x \in A$ y $x \in B$; en el segundo, $x \in A$ y $x \in C$. En ambos casos,

$$x \in A$$

y además

$$x \in B \cup C.$$

Por tanto,

$$x \in A \cap (B \cup C).$$

Así,

$$(A \cap B) \cup (A \cap C) \subseteq A \cap (B \cup C).$$

Las dos inclusiones prueban la igualdad.

La segunda dirección no es «obvia» porque se parece a la primera. Cuando dos direcciones tienen estructuras parecidas, aparece la tentación de escribir:

> La otra inclusión es análoga.

A veces esa abreviación es legítima para un lector experimentado. En un capítulo dedicado precisamente a aprender arquitectura de pruebas, conviene ser más cuidadosos.

Hay que preguntar:

- ¿qué objetos cambian de papel?;
- ¿qué definiciones se despliegan en cada dirección?;
- ¿la misma transformación realmente funciona al revés?;
- ¿hay algún paso que en una dirección usa sólo implicación y no equivalencia?

En el ejemplo distributivo anterior, las dos inclusiones son efectivamente cercanas, pero no idénticas. En una dirección comenzamos con una intersección que contiene una unión; en la otra, con una unión de dos intersecciones.

Hasta dominar la técnica, escribir ambas direcciones evita ocultar un posible salto.

Una inclusión puede ser estricta. ¿Por qué insistimos tanto en la segunda inclusión? Porque una sola inclusión es compatible con una diferencia real entre los conjuntos.

Tomemos

$$A = \{1\} \qquad \text{y} \qquad B = \{1, 2\}.$$

Tenemos

$$A \subseteq B,$$

pero

$$A \neq B.$$

Para demostrar igualdad, conocer una inclusión sólo descarta un tipo de diferencia. Todavía podría haber elementos adicionales en el conjunto mayor.

Lo mismo ocurre con una implicación. De

$$P \Rightarrow Q$$

no podemos concluir

$$P \Longleftrightarrow Q.$$

La recíproca

$$Q \Rightarrow P$$

es una afirmación adicional que necesita justificación.

Una caracterización: convertir una propiedad en una equivalencia. Las equivalencias son especialmente valiosas porque permiten cambiar de descripción.

Consideremos una función

$$f : A \to B.$$

Demostraremos:

f es sobreyectiva si y sólo si

$$f(A) = B.$$

Esta equivalencia conecta una definición cuantificada con una igualdad de conjuntos.

Primera dirección. Supongamos que f es sobreyectiva. Siempre tenemos

$$f(A) \subseteq B$$

porque f toma valores en B.

Para la otra inclusión, tomemos $y \in B$. Por sobreyectividad, existe $x \in A$ tal que

$$f(x) = y.$$

Entonces

$$y \in f(A).$$

Por tanto,

$$B \subseteq f(A),$$

y las dos inclusiones dan

$$f(A) = B.$$

Segunda dirección. Supongamos ahora

$$f(A) = B.$$

Sea $y \in B$. La igualdad implica

$$y \in f(A).$$

Por definición de imagen, existe $x \in A$ tal que

$$f(x) = y.$$

Como y era arbitrario, f es sobreyectiva.

La primera dirección contiene internamente una doble inclusión; la segunda convierte inmediatamente la igualdad de conjuntos en una existencia. Una equivalencia puede, por tanto, contener arquitecturas más pequeñas dentro de cada dirección.

Equivalencias múltiples. A veces queremos demostrar que tres o más condiciones son equivalentes:

$$P, \qquad Q, \qquad R.$$

Una posibilidad es demostrar todas las implicaciones por pares. Para tres condiciones eso produciría hasta seis direcciones:

$$P \Rightarrow Q, \quad Q \Rightarrow P,$$

$$P \Rightarrow R, \quad R \Rightarrow P,$$

$$Q \Rightarrow R, \quad R \Rightarrow Q.$$

Pero suele bastar mucho menos.

Si demostramos un ciclo

$$P \Rightarrow Q \Rightarrow R \Rightarrow P,$$

entonces cada condición implica las otras dos. Por ejemplo,

$$Q \Rightarrow R \Rightarrow P$$

da $Q \Rightarrow P$, y

$$R \Rightarrow P \Rightarrow Q$$

da $R \Rightarrow Q$.

También puede ser conveniente escoger una condición central y demostrar

$$P \Longleftrightarrow Q$$

y

$$Q \Longleftrightarrow R.$$

Entonces la transitividad de la implicación conecta P y R.

La elección entre una estrella, un ciclo u otra organización depende de qué direcciones sean más naturales. Aquí vuelve a aparecer una idea importante: **la forma del objetivo fija qué debe quedar probado, pero permite varias arquitecturas suficientes**.

Cuidado con cadenas de equivalencias. Una escritura como

$$P \iff Q \iff R$$

afirma más que una cadena de implicaciones.

Cada símbolo $\iff$ exige equivalencia en ambos sentidos. Si sólo sabemos

$$P \Rightarrow Q$$

y

$$Q \Rightarrow R,$$

podemos concluir

$$P \Rightarrow R,$$

pero no

$$P \iff R.$$

Esto es especialmente importante en cálculos algebraicos. Supongamos que partimos de

$$x = 1$$

y elevamos al cuadrado:

$$x^2 = 1.$$

El paso

$$x = 1 \Rightarrow x^2 = 1$$

es correcto. Pero no es una equivalencia sobre los reales, porque

$$x^2 = 1$$

también permite

$$x = -1.$$

Por tanto, escribir

$$x = 1 \iff x^2 = 1$$

sería falso.

Cuando transformamos una condición durante una demostración de equivalencia, debemos comprobar que los pasos señalados como reversibles realmente lo sean.

Doble inclusión sin perseguir elementos. La doble inclusión mediante elementos es la técnica más robusta para comenzar, pero no es la única manera de demostrar igualdad de conjuntos.

Si ya conocemos resultados que producen inclusiones, podemos utilizarlos directamente.

Por ejemplo, si sabemos

$$A \subseteq B$$

y

$$B \subseteq C,$$

entonces la transitividad de la inclusión da

$$A \subseteq C.$$

Si además sabemos

$$C \subseteq A,$$

podemos concluir

$$A = C.$$

No hace falta reiniciar desde un elemento arbitrario si las inclusiones necesarias ya están justificadas.

El principio general es el mismo que en §6.2: **usar resultados anteriores es legítimo cuando sus hipótesis se verifican y su conclusión coincide con la obligación actual**.

Comprobación. Sea $f : A \to B$. Demuestra la equivalencia:

$$f \text{ es inyectiva} \iff \forall S, T \subseteq A,\ f(S \cap T) = f(S) \cap f(T).$$

Antes de leer la respuesta, identifica qué parte de la igualdad vale para toda función y en qué dirección será necesaria la inyectividad.

Respuesta razonada.

Primero observemos que, para cualquier función,

$$f(S \cap T) \subseteq f(S) \cap f(T).$$

En efecto, si

$$y \in f(S \cap T),$$

existe $x \in S \cap T$ con $f(x) = y$. Como $x \in S$ y $x \in T$, tenemos

$$y \in f(S)$$

y

$$y \in f(T).$$

Por tanto,

$$y \in f(S) \cap f(T).$$

Ahora supongamos que f es inyectiva. Queremos demostrar la inclusión inversa. Sea

$$y \in f(S) \cap f(T).$$

Entonces existen $s \in S$ y $t \in T$ tales que

$$f(s) = y = f(t).$$

La inyectividad implica

$$s = t.$$

Ese elemento común pertenece a $S \cap T$, y su imagen es y. Así,

$$y \in f(S \cap T).$$

Hemos probado

$$f(S \cap T) = f(S) \cap f(T).$$

Esto demuestra una dirección de la equivalencia.

Para la recíproca, supongamos que la igualdad vale para todos los subconjuntos $S, T \subseteq A$. Tomemos $x_1, x_2 \in A$ y supongamos

$$f(x_1) = f(x_2).$$

Elegimos

$$S = \{x_1\}, \qquad T = \{x_2\}.$$

Si $x_1 \neq x_2$, entonces

$$S \cap T = \varnothing,$$

por lo que

$$f(S \cap T) = f(\varnothing) = \varnothing.$$

Pero la igualdad de imágenes $f(x_1) = f(x_2)$ muestra que

$$f(S) \cap f(T)$$

contiene ese valor común y, por tanto, no es vacío. Esto contradice la hipótesis

$$f(S \cap T) = f(S) \cap f(T).$$

Luego

$$x_1 = x_2.$$

Así, f es inyectiva.

Este ejemplo muestra algo que conviene conservar: las dos direcciones pueden tener naturalezas muy distintas. La primera usa doble inclusión; la segunda elige subconjuntos especialmente informativos y obtiene una contradicción local.

Dos direcciones

Para demostrar

$$P \iff Q,$$

deben cerrarse ambas obligaciones:

$$P \Rightarrow Q \qquad \text{y} \qquad Q \Rightarrow P.$$

Para demostrar

$$A = B,$$

la doble inclusión exige:

$$A \subseteq B \qquad \text{y} \qquad B \subseteq A.$$

Las dos direcciones no tienen por qué usar la misma idea ni tener la misma dificultad.

Hacia una forma equivalente más conveniente. Hasta ahora hemos mantenido el objetivo y trabajado cada dirección desde sus hipótesis. Pero una implicación puede tener otra forma lógicamente equivalente que resulte más manejable.

En particular,

$$P \Rightarrow Q$$

es equivalente a

$$\neg Q \Rightarrow \neg P.$$

Demostrar esta segunda implicación no significa probar «la otra dirección» del bicondicional. Significa demostrar la **contraposición** de la misma implicación original.

Ésa será nuestra siguiente arquitectura.

6.4. Contraposición: cambiar el objetivo sin cambiar la afirmación

Para una implicación

$$P \Rightarrow Q,$$

la lógica clásica garantiza la equivalencia

$$P \Rightarrow Q \quad \iff \quad \neg Q \Rightarrow \neg P.$$

La segunda implicación se llama **contraposición** de la primera.

Demostrar una implicación por contraposición consiste, por tanto, en demostrar

$$\neg Q \Rightarrow \neg P$$

en lugar de trabajar directamente desde P hacia Q.

La afirmación que estamos demostrando no cambia. Cambia la forma desde la cual intentamos alcanzarla.

Contraposición no es recíproco. La implicación recíproca de

$$P \Rightarrow Q$$

es

$$Q \Rightarrow P.$$

En general, la recíproca no es equivalente a la implicación original.

En cambio, la contraposición

$$\neg Q \Rightarrow \neg P$$

sí lo es.

Conviene fijar las cuatro formas:

Forma	Expresión
original	$P \Rightarrow Q$
recíproca	$Q \Rightarrow P$
contraria	$\neg P \Rightarrow \neg Q$
contraposición	$\neg Q \Rightarrow \neg P$

La original y la contraposición son lógicamente equivalentes. La recíproca y la contraria también son equivalentes entre sí, pero no tienen por qué coincidir con la original.

Por ejemplo,

> Si un entero es múltiplo de 4, entonces es par

es verdadero.

Su contraposición

> Si un entero no es par, entonces no es múltiplo de 4

también es verdadera.

Pero su recíproca

> Si un entero es par, entonces es múltiplo de 4

es falsa: 2 es par y no es múltiplo de 4.

¿Por qué puede ayudar?. A veces la conclusión Q tiene una forma difícil de construir directamente, mientras que su negación $\neg Q$ proporciona información concreta.

Consideremos:

> Si una función $f : A \to B$ es inyectiva, entonces para todos $x, y \in A$,

$$x \neq y \Rightarrow f(x) \neq f(y).$$

Podemos demostrarlo directamente, pero la contraposición de la implicación interior es

$$f(x) = f(y) \Rightarrow x = y,$$

que es exactamente la definición de inyectividad.

Aquí la contraposición no introduce ninguna idea nueva: simplemente revela que la forma equivalente coincide con una hipótesis disponible.

Primer ejemplo: una propiedad de conjuntos. Sea $f : A \to B$ y sean $S, T \subseteq A$.

Queremos demostrar:

Si $S \cap T \neq \varnothing$, entonces

$$f(S) \cap f(T) \neq \varnothing.$$

Una prueba directa también es natural: tomar un elemento común $x \in S \cap T$ y observar que $f(x)$ pertenece a ambas imágenes.

Pero si nos piden demostrar la implicación equivalente

Si

$$f(S) \cap f(T) = \varnothing,$$

entonces

$$S \cap T = \varnothing,$$

estamos viendo la contraposición exacta.

Supongamos que

$$f(S) \cap f(T) = \varnothing.$$

Si hubiera $x \in S \cap T$, entonces

$$f(x) \in f(S)$$

y

$$f(x) \in f(T),$$

por lo que

$$f(x) \in f(S) \cap f(T),$$

contradiciendo que esa intersección es vacía.

Luego

$$S \cap T = \varnothing.$$

Obsérvese un detalle: dentro de una prueba por contraposición puede aparecer una pequeña contradicción local. Eso no convierte automáticamente toda la prueba en "prueba por contradicción". La arquitectura global sigue siendo:

$$\neg Q \Rightarrow \neg P.$$

Segundo ejemplo: una función y su imagen. Consideremos una función

$$f : A \to B$$

y la afirmación:

Si f no es sobreyectiva, entonces existe $b \in B$ tal que

$$f(a) \neq b$$

para todo $a \in A$.

En realidad, esta afirmación es casi una lectura directa de la negación de sobreyectividad. Recordemos:

$$f \text{ es sobreyectiva} \iff \forall b \in B, \; \exists a \in A \; (f(a) = b).$$

Negando,

$$f \text{ no es sobreyectiva} \iff \exists b \in B, \; \forall a \in A \; (f(a) \neq b).$$

Aquí la forma negada no complica el problema: lo aclara.

Este tipo de situación es una buena señal para la contraposición en problemas más generales: **cuando negar la conclusión produce una hipótesis más manejable que la conclusión original**.

Un ejemplo donde la forma directa es incómoda. Queremos demostrar:

Si n^2 es impar, entonces n es impar.

La prueba directa comienza con

$$n^2 = 2k + 1$$

para algún entero k y debe terminar mostrando

$$n = 2m + 1$$

para algún entero m.

Esa ruta no es imposible, pero no conecta de manera evidente ambas representaciones.

La contraposición dice:

Si n no es impar, entonces n^2 no es impar.

Sobre los enteros, "no impar" equivale a "par". Así, basta demostrar:

Si n es par, entonces n^2 es par.

Sea

$$n = 2m.$$

Entonces

$$n^2 = 4m^2 = 2(2m^2),$$

y como $2m^2 \in \mathbb{Z}$, concluimos que n^2 es par.

Por contraposición,

$$n^2 \text{ impar} \Rightarrow n \text{ impar}.$$

Este ejemplo es clásico, pero aquí nos interesa la arquitectura: la negación de la conclusión transforma el objetivo en una propiedad que interactúa de forma inmediata con el cuadrado.

La negación debe ser correcta. La contraposición exige negar Q y P con precisión. Si

$$P: \quad x > 3,$$

entonces

$$\neg P: \quad x \leq 3,$$

no "$x < 3$".
Si

$$Q: \quad \forall y \in B\ R(y),$$

entonces

$$\neg Q: \quad \exists y \in B\ \neg R(y).$$

Una negación incorrecta produce una implicación distinta.
Consideremos:

Si todos los elementos de S pertenecen a T, entonces

$$S \subseteq T.$$

La negación de "$S \subseteq T$" es

$$\exists x \in S\ (x \notin T),$$

no "ningún elemento de S pertenece a T".
Negar bien no es un detalle previo a la técnica: es parte de la técnica.

Contraposición con cuantificadores. Supongamos que queremos demostrar:

Para todo $x \in A$, si $P(x)$, entonces $Q(x)$.

La forma completa es

$$\forall x \in A\ \big(P(x) \Rightarrow Q(x)\big).$$

Podemos demostrar para un x arbitrario la contraposición

$$\neg Q(x) \Rightarrow \neg P(x).$$

Eso produce

$$\forall x \in A\ \big(\neg Q(x) \Rightarrow \neg P(x)\big),$$

que es equivalente a la afirmación original.

Es importante mantener el mismo dominio y la misma dependencia del elemento x. No estamos negando toda la proposición universal; estamos contraponiendo la implicación que aparece dentro de ella.

Negar toda la universal produciría

$$\exists x \in A\ \big(P(x) \wedge \neg Q(x)\big),$$

que representa un contraejemplo, no una contraposición.

Contraposición frente a contradicción. Estas dos técnicas suelen confundirse porque ambas utilizan negaciones.

Para demostrar

$$P \Rightarrow Q$$

por **contraposición**, el objetivo se reemplaza por la implicación equivalente

$$\neg Q \Rightarrow \neg P.$$

La prueba tiene un objetivo positivo y explícito: desde $\neg Q$ debemos llegar a $\neg P$.

En una **prueba por contradicción**, en cambio, suponemos que la afirmación que queremos demostrar es falsa y tratamos de derivar una imposibilidad.

En forma esquemática:

$$\text{Contraposición: } \neg Q \longrightarrow \cdots \longrightarrow \neg P.$$

$$\text{Contradicción: } \neg(P \Rightarrow Q) \longrightarrow \cdots \longrightarrow \bot.$$

Como

$$\neg(P \Rightarrow Q) \Longleftrightarrow P \wedge \neg Q,$$

una contradicción para una implicación suele comenzar suponiendo simultáneamente P y $\neg Q$.

La diferencia conceptual es importante: en contraposición queremos **probar una nueva conclusión**; en contradicción queremos **mostrar que una suposición es imposible**.

Un mismo teorema puede admitir ambas. Tomemos la afirmación:

Si x^2 es par, entonces x es par.

Por contraposición:

$$x \text{ impar} \Rightarrow x^2 \text{ impar}.$$

Eso se prueba directamente escribiendo

$$x = 2k + 1$$

y desarrollando

$$x^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1.$$

Por contradicción, podríamos suponer a la vez:

$$x^2 \text{ es par}$$

y

$$x \text{ es impar},$$

y derivar que x^2 es impar, contradiciendo la primera hipótesis.

Ambos argumentos son válidos. Pero no son la misma arquitectura.

¿Cuándo conviene intentar contraposición?. No existe una regla infalible, pero hay señales útiles.

La contraposición puede ser especialmente natural cuando:

- Q es difícil de construir directamente, pero $\neg Q$ tiene una forma concreta;
- la hipótesis P es una propiedad positiva y su negación es fácil de reconocer;
- la conclusión implica una restricción fuerte, como inyectividad, divisibilidad, no vacuidad o pertenencia;
- una definición disponible se parece más a $\neg Q \Rightarrow \neg P$ que a $P \Rightarrow Q$.

Sin embargo, estas señales no convierten la contraposición en una receta. Si la prueba directa es más clara, no hay ninguna obligación de usar la forma indirecta.

Un ejemplo con relaciones. Sea R una relación transitiva sobre A. Queremos demostrar:

Si xRy y yRz, entonces xRz.

Aquí la prueba directa es inmediata porque ésa es exactamente la propiedad de transitividad.

Intentar una contraposición produciría:

Si no xRz, entonces no ocurre que xRy y yRz.

Equivalente a:

$$\neg(xRz) \Rightarrow \big(\neg(xRy) \vee \neg(yRz)\big).$$

La afirmación es correcta, pero pedagógicamente peor: introduce una disyunción que la definición original no necesitaba.

Esto muestra el otro lado de la técnica: una forma equivalente puede ser lógicamente válida y, sin embargo, menos útil.

Comprobación. Sea $f : A \to B$. Demuestra por contraposición:

Si f es inyectiva y

$$f(S) \cap f(T) \neq \varnothing,$$

entonces

$$S \cap T \neq \varnothing$$

para todos $S, T \subseteq A$.

Respuesta razonada.

La conclusión es

$$S \cap T \neq \varnothing.$$

Su negación es

$$S \cap T = \varnothing.$$

La hipótesis relevante que queremos negar es

$$f(S) \cap f(T) \neq \varnothing,$$

cuya negación es

$$f(S) \cap f(T) = \varnothing.$$

Por contraposición, basta demostrar:

Si

$$S \cap T = \varnothing,$$

entonces

$$f(S) \cap f(T) = \varnothing.$$

Supongamos

$$S \cap T = \varnothing.$$

Si hubiera

$$y \in f(S) \cap f(T),$$

existirían $s \in S$ y $t \in T$ tales que

$$f(s) = y = f(t).$$

Como f es inyectiva,

$$s = t.$$

Ese elemento pertenecería tanto a S como a T, de modo que

$$s \in S \cap T,$$

contradiciendo que la intersección es vacía.

Por tanto,

$$f(S) \cap f(T) = \varnothing.$$

La contraposición queda demostrada y, con ella, la implicación original.

Este ejemplo muestra además que la inyectividad es exactamente la hipótesis que impide que dos elementos distintos, uno de cada conjunto, colapsen en la misma imagen.

Contraposición

Para demostrar

$$P \Rightarrow Q,$$

podemos demostrar la forma equivalente

$$\neg Q \Rightarrow \neg P.$$

Esto no es demostrar la recíproca $Q \Rightarrow P$ ni es automáticamente una prueba por contradicción. La clave es negar correctamente y conservar las mismas dependencias y dominios.

Hacia la contradicción. La contraposición sustituye una implicación por otra equivalente y conserva una meta explícita:

$$\neg Q \Rightarrow \neg P.$$

Hay problemas en los que incluso esa meta no es la forma más natural. A veces lo útil es asumir que la afirmación que queremos demostrar falla y seguir las consecuencias hasta llegar a una imposibilidad.

Ésa es la arquitectura de la **prueba por contradicción**.

6.5. Prueba por contradicción: mostrar que la alternativa es imposible

La prueba por contradicción parte de una idea sencilla:

> Si negar la afirmación que queremos demostrar nos obliga a aceptar algo imposible, entonces esa negación no puede ser correcta.

En símbolos, si queremos demostrar una proposición P, asumimos temporalmente

$$\neg P$$

y tratamos de derivar una contradicción:

$$\neg P \implies \cdots \implies \bot.$$

Aquí $\bot$ representa una imposibilidad lógica: una situación en la que obtenemos simultáneamente una afirmación y su negación, o cualquier otra incompatibilidad con hechos ya establecidos.

Una vez alcanzada la contradicción, concluimos que la suposición $\neg P$ era falsa. Por tanto,

$$P$$

debe ser verdadera.

Qué hay que negar. La parte más delicada suele aparecer antes de comenzar la prueba: debemos negar **la afirmación completa que queremos demostrar**.

Si nuestro objetivo es

$$P \Rightarrow Q,$$

entonces su negación no es simplemente $\neg Q$.

Sabemos que

$$\neg(P \Rightarrow Q) \iff P \wedge \neg Q.$$

Por tanto, una prueba por contradicción de una implicación comienza típicamente suponiendo al mismo tiempo:

$$P$$

y

$$\neg Q.$$

Ésta es una diferencia estructural con la contraposición.

En contraposición comenzamos con

$$\neg Q$$

y debemos demostrar

$$\neg P.$$

En contradicción comenzamos con

$$P \wedge \neg Q$$

y debemos demostrar que esa combinación es imposible.

Un ejemplo elemental con funciones. Consideremos la afirmación:

Si $f : A \to B$ es inyectiva y $x, y \in A$ satisfacen

$$f(x) = f(y),$$

entonces

$$x = y.$$

Una prueba directa sería inmediata, porque la conclusión es exactamente lo que proporciona la definición de inyectividad.

Pero podemos escribir también una prueba por contradicción.

Supongamos que

$$f(x) = f(y)$$

y que, contra lo que queremos demostrar,

$$x \neq y.$$

Como f es inyectiva, de

$$f(x) = f(y)$$

se sigue

$$x = y.$$

Hemos obtenido simultáneamente

$$x = y$$

y

$$x \neq y.$$

La suposición $x \neq y$ es imposible. Por tanto,

$$x = y.$$

El argumento es correcto, aunque más largo que la prueba directa. Esto ilustra un principio importante: **la contradicción es una arquitectura válida, no necesariamente la más conveniente**.

Una contradicción debe depender de la suposición. Supongamos que queremos demostrar P y comenzamos suponiendo $\neg P$.

Durante la prueba recordamos que

$$1 \neq 2.$$

Eso es verdadero, pero no es una contradicción.

Tampoco serviría escribir simplemente una afirmación falsa como

$$1 = 2$$

sin haberla deducido de nuestras hipótesis.

La contradicción tiene que surgir mediante una cadena de inferencias válidas a partir de:

- las hipótesis originales;
- la negación temporal del objetivo;
- definiciones o resultados previamente disponibles.

De otro modo no hemos demostrado que la negación del objetivo conduzca a una imposibilidad.

La estructura correcta es:

$$\begin{aligned}\text{hipótesis} + \neg P &\Longrightarrow \cdots \\ &\Longrightarrow R \wedge \neg R.\end{aligned}$$

La contradicción final funciona porque muestra que el conjunto de supuestos con el que comenzamos no puede ser simultáneamente verdadero.

La contradicción puede adoptar formas distintas. No siempre veremos literalmente una expresión del tipo

$$R \wedge \neg R.$$

También cuentan como contradicciones situaciones incompatibles con propiedades ya establecidas.

Por ejemplo:

$$0 = 1$$

es imposible en un cuerpo, porque la definición exige

$$0 \neq 1.$$

Del mismo modo,

$$x < x$$

es imposible en un orden estricto.

En los reales,

$$x^2 < 0$$

es imposible para todo $x \in \mathbb{R}$.

Y para conjuntos,

$$a \in \varnothing$$

es imposible por definición.

En todos los casos, la fuerza de la contradicción procede de que hemos alcanzado una afirmación incompatible con el contexto matemático vigente.

Un ejemplo de no existencia. Demostremos:

No existe un entero n que sea simultáneamente par e impar.

Podemos expresar el objetivo como

$$\neg\exists n \in \mathbb{Z}\ (n \text{ par} \wedge n \text{ impar}).$$

Para una prueba por contradicción suponemos la negación:

$$\exists n \in \mathbb{Z}\ (n \text{ par} \wedge n \text{ impar}).$$

Sea entonces n un entero par e impar.

Por paridad, existen $a, b \in \mathbb{Z}$ tales que

$$n = 2a$$

y

$$n = 2b + 1.$$

Igualando,

$$2a = 2b + 1.$$

Por tanto,

$$2(a - b) = 1.$$

El miembro izquierdo es par y el derecho es impar. Equivalentemente,

$$1$$

sería divisible por 2, lo cual es imposible en $\mathbb{Z}$.

Así no existe un entero que sea par e impar a la vez.

La prueba no consistió en comprobar muchos enteros. La negación del objetivo nos dio un candidato hipotético y las dos definiciones incompatibles produjeron la contradicción.

Contradicción en una afirmación universal. Consideremos:

Para todo real x,

$$x^2 + 1 > 0.$$

Queremos demostrar

$$\forall x \in \mathbb{R}\ (x^2 + 1 > 0).$$

Tomemos $x \in \mathbb{R}$ arbitrario. Supongamos, para obtener una contradicción, que

$$x^2 + 1 \leq 0.$$

Entonces

$$x^2 \leq -1.$$

Pero para todo real,

$$x^2 \geq 0.$$

Así tendríamos simultáneamente

$$x^2 \geq 0$$

y

$$x^2 \leq -1.$$

Eso es imposible porque $0 > -1$.

Por tanto,

$$x^2 + 1 > 0.$$

Como x era arbitrario, la afirmación vale para todo real.

Aquí la contradicción ocurre dentro de una prueba universal: primero fijamos un elemento arbitrario y luego negamos la propiedad que queremos demostrar para ese elemento.

El papel de la negación en existencia y universalidad. La estructura de la negación determina qué objetos aparecen en la prueba.

Si queremos demostrar

$$\forall x \in A\ P(x),$$

su negación es

$$\exists x \in A\ \neg P(x).$$

Por contradicción supondríamos que existe un contraejemplo.

En cambio, si queremos demostrar

$$\exists x \in A\ P(x),$$

su negación es

$$\forall x \in A\ \neg P(x).$$

Entonces una prueba por contradicción de existencia comienza suponiendo que **ningún** elemento del dominio cumple la propiedad.

Estas dos formas no son intercambiables.

Un ejemplo de existencia por contradicción. Supongamos que A es un conjunto finito no vacío y que

$$f : A \to A$$

es inyectiva.

Queremos demostrar que f es sobreyectiva.

Una prueba directa requeriría mostrar que cada elemento de A tiene una preimagen. En conjuntos finitos, podemos usar un hecho estructural: una función inyectiva de un conjunto finito en sí mismo no puede "perder" un valor sin dejar espacio insuficiente para imágenes distintas.

Dependencia matemática: conteo finito

La demostración que sigue utiliza el principio elemental de conteo según el cual **no pueden colocarse n elementos distintos en sólo $n-1$ posiciones**. Aquí lo aceptamos explícitamente como hecho previo de aritmética finita, no como consecuencia automática de la palabra «inyectiva». En §11.8 demostraremos el lema general: si existe una inyección $I_m \to I_n$, entonces $m \leq n$, y justificaremos por qué un subconjunto propio de un conjunto finito tiene menos elementos. Ese resultado cierra formalmente la dependencia que estamos anticipando.

Veamos la forma por contradicción.

Supongamos que f no es sobreyectiva. Entonces existe

$$a \in A$$

que no pertenece a la imagen:

$$a \notin f(A).$$

Así,

$$f(A) \subseteq A \setminus \{a\}.$$

Pero f es inyectiva, de modo que los $|A|$ elementos de A tienen imágenes distintas. Eso obligaría a que el subconjunto

$$A \setminus \{a\}$$

contuviera al menos $|A|$ elementos distintos.

Sin embargo,

$$|A \setminus \{a\}| = |A| - 1.$$

La desigualdad

$$|A| \leq |A| - 1$$

es imposible.

Por tanto, f debe ser sobreyectiva.

La contradicción no es puramente lógica; utiliza la finitud del conjunto. Si A fuera infinito, la conclusión podría fallar. Por ejemplo,

$$f : \mathbb{N} \to \mathbb{N}, \qquad f(n) = n + 1,$$

es inyectiva pero no sobreyectiva si $\mathbb{N} = \{0, 1, 2, \dots\}$.

Esto muestra por qué las hipótesis importan incluso en una técnica indirecta: la contradicción sólo es válida dentro del contexto que la produce.

Contradicción y principio del tercero excluido. En lógica clásica, la prueba por contradicción utiliza el principio de que una proposición no puede ser falsa si su negación conduce a una imposibilidad.

En particular, para demostrar P asumimos $\neg P$ y derivamos $\bot$.

Este razonamiento es estándar en la matemática clásica que usamos en el libro.

No desarrollaremos aquí lógicas alternativas ni fundamentos constructivos. Basta registrar que nuestras reglas de negación, contraposición y contradicción se entienden dentro de ese marco clásico.

Una contradicción local dentro de una prueba directa. Supongamos que queremos demostrar directamente que una función $f : A \to B$ es inyectiva.

Tomamos $x, y \in A$ y suponemos

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

Tal vez el argumento continúe:

> Si $x \neq y$, entonces se produciría una incompatibilidad con otra propiedad;
> por tanto $x = y$.

El marco global sigue siendo una prueba directa de la inyectividad: comenzamos con la hipótesis requerida por la definición y terminamos en la conclusión requerida.

Dentro de esa arquitectura hemos usado una contradicción para justificar un paso.

Esto es normal. Las técnicas de prueba pueden anidarse.

Por eso es más preciso describir la arquitectura global por la obligación principal que organiza el argumento, y reconocer técnicas locales allí donde aparezcan.

Una contradicción no autoriza cualquier conclusión intermedia. Existe una regla lógica importante: si un conjunto de hipótesis lleva a una contradicción, entonces esas hipótesis no pueden ser todas verdaderas a la vez.

Pero durante la escritura debemos mantener claro **qué supuesto está siendo descargado**.

Si comenzamos con hipótesis aceptadas H y añadimos temporalmente $\neg P$, y de

$$H \wedge \neg P$$

derivamos una contradicción, entonces concluimos P bajo las hipótesis H.

No concluimos que alguna de las hipótesis originales de H sea falsa. El supuesto temporal que introdujimos precisamente para la contradicción es $\neg P$.

Esta disciplina evita confundir qué parte del argumento está en cuestión.

Un ejemplo donde la contradicción es natural. Demostremos:

No existe un número racional cuyo cuadrado sea 2.

Equivalente a:

$$\sqrt{2} \notin \mathbb{Q}.$$

Supondremos lo contrario:

$$\sqrt{2} \in \mathbb{Q}.$$

Entonces existen enteros a, b, con $b \neq 0$, tales que

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

Podemos elegir la fracción en términos mínimos, es decir,

$$\gcd(a, b) = 1.$$

Al elevar al cuadrado,

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

de donde

$$a^2 = 2b^2.$$

Así a^2 es par, y por el resultado sobre la paridad de los cuadrados, a es par. Escribimos

$$a = 2k.$$

Sustituyendo,

$$4k^2 = 2b^2,$$

por lo que

$$b^2 = 2k^2.$$

Entonces b también es par.

Así a y b son ambos divisibles por 2, contradiciendo que la fracción a/b estuviera en términos mínimos.

Por tanto,

$$\sqrt{2} \notin \mathbb{Q}.$$

Este ejemplo es clásico y aparece en muchos manuales, por lo que no lo usaremos como rasgo distintivo del capítulo. Aquí cumple una función precisa: mostrar una contradicción que no adopta inicialmente la forma $R \wedge \neg R$, sino que termina violando una condición de coprimalidad elegida al comienzo.

¿Cuándo conviene intentar contradicción?. Como antes, no hay algoritmo infalible. Puede ser una buena arquitectura cuando:

- la negación del objetivo produce una hipótesis concreta;
- queremos demostrar una no existencia;
- el objetivo afirma unicidad y su negación produce dos candidatos distintos;
- una propiedad global se vuelve difícil de construir directamente;
- el contexto contiene una restricción fuerte que una negación puede violar;
- un intento de contraposición no ofrece una meta más manejable.

Pero usar contradicción no debe ser una reacción automática ante una prueba difícil. Una prueba directa suele revelar mejor la dependencia entre hipótesis y conclusión cuando está disponible.

Contraposición y contradicción, lado a lado. Para demostrar

$$P \Rightarrow Q,$$

comparemos las dos arquitecturas:

Contraposición

$$\neg Q \quad \Longrightarrow \quad \cdots \quad \Longrightarrow \quad \neg P.$$

Objetivo local: demostrar $\neg P$.

Contradicción

$$P \wedge \neg Q \quad \Longrightarrow \quad \cdots \quad \Longrightarrow \quad \bot.$$

Objetivo local: demostrar que la combinación $P \wedge \neg Q$ es imposible.

Las dos técnicas pueden usar pasos semejantes, pero no descargan las mismas suposiciones ni persiguen la misma meta intermedia.

Comprobación. Sea $f : A \to A$ una función y supongamos que

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

para todo $x \in A$.

Demuestra por contradicción que f es inyectiva.

Respuesta razonada.

Queremos demostrar que

$$f(x) = f(y) \Rightarrow x = y.$$

Supongamos, para obtener una contradicción, que existen $x, y \in A$ tales que

$$f(x) = f(y)$$

pero

$$x \neq y.$$

Aplicamos f a ambos lados de la igualdad:

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

Por la hipótesis,

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

y

$$f(f(y)) = y.$$

Por tanto,

$$x = y.$$

Esto contradice la suposición

$$x \neq y.$$

Luego no pueden existir dos entradas distintas con la misma imagen. Por tanto, f es inyectiva.

El argumento podría escribirse también de manera directa: de $f(x) = f(y)$ aplicamos f y obtenemos $x = y$. La versión por contradicción es deliberadamente más larga porque aquí queremos hacer visible la arquitectura.

Prueba por contradicción

Para demostrar una afirmación P por contradicción:

1. suponemos temporalmente $\neg P$;
2. combinamos esa suposición con las hipótesis disponibles;
3. derivamos una incompatibilidad real;
4. concluimos que $\neg P$ es imposible y, por tanto, P es verdadera.

La contradicción debe ser consecuencia de la suposición añadida y de inferencias válidas; no basta introducir una afirmación falsa sin justificación.

Hacia las pruebas por casos. La contradicción explota una alternativa imposible. La siguiente arquitectura trabaja con una situación diferente: varias alternativas pueden ser posibles, pero juntas cubren todos los objetos relevantes.

Cuando el problema cambia de comportamiento según una propiedad —por ejemplo, según un número sea positivo, cero o negativo— podemos dividir el dominio y demostrar la conclusión en cada región.

Ésa es la **demostración por casos**.

6.6. Demostración por casos: dividir sin perder nada

A veces una afirmación debe demostrarse para objetos que pueden comportarse de maneras distintas. En lugar de forzar una única cadena de razonamiento, dividimos el problema en situaciones y demostramos la conclusión en cada una.

Ésa es la idea de una **demostración por casos**.

En su forma más simple, si sabemos que siempre ocurre

$$P \vee Q,$$

y podemos demostrar

$$P \Rightarrow R$$

y también

$$Q \Rightarrow R,$$

entonces concluimos

$$R.$$

La estructura es:

$$\boxed{\begin{aligned} P \vee Q, \quad P \Rightarrow R, \quad Q \Rightarrow R \\ \Longrightarrow \quad R. \end{aligned}}$$

La dificultad principal no suele estar en escribir «Caso 1» y «Caso 2». Está en justificar que **no hemos dejado ningún objeto fuera**.

Exhaustividad antes que exclusividad. Los casos deben ser **exhaustivos**: todo objeto relevante debe caer en al menos uno.

No es necesario que sean mutuamente excluyentes.

Por ejemplo, si queremos demostrar una propiedad para todo real x, podemos separar:

- Caso 1: $x \geq 0$;
- Caso 2: $x \leq 0$.

Los casos se solapan cuando $x = 0$, pero juntos cubren todos los reales. Si la misma conclusión se obtiene en ambos, el solapamiento no causa ningún problema.

En cambio, los casos

- $x > 0$;
- $x < 0$

no son exhaustivos: dejan fuera $x = 0$.

La regla esencial es:

> **Los casos pueden solaparse; lo que no pueden hacer es dejar un hueco.**

Un primer ejemplo: valor absoluto. Demostremos que para todo real x,

$$|x| \geq x.$$

La definición del valor absoluto ya sugiere los casos:

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

Caso 1: $x \geq 0$**.** Entonces

$$|x| = x,$$

de modo que

$$|x| \geq x.$$

Caso 2: $x < 0$**.** Entonces

$$|x| = -x.$$

Como $x < 0$, tenemos

$$-x > x,$$

y por tanto

$$|x| > x.$$

En ambos casos,

$$|x| \geq x.$$

La división no fue arbitraria. La definición del objeto ya estaba dada por casos y, por tanto, proporcionaba una partición natural del problema.

Los casos deben venir de algún dato matemático. Una prueba por casos es más convincente cuando la división responde a una estructura real del problema:

- signo de una cantidad;
- paridad;
- pertenencia a conjuntos;
- disyunción en una hipótesis;
- comportamiento definido por tramos;
- posibilidad de que una cantidad sea cero o no cero.

Dividir porque sí puede multiplicar el trabajo sin aportar claridad.

Supongamos que queremos demostrar para todo entero n que

$$n(n+1)$$

es par.

Una división natural es:

- n es par;
- n es impar.

Estos casos son exhaustivos para los enteros.

Si n es par, entonces $n(n+1)$ es par porque contiene el factor par n.

Si n es impar, entonces $n+1$ es par, así que el producto vuelve a tener un factor par.

La conclusión queda demostrada.

Aquí la clasificación par/impar no es una decoración: identifica cuál de los dos factores consecutivos aporta la divisibilidad por 2.

Una disyunción en la hipótesis produce casos. En §6.1 apareció el ejemplo

$$x \in B \cup C.$$

Por definición,

$$x \in B \vee x \in C.$$

Eso da inmediatamente dos casos.

De forma general, si una hipótesis tiene la forma

$$P \vee Q,$$

y queremos demostrar R, podemos intentar:

- bajo P, demostrar R;
- bajo Q, demostrar R.

No necesitamos decidir de antemano cuál de los dos ocurre. Basta demostrar que cualquiera de las alternativas conduce a la misma conclusión.

Casos que se solapan. Consideremos la afirmación:

Si $x \in A \cup B$, entonces

$$x \in (A \cup C) \cup (B \cup C).$$

Podemos dividir según:

- $x \in A$;
- $x \in B$.

Si x pertenece a ambos, ambos casos se aplican. Eso no invalida el argumento.

En el primer caso,

$$x \in A \subseteq A \cup C$$

y por tanto

$$x \in (A \cup C) \cup (B \cup C).$$

En el segundo caso,

$$x \in B \subseteq B \cup C,$$

y llegamos a la misma conclusión.

La prueba funciona porque la unión garantiza exhaustividad, no exclusividad.

Casos definidos por una dicotomía real. Hay divisiones donde sí obtenemos casos mutuamente excluyentes.

Por ejemplo, para un real x:

$$x = 0 \qquad \text{o} \qquad x \neq 0.$$

No hay solapamiento y no hay hueco.

Esta dicotomía es especialmente útil cuando una operación exige una condición de no nulidad.

Supongamos que trabajamos en un cuerpo F y queremos analizar las soluciones de

$$ax = b.$$

La división natural es:

- Caso 1: $a = 0$;
- Caso 2: $a \neq 0$.

En el segundo caso existe a^{-1} y podemos escribir

$$x = a^{-1}b.$$

En el primero, no podemos multiplicar por a^{-1} porque ese inverso no existe. Debemos distinguir además:

- si $b = 0$, todo $x \in F$ satisface la ecuación;
- si $b \neq 0$, no hay soluciones.

La prueba por casos impide precisamente aplicar una operación fuera de su dominio de validez.

Casos anidados. Una demostración puede contener casos dentro de casos.

El ejemplo anterior ya tiene esta forma:

$$a = 0 \quad \text{o} \quad a \neq 0.$$

Dentro de $a = 0$ distinguimos:

$$b = 0 \quad \text{o} \quad b \neq 0.$$

La estructura queda:

- Caso 1: $a = 0$.
  - Subcaso 1a: $b = 0$.
  - Subcaso 1b: $b \neq 0$.
- Caso 2: $a \neq 0$.

La única exigencia es que cada nivel de división sea exhaustivo respecto del caso en que se encuentra.

No repetir trabajo innecesariamente. Una mala organización por casos puede duplicar pasos que son comunes a todas las ramas.

Supongamos que para demostrar R desde $P \vee Q$ necesitamos primero una consecuencia S que se deduce de una hipótesis común H y después:

- bajo P, usar S para llegar a R;
- bajo Q, usar S para llegar a R.

No conviene volver a demostrar S dentro de cada caso. Es más claro obtener primero

$$H \Rightarrow S$$

y abrir la división después.

La arquitectura debería mostrar qué parte del argumento es común y dónde comienza realmente la bifurcación.

Un ejemplo con funciones. Sea

$$f : \mathbb{R} \to \mathbb{R}, \qquad f(x) = |x|.$$

Queremos demostrar que

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

para todo real x.

La definición de $|x|$ sugiere dos casos.

Caso 1: $x \geq 0$**.** Entonces

$$f(x) = x,$$

y

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

Caso 2: $x < 0$**.** Entonces

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

por lo que

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

La conclusión es la misma en ambos casos.

Podríamos también demostrarlo usando la identidad conocida

$$|x|^2 = x^2,$$

si ésta ya estuviera disponible. Pero si precisamente estamos justificando esa identidad, la definición por tramos hace natural la prueba por casos.

El orden de los casos puede importar para la claridad. Cuando un caso es excepcional y los demás siguen una regla general, suele ser mejor aislar primero la excepción.

Por ejemplo, al estudiar

$$\frac{x^2 - 1}{x - 1},$$

el valor $x = 1$ requiere atención porque el denominador se anula.

Una organización clara es:

- Caso 1: $x = 1$;
- Caso 2: $x \neq 1$.

En el segundo caso podemos factorizar y cancelar:

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

Pero esta identidad no puede aplicarse en $x = 1$ a la expresión original, porque allí no está definida.

La división por casos protege las condiciones de validez de cada paso.

Probar por casos no significa probar ejemplos. Hay una diferencia esencial entre:

- dividir un dominio en clases exhaustivas;
- comprobar algunos valores concretos.

Para demostrar una propiedad de todos los enteros no basta verificar $n = 0, 1, 2, 3$.

En cambio, demostrarla para los dos casos

$$n \text{ par} \qquad \text{o} \qquad n \text{ impar}$$

sí puede bastar, porque esos dos casos cubren todos los enteros.

Cada caso sigue siendo universal dentro de su clase.

Casos y contradicción. Una rama puede cerrarse por contradicción.

Supongamos que queremos demostrar una propiedad R para todo objeto y dividimos:

- Caso 1: P;
- Caso 2: $\neg P$.

Tal vez en el segundo caso las hipótesis globales hagan imposible $\neg P$. Entonces esa rama queda cerrada porque no contiene objetos admisibles.

No hay obligación de producir una construcción positiva en cada caso si una rama resulta vacía bajo las hipótesis.

Comprobación. Sea $f : \mathbb{R} \to \mathbb{R}$ definida por

$$f(x) = \begin{cases} x + 1, & x \geq 0, \\ 1 - x, & x < 0. \end{cases}$$

Demuestra que

$$f(x) \geq 1$$

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

Respuesta razonada.

Tomamos $x \in \mathbb{R}$ arbitrario.

Caso 1: $x \geq 0$**.** Entonces

$$f(x) = x + 1 \geq 1.$$

Caso 2: $x < 0$**.** Entonces

$$-x > 0,$$

y por tanto

$$f(x) = 1 - x = 1 + (-x) > 1.$$

En ambos casos,

$$f(x) \geq 1.$$

Los casos son exhaustivos porque todo real satisface exactamente una de las condiciones

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

La prueba no depende de probar valores particulares; en cada rama trabajamos con un elemento arbitrario sujeto sólo a la condición del caso.

Demostración por casos

Una prueba por casos es válida cuando:

1. las alternativas cubren todos los objetos permitidos;
2. en cada alternativa se demuestra la misma conclusión;
3. cada paso utilizado dentro de un caso respeta las hipótesis de ese caso.

Los casos pueden solaparse. La exhaustividad es obligatoria; la exclusividad no.

Hacia la existencia. Hasta aquí hemos estudiado arquitecturas cuyo objetivo principal era una implicación, una equivalencia o una afirmación universal.

Cuando el objetivo cambia a

$$\exists x \in A \; P(x),$$

la tarea ya no consiste en controlar todos los objetos, sino en justificar que **al menos uno existe**.

Eso abre una nueva distinción: encontrar un testigo, construirlo o demostrar que debe existir aunque no lo exhibamos explícitamente.

Ése será el siguiente paso.

6.7. Existencia: encontrar, construir o garantizar un objeto

Una afirmación existencial tiene la forma

$$\exists x \in A \; P(x).$$

Para demostrarla debemos justificar que **al menos un** elemento del dominio satisface la propiedad.

En su forma más visible, la prueba consiste en exhibir un testigo:

$$a \in A \qquad \text{y} \qquad P(a).$$

Pero no todas las pruebas de existencia se parecen. A veces el objeto está delante de nosotros y sólo hay que verificarlo. A veces debemos construirlo a partir de las hipótesis. Y, en otros casos, un argumento demuestra que necesariamente existe algún objeto, aunque la prueba no nos diga cuál es de manera explícita.

Por eso conviene separar tres preguntas:

1. ¿qué objeto necesitamos?
2. ¿cómo sabemos que pertenece al dominio correcto?
3. ¿cómo verificamos que satisface todas las condiciones?

Un testigo no es sólo un nombre. Consideremos:

Existe un real x tal que

$$x^2 - 5x + 6 = 0.$$

Podemos proponer

$$x = 2.$$

Pero una prueba no termina con «tomemos $x = 2$». Debemos verificar:

$$2 \in \mathbb{R}$$

y

$$2^2 - 5 \cdot 2 + 6 = 4 - 10 + 6 = 0.$$

Ahora sí hemos demostrado la existencia.

La estructura es:

proponer un testigo
+ verificar dominio
+ verificar propiedad.

Si el objetivo incluye varias condiciones,

$$\exists x \in A \, \big(P(x) \land Q(x) \land R(x)\big),$$

el testigo debe cumplirlas todas.

Un ejemplo con conjuntos. Sean A y B conjuntos no vacíos. Queremos demostrar:

Si $A \cap B \neq \varnothing$, entonces existe un elemento que pertenece simultáneamente a A y a B.

En símbolos,

$$A \cap B \neq \varnothing \Rightarrow \exists x \, (x \in A \land x \in B).$$

Aquí la existencia ya está contenida en la hipótesis.

La condición

$$A \cap B \neq \varnothing$$

significa precisamente que existe algún

$$x \in A \cap B.$$

Y por definición de intersección,

$$x \in A \qquad \text{y} \qquad x \in B.$$

No hemos construido el elemento mediante una fórmula; lo hemos obtenido al desplegar el significado de «no vacío».

Construir un testigo a partir de las hipótesis. Ahora consideremos una situación donde el testigo no viene dado.

Sean $a, b \in \mathbb{R}$ con

$$a < b.$$

Queremos demostrar:

Existe un real c tal que

$$a < c < b.$$

Una elección natural es el punto medio:

$$c = \frac{a+b}{2}.$$

Debemos verificar las dos desigualdades.

De

$$a < b$$

obtenemos

$$2a < a + b,$$

y dividiendo por $2 > 0$,

$$a < \frac{a+b}{2} = c.$$

También,

$$a + b < 2b,$$

de donde

$$c = \frac{a+b}{2} < b.$$

Por tanto,

$$a < c < b.$$

El testigo fue **construido** utilizando los datos del problema.

La idea del punto medio es contenido matemático; la forma existencial sólo nos decía que necesitábamos producir algún objeto entre a y b.

La fórmula del testigo debe estar bien definida. Construir un objeto exige comprobar que la expresión usada tiene sentido.

Supongamos que queremos resolver en un cuerpo F

$$ax = b.$$

Si $a \neq 0$, podemos proponer

$$x = a^{-1}b.$$

Esta fórmula es legítima porque la hipótesis $a \neq 0$ garantiza la existencia de a^{-1}.

Después verificamos:

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

Así hemos probado existencia de solución.

Pero si no tuviéramos $a \neq 0$, escribir $a^{-1}b$ sería ilegítimo. Una fórmula familiar no constituye un testigo si alguna de sus partes no está definida bajo las hipótesis disponibles.

Existencia por elección entre casos. A veces podemos demostrar existencia sin una fórmula única.

Queremos demostrar:

Para todo real x, existe un real $y \geq 0$ tal que

$$y^2 = x^2.$$

Podemos dividir en casos.

Si

$$x \geq 0,$$

elegimos

$$y = x.$$

Entonces $y \geq 0$ y $y^2 = x^2$.

Si

$$x < 0,$$

elegimos

$$y = -x.$$

Entonces $y > 0$ y

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

En ambos casos existe un testigo.

Aquí la existencia se demuestra mediante una construcción dependiente del caso:

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

Es, naturalmente, el valor absoluto:

$$y = |x|.$$

La prueba muestra cómo una construcción por casos puede terminar condensándose en una definición familiar.

Existencia sin exhibir el objeto de forma explícita. No toda prueba existencial termina con una fórmula como

$$x = \frac{a+b}{2}.$$

Consideremos una función continua

$$f : [0,1] \to \mathbb{R}$$

tal que

$$f(0) < 0 \qquad \text{y} \qquad f(1) > 0.$$

Un resultado del análisis —el teorema del valor intermedio— afirma que existe

$$c \in (0,1)$$

tal que

$$f(c) = 0.$$

Si utilizamos ese teorema, podemos demostrar existencia de c sin disponer necesariamente de una expresión cerrada para él.

La prueba sería:

- f es continua en $[0,1]$;
- 0 está entre $f(0)$ y $f(1)$;
- por el teorema del valor intermedio existe $c \in (0,1)$ con $f(c) = 0$.

El objeto queda garantizado, aunque no se calcule.

No desarrollaremos aquí el teorema del valor intermedio; pertenece al análisis. El ejemplo sólo muestra una distinción lógica:

> **demostrar existencia no siempre significa proporcionar una fórmula explícita para el objeto.**

Existencia constructiva y existencia no constructiva. En sentido amplio, llamamos **constructiva** a una prueba que proporciona un procedimiento, una fórmula o un objeto verificable que satisface la propiedad.

Por ejemplo,

$$c = \frac{a+b}{2}$$

construye explícitamente un real entre a y b.

Una prueba **no constructiva**, en el sentido elemental que usaremos aquí, demuestra que algún objeto debe existir sin identificarlo necesariamente de forma concreta.

No entraremos en las distinciones más profundas de lógica constructiva. Nos basta la diferencia práctica entre:

"aquí está un testigo"

y

"algún testigo debe existir".

Ambas pueden ser demostraciones matemáticas válidas dentro del marco clásico.

Un ejemplo clásico de existencia no constructiva elemental. Demostremos:

Existen números irracionales a y b tales que

$$a^b$$

es racional.

Consideremos

$$\alpha = (\sqrt{2})^{\sqrt{2}}.$$

Hay dos posibilidades.

Caso 1: α es racional.

Entonces tomamos

$$a = \sqrt{2}, \qquad b = \sqrt{2}.$$

Ambos son irracionales y

$$a^b = \alpha$$

es racional.

Caso 2: α es irracional.

Tomamos entonces

$$a = \alpha, \qquad b = \sqrt{2}.$$

Ambos son irracionales y

$$a^b = \left((\sqrt{2})^{\sqrt{2}}\right)^{\sqrt{2}} = (\sqrt{2})^2 = 2,$$

que es racional.

En cualquiera de los dos casos existen irracionales a y b con potencia racional.

La prueba establece existencia, pero no decide cuál de las dos parejas desempeña el papel de testigo sin información adicional sobre α.

Éste es precisamente el rasgo no constructivo elemental que queríamos observar.

No disponer de una fórmula no demuestra que sea imposible construir un testigo. La distinción anterior describe **lo que proporciona un argumento**, no una propiedad inalterable del objeto cuya existencia se afirma. Una demostración puede garantizar un testigo sin nombrarlo; otra demostración del mismo resultado podría llegar a exhibirlo. Tampoco debemos identificar automáticamente «no tenemos una expresión cerrada» con «no hay construcción posible»: un procedimiento justificado que determine el objeto buscado puede ser una construcción, aunque no termine en una fórmula elemental.

En la prueba por casos con $\alpha = (\sqrt{2})^{\sqrt{2}}$ no se ha decidido cuál de las dos parejas presentadas cumple la condición. Eso basta para que *ese argumento*, tal como está escrito, no seleccione un testigo de modo explícito. No demuestra que nadie pueda encontrar uno mediante otra idea. De modo parecido, cuando invocamos el teorema del valor intermedio, sólo vimos que nuestra invocación garantizaba un cero de f; no examinamos si, bajo datos adicionales sobre f, sería posible calcular o aproximar tal cero mediante un procedimiento.

Comprobación. Al leer «existe un elemento x con la propiedad $P(x)$» sin que aparezca una fórmula para x, ¿podemos concluir que ninguna prueba dará jamás un testigo explícito?

Respuesta razonada. No. La existencia se ha establecido, pero la falta de un testigo exhibido caracteriza la información aportada *hasta ese punto* de la argumentación. Para concluir que no existe ningún procedimiento de construcción habría que formular y justificar otra afirmación, mucho más fuerte. En este capítulo no necesitamos esa teoría: debemos identificar qué nos autoriza a decir la prueba que realmente tenemos.

Existencia no significa unicidad. Si encontramos un testigo, hemos probado

$$\exists x \in A\ P(x).$$

No hemos probado

$$\exists! x \in A\ P(x).$$

Por ejemplo,

$$x^2 = 1$$

tiene al menos una solución real:

$$x = 1.$$

Pero también

$$x = -1$$

es solución.

Exhibir 1 demuestra existencia y nada más.

Esta separación será el centro de §6.8.

El dominio forma parte de la existencia. Consideremos la ecuación

$$x^2 + 1 = 0.$$

Sobre los reales,

$$\nexists x \in \mathbb{R}\ (x^2 + 1 = 0).$$

Sobre los complejos,

$$\exists x \in \mathbb{C}\ (x^2 + 1 = 0),$$

pues

$$i^2 + 1 = 0.$$

La misma expresión algebraica puede definir un problema existencial distinto cuando cambia el dominio.

Por eso, una prueba de existencia debe indicar con precisión **dónde vive el testigo**.

Un error frecuente: verificar la propiedad y olvidar el dominio. Supongamos que se pide demostrar:

Existe un entero n tal que

$$n^2 = 9.$$

Elegimos $n = 3$ y verificamos

$$3^2 = 9.$$

También sabemos

$$3 \in \mathbb{Z}.$$

Aquí ambas obligaciones son triviales, pero ambas existen.

En problemas más abstractos, comprobar sólo una ecuación puede no bastar. El candidato puede satisfacer la identidad y, sin embargo, no pertenecer al conjunto permitido.

Existencia parametrizada. Muchas afirmaciones tienen la forma

$$\forall x \in A \, \exists y \in B \; P(x, y).$$

Aquí el testigo y puede depender de x.

Para demostrar la afirmación debemos:

1. fijar $x \in A$ arbitrario;
2. construir o garantizar un $y \in B$ apropiado;
3. verificar $P(x, y)$.

Por ejemplo:

Para todo real x, existe un real y tal que

$$y > x.$$

Tomamos un real arbitrario x y elegimos

$$y = x + 1.$$

Entonces

$$y \in \mathbb{R}$$

y

$$y = x + 1 > x.$$

La elección depende de x, lo cual es perfectamente legítimo porque el cuantificador existencial aparece **después** del universal.

Esto contrasta con

$$\exists y \in \mathbb{R} \, \forall x \in \mathbb{R} \; (y > x),$$

que pediría un único real mayor que todos los reales y es falso.

La arquitectura de existencia debe respetar el orden de cuantificadores.

Construir hacia atrás, verificar hacia adelante. En muchas pruebas existenciales ocurre una diferencia entre descubrimiento y presentación.

Para encontrar un testigo podemos trabajar hacia atrás.

Si queremos

$$f(x) = b,$$

quizá resolvamos informalmente la ecuación y obtengamos un candidato.

Pero la prueba final debe verificar hacia adelante:

1. definimos el candidato;
2. comprobamos que está permitido;
3. sustituimos;
4. demostramos que cumple la propiedad.

Por ejemplo, para

$$ax + b = c$$

con $a \neq 0$, podemos descubrir

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

La prueba no consiste en decir «despejando obtenemos». Debemos verificar:

$$a\big(a^{-1}(c - b)\big) + b = c.$$

Esta verificación convierte una intuición algebraica en una prueba existencial completa.

Comprobación. Sea $f : A \to B$ una función sobreyectiva y sea

$$S \subseteq B.$$

Demuestra que para todo $y \in S$ existe $x \in A$ tal que

$$f(x) = y.$$

Después identifica exactamente dónde se usa la hipótesis de sobreyectividad.

Respuesta razonada.

Tomemos un elemento arbitrario

$$y \in S.$$

Como

$$S \subseteq B,$$

tenemos

$$y \in B.$$

La sobreyectividad de

$$f : A \to B$$

afirma:

$$\forall b \in B \; \exists a \in A \; (f(a) = b).$$

Podemos aplicarla al elemento $y \in B$. Por tanto, existe

$$x \in A$$

tal que

$$f(x) = y.$$

La hipótesis de sobreyectividad se usa exactamente al pasar de

$$y \in B$$

a la existencia de una preimagen $x \in A$.

No necesitamos construir una fórmula para x. La definición de sobreyectividad ya garantiza su existencia.

Pruebas de existencia

Para demostrar

$$\exists x \in A\ P(x),$$

hay que justificar al menos:

1. que algún candidato pertenece a A;
2. que satisface P.

El testigo puede exhibirse, construirse a partir de los datos o quedar garantizado por un resultado previo. Existencia no implica unicidad.

Hacia la unicidad. Una prueba existencial responde:

¿hay al menos un objeto con esta propiedad?

La unicidad añade otra pregunta:

¿puede haber más de uno?

Estas obligaciones son independientes. En la siguiente sección aprenderemos a demostrar que dos candidatos cualesquiera deben coincidir.

6.8. Unicidad: existir no basta

Una afirmación de unicidad tiene la forma

$$\exists! x \in A\ P(x).$$

Como vimos antes, esta notación contiene dos obligaciones independientes:

existencia + a lo sumo uno.

La primera pide demostrar

$$\exists x \in A\ P(x).$$

La segunda pide demostrar que no puede haber dos objetos distintos con la misma propiedad. Una forma estándar de escribirla es

$$\forall x, y \in A\ \big(P(x) \wedge P(y) \Rightarrow x = y\big).$$

Por tanto,

$$\exists! x \in A\ P(x)$$

puede descomponerse como

$$(\exists x \in A\ P(x))$$
$$\wedge\, (\forall x, y \in A\ (P(x) \wedge P(y) \Rightarrow x = y))\,.$$

Esta separación es esencial porque encontrar un ejemplo demuestra existencia, pero no unicidad.

Un primer ejemplo. Queremos demostrar:

Existe un único real x tal que

$$x + 3 = 7.$$

La existencia es inmediata: tomamos

$$x = 4$$

y verificamos

$$4 + 3 = 7.$$

Para la unicidad, supongamos que x e y son dos soluciones:

$$x + 3 = 7$$

y

$$y + 3 = 7.$$

Entonces

$$x + 3 = y + 3.$$

Cancelando 3,

$$x = y.$$

Por tanto, la solución es única.

El argumento completo no termina cuando encontramos 4. La segunda parte explica por qué **cualquier** solución debe coincidir con 4, o equivalentemente, por qué dos soluciones cualesquiera deben coincidir entre sí.

Dos formas habituales de demostrar unicidad. Hay dos arquitecturas muy frecuentes.

Forma A: comparar dos candidatos. Suponemos que

$$P(x) \qquad \text{y} \qquad P(y)$$

y demostramos

$$x = y.$$

Ésta es la forma más simétrica:

$$\boxed{P(x) + P(y) \longrightarrow x = y.}$$

Forma B: fijar el candidato construido. Si en la parte de existencia ya hemos construido un objeto concreto a, podemos demostrar que cualquier otro candidato x debe satisfacer

$$x = a.$$

Esquemáticamente:

$$P(x) \longrightarrow x = a.$$

Ambas formas prueban "a lo sumo uno". La elección depende de cuál produzca una prueba más clara.

Unicidad del neutro. En el capítulo 5 vimos este patrón en una estructura algebraica.

Supongamos que una operación binaria $\star$ sobre un conjunto A tiene dos elementos neutros e y e'.

Por ser e neutro,

$$e \star e' = e'.$$

Por ser e' neutro,

$$e \star e' = e.$$

Por tanto,

$$e = e'.$$

La prueba es notablemente breve. No necesita construir el neutro; sólo muestra que, si existen dos, deben coincidir.

Así hemos probado **a lo sumo uno**, no existencia.

Para demostrar que existe un neutro único necesitamos además saber que al menos uno existe.

Unicidad de inversos. Supongamos que A tiene una operación asociativa con neutro e, y que un elemento $a \in A$ tiene dos inversos b y c:

$$a \star b = e, \qquad b \star a = e,$$

y

$$a \star c = e, \qquad c \star a = e.$$

Entonces

$$b = b \star e = b \star (a \star c).$$

Por asociatividad,

$$b \star (a \star c) = (b \star a) \star c.$$

Como $b \star a = e$,

$$(b \star a) \star c = e \star c = c.$$

Luego

$$b = c.$$

La prueba muestra otra característica frecuente: la unicidad se obtiene haciendo interactuar dos candidatos mediante la estructura disponible.

La unicidad puede depender de una hipótesis decisiva. Consideremos la ecuación

$$ax = b$$

en un cuerpo F.

Si

$$a \neq 0,$$

sabemos que existe

$$a^{-1}.$$

La solución

$$x = a^{-1}b$$

existe.

Ahora probemos unicidad.

Sean $x, y \in F$ dos soluciones:

$$ax = b$$

y

$$ay = b.$$

Entonces

$$ax = ay.$$

Como $a \neq 0$, podemos cancelar a y obtener

$$x = y.$$

La hipótesis $a \neq 0$ es decisiva.

Si $a = 0$ y $b = 0$, la ecuación se convierte en

$$0x = 0,$$

que es satisfecha por todo $x \in F$.

Hay existencia, pero no unicidad siempre que el cuerpo tenga al menos dos elementos, lo que está garantizado por

$$0 \neq 1.$$

La unicidad no es una propiedad automática de una ecuación; depende de las hipótesis.

Existencia y unicidad pueden probarse en órdenes distintos. A veces resulta natural demostrar primero existencia y después unicidad.

Otras veces es más fácil comenzar con "a lo sumo uno" y construir el objeto al final.

Por ejemplo, si queremos demostrar que una función $f : A \to B$ tiene una única inversa, podemos primero mostrar:

Si g y h son inversas de f, entonces $g = h$.

Después, en otro argumento, demostramos que una inversa existe bajo ciertas hipótesis.

El orden de presentación no altera la obligación lógica:

$$\begin{aligned}\text{exactamente uno} = {} & \text{al menos uno} \\ & + \text{a lo sumo uno.}\end{aligned}$$

Unicidad mediante inyectividad. La inyectividad es una herramienta natural para unicidad.

Supongamos que

$$f : A \to B$$

es inyectiva y queremos resolver

$$f(x) = b.$$

Si existen dos soluciones $x_1, x_2 \in A$, entonces

$$f(x_1) = b = f(x_2).$$

La inyectividad da

$$x_1 = x_2.$$

Por tanto, para cada $b \in B$ hay **a lo sumo una** solución de

$$f(x) = b.$$

Si además f es sobreyectiva, entonces para cada $b \in B$ hay al menos una solución. Juntando ambas propiedades, para cada $b \in B$ existe una única preimagen.

Así, una función biyectiva satisface:

$$\forall b \in B \; \exists! x \in A \; (f(x) = b).$$

Ésta es precisamente la base de la existencia de la función inversa.

Unicidad mediante una desigualdad. No toda prueba de unicidad usa igualdad algebraica o inyectividad.

Supongamos que queremos demostrar que un conjunto ordenado tiene a lo sumo un máximo.

Sean m y n dos máximos de un mismo conjunto S.

Como m es máximo y $n \in S$,

$$n \leq m.$$

Como n es máximo y $m \in S$,

$$m \leq n.$$

Por antisimetría,

$$m = n.$$

La estructura es la misma: dos candidatos se comparan hasta forzar coincidencia.

Unicidad mediante contradicción. También podemos demostrar "a lo sumo uno" suponiendo que existen dos candidatos distintos.

Supongamos que $P(x)$ describe la solución de un problema y queremos probar unicidad.

Podemos asumir:

$$P(x), \qquad P(y), \qquad x \neq y,$$

y derivar una contradicción.

Entonces no pueden existir dos candidatos distintos.

Por ejemplo, sea $f : A \to B$ inyectiva. Para un $b \in B$, supongamos que hay dos preimágenes distintas:

$$f(x) = b, \qquad f(y) = b, \qquad x \neq y.$$

De las dos primeras igualdades obtenemos

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

La inyectividad da

$$x = y,$$

contradicción.

Por tanto, la preimagen, si existe, es única.

"A lo sumo uno" admite el caso de ninguno. La afirmación

$$\forall x, y \in A \; (P(x) \land P(y) \Rightarrow x = y)$$

no garantiza que exista algún x con $P(x)$.

Puede cumplirse vacíamente.

Por ejemplo, en los reales la ecuación

$$x^2 + 1 = 0$$

no tiene soluciones.

Sin embargo, es cierto que tiene **a lo sumo una** solución si formulamos ingenuamente esa condición como

$$\forall x, y \in \mathbb{R}, \quad (x^2 + 1 = 0) \land (y^2 + 1 = 0) \Rightarrow x = y.$$

No hay candidatos que puedan violarla.

De hecho, la ecuación tiene cero soluciones, no una.

Este ejemplo muestra por qué "a lo sumo uno" no debe confundirse con "exactamente uno".

Unicidad parametrizada. Muchas afirmaciones tienen la forma

$$\forall a \in A \; \exists! b \in B \; P(a, b).$$

Para cada entrada a, debemos demostrar:

1. existencia de algún b;
2. unicidad de ese b.

El objeto único puede depender de a.

Por ejemplo, para todo real a existe un único real b tal que

$$a + b = 0.$$

La existencia se obtiene tomando

$$b = -a.$$

Para la unicidad, supongamos

$$a + b = 0$$

y

$$a + c = 0.$$

Entonces

$$a + b = a + c.$$

Cancelando a,

$$b = c.$$

Así,

$$\forall a \in \mathbb{R} \; \exists! b \in \mathbb{R} \; (a + b = 0).$$

La notación $-a$ queda legitimada como nombre de ese objeto único.

La unicidad legitima notación. Éste es un papel conceptual importante.

Antes de demostrar unicidad, una descripción puede referirse a varios objetos.

Después de demostrar existencia y unicidad, podemos introducir una notación definida sin ambigüedad.

Por ejemplo:

- "el neutro aditivo";
- "el inverso multiplicativo de a";
- "la imagen inversa" en un contexto apropiado;
- "la solución" de una ecuación bajo hipótesis que garantizan unicidad.

La palabra **el** tiene contenido matemático: presupone que el objeto está determinado de manera única.

Un ejemplo con funciones. Supongamos que $f : A \to B$ es biyectiva.
Queremos demostrar:

Para todo $b \in B$ existe un único $a \in A$ tal que

$$f(a) = b.$$

Existencia. Como f es sobreyectiva, para todo $b \in B$ existe al menos un $a \in A$ con

$$f(a) = b.$$

Unicidad. Supongamos que $a_1, a_2 \in A$ satisfacen

$$f(a_1) = b$$

y

$$f(a_2) = b.$$

Entonces

$$f(a_1) = f(a_2).$$

Como f es inyectiva,

$$a_1 = a_2.$$

Por tanto,

$$\forall b \in B \; \exists! a \in A \; (f(a) = b).$$

Ahora podemos definir

$$f^{-1}(b)$$

como ese único elemento de A cuya imagen es b.

La función inversa no aparece mágicamente por invertir una flecha; su definición se apoya en una existencia y una unicidad previamente justificadas.

Comprobación. Sea F un cuerpo y fija $a, b \in F$ con

$$a \neq 0.$$

Demuestra que existe un único $x \in F$ tal que

$$ax + b = 0.$$

Respuesta razonada.
Primero probamos existencia.
Como $a \neq 0$, existe a^{-1}. Tomamos

$$x = -a^{-1}b.$$

Entonces

$$\begin{aligned} ax + b &= a(-a^{-1}b) + b \\ &= -(aa^{-1})b + b \\ &= -b + b \\ &= 0. \end{aligned}$$

Por tanto, existe al menos una solución.

Ahora probamos unicidad.

Sean $x, y \in F$ dos soluciones:

$$ax + b = 0$$

y

$$ay + b = 0.$$

Entonces

$$ax + b = ay + b.$$

Cancelando aditivamente b,

$$ax = ay.$$

Como $a \neq 0$, la cancelación multiplicativa da

$$x = y.$$

Por tanto, existe una única solución.

Podemos escribirla como

$$x = -a^{-1}b.$$

La fórmula obtenida en la parte de existencia queda ahora identificada como **la** solución, no sólo como una solución.

Unicidad

Para demostrar

$$\exists! x \in A \; P(x),$$

hay dos obligaciones:

1. **existencia:** demostrar que al menos un objeto satisface P;
2. **unicidad:** demostrar que dos objetos cualesquiera que satisfacen P deben ser iguales.

"A lo sumo uno" no implica "existe uno". Ambas partes deben estar justificadas.

Hacia la integración de arquitecturas. Ya tenemos las formas fundamentales del capítulo:

- prueba directa;
- dos direcciones;
- contraposición;
- contradicción;
- casos;
- existencia;
- unicidad.

La siguiente tarea será aprender a reconocer cómo pueden combinarse y cuándo una afirmación debe ser refutada en lugar de demostrada.

Eso nos permitirá cerrar el mapa de arquitecturas antes de pasar a los ejercicios.

6.9. Integrar arquitecturas: elegir, combinar y saber cuándo refutar

Ya conocemos varias formas fundamentales de prueba. El paso siguiente no consiste en memorizar una tabla más grande, sino en aprender a reconocer **qué obligación domina el problema** y qué arquitectura permite cumplirla con mayor claridad.

No existe un algoritmo que, leyendo únicamente los símbolos del enunciado, produzca la demostración. La forma lógica orienta; el contenido matemático decide qué camino funciona.

Por eso, ante un nuevo problema, una pregunta más útil que

> «¿Qué método tengo que usar?»

es

> **«¿Qué debo dejar demostrado y qué información me dan realmente las hipótesis?»**

A partir de ahí podemos decidir si conviene avanzar directamente, cambiar a una forma equivalente, dividir en casos, construir un objeto o buscar una refutación.

La forma exterior organiza primero. Consideremos una afirmación con varias capas:

$$\forall x \in A\ (P(x) \Rightarrow \exists! y \in B\ Q(x,y)).$$

Antes de pensar en detalles, podemos leer sus obligaciones de fuera hacia dentro:

1. tomar un $x \in A$ arbitrario;
2. suponer $P(x)$;
3. demostrar existencia de algún $y \in B$ con $Q(x,y)$;
4. demostrar que cualquier otro $y' \in B$ con $Q(x,y')$ coincide con y.

El argumento final puede contener una construcción para la existencia y una prueba directa o por contradicción para la unicidad.

La afirmación no tiene «un método». Tiene una **arquitectura compuesta**.

Un ejemplo integrador: una involución es biyectiva. Sea

$$f : A \to A$$

una función que satisface

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

para todo $x \in A$.

Queremos demostrar que f es biyectiva.

El objetivo “biyectiva” se descompone en dos obligaciones:

$$\text{inyectiva} \qquad \text{y} \qquad \text{sobreyectiva}.$$

Las dos partes requieren recorridos distintos.

Inyectividad. Tomemos $x, y \in A$ y supongamos

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

Aplicando f a ambos lados,

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

La hipótesis de involución da

$$x = y.$$

Por tanto, f es inyectiva.

Sobreyectividad. Tomemos ahora un elemento arbitrario

$$y \in A.$$

Para demostrar sobreyectividad necesitamos construir una preimagen.

Elegimos

$$x = f(y).$$

Entonces

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

Por tanto, todo $y \in A$ tiene una preimagen y f es sobreyectiva.

Hemos concluido que f es biyectiva.

La demostración combina dos arquitecturas:

- una prueba directa para la inyectividad;
- una prueba existencial para la sobreyectividad.

Además, la misma hipótesis

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

cumple dos funciones distintas según el objetivo que estamos resolviendo.

La técnica se elige localmente. En una prueba larga conviene decidir la arquitectura **obligación por obligación**.

Supongamos que queremos demostrar

$$P \iff Q.$$

Podemos probar

$$P \Rightarrow Q$$

directamente y

$$Q \Rightarrow P$$

por contraposición.

No hay ninguna exigencia de simetría metodológica.

Del mismo modo, una doble inclusión puede tener:

- una dirección inmediata por definición;
- otra que necesite una prueba por casos.

Una prueba de unicidad puede comenzar directamente comparando candidatos y cerrar un paso intermedio por contradicción.

La combinación es normal. Lo importante es que cada transición esté justificada y que todas las obligaciones del objetivo queden cerradas.

Elegir una arquitectura no es etiquetar después. Podríamos leer una demostración terminada y decir:

«Aquí hubo contradicción.»

Eso describe una parte de su forma, pero todavía no explica por qué esa forma era útil.

Una buena lectura estratégica pregunta:

- ¿qué hacía difícil la prueba directa?;
- ¿qué información apareció al negar la conclusión?;
- ¿qué disyunción obligó a separar casos?;
- ¿qué objeto necesitaba construirse?;
- ¿qué comparación forzó la unicidad?

Así empezamos a pasar de reconocer pruebas a entender por qué funcionan.

El capítulo 7 llevará esta pregunta mucho más lejos. Aquí sólo necesitamos que la elección de arquitectura sea consciente y verificable.

Un mapa inicial, no un algoritmo. La siguiente tabla puede servir como primera orientación:

Forma del objetivo o situación	Primera pregunta útil
$P \Rightarrow Q$	¿Puedo avanzar directamente desde P hasta Q?
$P \Rightarrow Q$ con Q difícil de alcanzar	¿Es más manejable $\neg Q \Rightarrow \neg P$?
negación del objetivo muy restrictiva	¿Conduce naturalmente a una contradicción?
$P \Longleftrightarrow Q$	¿He separado y probado las dos direcciones?
$A = B$	¿Puedo demostrar las dos inclusiones?
hipótesis $P \vee Q$ o dominio naturalmente dividido	¿Los casos cubren todo lo posible?
$\exists x\, P(x)$	¿Puedo exhibir, construir o garantizar un testigo?
$\exists! x\, P(x)$	¿He separado existencia y unicidad?
afirmación universal sospechosa	¿Existe un contraejemplo que satisfaga las hipótesis y viole la conclusión?

La tabla no decide la prueba. Sólo transforma una consigna vaga en preguntas matemáticas más precisas.

Cuando la afirmación es falsa. Hasta ahora la mayoría de los ejercicios nos han pedido demostrar afirmaciones verdaderas. Pero ante una afirmación nueva no debemos presuponer que existe una prueba.

Consideremos:

Para toda función $f : A \to B$ y todos los subconjuntos $S, T \subseteq A$,

$$f(S \cap T) = f(S) \cap f(T).$$

Sabemos que siempre vale una inclusión:

$$f(S \cap T) \subseteq f(S) \cap f(T).$$

¿Vale necesariamente la igualdad?

No.

Tomemos

$$A = \{1, 2\}, \qquad B = \{0\},$$

y definamos

$$f(1) = 0, \qquad f(2) = 0.$$

Elijamos

$$S = \{1\}, \qquad T = \{2\}.$$

Entonces

$$S \cap T = \varnothing,$$

de modo que

$$f(S \cap T) = f(\varnothing) = \varnothing.$$

Pero

$$f(S) = \{0\}$$

y

$$f(T) = \{0\},$$

por lo que

$$f(S) \cap f(T) = \{0\}.$$

Así,

$$f(S \cap T) \neq f(S) \cap f(T).$$

La afirmación universal queda refutada.

Qué debe verificar un contraejemplo. Un contraejemplo completo no es un objeto que «parece problemático». Debe pasar tres controles:

1. **pertenecer al dominio correcto**;
2. **satisfacer todas las hipótesis**;
3. **hacer falsa la conclusión**.

En el ejemplo anterior:

- f es efectivamente una función $A \to B$;
- S y T son subconjuntos de A;
- la igualdad propuesta falla.

Sólo después de esas verificaciones tenemos una refutación.

La búsqueda sistemática de contraejemplos se estudiará en el capítulo 9. Aquí nos basta entender su función lógica.

Un solo contraejemplo basta para una universal. Si una afirmación tiene la forma

$$\forall x \in A \; P(x),$$

su negación es

$$\exists x \in A \; \neg P(x).$$

Por tanto, basta **un** testigo de fallo.

No necesitamos demostrar que la propiedad falla muchas veces.

Esto produce una asimetría fundamental:

- para probar una universal debemos controlar todos los elementos;
- para refutarla basta uno.

En cambio, encontrar muchos ejemplos favorables nunca demuestra por sí solo una universal.

Un ejemplo favorable no es una prueba. Supongamos que alguien conjetura:

Para todo entero positivo n,

$$n^2 + n + 41$$

es primo.

Podemos comprobar muchos valores y encontrar números primos.

Eso puede apoyar la conjetura, pero no demuestra la universal.

De hecho, para

$$n = 41,$$

obtenemos

$$41^2 + 41 + 41 = 41(41 + 1 + 1) = 41 \cdot 43,$$

que no es primo.

Un solo valor destruye la afirmación universal.

La lección no es buscar siempre números grandes o fórmulas famosas. Es distinguir entre:

evidencia por ejemplos

y

demostración.

Antes de refutar, identificar exactamente la universal. Un contraejemplo debe atacar la afirmación correcta.

Si el enunciado dice

Para todo $x \in A$, **si** $P(x)$, entonces $Q(x)$,

el contraejemplo debe satisfacer

$$P(x)$$

y violar

$$Q(x).$$

Un elemento con $\neg P(x)$ no refuta la implicación.

Por ejemplo, para refutar

Si un entero es múltiplo de 4, entonces es múltiplo de 8,

no sirve elegir 6: no es múltiplo de 4.

Sí sirve elegir

4,

porque satisface la hipótesis y no la conclusión.

Así, un contraejemplo a

$$P \Rightarrow Q$$

debe realizar exactamente

$$P \wedge \neg Q.$$

Equivalencias y cadenas: conservar el sentido de cada flecha. Al integrar técnicas es fácil escribir demasiado deprisa una cadena como

$$P \iff Q \iff R.$$

Cada doble flecha afirma dos implicaciones.

Si alguna transformación sólo está justificada en una dirección, debemos escribir

$$P \Rightarrow Q$$

y conservar esa orientación.

Por ejemplo, sobre los reales,

$$x = y \Rightarrow x^2 = y^2,$$

pero la recíproca es falsa:

$$x^2 = y^2$$

permite

$$x = y \qquad \text{o} \qquad x = -y.$$

Cambiar una flecha simple por una doble puede convertir un razonamiento correcto en uno falso.

Equivalencias múltiples: escoger suficientes conexiones. Si queremos demostrar equivalencia entre tres condiciones

$$P, \quad Q, \quad R,$$

no necesitamos demostrar todas las implicaciones posibles.

Un ciclo

$$P \Rightarrow Q, \qquad Q \Rightarrow R, \qquad R \Rightarrow P$$

es suficiente.

También puede bastar demostrar

$$P \iff Q$$

y

$$Q \iff R.$$

La transitividad completa las conexiones restantes.

La elección debe minimizar trabajo sin ocultar dependencias.

Diagnosticar una prueba que no avanza. Si una prueba se atasca, conviene revisar la arquitectura antes de hacer más cálculos.

Podemos preguntar:

- ¿he desplegado correctamente la definición del objetivo?;
- ¿estoy intentando demostrar una hipótesis en vez de usarla?;
- ¿me falta una dirección de una equivalencia?;
- ¿he elegido casos exhaustivos?;
- ¿estoy intentando probar una universal que quizá sea falsa?;
- ¿la contraposición ofrece información más concreta?;
- ¿la negación del objetivo produce una contradicción útil?;
- ¿en una existencia, sé qué tipo de testigo necesito?;
- ¿en una unicidad, estoy comparando realmente dos candidatos?

Estas preguntas no resuelven el problema, pero pueden localizar el tipo de bloqueo.

La demostración final y la búsqueda de la demostración. Hay una diferencia entre

encontrar una prueba

y

presentar una prueba.

Para encontrarla podemos probar ejemplos, trabajar hacia atrás, cambiar de representación, formular conjeturas auxiliares y abandonar intentos.

La prueba final elimina la mayor parte de ese recorrido y conserva una cadena verificable.

En este capítulo hemos estudiado principalmente esa cadena final.

En el capítulo 7 nos preguntaremos qué hacer cuando todavía no sabemos cuál debería ser esa cadena.

Comprobación integradora. Sea

$$f : A \to B.$$

Considera las siguientes afirmaciones:

1. f es biyectiva.
2. Para todo $b \in B$ existe un único $a \in A$ tal que $f(a) = b$.

Demuestra que son equivalentes e identifica qué parte usa sobreyectividad y cuál usa inyectividad.

Respuesta razonada.

De 1 a 2. Supongamos que f es biyectiva.

Sea $b \in B$ arbitrario.

Como f es sobreyectiva, existe

$$a \in A$$

tal que

$$f(a) = b.$$

Esto prueba existencia.

Para unicidad, supongamos que

$$f(a_1) = b$$

y

$$f(a_2) = b.$$

Entonces

$$f(a_1) = f(a_2).$$

Como f es inyectiva,

$$a_1 = a_2.$$

Por tanto, para cada $b \in B$ existe una única preimagen.

De 2 a 1. Supongamos ahora que para todo $b \in B$ existe un único $a \in A$ con

$$f(a) = b.$$

La existencia de una preimagen para cada b demuestra directamente que f es sobreyectiva.

Para probar inyectividad, sean $x, y \in A$ y supongamos

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

Llamemos

$$b = f(x) = f(y).$$

Entonces x e y son dos preimágenes de b.

Por la unicidad supuesta,

$$x = y.$$

Así, f es inyectiva.

Por tanto, f es biyectiva.

La equivalencia integra varias obligaciones: dos direcciones, existencia, unicidad, sobreyectividad e inyectividad. Ninguna etiqueta aislada describe por completo la demostración.

Elegir arquitectura

La forma del objetivo indica obligaciones, no una receta única.

Una prueba puede combinar varias arquitecturas. Antes de calcular, identifica qué debes demostrar; después decide localmente qué forma de argumento hace visible el paso siguiente.

Si una afirmación universal es falsa, la respuesta correcta no es una prueba más ingeniosa, sino un contraejemplo válido.

Cierre del mapa conceptual. Con esto queda completo el repertorio básico de arquitecturas que necesitaremos:

$$\begin{aligned} \text{directa} &\longrightarrow \text{dos direcciones} \\ &\longrightarrow \text{contraposición} \\ &\longrightarrow \text{contradicción} \longrightarrow \text{casos} \\ &\longrightarrow \text{existencia} \longrightarrow \text{unicidad}. \end{aligned}$$

No forman una jerarquía ni un orden obligatorio. Son formas recurrentes de organizar obligaciones matemáticas.

El paso final del capítulo será practicarlas hasta poder reconocer no sólo **cómo** se ejecutan, sino también **qué error comete una prueba defectuosa y por qué otra arquitectura resulta más adecuada**.

6.10. Ejercicios y soluciones desarrolladas

Los ejercicios están organizados para retirar gradualmente el andamiaje. En los primeros importa reconocer la obligación antes de demostrar; en los últimos hay que decidir qué arquitectura conviene y justificar esa elección.

Familia I. Reconocer la obligación. Ejercicio 1. Antes de empezar. Para cada objetivo, escribe qué tendría que quedar demostrado antes de intentar resolverlo:

1. $A \subseteq B$;
2. $P \iff Q$;
3. $\exists x \in A\ P(x)$;
4. $\exists! x \in A\ P(x)$;
5. refutar $\forall x \in A\ P(x)$;
6. demostrar $P \Rightarrow Q$ por contraposición.

No hagas todavía ninguna prueba concreta.

Ejercicio 2. Original, recíproca y contraposición. Considera:

> Si $f : A \to B$ es inyectiva, entonces cada elemento de B tiene a lo sumo una preimagen.

Escribe:

1. la afirmación en cuantificadores;
2. su contraposición;
3. una afirmación recíproca natural.

Decide cuáles son equivalentes a la original y cuáles requieren una hipótesis adicional para ser verdaderas.

Ejercicio 3. ¿Cubren realmente todos los casos? Para cada división del dominio $\mathbb{R}$, decide si los casos son exhaustivos y si son mutuamente excluyentes:

1. $x > 0$ / $x < 0$;
2. $x \geq 0$ / $x \leq 0$;
3. $x < 1$ / $x \geq 1$;
4. $x \leq 2$ / $x \geq 1$.

Explica cuáles podrían sostener una prueba por casos de una afirmación universal sobre todos los reales.

Familia II. Ejecutar una arquitectura. Ejercicio 4. Preimagen y monotonía. Sea $f : A \to B$ y sean $U, V \subseteq B$ con

$$U \subseteq V.$$

Demuestra directamente que

$$f^{-1}(U) \subseteq f^{-1}(V).$$

Haz explícito en qué momento usas la definición de preimagen y en qué momento la hipótesis $U \subseteq V$.

Ejercicio 5. Una equivalencia sobre inyectividad. Sea $f : A \to B$. Demuestra que son equivalentes:

1. f es inyectiva;

2. para todo $b \in B$, el conjunto

$$f^{-1}(\{b\})$$

tiene a lo sumo un elemento.

Separa claramente las dos direcciones.

Ejercicio 6. Contraposición funcional. Sean

$$f : A \to B, \qquad g : B \to C.$$

Demuestra **por contraposición**:

Si $g \circ f$ es inyectiva, entonces f es inyectiva.

No uses como prueba principal la demostración directa ya vista en el capítulo.

Ejercicio 7. Una prueba por casos. Demuestra que para todo real x,

$$|x - 2| \geq 2 - x.$$

Elige casos exhaustivos a partir de la definición del valor absoluto. Indica si son también mutuamente excluyentes.

Familia III. Existencia y unicidad. Ejercicio 8. Construir un punto intermedio distinto. Sean $a, b \in \mathbb{R}$ con

$$a \neq b.$$

Demuestra que existe $c \in \mathbb{R}$ tal que

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

Construye un testigo a partir de a y b y verifica ambas desigualdades.

Ejercicio 9. Unicidad de la inversa funcional. Sea $f : A \to B$. Supón que $g, h : B \to A$ satisfacen

$$g \circ f = \mathrm{id}_A, \qquad f \circ g = \mathrm{id}_B,$$

y también

$$h \circ f = \mathrm{id}_A, \qquad f \circ h = \mathrm{id}_B\,.$$

Demuestra que

$$g = h.$$

No presupongas que “la inversa” es única antes de demostrarlo.

Ejercicio 10. Existencia única parametrizada. Sea F un cuerpo y fija $a \in F$. Demuestra:

$$\forall b \in F\ \exists! x \in F\ (x + a = b).$$

Construye el candidato y después compara dos soluciones.

Familia IV. Reparar argumentos. Ejercicio 11. Confundir recíproca con contraposición. Alguien quiere demostrar:

> Si n es múltiplo de 6, entonces n es múltiplo de 3.

Escribe:

> «Por contraposición, supongo que n es múltiplo de 3 y demostraré que es múltiplo de 6».

Explica exactamente qué error lógico contiene. Escribe la contraposición correcta y demuestra el resultado.

Ejercicio 12. Un caso desaparecido. Se quiere demostrar que

$$x^2 \geq 0$$

para todo real x. Se presenta:

> «Si $x > 0$, entonces $x^2 > 0$. Si $x < 0$, entonces $-x > 0$ y también $x^2 = (-x)^2 > 0$. Por tanto, $x^2 \geq 0$ para todo real x».

Diagnostica el argumento y repáralo con una partición exhaustiva.

Ejercicio 13. Un testigo fuera del dominio. Se quiere probar:

> Para todo $a \in \mathbb{Z}$ existe $x \in \mathbb{Z}$ tal que
>
> $$2x = a.$$

Alguien propone

$$x = \frac{a}{2}.$$

Explica por qué esta fórmula no demuestra la afirmación. Decide si la afirmación universal es verdadera o falsa y actúa en consecuencia.

Familia V. Transferencia y elección autónoma. Ejercicio 14. Elegir arquitectura para una identidad de conjuntos. Demuestra, para conjuntos arbitrarios A, B, C,

$$A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C).$$

Antes de la prueba, indica qué arquitectura eliges y por qué. No cites simplemente una ley de De Morgan: desarrolla la justificación a partir de pertenencia.

Ejercicio 15. Una caracterización global de la inyectividad. Sea $f : A \to B$. Demuestra que son equivalentes:

1. f es inyectiva;
2. para todo subconjunto $S \subseteq A$,
$$f^{-1}(f(S)) = S.$$

Recuerda que para toda función siempre se cumple

$$S \subseteq f^{-1}(f(S)).$$

Determina dónde es necesaria la inyectividad para obtener la inclusión inversa y, en la recíproca, elige un subconjunto que permita recuperar la igualdad de dos entradas con la misma imagen.

Soluciones desarrolladas.

Solución 1. Antes de empezar.

1. Para $A \subseteq B$, debemos tomar un elemento arbitrario $x \in A$ y demostrar $x \in B$.
2. Para $P \iff Q$, debemos demostrar las dos implicaciones:
$$P \Rightarrow Q \qquad \text{y} \qquad Q \Rightarrow P.$$
3. Para $\exists x \in A\ P(x)$, debemos justificar algún testigo $a \in A$ y verificar $P(a)$.
4. Para $\exists! x \in A\ P(x)$, necesitamos existencia y además demostrar que dos candidatos cualesquiera coinciden:
$$P(x) \wedge P(y) \Rightarrow x = y.$$
5. Para refutar $\forall x \in A\ P(x)$, basta encontrar $a \in A$ con $\neg P(a)$.
6. Para demostrar $P \Rightarrow Q$ por contraposición, debemos demostrar
$$\neg Q \Rightarrow \neg P.$$

El ejercicio no pide todavía contenido matemático. Su objetivo es separar la lectura del problema de la ejecución de una técnica.

Solución 2. Original, recíproca y contraposición. La afirmación original puede escribirse como

$$\begin{aligned} f \text{ inyectiva} \Rightarrow\ & \forall b \in B, \\ & \forall x, y \in A, \\ & f(x) = b \wedge f(y) = b \\ & \Rightarrow x = y. \end{aligned}$$

La segunda parte expresa que cada b tiene a lo sumo una preimagen.

Una contraposición natural de la afirmación completa es:

Si existe $b \in B$ con dos preimágenes distintas, entonces f no es inyectiva.

En símbolos, si existen $x \neq y$ tales que

$$f(x) = b = f(y),$$

entonces

$$f(x) = f(y)$$

con $x \neq y$, lo que viola la inyectividad. Esta contraposición es equivalente a la original.

Una recíproca natural es:

Si cada $b \in B$ tiene a lo sumo una preimagen, entonces f es inyectiva.

En este caso la recíproca también es verdadera, pero no por ser recíproca: necesita prueba. Si

$$f(x) = f(y),$$

llamamos $b = f(x) = f(y)$. Entonces x e y son dos preimágenes de b y, por la hipótesis "a lo sumo una", deben coincidir. Por tanto, f es inyectiva.

No hace falta sobreyectividad. "A lo sumo una preimagen" permite que algunos $b \in B$ no tengan ninguna.

Solución 3. ¿Cubren realmente todos los casos?.

1. $x > 0\ /\ x < 0$ no es exhaustivo: falta $x = 0$. Sí son mutuamente excluyentes.
2. $x \geq 0\ /\ x \leq 0$ es exhaustivo, pero no excluyente: $x = 0$ pertenece a ambos casos.
3. $x < 1\ /\ x \geq 1$ es exhaustivo y mutuamente excluyente.
4. $x \leq 2\ /\ x \geq 1$ es exhaustivo, pero se solapa en todo el intervalo $[1, 2]$.

Las divisiones 2, 3 y 4 pueden sostener una prueba universal por casos si en cada rama se demuestra la misma conclusión. La primera no basta sin añadir el caso $x = 0$.

Solución 4. Preimagen y monotonía. Sea

$$x \in f^{-1}(U).$$

Por definición de preimagen,

$$f(x) \in U.$$

Como

$$U \subseteq V,$$

se sigue que

$$f(x) \in V.$$

Aplicando de nuevo la definición de preimagen,

$$x \in f^{-1}(V).$$

Puesto que x era arbitrario,

$$f^{-1}(U) \subseteq f^{-1}(V).$$

La definición de preimagen se usa al principio y al final; la inclusión $U \subseteq V$ constituye el paso intermedio.

Solución 5. Una equivalencia sobre inyectividad. **De 1 a 2.** Supongamos que f es inyectiva. Fijemos $b \in B$ y supongamos que $x, y \in f^{-1}(\{b\})$. Entonces

$$f(x) = b \qquad \text{y} \qquad f(y) = b,$$

de modo que

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

Por inyectividad,

$$x = y.$$

Así, $f^{-1}(\{b\})$ tiene a lo sumo un elemento.

De 2 a 1. Supongamos ahora que toda fibra

$$f^{-1}(\{b\})$$

tiene a lo sumo un elemento. Sean $x, y \in A$ con

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

Llamemos

$$b = f(x) = f(y).$$

Entonces

$$x, y \in f^{-1}(\{b\}).$$

Como esa fibra tiene a lo sumo un elemento,

$$x = y.$$

Por tanto, f es inyectiva.

Solución 6. Contraposición funcional. Queremos demostrar:

$$g \circ f \text{ inyectiva} \Rightarrow f \text{ inyectiva}.$$

La contraposición es:

$$f \text{ no inyectiva} \Rightarrow g \circ f \text{ no inyectiva}.$$

Supongamos que f no es inyectiva. Entonces existen $x, y \in A$ tales que

$$x \neq y$$

y

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

Aplicando g,

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

Es decir,

$$(g \circ f)(x) = (g \circ f)(y).$$

Como $x \neq y$, la composición no es inyectiva.

Hemos demostrado la contraposición y, por equivalencia lógica, la implicación original.

Solución 7. Una prueba por casos. La definición de valor absoluto aplicada a $x-2$ sugiere separar:

$$x - 2 \geq 0 \qquad \text{o} \qquad x - 2 < 0,$$

es decir,

$$x \geq 2 \qquad \text{o} \qquad x < 2.$$

Son casos exhaustivos y mutuamente excluyentes.

Caso 1: $x \geq 2$**.** Entonces

$$|x - 2| = x - 2.$$

Queremos comprobar

$$x - 2 \geq 2 - x.$$

Esto equivale a

$$2x \geq 4,$$

que es cierto porque $x \geq 2$.

Caso 2: $x < 2$**.** Entonces

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

Por tanto,

$$|x-2| = 2-x,$$

y en particular

$$|x-2| \geq 2-x.$$

Así la desigualdad vale para todo real x.

Solución 8. Construir un punto intermedio distinto. Tomemos

$$c = \frac{a+b}{2}.$$

Claramente $c \in \mathbb{R}$.

Si $c = a$, entonces

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

de donde

$$a + b = 2a$$

y por tanto

$$b = a,$$

contradiciendo $a \neq b$.

Del mismo modo, si $c = b$, entonces

$$a = b.$$

Por tanto,

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

Así existe un real distinto de ambos.

No hemos necesitado suponer un orden entre a y b; sólo que sean distintos.

Solución 9. Unicidad de la inversa funcional. Queremos demostrar que para todo $b \in B$,

$$g(b) = h(b).$$

Usamos las identidades disponibles:

$$g = g \circ \mathrm{id}_B\,.$$

Como

$$f \circ h = \mathrm{id}_B,$$

obtenemos

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

Por asociatividad de la composición,

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

Pero

$$g \circ f = \mathrm{id}_A \,.$$

Luego

$$g = \mathrm{id}_A \circ h = h.$$

Por tanto,

$$g = h.$$

La prueba compara dos candidatos a inversa sin usar previamente la frase “la inversa”.

Solución 10. Existencia única parametrizada. Fijemos $b \in F$ arbitrario.

Queremos resolver

$$x + a = b.$$

Existencia. Elegimos

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

Entonces

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

Así existe al menos una solución.

Unicidad. Supongamos que x e y satisfacen

$$x + a = b$$

y

$$y + a = b.$$

Entonces

$$x + a = y + a.$$

Por cancelación aditiva,

$$x = y.$$

Como b era arbitrario,

$$\forall b \in F \; \exists! x \in F \; (x + a = b).$$

Solución 11. Confundir recíproca con contraposición. La afirmación original es

$$P \Rightarrow Q,$$

donde

$$P: \ n \text{ es múltiplo de } 6,$$

y

$$Q: \ n \text{ es múltiplo de } 3.$$

El texto propuesto comienza con Q e intenta demostrar P. Eso es la **recíproca**

$$Q \Rightarrow P,$$

no la contraposición.

Además, la recíproca es falsa: 3 es múltiplo de 3 y no de 6.

La contraposición correcta es

Si n no es múltiplo de 3, entonces n no es múltiplo de 6.

Para probarla, supongamos que n no es múltiplo de 3. Si n fuera múltiplo de 6, existiría $k \in \mathbb{Z}$ con

$$n = 6k = 3(2k),$$

lo cual haría a n múltiplo de 3, contradiciendo la hipótesis.

Por tanto, n no es múltiplo de 6. La contraposición queda demostrada y, con ella, la afirmación original.

Solución 12. Un caso desaparecido. Los casos

$$x > 0 \qquad \text{y} \qquad x < 0$$

no cubren $x = 0$.

Podemos reparar la prueba usando tres casos exhaustivos:

Caso 1: $x > 0$**.** Entonces

$$x^2 > 0.$$

Caso 2: $x = 0$**.** Entonces

$$x^2 = 0.$$

Caso 3: $x < 0$**.** Entonces

$$-x > 0$$

y

$$x^2 = (-x)^2 > 0.$$

En todos los casos,

$$x^2 \geq 0.$$

También podríamos usar dos casos exhaustivos, por ejemplo $x \geq 0$ y $x < 0$.

Solución 13. Un testigo fuera del dominio. La fórmula

$$x = \frac{a}{2}$$

siempre produce un número racional o real, pero no necesariamente un entero.

Si

$$a = 1,$$

entonces

$$x = \frac{1}{2} \notin \mathbb{Z}.$$

De hecho, el caso $a = 1$ refuta toda la afirmación:

no existe $x \in \mathbb{Z}$ tal que $2x = 1$.

Por tanto, la universal es falsa.

El error no consiste sólo en una verificación incompleta del candidato; revela que no existe testigo en el dominio para ciertos valores de a.

Una versión verdadera sería:

Para todo entero par a existe $x \in \mathbb{Z}$ tal que $2x = a$.

Si $a = 2k$, basta elegir $x = k$.

Solución 14. Elegir arquitectura para una identidad de conjuntos. Como el objetivo es una igualdad de conjuntos, elegimos **doble inclusión**.

Primera inclusión. Sea

$$x \in A \setminus (B \cup C).$$

Entonces

$$x \in A$$

y

$$x \notin B \cup C.$$

La segunda condición significa

$$x \notin B \qquad \text{y} \qquad x \notin C.$$

Por tanto,

$$x \in A \setminus B$$

y

$$x \in A \setminus C.$$

Luego

$$x \in (A \setminus B) \cap (A \setminus C).$$

Así,

$$A \setminus (B \cup C) \subseteq (A \setminus B) \cap (A \setminus C).$$

Segunda inclusión. Sea ahora

$$x \in (A \setminus B) \cap (A \setminus C).$$

Entonces

$$x \in A \setminus B$$

y

$$x \in A \setminus C.$$

De ambas pertenencias obtenemos

$$x \in A, \qquad x \notin B, \qquad x \notin C.$$

Por tanto,

$$x \notin B \cup C.$$

Como además $x \in A$,

$$x \in A \setminus (B \cup C).$$

Así,

$$(A \setminus B) \cap (A \setminus C) \subseteq A \setminus (B \cup C).$$

Las dos inclusiones prueban la igualdad.

Solución 15. Una caracterización global de la inyectividad. Primero recordemos que, para cualquier función y cualquier $S \subseteq A$,

$$S \subseteq f^{-1}(f(S)).$$

En efecto, si $x \in S$, entonces

$$f(x) \in f(S),$$

por lo que

$$x \in f^{-1}(f(S)).$$

De 1 a 2. Supongamos que f es inyectiva y fijemos $S \subseteq A$.

Ya sabemos

$$S \subseteq f^{-1}(f(S)).$$

Para la inclusión inversa, sea

$$x \in f^{-1}(f(S)).$$

Entonces

$$f(x) \in f(S).$$

Por definición de imagen, existe $s \in S$ tal que

$$f(s) = f(x).$$

Como f es inyectiva,

$$s = x.$$

Puesto que $s \in S$,

$$x \in S.$$

Así,

$$f^{-1}(f(S)) \subseteq S.$$

Por doble inclusión,

$$f^{-1}(f(S)) = S.$$

De 2 a 1. Supongamos ahora que

$$f^{-1}(f(S)) = S$$

para todo $S \subseteq A$.
Sean $x, y \in A$ y supongamos

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

Elegimos el subconjunto

$$S = \{x\}.$$

Como

$$f(y) = f(x) \in f(\{x\}),$$

tenemos

$$y \in f^{-1}(f(\{x\})).$$

Por la hipótesis,

$$f^{-1}(f(\{x\})) = \{x\}.$$

Luego

$$y \in \{x\},$$

y por tanto

$$y = x.$$

Así, f es inyectiva.

La elección del singleton fue decisiva: convierte la igualdad global para todos los subconjuntos en información sobre dos entradas particulares.

Del repertorio a la práctica. Los ejercicios del capítulo han obligado a distinguir varias decisiones que al principio pueden parecer similares:

- usar una hipótesis frente a intentar demostrarla;
- demostrar una implicación frente a demostrar su recíproca;
- cambiar a la contraposición frente a iniciar una contradicción;
- separar casos exhaustivos frente a enumerar ejemplos;
- exhibir existencia frente a demostrar unicidad;
- intentar probar una universal frente a buscar un contraejemplo.

El objetivo no es que cada demostración lleve una etiqueta. Es que, al leer una afirmación, podamos reconstruir qué obligaciones contiene y verificar que el argumento realmente las cumple.

En el capítulo siguiente cambiaremos el punto de vista. Ya no partiremos siempre de una arquitectura evidente. La pregunta será qué hacer **cuando sabemos qué debemos demostrar, pero todavía no sabemos cómo encontrar la idea que lo demuestra**.

Capítulo 7 — ¿Cómo se descubre una demostración que todavía no conocemos?

En el capítulo anterior aprendimos a reconocer varias formas recurrentes de demostración. Ante una implicación sabemos qué significa demostrarla directamente; ante una equivalencia sabemos que aparecen dos direcciones; ante una afirmación existencial sabemos que necesitamos justificar un testigo; ante una igualdad de conjuntos podemos pensar en dos inclusiones.

Todo eso responde a una pregunta fundamental:

> **¿Qué tendría que quedar demostrado para que la prueba estuviera completa?**

Pero todavía queda otra pregunta, distinta y más difícil:

> **¿Cómo encontramos las ideas matemáticas que permiten demostrarlo?**

Conocer la arquitectura de una prueba no significa conocer ya la prueba.

Podemos saber perfectamente dónde empezar y qué conclusión necesitamos alcanzar y, sin embargo, no ver todavía qué debe ocurrir entre ambos extremos.

Este capítulo estudia precisamente ese espacio intermedio.

7.1. Cuando conocer la arquitectura no basta

Consideremos la afirmación siguiente:

> Sean $x, y \in \mathbb{R}$. Si
>
> $$x + y = 1,$$
>
> entonces
>
> $$xy \leq \frac{1}{4}.$$

La forma lógica no ofrece ninguna dificultad especial. Tenemos una implicación:

$$x + y = 1 \quad \Longrightarrow \quad xy \leq \frac{1}{4}.$$

Por tanto, una prueba directa debería comenzar suponiendo

$$x + y = 1$$

y terminar demostrando

$$xy \leq \frac{1}{4}.$$

La arquitectura está completamente determinada.

Pero eso no nos dice qué hacer.

De la hipótesis podemos despejar, por ejemplo,

$$y = 1 - x.$$

Al sustituir,

$$xy = x(1 - x).$$

Ahora queremos demostrar

$$x(1 - x) \leq \frac{1}{4}.$$

El problema ha cambiado de forma, pero todavía puede no ser evidente por qué la desigualdad es verdadera.

Podemos expandir:

$$x - x^2 \leq \frac{1}{4}.$$

Podemos mover todos los términos:

$$x^2 - x + \frac{1}{4} \geq 0.$$

Y entonces aparece una estructura reconocible:

$$x^2 - x + \frac{1}{4} = \left(x - \frac{1}{2}\right)^2.$$

Como todo cuadrado real es no negativo,

$$\left(x - \frac{1}{2}\right)^2 \geq 0.$$

De ahí obtenemos

$$x^2 - x + \frac{1}{4} \geq 0,$$

y por tanto

$$x - x^2 \leq \frac{1}{4}.$$

Así,

$$xy \leq \frac{1}{4}.$$

La prueba es breve. El descubrimiento de la prueba puede no haberlo sido.

La arquitectura y el puente. Podemos representar la situación inicial así:

$$\boxed{x + y = 1} \qquad \longrightarrow \qquad ? \qquad \longrightarrow \qquad \boxed{xy \leq \frac{1}{4}}.$$

La forma lógica nos dice que debemos construir una cadena desde la hipótesis hasta la conclusión.

No nos proporciona automáticamente el puente.

En este ejemplo, una forma especialmente útil de ese puente es observar que, usando $x + y = 1$,

$$\frac{1}{4} - xy = \frac{(x+y)^2}{4} - xy = \frac{(x-y)^2}{4}.$$

El término de la derecha es no negativo. Eso es contenido matemático: no procede solamente de saber que estamos ante una prueba directa.

Dos problemas pueden tener exactamente la misma arquitectura lógica y requerir ideas completamente diferentes.

Por ejemplo, las afirmaciones

$$P \Rightarrow Q$$

y

$$R \Rightarrow S$$

comparten forma lógica. Pero conocer cómo demostrar la primera puede no ofrecer ninguna información sobre qué objeto auxiliar, identidad, definición o teorema necesitamos para la segunda.

Ésta será una distinción central:

$$\boxed{\text{forma de la obligación}} \neq \boxed{\text{idea que permite satisfacerla}}.$$

Saber qué falta ya es progreso. Cuando una prueba no aparece inmediatamente, es fácil sentir que no sabemos por dónde avanzar. Pero normalmente conocemos al menos los dos extremos del problema.

En el ejemplo anterior tenemos:

Datos disponibles

$$x, y \in \mathbb{R}, \qquad x + y = 1.$$

Objetivo

$$xy \leq \frac{1}{4}.$$

La dificultad consiste en que todavía no conocemos una cadena que conecte esos dos estados.

Podemos pensar provisionalmente:

$$\text{datos} \longrightarrow \boxed{\text{hueco}} \longrightarrow \text{objetivo}.$$

Nombrar ese hueco cambia la naturaleza del problema.

Ya no preguntamos vagamente

> «¿Cómo demuestro esto?»

sino preguntas más precisas:

- ¿qué puedo obtener de la hipótesis?;
- ¿qué bastaría para alcanzar la conclusión?;
- ¿hay una definición que vuelva más explícito alguno de los extremos?;
- ¿puedo representar el objetivo de otra forma?;
- ¿necesito construir un objeto intermedio?;
- ¿hay alguna hipótesis que todavía no he utilizado?;
- ¿puedo formular una afirmación intermedia que conecte las dos partes?

No responderemos todas estas preguntas en esta sección. Cada una abrirá una técnica que estudiaremos más adelante.

Por ahora importa comprender que **descubrir una prueba significa trabajar sobre el hueco**.

El orden en que encontramos una prueba puede ser distinto del orden en que la escribimos. Volvamos al ejemplo.

Durante la búsqueda podríamos partir del objetivo

$$xy \leq \frac{1}{4}$$

y preguntarnos qué forma equivalente o qué condición suficiente haría evidente la desigualdad. Por ejemplo:

$$xy \leq \frac{1}{4}$$

sugiere estudiar

$$\frac{1}{4} - xy \geq 0.$$

Como $x + y = 1$, podemos explorar

$$\frac{1}{4} - xy = \frac{(x+y)^2}{4} - xy = \frac{x^2 - 2xy + y^2}{4} = \frac{(x-y)^2}{4}.$$

Ahora la no negatividad es inmediata.

El razonamiento de descubrimiento ha comenzado, en cierto sentido, desde la conclusión: hemos preguntado qué forma sería suficiente para reconocer por qué debe cumplirse.

Sin embargo, la prueba final puede escribirse en el sentido de las dependencias justificadas:

Demostración. Sean $x, y \in \mathbb{R}$ y supongamos que

$$x + y = 1.$$

Como

$$(x-y)^2 \geq 0,$$

tenemos

$$x^2 - 2xy + y^2 \geq 0.$$

Sumando $4xy$ a ambos lados,

$$x^2 + 2xy + y^2 \geq 4xy.$$

Por tanto,

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

Usando la hipótesis $x + y = 1$,

$$1 \geq 4xy.$$

Finalmente,

$$xy \leq \frac{1}{4}.$$

□

Ahora cada afirmación aparece después de algo que la justifica.

La prueba final no necesita reproducir toda la búsqueda que permitió encontrarla.

El borrador de descubrimiento. Mientras buscamos una demostración podemos:

- despejar una variable;
- probar una representación alternativa;
- comenzar desde la conclusión;
- escribir condiciones que serían suficientes;
- ensayar ejemplos;
- introducir una cantidad auxiliar;
- abandonar un cálculo;
- volver a una definición;
- preguntarnos qué hipótesis todavía no hemos usado.

Ese trabajo puede ser no lineal.

Una hoja de exploración podría contener algo como:

$$x + y = 1$$

$$xy \leq \frac{1}{4} \quad ?$$

$$4xy \leq 1$$

$$4xy \leq (x+y)^2$$

$$0 \leq x^2 - 2xy + y^2$$

$$0 \leq (x-y)^2.$$

Esta secuencia es informativa para quien está descubriendo la prueba.

Pero debemos observar cuidadosamente lo que ha ocurrido. Comenzamos con el objetivo y lo transformamos hasta llegar a una afirmación que sabemos justificar. Durante la exploración todavía debemos comprobar si esas transformaciones permiten regresar legítimamente al objetivo.

No podemos asumir que toda cadena escrita hacia atrás es automáticamente reversible.

La prueba final y el borrador no son el mismo objeto. Conviene distinguir entonces:

$$\boxed{\text{borrador de descubrimiento}} \qquad \text{y} \qquad \boxed{\text{prueba final}}.$$

El borrador puede contener flechas en ambos sentidos, ejemplos, cálculos incompletos, intentos descartados y condiciones provisionales.

La prueba final debe reconstruirse según dependencias legítimas.

Esta diferencia explica algo que puede resultar desconcertante cuando se leen buenos textos matemáticos: una demostración publicada puede ocupar diez líneas aunque encontrarla haya requerido mucho más trabajo.

La brevedad de la versión final no es evidencia de que el camino de descubrimiento haya sido igualmente breve.

Descubrir no significa adivinar. La diferencia entre descubrimiento y justificación tampoco significa que encontrar una prueba sea un acto misterioso reservado a una intuición inexplicable.

No existe, en general, un algoritmo que garantice encontrar cualquier demostración. Pero entre una receta automática y una adivinación sin método hay un territorio muy amplio de decisiones que pueden aprenderse.

Podemos entrenarnos para:

extraer consecuencias,

examinar el objetivo,

cambiar de representación,

buscar una condición intermedia,

formular un lema,

examinar por qué un intento se atasca.

Cada uno de esos movimientos pertenece al espacio de descubrimiento. Todos deberán someterse después al control lógico de una demostración válida.

La exploración puede ser flexible; la prueba final no puede ser arbitraria.

Una misma prueba tiene dos historias. Podemos distinguir dos historias diferentes.

La primera es la historia de cómo encontramos la demostración:

$$\text{intento} \to \text{reformulación} \to \text{callejón} \to \text{nueva representación} \to \text{idea}.$$

La segunda es la historia matemática que finalmente justifica el resultado:

$$\text{hipótesis} \to \text{inferencias válidas} \to \text{conclusión}.$$

La primera puede contener retrocesos.

La segunda debe poder auditarse paso por paso.

Confundirlas produce dos errores opuestos.

Uno consiste en pensar que, como la prueba final es corta, encontrarla debía ser igualmente inmediato.

El otro consiste en escribir todo el proceso de tanteo como si cada paso exploratorio perteneciera a la demostración.

Ninguna de las dos conclusiones es correcta.

Dos tareas distintas

Al comenzar una prueba debemos distinguir:

1. **qué tendría que demostrarse** para completar el argumento;
2. **cómo encontrar el contenido matemático** que permite demostrarlo.

C06 estudió principalmente la primera tarea.

C07 estudiará la segunda.

Comprobación. Supón que debes demostrar una afirmación de la forma

$$P \Rightarrow Q.$$

Sabes ya que una prueba directa puede comenzar suponiendo P y debe terminar estableciendo Q.

¿Significa esto que ya sabes cómo demostrar la afirmación?

Respuesta razonada.

No.

La forma

$$P \Rightarrow Q$$

determina una obligación lógica: bajo la hipótesis P, debemos justificar Q.

Pero no determina qué consecuencias de P serán relevantes, qué definiciones conviene desplegar, qué objetos auxiliares pueden necesitarse ni qué resultado intermedio permitirá alcanzar Q.

Por tanto,

conocer la arquitectura

es una condición importante para organizar la prueba, pero no equivale a

conocer el puente matemático.

El resto del capítulo estudiará cómo buscar ese puente de manera controlada.

Lo que cambia a partir de ahora. Hasta aquí hemos preguntado principalmente:

> «Dada una demostración, ¿qué obligaciones debe cumplir?»

A partir de la siguiente sección añadiremos otra pregunta:

> «Dadas las hipótesis y el objetivo, ¿qué información conviene producir primero?»

Comenzaremos por uno de los movimientos más naturales: trabajar **hacia adelante desde las hipótesis**, pero no acumulando consecuencias al azar. La cuestión será aprender a distinguir una consecuencia disponible de una consecuencia que realmente acerca la prueba a su objetivo.

7.2. Trabajar hacia adelante desde las hipótesis

Una de las maneras más naturales de comenzar una búsqueda consiste en preguntar:

> **¿Qué puedo obtener ya, de manera legítima, a partir de los datos que tengo?**

Éste es el sentido básico de trabajar **hacia adelante**.

Partimos de las hipótesis y producimos consecuencias cuya justificación conocemos. Si una hipótesis dice

$$P,$$

y sabemos que

$$P \Rightarrow R,$$

podemos añadir

$$R$$

a nuestro conjunto de información disponible.

Después quizá podamos usar R junto con otro dato para obtener una consecuencia nueva.

La imagen más sencilla es

$$\text{hipótesis} \longrightarrow \text{consecuencia} \longrightarrow \text{consecuencia} \longrightarrow \cdots$$

Pero hay una dificultad inmediata: de unas mismas hipótesis suelen seguir muchísimas cosas.

Por eso trabajar hacia adelante no significa **deducir todo lo que podamos**. Significa producir consecuencias que tengan alguna posibilidad de acercarnos al objetivo.

Un primer ejemplo: hacer avanzar un dato. Sean A, B, C conjuntos y supongamos

$$A \subseteq B,$$

$$B \cap C = \varnothing,$$

y

$$x \in A.$$

Queremos demostrar

$$x \notin C.$$

La hipótesis

$$x \in A$$

puede combinarse inmediatamente con

$$A \subseteq B.$$

Por definición de inclusión,

$$x \in A \quad \Longrightarrow \quad x \in B.$$

Ya hemos producido una consecuencia nueva:

$$\boxed{x \in B}.$$

Ahora la segunda hipótesis adquiere utilidad. La igualdad

$$B \cap C = \varnothing$$

nos dice que ningún elemento puede pertenecer simultáneamente a B y a C.

Como ya sabemos que

$$x \in B,$$

si también tuviéramos

$$x \in C,$$

entonces

$$x \in B \cap C,$$

lo cual sería imposible.

Por tanto,

$$x \notin C.$$

La cadena útil ha sido

$$\boxed{x \in A} \overset{A \subseteq B}{\longrightarrow} \boxed{x \in B} \overset{B \cap C = \varnothing}{\longrightarrow} \boxed{x \notin C}.$$

Cada flecha tiene una procedencia identificable.

Ésta es una característica importante del trabajo hacia adelante: no basta con escribir una lista de afirmaciones verdaderas. Debemos poder responder, para cada paso,

«¿De qué dato anterior proviene esta nueva información?»

Consecuencias disponibles y consecuencias útiles. De las mismas hipótesis podríamos obtener muchas otras afirmaciones verdaderas.

Por ejemplo, de

$$x \in A$$

podemos concluir

$$x \in A \cup C.$$

También, de

$$A \subseteq B,$$

podemos obtener

$$A \cup C \subseteq B \cup C.$$

Y de

$$B \cap C = \varnothing$$

se sigue, entre otras cosas, que

$$C \subseteq B^c$$

si estamos trabajando dentro de un universo donde ese complemento está definido.

Todas estas consecuencias pueden ser correctas.

Pero para demostrar

$$x \notin C,$$

no todas tienen el mismo valor estratégico.

La distinción que necesitamos es

$$\boxed{\text{consecuencia disponible}} \qquad \text{frente a} \qquad \boxed{\text{consecuencia útil para este objetivo}}.$$

Una consecuencia es estratégicamente útil cuando reduce de algún modo la distancia entre lo que ya sabemos y lo que necesitamos demostrar.

En nuestro ejemplo,

$$x \in B$$

es útil porque permite activar precisamente la información contenida en

$$B \cap C = \varnothing.$$

En cambio,

$$x \in A \cup C$$

es verdadera pero no hace visible ningún camino inmediato hacia

$$x \notin C.$$

Las hipótesis no son una lista decorativa. Cuando un problema ofrece varias hipótesis conviene tratarlas como **recursos potenciales**.

Supongamos que tenemos

$$H_1, H_2, H_3$$

y queremos demostrar Q.

Una búsqueda hacia adelante puede preguntar:

- ¿qué produce H_1 por sí sola?;
- ¿qué produce H_2?;
- ¿puedo combinar H_1 y H_3?;
- ¿alguna consecuencia obtenida permite utilizar una hipótesis que antes parecía inerte?;
- ¿qué dato todavía no he usado?

La última pregunta es especialmente valiosa.

Una hipótesis sin usar puede indicar varias cosas distintas:

1. todavía no hemos encontrado el punto donde se vuelve necesaria;
2. hemos tomado una ruta que no aprovecha la estructura completa del problema;
3. la hipótesis es redundante y el resultado es en realidad más fuerte de lo que se anunció.

No debemos decidir demasiado pronto cuál de estas posibilidades ocurre.

Una hipótesis puede resultar innecesaria. Sea

$$f : A \to B$$

una función inyectiva y sean

$$X \subseteq Y \subseteq A.$$

Queremos demostrar

$$f(X) \subseteq f(Y).$$

Tomemos

$$z \in f(X).$$

Por definición de imagen, existe $x \in X$ tal que

$$f(x) = z.$$

Como

$$X \subseteq Y,$$

obtenemos

$$x \in Y.$$

Por tanto,

$$f(x) \in f(Y),$$

y como $f(x) = z$,

$$z \in f(Y).$$

Así,

$$f(X) \subseteq f(Y).$$

¿Dónde utilizamos que f era inyectiva?

En ninguna parte.

Esto no invalida la prueba. Revela algo mejor: la afirmación sigue siendo verdadera sin esa hipótesis.

Podemos fortalecer el resultado:

Para cualquier función $f : A \to B$, si $X \subseteq Y \subseteq A$, entonces

$$f(X) \subseteq f(Y).$$

Por eso una hipótesis no utilizada merece atención, pero no autoriza por sí sola a declarar que la prueba está equivocada.

Puede ser una señal de que todavía falta una idea; también puede revelar redundancia en el enunciado.

Mantener la trazabilidad. Durante una búsqueda algo más larga conviene conservar la procedencia de las afirmaciones importantes.

Por ejemplo, si sabemos

$$H_1 : x \in A,$$

$$H_2 : A \subseteq B,$$

$$H_3 : B \subseteq D,$$

podemos registrar mentalmente —o en el borrador—

$$x \in A \overset{H_2}{\Longrightarrow} x \in B \overset{H_3}{\Longrightarrow} x \in D.$$

Las etiquetas no son necesarias en la prueba final, pero durante la exploración cumplen una función: evitan que una consecuencia aparezca desligada de sus premisas.

Esto es especialmente importante cuando intervienen varios objetos o varias hipótesis parecidas. Una afirmación obtenida legítimamente puede dejar de ser utilizable si olvidamos bajo qué condiciones fue obtenida.

La búsqueda debe permitir siempre reconstruir una cadena de dependencias.

Avanzar no es manipular símbolos sin dirección. Volvamos por un momento a la relación

$$x + y = 1.$$

A partir de ella podemos producir indefinidamente expresiones equivalentes:

$$y = 1 - x,$$

$$x = 1 - y,$$

$$2x + 2y = 2,$$

$$x + y - 1 = 0,$$

$$(x + y)^2 = 1,$$

$$x^2 + 2xy + y^2 = 1,$$

y muchas más.

Todas proceden legítimamente de la hipótesis.

Pero una lista cada vez más larga no constituye necesariamente progreso.

Si nuestro objetivo fuese

$$x^2 + 2xy + y^2 = 1,$$

entonces elevar al cuadrado sería decisivo.

Si el objetivo fuese

$$y = 1 - x,$$

bastaría despejar.

Si buscamos una cota para xy, quizá la expansión

$$x^2 + 2xy + y^2 = 1$$

resulte útil sólo cuando podamos relacionar los términos cuadrados con alguna cantidad no negativa.

La misma hipótesis admite muchas transformaciones. El objetivo es lo que permite valorar su pertinencia.

Por eso el trabajo hacia adelante tiene siempre dos componentes:

$$\boxed{\text{corrección}} \quad \text{y} \quad \boxed{\text{dirección}}.$$

Una consecuencia debe ser legítima, pero además queremos saber por qué la estamos produciendo.

Consecuencias inmediatas antes de consecuencias sofisticadas. Cuando una prueba se resiste, es tentador buscar enseguida un teorema potente o una transformación ingeniosa.

Antes conviene agotar las consecuencias inmediatas de las hipótesis que parecen estructuralmente relevantes.

Si sabemos

$$x \in A \cap B,$$

extraemos

$$x \in A \qquad \text{y} \qquad x \in B.$$

Si sabemos que f es inyectiva, conservamos disponible la regla

$$f(u) = f(v) \Longrightarrow u = v.$$

Si sabemos

$$a < b,$$

podemos preguntarnos qué operaciones preservan esa desigualdad bajo las condiciones presentes.

Si una hipótesis es una existencia,

$$\exists x \in A \; P(x),$$

podemos introducir un testigo del que sabemos

$$x \in A \qquad \text{y} \qquad P(x).$$

Estas operaciones no son espectaculares. Precisamente por eso son fáciles de omitir.

Muchas pruebas se bloquean no porque falte una idea profunda, sino porque una hipótesis todavía no ha sido desempaquetada hasta la forma en que puede interactuar con las demás.

Un mapa de información. Podemos representar el trabajo hacia adelante no como una única cadena, sino como una pequeña red:

$$\begin{array}{ccccc} H_1 & \longrightarrow & R_1 & \longrightarrow & R_3 \\ & & \uparrow & & \\ H_2 & \longrightarrow & R_2 & & \\ H_3 & \longrightarrow & R_4. & & \end{array}$$

Algunas ramas se conectarán.

Otras terminarán sin utilidad.

La tarea no consiste en desarrollar todas las ramas hasta el infinito, sino en reconocer cuál de ellas parece aproximarse a la región donde vive el objetivo.

Esta imagen también explica por qué un intento fallido puede ser informativo: quizá una rama produzca consecuencias correctas pero nunca interactúe con el resto de los datos. En §7.7 aprenderemos a utilizar ese fracaso como diagnóstico.

¿Cuándo conviene dejar de avanzar?. Trabajar hacia adelante tiene un límite práctico.

Si hemos obtenido varias consecuencias correctas pero ninguna parece reducir el hueco hacia el objetivo, seguir produciendo fórmulas puede convertirse en actividad sin estrategia.

Una señal de detención es:

> **sé cada vez más cosas, pero no sé mejor por qué la conclusión debería seguirse.**

En ese momento no debemos concluir que el teorema sea falso ni que la prueba sea imposible.

Puede ser necesario cambiar de pregunta.

En lugar de seguir preguntando

> «¿qué más puedo obtener de las hipótesis?»,

podemos preguntar

> «¿qué tendría que obtener para que el objetivo quedara demostrado?»

Ese cambio de dirección será el tema de la sección siguiente.

Avanzar con propósito

Trabajar hacia adelante significa extraer consecuencias justificadas de las hipótesis y conservar su procedencia.

Pero una consecuencia verdadera no es automáticamente una consecuencia útil. El objetivo funciona como criterio de selección: buscamos información que pueda reducir el hueco entre los datos y lo que todavía falta demostrar.

Comprobación. Supón que

$$A \subseteq B, \qquad B \subseteq C, \qquad x \in A.$$

Quieres demostrar

$$x \in C.$$

¿Qué cadena hacia adelante es la más directa y qué papel cumple cada hipótesis?

Respuesta razonada.

De

$$x \in A$$

y

$$A \subseteq B$$

obtenemos

$$x \in B.$$

Esta nueva información activa la segunda inclusión:

$$x \in B \qquad \text{y} \qquad B \subseteq C$$

implican

$$x \in C.$$

Por tanto, la cadena es

$$\boxed{x \in A} \overset{A \subseteq B}{\Longrightarrow} \boxed{x \in B} \overset{B \subseteq C}{\Longrightarrow} \boxed{x \in C}.$$

La primera inclusión permite abandonar A y entrar en B; la segunda permite pasar de B a C. Ninguna de las dos inclusiones produce por sí sola la conclusión: la cadena aparece al combinar los datos en el orden adecuado.

Del avance a la pregunta inversa. Trabajar hacia adelante nos enseña a explotar lo que ya tenemos.

Pero puede ocurrir que, después de varias consecuencias razonables, todavía no sepamos cuál de ellas necesitamos.

Entonces el objetivo debe participar de manera más activa en la búsqueda.

En §7.3 aprenderemos a trabajar **hacia atrás desde lo que queremos demostrar**: no para invertir ilegítimamente las implicaciones, sino para identificar condiciones suficientes que, si logramos alcanzarlas desde las hipótesis, cerrarían el argumento.

7.3. Trabajar hacia atrás desde el objetivo

En la sección anterior comenzamos por las hipótesis y preguntamos qué consecuencias podían producir. Ahora cambiaremos el punto de vista.

Supongamos que el objetivo es una afirmación Q y que todavía no vemos cómo alcanzarla. En vez de seguir generando consecuencias desde los datos, podemos preguntar:

¿Qué me bastaría demostrar para que Q quedara establecido?

Si encontramos una condición R tal que

$$R \Rightarrow Q,$$

podemos convertir provisionalmente el problema en otro:

$$\boxed{\text{demostrar } R}.$$

Eso es trabajar hacia atrás.

No hemos demostrado todavía Q. Hemos identificado una condición suficiente para obtenerla.

Cambiar el problema de búsqueda. La operación central puede representarse así:

$$\boxed{Q} \quad \Longleftarrow \quad \boxed{R}.$$

La flecha apunta hacia Q porque lo que sabemos es

$$R \Rightarrow Q.$$

Durante la búsqueda, sin embargo, nuestra atención se desplaza en sentido contrario: partimos visualmente de Q y buscamos una afirmación R que sería suficiente.

Por eso conviene distinguir dos direcciones:

$$\text{dirección de búsqueda:} \qquad Q \rightsquigarrow R,$$

$$\text{dirección de justificación:} \qquad R \Rightarrow Q.$$

La primera flecha no representa una inferencia lógica. Representa una decisión estratégica.

Esta distinción evita uno de los errores más peligrosos del trabajo hacia atrás: escribir el razonamiento de descubrimiento como si ya fuera una demostración.

Un ejemplo: descubrir una desigualdad. Sea $x > 0$. Queremos demostrar

$$x + \frac{1}{x} \geq 2.$$

Supongamos que la idea no es inmediata.

Podemos comenzar desde el objetivo y preguntar qué desigualdad equivalente aparecería si eliminásemos el denominador. Como $x > 0$, multiplicar por x conserva el sentido de la desigualdad. Así,

$$x + \frac{1}{x} \geq 2$$

es equivalente a

$$x^2 + 1 \geq 2x.$$

Reordenando,

$$x^2 - 2x + 1 \geq 0,$$

es decir,

$$(x-1)^2 \geq 0.$$

Ahora hemos llegado a una afirmación que conocemos.

El borrador de descubrimiento puede verse como

$$\boxed{x + \frac{1}{x} \geq 2} \Longleftarrow \boxed{x^2 + 1 \geq 2x} \Longleftarrow \boxed{(x-1)^2 \geq 0}.$$

En este caso cada transformación es reversible bajo la hipótesis $x > 0$. Pero la prueba final conviene escribirla en la dirección justificativa:

Demostración. Como $x \in \mathbb{R}$,

$$(x-1)^2 \geq 0.$$

Por tanto,

$$x^2 - 2x + 1 \geq 0,$$

de donde

$$x^2 + 1 \geq 2x.$$

Como $x > 0$, podemos dividir por x sin cambiar el sentido de la desigualdad. Obtenemos

$$x + \frac{1}{x} \geq 2.$$

□

El trabajo hacia atrás descubrió el punto de partida útil. La prueba final no necesitó fingir que habíamos inferido $(x-1)^2 \geq 0$ desde la conclusión.

No toda transformación hacia atrás es reversible. El ejemplo anterior puede producir una falsa sensación de seguridad porque todas las transformaciones resultaron equivalentes.

En general, al trabajar hacia atrás sólo necesitamos una condición suficiente.

Supongamos que queremos demostrar Q y descubrimos

$$R \Rightarrow Q.$$

Entonces demostrar R basta, aunque Q no implique R.

Por ejemplo, si queremos demostrar que un entero n es par, sería suficiente demostrar que n es múltiplo de 4:

$$4 \mid n \Rightarrow 2 \mid n.$$

Pero la condición es más fuerte de lo necesario. Un entero puede ser par sin ser múltiplo de 4.

Esto introduce una consideración estratégica: una condición suficiente puede ser correcta y, sin embargo, resultar demasiado difícil de alcanzar desde las hipótesis.

Trabajar hacia atrás no consiste sólo en encontrar **alguna** condición suficiente, sino en encontrar una que tenga posibilidades razonables de conectarse con los datos disponibles.

Una flecha en el sentido incorrecto. Hay una diferencia esencial entre encontrar una condición suficiente para el objetivo y encontrar simplemente una consecuencia del objetivo.

Supongamos que queremos demostrar

$$x = y.$$

Es cierto que

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

Pero de aquí no podemos concluir que bastaría demostrar

$$x^2 = y^2.$$

En los reales,

$$x^2 = y^2$$

permite también

$$x = -y.$$

La relación disponible tiene la orientación

$$Q \Rightarrow R,$$

cuando para trabajar hacia atrás necesitamos

$$R \Rightarrow Q.$$

Por tanto, ante cada paso regresivo debemos preguntar:

> **Si logro demostrar esta nueva meta, ¿la meta anterior se sigue realmente de ella?**

Esta pregunta controla la dirección lógica del procedimiento.

Equivalencia, suficiencia y mera consecuencia. Cuando reformulamos un objetivo Q durante la búsqueda pueden darse tres situaciones distintas.

Primera situación: equivalencia.

$$R \Longleftrightarrow Q.$$

Podemos reemplazar el objetivo por R sin perder ni ganar casos.

Segunda situación: condición suficiente.

$$R \Rightarrow Q.$$

Demostrar R bastará, aunque quizá estemos imponiendo una condición más fuerte de la necesaria.

Tercera situación: sólo una consecuencia del objetivo.

$$Q \Rightarrow R.$$

Demostrar R no basta en general para recuperar Q. Esta transformación no es válida como sustitución regresiva del objetivo.

Distinguir estas tres posibilidades es una disciplina básica del trabajo hacia atrás.

Trabajar hacia atrás con una meta de conjuntos. La técnica no pertenece sólo al álgebra.

Supongamos que queremos demostrar

$$X \cup Z \subseteq Y \cup Z$$

bajo la hipótesis

$$X \subseteq Y.$$

La arquitectura de inclusión nos lleva a tomar

$$u \in X \cup Z$$

y fijar como objetivo

$$u \in Y \cup Z.$$

Ahora podemos mirar el objetivo desde atrás.

Para demostrar

$$u \in Y \cup Z,$$

sería suficiente demostrar una de las dos condiciones

$$u \in Y \qquad \text{o} \qquad u \in Z.$$

Esto nos indica qué información debemos buscar en la hipótesis

$$u \in X \cup Z.$$

Al desplegarla aparecen justamente dos posibilidades:

$$u \in X \qquad \text{o} \qquad u \in Z.$$

Si $u \in X$, la inclusión $X \subseteq Y$ produce $u \in Y$. Si $u \in Z$, ya tenemos directamente la segunda condición suficiente.

El análisis del objetivo no hizo la prueba por sí solo. Nos dijo qué clase de información sería útil encontrar.

Metas intermedias encadenadas. A veces una única transformación regresiva no basta.

Podemos construir una cadena de condiciones suficientes:

$$Q \Longleftarrow R \Longleftarrow S \Longleftarrow T.$$

Lo que esta escritura significa es

$$T \Rightarrow S, \qquad S \Rightarrow R, \qquad R \Rightarrow Q.$$

Si finalmente logramos obtener T desde las hipótesis, la prueba podrá recorrer la cadena en el sentido legítimo:

$$T \Rightarrow S \Rightarrow R \Rightarrow Q.$$

La exploración puede, por tanto, avanzar visualmente de derecha a izquierda mientras la justificación final avanza de izquierda a derecha.

Esta inversión aparente no es un problema siempre que mantengamos clara la orientación de cada implicación.

Una nueva meta puede ser peor que la anterior. El trabajo hacia atrás tampoco garantiza progreso.

Podemos reemplazar un objetivo por una condición suficiente que sea:

- demasiado fuerte;
- más complicada que el objetivo original;
- desconectada de las hipótesis;
- o dependiente de un resultado que todavía no tenemos.

Por ejemplo, para demostrar que un entero n es par podríamos proponernos demostrar

$$n = 0.$$

La condición es suficiente:

$$n = 0 \Rightarrow 2 \mid n.$$

Pero salvo que las hipótesis apunten realmente a $n = 0$, esta sustitución probablemente ha empeorado el problema.

Una buena meta intermedia no sólo debe implicar el objetivo. También debe ser **alcanzable** desde la información disponible.

Todavía no estudiaremos sistemáticamente cómo reconocer esa conexión. Ésa será precisamente la función de §7.4.

Mantener visibles las obligaciones. Durante un borrador puede ser útil registrar las metas regresivas con una notación que no las confunda con afirmaciones ya demostradas.

Por ejemplo:

OBJETIVO: Q

bastaría: R

bastaría para R : S.

Eso es muy distinto de escribir sin más

$$Q \Rightarrow R \Rightarrow S,$$

porque esa cadena podría afirmar implicaciones que no poseemos.

El borrador debe permitir distinguir al menos tres estados:

$$\boxed{\text{dato disponible}}, \qquad \boxed{\text{meta pendiente}}, \qquad \boxed{\text{relación demostrada entre ambas}}.$$

Esta separación será especialmente importante cuando combinemos trabajo hacia adelante y hacia atrás.

Trabajar hacia atrás

Trabajar hacia atrás no significa inferir las causas desde las consecuencias.

Significa buscar una condición R cuya verdad sea suficiente para el objetivo Q:

$$R \Rightarrow Q.$$

Durante el descubrimiento podemos mover nuestra atención de Q hacia R. La prueba final deberá justificar la dirección contraria, desde aquello que realmente sabemos hasta la conclusión.

Comprobación. Quieres demostrar una afirmación Q. Consideras tres posibles metas auxiliares R, S, T y sabes que

$$Q \Rightarrow R,$$

$$S \Rightarrow Q,$$

$$T \Longleftrightarrow Q.$$

¿Cuáles pueden sustituir legítimamente a Q como metas durante un trabajo hacia atrás?

Respuesta razonada.

La condición S puede utilizarse porque

$$S \Rightarrow Q.$$

Si conseguimos demostrar S, el objetivo original queda demostrado.

También podemos reemplazar Q por T, porque

$$T \Longleftrightarrow Q$$

incluye en particular

$$T \Rightarrow Q.$$

En cambio, conocer sólo

$$Q \Rightarrow R$$

no permite reemplazar el objetivo por R. R puede ser una consecuencia necesaria de Q y, sin embargo, ser demasiado débil para recuperarlo.

Por tanto, las sustituciones regresivas válidas son

$$Q \Longleftarrow S$$

y

$$Q \Longleftrightarrow T.$$

La dirección de la implicación decide si la nueva meta realmente basta.

Dos frentes todavía separados. Ahora disponemos de dos movimientos complementarios.

Desde las hipótesis podemos avanzar:

$$\text{datos} \longrightarrow \text{consecuencias disponibles.}$$

Desde el objetivo podemos retroceder en la búsqueda:

$$\text{condiciones suficientes} \longrightarrow \text{objetivo.}$$

Pero todavía pueden quedar separados:

$$\text{datos} \longrightarrow \boxed{\text{lo que sabemos obtener}} \quad \Big| \quad \boxed{\text{lo que necesitaríamos}} \longrightarrow \text{objetivo.}$$

La siguiente sección estudiará el momento central del descubrimiento: **hacer que esos dos frentes se encuentren**.

7.4. Hacer que los dos extremos se encuentren

En las dos secciones anteriores hemos aprendido a movernos desde extremos opuestos del problema.

Desde las hipótesis podemos avanzar:

$$\text{datos} \longrightarrow \text{consecuencias disponibles.}$$

Desde el objetivo podemos trabajar hacia atrás y preguntar qué condiciones serían suficientes:

$$\text{condiciones suficientes} \longrightarrow \text{objetivo.}$$

El momento decisivo aparece cuando esas dos búsquedas dejan de ser independientes.

Supongamos que hacia adelante hemos llegado a una afirmación R y que, al trabajar desde el objetivo Q, descubrimos que sería suficiente demostrar una afirmación S:

$$\text{hipótesis} \longrightarrow R \quad \Big| \quad S \longrightarrow Q.$$

La pregunta estratégica ya no es simplemente

> «¿qué más puedo deducir?»

ni tampoco

> «¿qué otra condición bastaría?».

Ahora podemos preguntar algo mucho más preciso:

> **¿cómo conecto R con S?**

Éste es el **hueco intermedio** de la prueba.

Descubrir una demostración consiste muchas veces en reducir un problema amplio a esta pregunta local.

Un ejemplo central: localizar el punto de encuentro. Sean $u, v, w \in \mathbb{R}$ y supongamos

$$u \leq v$$

y

$$w \geq 0.$$

Queremos demostrar

$$uw \leq vw.$$

Sabemos desde C06 qué arquitectura tiene el argumento: partimos de las dos hipótesis y debemos justificar la desigualdad final.

Pero examinemos el problema desde ambos extremos.

Hacia adelante. De

$$u \leq v$$

obtenemos

$$v - u \geq 0.$$

Como además

$$w \geq 0,$$

el producto de dos reales no negativos es no negativo. Por tanto,

$$w(v - u) \geq 0.$$

Nuestro frente hacia adelante ha llegado a

$$\boxed{w(v - u) \geq 0}.$$

Hacia atrás. Ahora miremos el objetivo:

$$uw \leq vw.$$

Sería suficiente demostrar

$$0 \leq vw - uw.$$

Factorizando,

$$vw - uw = w(v - u).$$

Por tanto, una condición suficiente para el objetivo es precisamente

$$\boxed{w(v - u) \geq 0}.$$

Los dos frentes han llegado a la misma afirmación:

$$u \leq v, \quad w \geq 0 \quad \longrightarrow \quad \boxed{w(v - u) \geq 0} \quad \longrightarrow \quad uw \leq vw.$$

El hueco ha desaparecido.

La prueba final puede escribirse ahora linealmente:

Demostración. De $u \leq v$ se sigue

$$v - u \geq 0.$$

Como $w \geq 0$, tenemos

$$w(v - u) \geq 0.$$

Por distributividad,

$$vw - uw \geq 0,$$

y por tanto

$$uw \leq vw.$$

□

La idea importante no es la desigualdad. Es que la búsqueda produjo una afirmación intermedia que era simultáneamente

- una consecuencia alcanzable desde las hipótesis;
- y una condición suficiente para el objetivo.

Eso es un punto de encuentro.

El puente puede ser una afirmación exacta. En el ejemplo anterior ambos frentes llegaron literalmente a la misma expresión:

$$w(v - u) \geq 0.$$

Éste es el caso más limpio:

$$\text{datos} \longrightarrow \boxed{M} \longrightarrow \text{objetivo}.$$

Una vez que M está justificado desde los datos y sabemos que M implica el objetivo, la búsqueda principal ha terminado.

Queda todavía revisar que cada paso sea válido y reconstruir la prueba final, pero ya no existe un hueco conceptual entre los extremos.

El puente también puede ser una implicación corta. No siempre los dos frentes llegan a la misma oración.

Puede ocurrir que hacia adelante obtengamos R y hacia atrás descubramos que necesitamos S:

$$\text{datos} \longrightarrow R \qquad \Big| \qquad S \longrightarrow Q.$$

Si encontramos una razón para afirmar

$$R \Rightarrow S,$$

la conexión queda cerrada:

$$\text{datos} \longrightarrow R \longrightarrow S \longrightarrow Q.$$

En este caso el puente no es una afirmación común, sino una inferencia entre dos estados intermedios.

Por eso conviene distinguir

$$\boxed{\text{lo que ya sabemos obtener}}$$

de

$$\boxed{\text{lo que todavía necesitamos}}.$$

La diferencia entre ambos puede ser mucho menor que la diferencia entre las hipótesis originales y el objetivo original.

Reducir el problema hasta que el hueco sea local. Esta reducción es una de las ventajas principales de trabajar con dos frentes.

Al comienzo podemos tener una pregunta amplia:

> «¿Cómo demuestro Q a partir de todas estas hipótesis?»

Después de avanzar y retroceder de manera controlada, la pregunta puede convertirse en algo mucho más específico:

> «Sé demostrar R. Para terminar me bastaría S. ¿Puedo obtener S a partir de R?»

El problema global se ha convertido en un problema local.

Podemos representarlo así:

$$\underbrace{\text{hipótesis} \longrightarrow \cdots \longrightarrow R}_{\text{parte ya accesible}} \quad \boxed{?} \quad \underbrace{S \longrightarrow \cdots \longrightarrow \text{objetivo}}_{\text{parte ya comprendida}}.$$

La caja central concentra ahora la búsqueda.

Este principio será importante más adelante: cuando una prueba parece difícil, no siempre necesitamos inventar una cadena completa de una sola vez. Puede bastar con **identificar exactamente qué enlace falta**.

Un ejemplo con funciones. Sea

$$f : A \to B$$

y supongamos que $g : B \to C$ es una función tal que la composición

$$g \circ f : A \to C$$

es sobreyectiva.

Queremos demostrar que g es sobreyectiva.

La conclusión significa que debemos tomar un elemento arbitrario

$$c \in C$$

y encontrar un elemento

$$b \in B$$

tal que

$$g(b) = c.$$

Éste es nuestro frente desde el objetivo:

$$\boxed{\text{necesitamos un } b \in B \text{ con } g(b) = c}.$$

Ahora avancemos desde la hipótesis. Como $g \circ f$ es sobreyectiva, para ese mismo $c \in C$ existe

$$a \in A$$

tal que

$$(g \circ f)(a) = c.$$

Es decir,

$$g(f(a)) = c.$$

Nuestro frente hacia adelante ha producido

$$\boxed{g(f(a)) = c}.$$

La conexión se vuelve visible en cuanto comparamos ambas expresiones.

Necesitamos un $b \in B$ cuya imagen por g sea c, y ya disponemos del objeto

$$f(a) \in B$$

con precisamente esa propiedad.

Por tanto elegimos

$$b = f(a).$$

Aquí el puente no fue una nueva identidad complicada. Fue reconocer que un objeto producido por el frente hacia adelante tiene exactamente el tipo que pedía el frente construido desde el objetivo.

Este ejemplo muestra que el encuentro puede ocurrir no sólo entre fórmulas, sino también entre **objetos requeridos y objetos disponibles**.

Comparar el tipo de información. Cuando los dos frentes parecen cercanos conviene comparar no sólo las expresiones, sino también qué clase de información contienen.

Podemos preguntar:

- ¿necesito una igualdad y ya tengo otra igualdad con los mismos objetos?;
- ¿necesito una desigualdad y he obtenido una cantidad no negativa relacionada?;
- ¿necesito un elemento de cierto conjunto y ya he construido un candidato?;
- ¿necesito una existencia y dispongo de otra existencia cuyo testigo puede reutilizarse?;
- ¿necesito demostrar que dos objetos coinciden y he obtenido una hipótesis que activa una condición de unicidad o inyectividad?;
- ¿necesito una pertenencia y ya tengo una cadena de inclusiones que puede transportarla?

Estas preguntas no son recetas independientes. Son maneras de comprobar si ambos lados del hueco están expresados en formas compatibles.

A veces los extremos casi se encuentran. Puede ocurrir que obtengamos

$$R$$

y necesitemos

$$S,$$

pero que la relación entre ambas no sea visible en la representación actual.

Por ejemplo, un lado puede estar escrito mediante una definición y el otro mediante una expresión equivalente; o un lado puede hablar de un objeto mientras el otro habla de sus componentes.

En ese caso la situación puede parecer

$$R \quad \Big| \quad S,$$

cuando en realidad una reescritura adecuada revelaría

$$R \Longleftrightarrow S$$

o al menos

$$R \Rightarrow S.$$

Ésta será la función de la sección siguiente: estudiar cuándo conviene **desplegar una definición, cambiar de representación o introducir un objeto auxiliar** para volver visible una conexión que ya está cerca.

No forzar un encuentro inexistente. La búsqueda de un puente también puede inducir errores.

Supongamos que desde las hipótesis obtenemos R y desde el objetivo descubrimos que S sería suficiente. No podemos declarar que la prueba está prácticamente terminada sólo porque R y S «se parecen».

Necesitamos una relación demostrable:

$$R \Rightarrow S,$$

o alguna cadena intermedia que produzca esa implicación.

La semejanza notacional no es una inferencia.

Tampoco debemos modificar ilegítimamente uno de los frentes para hacerlos coincidir. Si para conectar R con S necesitamos suponer precisamente lo que intentamos demostrar, hemos introducido circularidad.

Por tanto, el encuentro debe ser **verificable**, no sólo visualmente convincente.

El hueco más pequeño es el problema actual. Durante la búsqueda puede ser útil formular de manera explícita:

¿Cuál es el hueco más pequeño que todavía no sé cerrar?

Al principio podría ser

$$\text{hipótesis} \quad \Big| \quad \text{objetivo}.$$

Después de varias transformaciones quizá sea

$$R_3 \quad \Big| \quad S_2.$$

Ese cambio representa progreso, aunque la demostración todavía no esté completa.

Hemos reemplazado una pregunta global por una obligación intermedia más precisa.

Más adelante podremos descubrir que el hueco se cierra desplegando una definición, cambiando la representación, formulando un lema o abandonando una ruta inadecuada.

Pero antes de hacer cualquiera de esas cosas necesitamos saber **qué enlace estamos intentando construir**.

Del mapa de búsqueda a la cadena de prueba. Supongamos que el borrador ha producido finalmente

$$H \Rightarrow R_1 \Rightarrow R_2,$$

y, desde el objetivo,

$$S_2 \Rightarrow S_1 \Rightarrow Q.$$

Si encontramos el puente

$$R_2 \Rightarrow S_2,$$

entonces ya podemos leer toda la estructura en una sola dirección:

$$H \Rightarrow R_1 \Rightarrow R_2 \Rightarrow S_2 \Rightarrow S_1 \Rightarrow Q.$$

Ésta es la cadena que deberá sobrevivir en la demostración final.

El borrador puede haber descubierto sus piezas en otro orden. La prueba debe conservar sólo las dependencias que realmente han quedado justificadas.

El punto de encuentro

El trabajo hacia adelante produce información alcanzable desde los datos.

El trabajo hacia atrás identifica información suficiente para alcanzar el objetivo.

La búsqueda se concentra en conectar ambos frentes:

$$\text{datos} \longrightarrow \boxed{R} \quad \Big| \quad \boxed{S} \longrightarrow \text{objetivo}.$$

Si demostramos $R \Rightarrow S$, o descubrimos que R y S son dos formas de la misma información, el hueco queda cerrado.

Comprobación. Supón que, durante una búsqueda, has obtenido desde las hipótesis

$$H \Rightarrow R,$$

y al trabajar hacia atrás desde el objetivo Q has comprobado

$$S \Rightarrow Q.$$

¿Qué falta exactamente para completar una ruta de prueba a través de esas dos afirmaciones?

Respuesta razonada.

Falta conectar la información disponible R con la condición suficiente S.

La forma más directa sería demostrar

$$R \Rightarrow S.$$

Entonces podríamos componer las implicaciones:

$$H \Rightarrow R \Rightarrow S \Rightarrow Q,$$

y por tanto

$$H \Rightarrow Q.$$

También podría ocurrir que descubramos

$$R \iff S,$$

o que ambos sean expresiones diferentes de una misma afirmación; eso también proporcionaría el enlace necesario.

En cambio, conocer sólo

$$S \Rightarrow R$$

no cerraría el hueco en general, porque la prueba necesita avanzar desde lo que ya sabemos, R, hacia lo que basta para terminar, S.

La orientación del puente sigue siendo parte de la prueba.

Cuando el puente todavía está oculto. Después de identificar el hueco puede ocurrir que no sepamos todavía cómo cerrarlo.

Eso no nos devuelve al punto de partida. Ahora conocemos con precisión qué relación falta.

La siguiente pregunta será:

«¿Están R y S escritos de la forma adecuada para reconocer su conexión?»

En §7.5 estudiaremos cómo las **definiciones, las representaciones y los objetos auxiliares** pueden hacer visible un puente que la formulación actual mantiene oculto.

7.5. Definiciones, representaciones y objetos auxiliares

En la sección anterior redujimos la búsqueda a un problema local. Desde las hipótesis sabemos llegar hasta cierta información R; desde el objetivo sabemos que cierta información S bastaría para terminar. El hueco es entonces

$$R \quad \Big| \quad S.$$

A veces el puente no aparece porque falte un teorema nuevo. Puede estar oculto porque R y S están escritos en lenguajes distintos.

Una afirmación puede presentarse como pertenencia a un conjunto, como igualdad, como existencia de un objeto, como propiedad definida o como relación entre varios objetos. Cambiar de una de esas representaciones a otra puede revelar una conexión que antes no era visible.

Por eso, cuando el hueco ya está localizado, conviene preguntar:

> **¿Qué forma debería tener la información para que ambos lados pudieran interactuar?**

Esta pregunta conduce a tres movimientos muy frecuentes:

1. desplegar una definición;
2. cambiar de representación;
3. introducir un objeto auxiliar.

Los tres pueden ser muy eficaces. También pueden producir ruido si se aplican sin un problema concreto que resolver.

Desplegar una definición en el momento adecuado. En el capítulo 5 aprendimos que una definición transforma una expresión matemática en condiciones operativas.

Aquí la pregunta es distinta.

No preguntamos simplemente

«¿Qué significa esta definición?»

sino

«¿Me conviene desplegarla **ahora** para cerrar el hueco que tengo delante?»

Consideremos un ejemplo.

Sea

$$f : A \to B$$

una función inyectiva, y sean

$$X, Y \subseteq A$$

tales que

$$X \cap Y = \varnothing.$$

Queremos demostrar

$$f(X) \cap f(Y) = \varnothing.$$

La conclusión habla de una intersección de imágenes. Si intentamos manipular directamente los nombres de los conjuntos, el puente puede no ser evidente.

La definición de igualdad con el vacío sugiere que no debe existir ningún elemento en esa intersección. Supongamos entonces que un elemento

$$z \in f(X) \cap f(Y)$$

estuviera disponible.

Desplegar la intersección produce dos datos:

$$z \in f(X) \qquad \text{y} \qquad z \in f(Y).$$

Ahora conviene desplegar la definición de imagen. De la primera pertenencia obtenemos que existe

$$x \in X$$

tal que

$$f(x) = z.$$

De la segunda obtenemos que existe

$$y \in Y$$

tal que

$$f(y) = z.$$

Por tanto,

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

La inyectividad, que en la representación inicial parecía separada de la intersección de imágenes, se vuelve ahora utilizable:

$$x = y.$$

Pero entonces ese mismo elemento pertenece a X y a Y, de modo que

$$x \in X \cap Y,$$

contradiciendo

$$X \cap Y = \varnothing.$$

La prueba no apareció porque repitiéramos todas las definiciones disponibles. Apareció porque desplegamos **las definiciones que convertían el hueco en información capaz de interactuar con la inyectividad**.

Podemos resumir el cambio así:

$$f(X) \cap f(Y) \quad \rightsquigarrow \quad z \in f(X),\ z \in f(Y) \quad \rightsquigarrow \quad f(x) = f(y) \quad \rightsquigarrow \quad x = y.$$

La definición actuó como instrumento de traducción.

No desplegar todo al mismo tiempo. El ejemplo anterior podría inducir otra mala costumbre: ante cualquier problema, expandir inmediatamente todas las definiciones.

Eso suele empeorar la búsqueda.

Supongamos que un enunciado contiene simultáneamente:

- una función inyectiva;
- dos imágenes de conjuntos;
- una intersección;
- una inclusión;
- una igualdad de conjuntos.

Cada una de esas expresiones admite un despliegue detallado. Si lo hacemos todo de una vez, podemos transformar una pregunta legible en una masa de cuantificadores y variables.

La selección debe responder al hueco actual.

Si necesitamos activar la inyectividad, probablemente buscaremos una igualdad

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

Si necesitamos demostrar una inclusión, quizá convenga introducir un elemento arbitrario.

Si necesitamos utilizar una existencia, quizá debamos extraer un testigo.

La pregunta estratégica es:

> **¿Qué definición produce el tipo de información que necesito en el siguiente enlace?**

Cambiar de representación. A veces el problema no contiene una definición especialmente oculta. El obstáculo consiste en que una misma información puede escribirse de varias maneras y sólo una de ellas hace visible la conexión.

Supongamos que queremos demostrar, para $a, b \in \mathbb{R}$,

$$a^2 + b^2 \geq 2ab.$$

En la forma original aparecen tres términos cuadráticos o bilineales. Podemos mover todo a un lado:

$$a^2 - 2ab + b^2 \geq 0.$$

Ahora reconocemos

$$a^2 - 2ab + b^2 = (a - b)^2.$$

La desigualdad se convierte en

$$(a - b)^2 \geq 0,$$

que es inmediata.

No hemos cambiado el problema. Hemos cambiado su representación hasta hacer visible una propiedad disponible: la no negatividad de los cuadrados reales.

El movimiento puede esquematizarse como

$$\boxed{a^2 + b^2 \geq 2ab} \quad \longleftrightarrow \quad \boxed{(a-b)^2 \geq 0}.$$

En otros problemas el cambio de representación puede consistir en:

- pasar de una igualdad de conjuntos a dos inclusiones;
- pasar de una pertenencia a una preimagen a una condición sobre la imagen del elemento;
- escribir una divisibilidad como existencia de un entero multiplicador;
- transformar una igualdad de funciones en una igualdad punto a punto;
- expresar una afirmación verbal mediante cuantificadores;
- sustituir una expresión por otra equivalente que haga visible una hipótesis.

El objetivo no es preferir siempre una representación «más algebraica» o «más simbólica». Es encontrar una representación en la que las dependencias relevantes se vuelvan operativas.

Controlar qué se conserva al cambiar de forma. Un cambio de representación sólo sirve si sabemos qué relación mantiene con la formulación original.

A veces tenemos equivalencia:

$$R \Longleftrightarrow S.$$

Entonces podemos movernos entre ambas formas sin perder información.

Otras veces sólo tenemos una implicación:

$$R \Rightarrow S.$$

En ese caso el cambio puede ser útil en una dirección y peligroso en la otra.

Por ejemplo,

$$x = y \Rightarrow x^2 = y^2,$$

pero no podemos reemplazar libremente una meta $x = y$ por $x^2 = y^2$.

Así, cambiar de representación no suspende las reglas del capítulo anterior. Cada reescritura debe conservar la orientación lógica que realmente poseemos.

Introducir un objeto auxiliar. A veces los dos lados del hueco no se conectan porque falta un objeto que permita transportar la información.

Consideremos funciones

$$f : A \to B, \qquad g, h : B \to C,$$

y supongamos que f es sobreyectiva y que

$$g \circ f = h \circ f.$$

Queremos demostrar

$$g = h.$$

La igualdad de funciones nos pide probar que para todo

$$b \in B$$

se cumple

$$g(b) = h(b).$$

Fijemos entonces un $b \in B$ arbitrario.

Nuestro objetivo local es

$$g(b) = h(b).$$

La hipótesis disponible, sin embargo, habla de valores de la forma

$$g(f(a)) \qquad \text{y} \qquad h(f(a)).$$

El hueco está claro:

$$b \qquad \Big| \qquad f(a).$$

Aquí la sobreyectividad indica exactamente qué objeto auxiliar necesitamos. Como f es sobreyectiva, existe

$$a \in A$$

tal que

$$f(a) = b.$$

Ese a no aparece en la conclusión del teorema. Su función es permitirnos llevar b al lenguaje en que actúa la hipótesis de composición.

Ahora,

$$g(b) = g(f(a)) = h(f(a)) = h(b).$$

Por tanto,

$$g(b) = h(b).$$

Como b era arbitrario,

$$g = h.$$

El objeto auxiliar a fue un **puente**. No era parte del objetivo final; fue introducido porque resolvía una incompatibilidad concreta entre las formas disponibles de la información.

Un objeto auxiliar debe tener una función. Durante la exploración es legítimo introducir nombres, puntos, conjuntos, variables o expresiones nuevas.

Pero cada introducción aumenta la complejidad del problema.

Por eso conviene poder responder:

> **¿Qué relación espero volver posible gracias a este objeto?**

En el ejemplo anterior, la respuesta era precisa:

> necesito representar el elemento arbitrario $b \in B$ como $f(a)$ para poder aplicar la igualdad $g \circ f = h \circ f$.

Sin una función identificable, introducir objetos adicionales puede generar nuevas obligaciones sin reducir el hueco original.

Un buen objeto auxiliar suele cumplir al menos una de estas tareas:

- convierte un objeto dado en uno al que puede aplicarse una hipótesis;
- reúne varias cantidades en una sola expresión;

- proporciona un testigo para una existencia;
- permite comparar dos objetos mediante una propiedad de unicidad;
- traduce una configuración a una representación más manejable.

Definición, representación y auxiliar pueden actuar juntos. Los tres movimientos no son independientes.

En el ejemplo de las imágenes de conjuntos:

1. desplegamos la definición de pertenencia a una imagen;
2. esa definición produjo objetos auxiliares x e y;
3. la igualdad de sus imágenes puso el problema en la representación exacta que necesitaba la inyectividad.

La secuencia fue

$$\text{intersección de imágenes}$$
$$\longrightarrow \text{existencia de preimágenes} \longrightarrow \text{igualdad de valores} \longrightarrow \text{igualdad de entradas.}$$

Una estrategia real puede combinar varios cambios de lenguaje antes de que aparezca el puente.

Elegir la representación que reduce el hueco. Podemos reformular el principio general de esta sección.

Supongamos que el trabajo previo nos ha dejado

$$R \quad \Big| \quad S.$$

Antes de buscar una idea completamente nueva podemos intentar transformar uno o ambos lados:

$$R \longrightarrow R' \quad \Big| \quad S' \longrightarrow S.$$

La transformación es útil si el nuevo hueco

$$R' \quad \Big| \quad S'$$

es más fácil de cerrar que el anterior.

Éste es el criterio relevante.

No importa cuántas reescrituras hagamos, sino si la nueva forma produce una relación que podamos justificar.

Cambiar la forma para revelar el puente

Una definición, una representación alternativa o un objeto auxiliar son útiles cuando convierten el hueco actual en uno más pequeño o más explícito.

No desplegamos definiciones por rutina ni introducimos objetos por decoración. Cada cambio debe responder a una pregunta concreta:

¿qué conexión se vuelve visible en esta nueva forma?

Comprobación. Sean

$$f : A \to B, \qquad g, h : B \to C,$$

con f sobreyectiva, y supón

$$g \circ f = h \circ f.$$

Quieres demostrar $g = h$.

¿Por qué, para un $b \in B$ arbitrario, introducir un $a \in A$ tal que $f(a) = b$ es una decisión estratégica y no una variable añadida sin propósito?

Respuesta razonada.

La meta local es

$$g(b) = h(b),$$

mientras que la hipótesis de composición proporciona información sobre expresiones de la forma

$$g(f(a)) \qquad \text{y} \qquad h(f(a)).$$

Por tanto, las dos partes del problema están escritas inicialmente en formas incompatibles.

La sobreyectividad de f permite elegir $a \in A$ con

$$f(a) = b.$$

Así podemos reescribir el objetivo como

$$g(f(a)) = h(f(a)),$$

que es precisamente la información suministrada por

$$g \circ f = h \circ f.$$

El objeto auxiliar a tiene, entonces, una función identificable: transforma el elemento b del objetivo en un argumento al que puede aplicarse la hipótesis disponible.

Cuando la nueva representación todavía no basta. Puede ocurrir que, después de desplegar las definiciones pertinentes, cambiar de representación e introducir los objetos naturales, el hueco siga abierto.

Entonces quizá no falte otra reescritura, sino una afirmación matemática intermedia que todavía no hemos demostrado.

En ese caso la pregunta cambia de nuevo:

> «¿Qué resultado intermedio resolvería exactamente este obstáculo?»

Esa pregunta conduce al siguiente paso: **construir un lema auxiliar**.

7.6. Construir un lema auxiliar

A veces el hueco ya está bien localizado y, sin embargo, ninguna reescritura elemental consigue cerrarlo.

Sabemos obtener una afirmación R desde las hipótesis. También sabemos que una afirmación S bastaría para alcanzar el objetivo. Pero todavía falta justificar

$$R \Rightarrow S.$$

En ese momento puede ser útil aislar esa obligación y convertirla en un problema propio.

Eso es, en esencia, construir un **lema auxiliar**.

La idea no consiste en añadir un resultado porque parezca interesante. Consiste en formular una afirmación intermedia cuya función dentro de la prueba pueda explicarse con precisión.

Podemos representar la situación así:

$$\text{hipótesis} \longrightarrow R \quad \Big| \quad S \longrightarrow \text{objetivo}.$$

Si el obstáculo es pasar de R a S, un lema útil debería producir precisamente ese enlace o uno suficientemente cercano:

$$\boxed{R \Rightarrow S}.$$

Antes de intentar demostrarlo conviene poder responder cuatro preguntas:

1. ¿qué obstáculo resuelve?;
2. ¿qué conclusión necesita producir?;
3. ¿sus hipótesis son alcanzables desde los datos disponibles?;
4. ¿realmente simplifica el problema principal?

Estas preguntas distinguen un lema estratégico de una afirmación lateral que sólo aumenta la cantidad de trabajo.

Del obstáculo a la formulación. Consideremos tres funciones

$$f : A \to B, \qquad g : B \to C, \qquad h : C \to D,$$

y supongamos que las tres son inyectivas.

Queremos demostrar que

$$h \circ g \circ f$$

es inyectiva.

Podemos atacar directamente la composición completa. Sean $x, y \in A$ y supongamos

$$(h \circ g \circ f)(x) = (h \circ g \circ f)(y).$$

Entonces, usando sucesivamente la inyectividad de h, de g y de f, obtendríamos

$$g(f(x)) = g(f(y)),$$

luego

$$f(x) = f(y),$$

y finalmente

$$x = y.$$

La prueba funciona. Pero contiene un patrón que merece aislarse: **la composición de dos funciones inyectivas vuelve a ser inyectiva**.

Ese patrón puede formularse como un lema.

Lema. Sean

$$p : X \to Y \qquad \text{y} \qquad q : Y \to Z.$$

Si p y q son inyectivas, entonces

$$q \circ p : X \to Z$$

es inyectiva.

¿Por qué esta afirmación tiene la forma adecuada?

Porque sus hipótesis son exactamente propiedades que ya poseemos para las funciones de la prueba principal, y su conclusión produce exactamente el tipo de objeto que necesitamos volver a utilizar.

Demostrar el lema. Sean $u, v \in X$ y supongamos

$$(q \circ p)(u) = (q \circ p)(v).$$

Entonces

$$q(p(u)) = q(p(v)).$$

Como q es inyectiva,

$$p(u) = p(v).$$

Como p es inyectiva,

$$u = v.$$

Por tanto,

$$q \circ p$$

es inyectiva.

□

El lema ya está disponible como una pieza independiente.

Volver al problema principal. Ahora podemos utilizarlo dos veces.

Como f y g son inyectivas, el lema implica que

$$g \circ f : A \to C$$

es inyectiva.

Como $g \circ f$ es inyectiva y h también lo es, una segunda aplicación del lema nos da que

$$h \circ (g \circ f) : A \to D$$

es inyectiva.

Por asociatividad de la composición,

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

y escribimos simplemente

$$h \circ g \circ f.$$

La prueba principal se ha reducido a dos aplicaciones de una afirmación intermedia ya demostrada.

El lema no introdujo una idea ajena al problema. Aisló una estructura que estaba presente dentro de él.

Un lema debe atacar el hueco correcto. Supongamos que nuestro problema tiene la forma

$$H \Rightarrow R \qquad \Big| \qquad S \Rightarrow Q.$$

Un lema que demuestre

$$R \Rightarrow T$$

puede ser matemáticamente correcto y, sin embargo, no servir de mucho si T no nos acerca a S.

De manera semejante, un resultado muy potente puede ser innecesario si el hueco concreto requiere algo mucho más pequeño.

Por eso, antes de formular un lema, conviene completar la frase:

> «Si pudiera demostrar que _____, entonces el problema principal quedaría reducido a _____».

La primera parte debe describir el resultado intermedio. La segunda debe mostrar su función.

Si no podemos completar esa frase con claridad, quizá todavía no hayamos identificado bien el obstáculo.

Hipótesis alcanzables. Un lema puede tener una conclusión perfecta y aun así ser inútil si sus hipótesis no pueden obtenerse desde los datos del problema.

Imaginemos que para cerrar un hueco necesitaríamos demostrar S y proponemos el lema

$$P \wedge T \Rightarrow S.$$

Antes de celebrarlo debemos preguntar:

¿puedo obtener P y T desde las hipótesis originales?

Si sólo disponemos de P y no existe ninguna ruta hacia T, hemos desplazado el hueco, no lo hemos reducido.

El nuevo esquema sería

$$\text{datos} \longrightarrow P \qquad \Big| \qquad P \wedge T \longrightarrow S \longrightarrow Q.$$

Ahora el problema es producir T.

Eso puede ser progreso si T es más concreto o más accesible que el obstáculo original. Pero no debemos confundir un cambio de nombre con una solución.

Un buen lema crea obligaciones que podamos imaginar cómo satisfacer.

La generalidad adecuada. Al descubrir un lema aparece otra decisión: ¿cuánto debemos generalizarlo?

En el ejemplo anterior podríamos haber formulado únicamente:

> Si f y g son inyectivas, entonces $g \circ f$ es inyectiva.

Eso habría bastado para la primera aplicación, pero el mismo razonamiento habría tenido que repetirse después con $g \circ f$ y h.

En cambio, escribir el lema con funciones arbitrarias

$$p : X \to Y, \qquad q : Y \to Z$$

captura exactamente el patrón reutilizable.

Pero también podríamos intentar una generalización mucho mayor:

> toda composición finita de funciones inyectivas es inyectiva.

La afirmación es natural, pero demostrarla en toda generalidad introduciría cuestiones adicionales que no necesitamos todavía —por ejemplo, una formulación indexada y una prueba inductiva— y anticiparía contenido del capítulo 8.

La regla práctica es:

> **generalizar lo suficiente para capturar el patrón, pero no tanto como para crear un problema nuevo más difícil que el original.**

La generalidad también es una decisión estratégica.

Lemas locales y lemas reutilizables. No todo resultado intermedio necesita recibir un nombre solemne ni convertirse en una pieza permanente del texto.

A veces basta con una afirmación local:

«Necesitamos comprobar primero que R implica S».

Se demuestra esa pequeña afirmación y se continúa con la prueba principal.

En otros casos conviene separarla formalmente como lema porque:

- se utilizará más de una vez;
- encapsula un argumento que distraería dentro de la prueba principal;
- expresa un patrón que probablemente volverá a aparecer;
- o permite que la dependencia de la prueba quede más clara.

La diferencia no es matemática sino arquitectónica.

Un lema formal es útil cuando mejora la estructura de dependencias.

El lema como interfaz. Podemos pensar un lema como una pequeña interfaz entre dos partes de la prueba.

Tiene **entradas**:

$$H_1, H_2, \ldots, H_k,$$

y produce una **salida**:

$$L.$$

Una vez demostrado,

$$H_1 \wedge \cdots \wedge H_k \Rightarrow L,$$

la prueba principal puede utilizar L sin repetir toda la derivación interna.

Esto permite separar dos preguntas:

1. ¿por qué el lema es verdadero?;
2. ¿cómo utiliza la prueba principal la conclusión del lema?

Mantenerlas separadas puede reducir considerablemente la complejidad de una demostración larga.

Evitar la circularidad. La formulación de un lema también puede ocultar un error serio: utilizar, dentro de su demostración, precisamente el resultado principal que el lema debía ayudar a probar.

Supongamos que queremos demostrar un teorema T y proponemos un lema L porque

$$L \Rightarrow T.$$

Si al intentar demostrar L utilizamos T, obtenemos

$$T \Rightarrow L \Rightarrow T.$$

No hemos construido una prueba de T desde las hipótesis originales. Hemos creado un círculo.

Por eso debemos vigilar el grafo de dependencias:

$$\text{hipótesis} \longrightarrow L \longrightarrow T$$

es legítimo;

$$T \longrightarrow L \longrightarrow T$$

no establece T a menos que dispongamos de otra razón independiente para iniciar la cadena.

Un lema auxiliar debe reducir la dependencia del teorema principal, no esconderla.

Un lema que falla también informa. Supongamos que hemos identificado cuidadosamente un lema candidato, pero no conseguimos demostrarlo.

Esto puede significar varias cosas:

- el lema es falso;
- sus hipótesis son insuficientes;
- lo hemos formulado con demasiada generalidad;
- su conclusión es más fuerte de lo necesario;
- necesitamos otra representación;
- o el camino principal que lo motivó no es el adecuado.

Por tanto, fallar al demostrar un lema no nos devuelve necesariamente al comienzo.

Puede localizar con mayor precisión el cuello de botella.

Por ejemplo, si el lema que necesitamos tiene forma

$$P \Rightarrow S$$

pero todos nuestros intentos parecen requerir además una hipótesis T, conviene preguntar si

$$P \wedge T \Rightarrow S$$

es la afirmación correcta y si T está disponible en el problema principal.

El fracaso puede revelar qué dato falta.

Del lema a una prueba modular. Cuando el lema funciona, la estructura de la demostración cambia.

En lugar de una única cadena difícil de leer,

$$H \Rightarrow \cdots \Rightarrow Q,$$

podemos tener dos módulos:

$$\boxed{H \Rightarrow L}$$

y

$$\boxed{L \Rightarrow Q}.$$

Cada módulo puede examinarse por separado.

Éste es uno de los motivos por los que los lemas son tan importantes en matemática: no sólo prueban afirmaciones verdaderas; también **organizan la dependencia de las ideas**.

Un lema con función precisa

Un lema auxiliar debe responder a un obstáculo identificado.

Antes de formularlo, pregunta:

- ¿qué hueco pretende cerrar?;
- ¿qué salida necesita producir?;
- ¿puedo satisfacer sus hipótesis con los datos disponibles?;
- ¿reduce realmente la dificultad de la prueba principal?

Un lema útil no desplaza el problema de manera arbitraria: transforma un hueco amplio en una obligación más pequeña y verificable.

Comprobación. Quieres demostrar un teorema de la forma

$$H \Rightarrow Q.$$

Durante la búsqueda has obtenido

$$H \Rightarrow R,$$

y sabes que

$$S \Rightarrow Q.$$

Propones como lema

$$R \wedge T \Rightarrow S.$$

¿Qué debes verificar antes de considerar que ese lema realmente mejora la situación?

Respuesta razonada.

Primero debemos comprobar que la conclusión del lema es útil. En este caso lo es, porque produce S, y ya sabemos

$$S \Rightarrow Q.$$

Pero eso no basta.

El lema exige dos hipótesis:

$$R \qquad \text{y} \qquad T.$$

La primera ya es alcanzable desde H, pues tenemos

$$H \Rightarrow R.$$

Todavía debemos comprobar que T también puede obtenerse desde las hipótesis originales o desde consecuencias ya disponibles.

Si podemos demostrar

$$H \Rightarrow T,$$

entonces la cadena puede completarse:

$$H \Rightarrow R \wedge T \Rightarrow S \Rightarrow Q.$$

Si no tenemos ninguna ruta hacia T, el lema no ha cerrado el hueco: sólo lo ha trasladado a una nueva obligación.

Por tanto, un lema mejora la búsqueda cuando su conclusión ataca el obstáculo correcto **y** sus hipótesis son efectivamente accesibles.

Cuando el lema revela el verdadero obstáculo. Después de formular un lema podemos descubrir que la dificultad principal no estaba donde creíamos.

Tal vez el resultado intermedio sea fácil, pero una de sus hipótesis no pueda justificarse. Tal vez la formulación sea demasiado fuerte. Tal vez una rama de la búsqueda no interactúe nunca con el objetivo.

En ese momento ya no conviene seguir añadiendo consecuencias o lemas sin control.

La pregunta siguiente será diagnóstica:

> **¿Cuál es exactamente el cuello de botella y qué nos dice el fracaso de los intentos anteriores?**

Ése será el tema de §7.7.

7.7. Detectar el cuello de botella y abandonar callejones

Una búsqueda puede producir muchas afirmaciones correctas y, sin embargo, no acercarse a una demostración.

En ese momento la pregunta útil ya no es

> «¿qué otra cosa puedo intentar?»

sino

> **«¿dónde se detiene exactamente la cadena que necesito y qué impide continuar?»**

Ese punto es el **cuello de botella** de la prueba.

No tiene por qué coincidir con la parte más larga del cálculo ni con la expresión más complicada. Es simplemente la obligación mínima que todavía no sabemos justificar y de la que depende el resto del argumento.

Si el mapa de búsqueda tiene la forma

$$H \Rightarrow R_1 \Rightarrow R_2 \qquad \Big| \qquad S_2 \Rightarrow S_1 \Rightarrow Q,$$

entonces el problema actual no es ya demostrar Q «desde cero». El problema es comprender por qué debería existir un puente entre

$$R_2 \qquad \text{y} \qquad S_2.$$

Localizar ese punto evita seguir trabajando en partes de la demostración que ya están resueltas.

Un ejemplo: el lugar exacto donde actúa una hipótesis. Sean

$$f, g : A \to B$$

y

$$h : B \to C$$

funciones. Supongamos que h es inyectiva y que

$$h \circ f = h \circ g.$$

Queremos demostrar

$$f = g.$$

La igualdad de funciones nos indica que debemos fijar un elemento arbitrario

$$a \in A$$

y demostrar

$$f(a) = g(a).$$

De la hipótesis

$$h \circ f = h \circ g$$

obtenemos, evaluando en a,

$$h(f(a)) = h(g(a)).$$

Ahora la búsqueda queda reducida a

$$\boxed{h(f(a)) = h(g(a))} \quad \Big| \quad \boxed{f(a) = g(a)}.$$

Éste es el cuello de botella.

No necesitamos seguir manipulando la composición ni introducir nuevas variables. Necesitamos una regla que permita pasar de la igualdad de los valores de h a la igualdad de sus argumentos.

La hipótesis de inyectividad tiene exactamente esa función:

$$h(u) = h(v) \Rightarrow u = v.$$

Aplicándola con

$$u = f(a), \qquad v = g(a),$$

obtenemos

$$f(a) = g(a).$$

Como a era arbitrario,

$$f = g.$$

La prueba es breve porque, una vez localizado el cuello de botella, la hipótesis decisiva se vuelve visible.

Una hipótesis no utilizada puede señalar el cuello de botella. En una búsqueda larga es frecuente haber usado varias hipótesis y dejar una aparentemente desconectada.

Eso merece una inspección especial.

En el ejemplo anterior, si hubiéramos llegado hasta

$$h(f(a)) = h(g(a))$$

sin utilizar todavía la inyectividad de h, la pregunta

> «¿qué hipótesis no he usado?»

habría señalado precisamente el recurso que faltaba.

Esto no significa que toda hipótesis deba utilizarse necesariamente. Ya vimos en §7.2 que algunas pueden ser redundantes.

Pero cuando la prueba se detiene en un punto concreto y existe una hipótesis todavía inactiva, conviene comparar **la forma de esa hipótesis** con **la forma del obstáculo**.

La coincidencia puede revelar la función para la cual fue incluida.

Un intento fallido no es todavía un diagnóstico. Supongamos que intentamos una ruta y no obtenemos la conclusión.

Decir

> «esto no funcionó»

contiene poca información.

En cambio, afirmar

> «la ruta llega hasta R, pero para continuar necesitaría S, y sólo he conseguido demostrar $S \Rightarrow R$»

localiza un problema lógico preciso.

O afirmar

> «el lema produciría la conclusión correcta, pero exige una hipótesis T que no puedo obtener de los datos»

identifica una obligación nueva.

El fracaso se vuelve matemáticamente útil cuando podemos describir **qué enlace falta**.

Podemos registrar un intento mediante cuatro preguntas:

1. ¿hasta dónde llegué legítimamente?;
2. ¿qué necesitaba a continuación?;
3. ¿qué paso concreto no pude justificar?;
4. ¿qué información nueva revela ese bloqueo?

Ésta es una forma de convertir un callejón sin salida en información sobre la estructura del problema.

Cinco diagnósticos frecuentes. Cuando una ruta se atasca, el bloqueo suele adoptar alguna de estas formas.

1. Falta activar una hipótesis disponible. Tenemos la información necesaria, pero todavía no está escrita en la forma en que puede utilizarse.

Por ejemplo, disponemos de inyectividad pero aún no hemos producido una igualdad de imágenes a la cual aplicarla.

La respuesta puede ser desplegar una definición o cambiar de representación.

2. La meta intermedia es demasiado fuerte. Al trabajar hacia atrás hemos elegido una condición suficiente R para el objetivo Q:

$$R \Rightarrow Q.$$

Pero R exige mucho más de lo necesario y no parece accesible desde las hipótesis.

En ese caso conviene buscar otra condición suficiente, quizá más débil:

$$S \Rightarrow Q,$$

con S más cercana a la información disponible.

La suficiencia lógica no garantiza utilidad estratégica.

3. La representación mantiene oculto el puente. Hemos obtenido R y necesitamos S, pero ambos describen esencialmente la misma estructura en lenguajes distintos.

Antes de buscar un teorema nuevo puede bastar con volver a §7.5 y preguntar si una definición o reescritura convierte

$$R \quad \Big| \quad S$$

en una relación explícita.

4. La flecha apunta en el sentido incorrecto. Durante la búsqueda encontramos

$$S \Rightarrow R,$$

cuando lo que la prueba necesita es

$$R \Rightarrow S.$$

No hay progreso sólo porque dos afirmaciones estén relacionadas.

La orientación debe coincidir con la cadena final de dependencias.

5. Hemos introducido circularidad. Queremos demostrar Q y proponemos un resultado intermedio L tal que

$$L \Rightarrow Q.$$

Pero al demostrar L utilizamos Q.

El esquema real es

$$Q \Rightarrow L \Rightarrow Q,$$

que no establece Q desde las hipótesis originales.

La circularidad puede esconderse dentro de un lema, una definición equivalente utilizada sin control o una frase como «por el resultado que queremos probar» formulada de manera menos evidente.

Por eso conviene revisar no sólo las afirmaciones, sino su **grafo de dependencias**.

Abandonar una ruta con una razón. Una estrategia improductiva no debe conservarse indefinidamente sólo porque ya hemos invertido trabajo en ella.

Podemos abandonar una rama cuando identificamos una razón estructural suficiente, por ejemplo:

- termina en información que no interactúa con el objetivo;
- exige una hipótesis que no está disponible;
- sustituye el objetivo por una condición mucho más fuerte sin ruta hacia ella;
- depende de invertir una implicación no reversible;
- entra en un ciclo de dependencias;
- repite transformaciones que no reducen el hueco.

Abandonar una ruta en estas condiciones no significa afirmar que el teorema sea falso.

Significa únicamente:

> **esta estrategia concreta no proporciona una cadena justificable desde los datos hasta el objetivo.**

La distinción es esencial.

Un método puede fracasar aunque el teorema sea verdadero.

Fracaso de una estrategia frente a falsedad del enunciado. Supongamos que todos nuestros intentos de demostrar una afirmación parecen requerir una hipótesis adicional T.

Eso puede sugerir preguntas importantes:

- ¿hemos pasado por alto una forma de obtener T?;
- ¿existe otra ruta que no necesite T?;
- ¿el enunciado fue formulado con hipótesis insuficientes?;
- ¿podría ser falso sin T?

Estas preguntas son legítimas.

Pero el hecho de que **nuestra** ruta necesite T no demuestra que el teorema necesite T.

Para concluir que una afirmación universal es falsa hace falta algo diferente: producir y verificar un caso que realmente la contradiga, o disponer de otra demostración de su falsedad.

La búsqueda sistemática de esos casos pertenece al capítulo 9.

Aquí nos interesa una etapa anterior: utilizar el punto de fallo para comprender qué parte de la estructura todavía no controlamos.

No seguir trabajando donde ya no está el problema. Una fuente frecuente de ineficiencia consiste en rehacer partes ya resueltas.

Si tenemos

$$H \Rightarrow R$$

y sabemos también

$$S \Rightarrow Q,$$

pero desconocemos

$$R \Rightarrow S,$$

no tiene sentido seguir produciendo cinco pruebas distintas de $H \Rightarrow R$ salvo que alguna cambie la información disponible en el punto de encuentro.

El cuello de botella sigue siendo

$$R \quad \Big| \quad S.$$

La búsqueda debe concentrarse allí.

Esto puede parecer obvio cuando se escribe simbólicamente, pero en problemas largos es fácil perderlo de vista porque las partes conocidas también contienen cálculos y decisiones.

Una disciplina útil es marcar provisionalmente:

$$\boxed{\text{RESUELTO}} \qquad \boxed{\text{BLOQUEO ACTUAL}} \qquad \boxed{\text{RESUELTO DESDE EL OBJETIVO}}.$$

El borrador deja así de ser una acumulación cronológica de intentos y se convierte en un mapa del estado actual de la prueba.

El callejón puede revelar una hipótesis suficiente. Consideremos una variante del ejemplo inicial.

Supongamos que sólo sabemos

$$h \circ f = h \circ g$$

y queremos demostrar

$$f = g,$$

pero no se nos ha dicho que h sea inyectiva.

El mismo intento produce

$$h(f(a)) = h(g(a)),$$

pero se detiene antes de

$$f(a) = g(a).$$

El bloqueo identifica con precisión una propiedad que sería suficiente:

$$h(u) = h(v) \Rightarrow u = v.$$

Es decir, la inyectividad de h cerraría la ruta.

Esto no prueba todavía que la inyectividad sea necesaria para el enunciado ni que, sin ella, la afirmación sea falsa en todos los casos.

Lo que sí hemos aprendido es algo concreto:

> **esta prueba necesita una forma de cancelar h, y la inyectividad proporciona exactamente esa cancelación.**

El intento fallido ha producido información estructural.

Cambiar de estrategia sin perder lo aprendido. Abandonar una ruta no obliga a descartar todo lo obtenido en ella.

Podemos conservar:

- consecuencias válidas de las hipótesis;
- definiciones que resultaron pertinentes;
- objetos auxiliares bien motivados;
- lemas ya demostrados;
- el lugar exacto donde apareció una obligación imposible de justificar.

Lo que se abandona es la **pretensión de que esa secuencia concreta cerrará la prueba**.

Después podemos cambiar de representación, buscar una meta intermedia distinta o explorar ejemplos sencillos para descubrir qué comportamiento debería esperarse.

Por eso un buen borrador no borra simplemente los intentos fallidos. Conserva, al menos durante la búsqueda, la razón de su descarte.

Un registro mínimo de diagnóstico. Cuando una prueba se bloquea, puede ser útil escribir sólo cuatro líneas:

DISPONIBLE:	R,
NECESITO:	S,
BLOQUEO:	no sé justificar $R \Rightarrow S$,
SIGUIENTE PREGUNTA:	¿qué haría visible o demostrable ese enlace?

Este registro impide que la dificultad vuelva a convertirse en la pregunta vaga

> «¿cómo se demuestra todo esto?»

La búsqueda permanece localizada.

Diagnosticar antes de multiplicar intentos

Un callejón sin salida es informativo sólo cuando podemos identificar el punto exacto donde la cadena deja de estar justificada.

Antes de intentar otra estrategia, pregunta:

- ¿qué parte ya está resuelta?;
- ¿cuál es la obligación mínima pendiente?;
- ¿qué hipótesis, representación o dirección lógica interviene en ese bloqueo?;
- ¿la ruta puede repararse o conviene abandonarla?

No confundas el fracaso de una estrategia con la falsedad del teorema.

Comprobación. Quieres demostrar

$$H \Rightarrow Q.$$

Ya has obtenido

$$H \Rightarrow R,$$

y sabes que

$$S \Rightarrow Q.$$

Consideras tres rutas:

1. demostrar $R \Rightarrow T$, pero no conoces ninguna relación entre T y S;
2. demostrar $R \Rightarrow S$ utilizando Q como paso intermedio;
3. demostrar primero $R \Rightarrow U$ y después $U \Rightarrow S$.

¿Qué diagnóstico corresponde a cada ruta?

Respuesta razonada.

La primera ruta puede producir una consecuencia verdadera, pero mientras T no tenga conexión con S no ataca el cuello de botella actual. Es una rama que, por ahora, no reduce el hueco.

La segunda ruta tiene la forma

$$R \Rightarrow Q \Rightarrow S$$

para intentar justificar precisamente el enlace que permitiría llegar a Q. Si Q no estaba disponible por una razón independiente, la ruta es circular y debe rechazarse como prueba.

La tercera ruta sí puede ser una descomposición legítima del cuello de botella, siempre que podamos demostrar independientemente

$$R \Rightarrow U$$

y

$$U \Rightarrow S.$$

En ese caso el puente se refina como

$$R \Rightarrow U \Rightarrow S,$$

y la cadena completa queda

$$H \Rightarrow R \Rightarrow U \Rightarrow S \Rightarrow Q.$$

Por tanto, el criterio no es cuántos pasos contiene una ruta, sino si cada paso reduce o resuelve la obligación pendiente sin introducir dependencias ilegítimas.

Del diagnóstico a la exploración controlada. Una vez localizado el cuello de botella podemos necesitar información que las manipulaciones formales todavía no revelan.

Una posibilidad es estudiar casos pequeños, ejemplos concretos o configuraciones extremas para observar qué estructura parece estar actuando.

Ese recurso puede ser muy productivo, pero introduce un nuevo riesgo: confundir lo que un ejemplo **sugiere** con lo que una demostración **establece**.

La siguiente sección estudiará precisamente cómo explorar con ejemplos sin convertir evidencia local en prueba universal.

7.8. Explorar con ejemplos sin confundirlos con pruebas

Cuando una prueba se bloquea, a veces lo que falta no es otra manipulación formal sino una pista sobre **qué estructura conviene buscar**.

En ese momento los ejemplos concretos pueden ser muy útiles.

Podemos probar valores sencillos, estudiar un caso simétrico, examinar una configuración límite o construir una instancia pequeña del problema. El objetivo de estas exploraciones no es reemplazar la demostración, sino hacer visible una regularidad que todavía no sabemos formular.

Conviene distinguir desde el comienzo tres tareas:

$$\boxed{\text{explorar}} \quad \longrightarrow \quad \boxed{\text{conjeturar}} \quad \longrightarrow \quad \boxed{\text{demostrar}}.$$

Una exploración produce información local.

Una conjetura propone qué afirmación general podría explicar esa información.

Una demostración justifica la afirmación para todos los casos cubiertos por sus hipótesis.

Los tres momentos pueden estar relacionados, pero no son intercambiables.

Un ejemplo puede revelar qué deberíamos intentar demostrar. Volvamos a una desigualdad que apareció al comienzo del capítulo.

Sean $x, y \in \mathbb{R}$ tales que

$$x + y = 1.$$

Queremos comprender por qué parece razonable esperar una cota para

$$xy.$$

Antes de conocer la prueba podemos mirar algunos casos:

x	y	xy
0	1	0
$\frac{1}{4}$	$\frac{3}{4}$	$\frac{3}{16}$
$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{4}$
$\frac{3}{4}$	$\frac{1}{4}$	$\frac{3}{16}$
1	0	0

La tabla no demuestra nada acerca de todos los pares reales cuya suma es 1.

Pero hace visible algo interesante: el producto parece aumentar cuando los dos números se aproximan y alcanza, al menos en los casos examinados, su mayor valor cuando

$$x = y = \frac{1}{2}.$$

Esto sugiere la conjetura

$$xy \leq \frac{1}{4}.$$

Más importante todavía, el caso simétrico sugiere mirar la **diferencia** entre x e y.

Si la igualdad ocurre cuando

$$x - y = 0,$$

quizá la distancia entre el producto y la cota esté relacionada con

$$(x - y)^2.$$

Y, en efecto, usando $x + y = 1$,

$$\begin{aligned}\frac{1}{4} - xy &= \frac{(x+y)^2}{4} - xy \\ &= \frac{x^2 - 2xy + y^2}{4} \\ &= \frac{(x-y)^2}{4}.\end{aligned}$$

Ahora sí tenemos una demostración, porque

$$(x-y)^2 \geq 0.$$

La exploración no probó la desigualdad. Ayudó a descubrir **qué expresión podía probarla**.

Podemos distinguir claramente los tres momentos:

$$\begin{aligned}\text{exploración:} &\quad \text{calcular algunos productos;} \\ \text{conjetura:} &\quad xy \leq \frac{1}{4}; \\ \text{demostración:} &\quad \frac{1}{4} - xy = \frac{(x-y)^2}{4} \geq 0.\end{aligned}$$

Esta separación será útil cada vez que un patrón observado sugiera una idea general.

Los ejemplos pueden ayudar a construir un testigo. Los ejemplos no sirven sólo para adivinar fórmulas o desigualdades. También pueden sugerir **qué objeto construir** cuando el objetivo es existencial.

Supongamos que

$$a, b \in \mathbb{R}, \qquad a < b,$$

y queremos demostrar que existe un real c tal que

$$a < c < b.$$

La arquitectura es existencial: necesitamos construir un testigo.

Pero quizá no sepamos inmediatamente cuál.

Probemos algunos casos.

Si

$$a = 2, \qquad b = 8,$$

un candidato natural es

$$c = 5.$$

Si

$$a = -4, \qquad b = 2,$$

podemos elegir

$$c = -1.$$

En ambos casos hemos tomado el punto medio.

Esto sugiere

$$c = \frac{a+b}{2}.$$

Los dos ejemplos no demuestran que este candidato funcione para todo par $a < b$. Su función fue revelar una construcción uniforme.

Ahora debemos verificarla.

De

$$a < b$$

obtenemos, sumando a,

$$2a < a + b.$$

Dividiendo por $2 > 0$,

$$a < \frac{a+b}{2}.$$

De manera análoga, sumando b a $a < b$,

$$a + b < 2b,$$

y por tanto

$$\frac{a+b}{2} < b.$$

Así,

$$a < \frac{a+b}{2} < b.$$

La exploración concreta descubrió el candidato; la demostración general verificó que funciona.

Este esquema aparece con frecuencia:

$$\text{casos concretos} \longrightarrow \text{candidato} \longrightarrow \text{verificación general.}$$

Elegir ejemplos que respondan una pregunta. No todos los ejemplos son igualmente informativos.

Si elegimos valores al azar sin saber qué estamos buscando, podemos acumular datos sin reducir el cuello de botella.

Un ejemplo es más útil cuando responde a una pregunta concreta.

Podemos buscar, por ejemplo:

- un caso donde varias hipótesis se vuelvan especialmente simples;
- una configuración simétrica;
- un caso en que dos objetos coincidan;
- un valor cero o identidad que elimine términos;
- una situación límite donde esperemos igualdad;
- un ejemplo pequeño en el que podamos enumerar todos los objetos relevantes;
- dos ejemplos que difieran sólo en una característica que queremos comprender.

La elección del ejemplo forma parte del razonamiento.

No preguntamos simplemente

> «¿qué pasa si pongo números?»

sino

«¿qué aspecto del problema quiero aislar con este ejemplo?»

Esta pregunta evita que la experimentación se convierta en una colección de cálculos desconectados.

Casos simétricos y configuraciones límite. Las configuraciones especialmente simples pueden señalar dónde buscar una igualdad o qué condición podría ser decisiva.

En la desigualdad

$$xy \leq \frac{1}{4}$$

bajo la restricción

$$x + y = 1,$$

el caso

$$x = y$$

no fue elegido al azar.

La expresión xy es simétrica en x e y: intercambiarlos no modifica el producto ni la restricción. Por ello es natural examinar qué ocurre cuando ambos tienen el mismo valor.

La condición

$$x = y, \qquad x + y = 1$$

obliga a

$$x = y = \frac{1}{2}.$$

Ese caso sugiere tanto la constante

$$\frac{1}{4}$$

como la condición de igualdad.

Una configuración especial puede, por tanto, proporcionar dos pistas distintas:

1. **qué afirmación general podría ser cierta**;
2. **qué mecanismo podría explicar el caso de igualdad**.

Ninguna de las dos pistas constituye todavía una prueba.

Probar una conjetura contra casos exigentes. Una vez que aparece una conjetura, no conviene buscar solamente ejemplos que la confirmen.

También podemos preguntarnos qué casos serían especialmente difíciles para ella.

Supongamos que, observando algunos pares con suma 1, formulamos demasiado deprisa la conjetura

$$xy < \frac{1}{4}.$$

El caso simétrico

$$x = y = \frac{1}{2}$$

produce

$$xy = \frac{1}{4}.$$

Así descubrimos inmediatamente que la desigualdad estricta no puede ser la formulación correcta.

La versión compatible con ese caso es

$$xy \leq \frac{1}{4}.$$

Aquí un ejemplo ha servido para **corregir una conjetura antes de invertir trabajo en demostrarla**.

Este uso es importante, pero debemos mantener la frontera conceptual.

En este capítulo nos interesa el ejemplo como instrumento de exploración y diagnóstico. La búsqueda sistemática de casos que refutan afirmaciones universales, y el diseño deliberado de contraejemplos, se estudiarán en el capítulo 9.

La asimetría entre confirmar y refutar. Hay una diferencia lógica que conviene mantener siempre visible.

Si una afirmación dice

$$\forall x \in A \; P(x),$$

comprobar que

$$P(a_1),\ P(a_2), \ldots,\ P(a_{100})$$

para cien elementos concretos no demuestra la universal.

Sólo sabemos que esos cien casos funcionan.

En cambio, si encontramos un elemento

$$a \in A$$

para el cual

$$\neg P(a),$$

la afirmación universal queda refutada.

Esta asimetría ya apareció al estudiar contraejemplos en C06.

Para el trabajo de descubrimiento significa lo siguiente:

- muchos ejemplos favorables pueden aumentar nuestra confianza o revelar un patrón, pero no sustituyen la prueba;
- un caso contrario obliga a revisar inmediatamente la conjetura universal.

La función de la experimentación debe interpretarse de acuerdo con la forma lógica del enunciado.

Una tabla puede revelar una regularidad, no su causa. Supongamos que observamos una secuencia de casos y todos cumplen la misma propiedad.

Podemos registrar

caso	resultado
1	funciona
2	funciona
3	funciona
4	funciona

La regularidad puede ser valiosa porque nos permite preguntar:

«¿qué tienen en común estos casos que pueda expresarse de manera general?»

Pero la tabla no responde esa pregunta.

La demostración debe identificar un **mecanismo** que no dependa de haber enumerado los ejemplos.

Por eso el paso decisivo después de detectar un patrón es dejar de preguntar

> «¿sigue funcionando en el siguiente caso?»

para preguntar

> **«¿qué propiedad compartida por todos los casos explicaría por qué funciona?»**

Éste es el movimiento que transforma observación en matemática demostrativa.

Variar una característica a la vez. Los ejemplos también pueden utilizarse comparativamente.

Supongamos que una afirmación depende de varias hipótesis:

$$H_1, \qquad H_2, \qquad H_3.$$

Si queremos comprender qué papel cumple H_2, puede ser útil comparar dos situaciones que mantengan aproximadamente fija la estructura relevante y difieran precisamente en esa condición.

El objetivo no es todavía demostrar que H_2 sea necesaria.

Queremos observar si el comportamiento del problema cambia cuando esa característica desaparece.

Este tipo de comparación puede sugerir:

- qué hipótesis activa cierto paso;
- cuál puede ser redundante;
- dónde buscar una construcción;
- qué formulación de un lema sería razonable.

Después habrá que justificar cualquiera de esas conclusiones por medios generales.

Ejemplos pequeños para comprender definiciones abstractas. En problemas con conjuntos, funciones o relaciones, una instancia finita puede permitir ver toda la estructura de una sola vez.

Supongamos que estamos estudiando una función

$$f : A \to B$$

y una afirmación que relaciona imágenes y preimágenes.

Elegir conjuntos pequeños, por ejemplo

$$A = \{1, 2, 3\}, \qquad B = \{p, q\},$$

y dibujar una función concreta puede ayudarnos a distinguir:

- qué información conserva una imagen;
- qué elementos pueden compartir valor;
- dónde interviene la inyectividad;
- qué produce realmente una preimagen.

El diagrama concreto puede revelar qué definición conviene desplegar en la prueba general.

Pero una vez identificada esa estructura, la prueba debe regresar a elementos arbitrarios y a las definiciones pertinentes.

La instancia finita funciona como **modelo de exploración**, no como sustituto del argumento general.

Del ejemplo a la conjetura: formular exactamente qué creemos haber visto. Una exploración es más útil cuando termina en una afirmación explícita.

En vez de escribir simplemente

«parece que funciona»,

conviene formular:

$$\boxed{\text{CONJETURA: } C}.$$

Después debemos preguntar:

1. ¿cuál es el dominio exacto de C?;
2. ¿qué hipótesis necesita?;
3. ¿es universal, existencial o una equivalencia?;
4. ¿qué ejemplos la motivaron?;
5. ¿qué tendría que demostrarse para convertirla en teorema?

Formular la conjetura con precisión impide que la evidencia empírica permanezca en una zona ambigua entre observación y resultado.

Del ejemplo al lema. A veces los casos concretos no sugieren directamente el teorema final, sino una pieza intermedia.

Podemos observar que en varios ejemplos el mismo paso resuelve el cuello de botella.

Entonces la pregunta apropiada puede ser:

«¿Puedo formular ese paso como un lema independiente?»

El proceso es

$$\text{ejemplos} \longrightarrow \text{patrón intermedio} \longrightarrow \text{lema candidato} \longrightarrow \text{demostración del lema}.$$

Esto conecta la exploración con §7.6.

Un lema descubierto experimentalmente sólo se convierte en una pieza legítima de la prueba después de haber sido demostrado con las hipótesis correctas.

Saber cuándo dejar de experimentar. Los ejemplos son especialmente útiles cuando todavía no vemos la estructura.

Pero llega un momento en que seguir calculando casos deja de aportar información nueva.

Una señal clara es que ya podemos formular:

- una conjetura precisa;
- un candidato para el objeto que necesitamos;
- una identidad que parece explicar los casos;
- un lema cuya demostración cerraría el hueco.

En ese momento debemos cambiar de tarea.

La pregunta deja de ser

«¿qué ocurre en más ejemplos?»

para convertirse en

«¿por qué tiene que ocurrir esto en general?»

Ese cambio marca el paso de la exploración a la demostración.

El borrador puede contener evidencia que la prueba final eliminará. Durante la búsqueda podemos conservar tablas, dibujos, cálculos numéricos y casos particulares porque documentan cómo apareció una idea.

Pero la prueba final no necesita incluirlos si no cumplen una función justificativa.

Por ejemplo, la demostración de

$$xy \leq \frac{1}{4}$$

bajo $x + y = 1$ no necesita enumerar los cinco pares que nos ayudaron a adivinar la constante.

Basta con la identidad general

$$\frac{1}{4} - xy = \frac{(x-y)^2}{4} \geq 0.$$

Esto vuelve a mostrar la diferencia entre

historia del descubrimiento

y

estructura de la justificación.

Los ejemplos pueden pertenecer plenamente a la primera sin aparecer en la segunda.

Explorar, conjeturar, demostrar

Un ejemplo concreto puede revelar un patrón, sugerir un testigo, indicar un caso de igualdad o poner a prueba una conjetura.

Pero los ejemplos favorables no establecen por sí solos una afirmación universal.

Después de explorar, formula con precisión qué crees haber descubierto y cambia de pregunta:

¿qué mecanismo general explica todos estos casos?

La respuesta a esa pregunta debe convertirse en una demostración.

Comprobación. Quieres demostrar una afirmación universal

$$\forall x \in A \; P(x).$$

Has comprobado $P(x)$ en numerosos ejemplos y, al observarlos, descubres una identidad general $I(x)$ que parece explicar todos los casos. ¿Qué parte del trabajo es evidencia exploratoria y qué tendría que ocurrir para obtener una demostración?

Respuesta razonada.

Los ejemplos concretos son evidencia exploratoria. Pueden justificar que formulemos una conjetura o que prestemos atención a la identidad $I(x)$, pero no permiten concluir

$$\forall x \in A \; P(x).$$

Para obtener una demostración debemos trabajar con un elemento arbitrario

$$x \in A$$

y justificar una cadena general que conecte la información disponible con $P(x)$.

Si la identidad descubierta permite demostrar

$$I(x) \Rightarrow P(x),$$

y además podemos establecer $I(x)$ para todo $x \in A$ bajo las hipótesis del problema, entonces tendremos una ruta válida:

$$x \in A \longrightarrow I(x) \longrightarrow P(x).$$

Lo que los ejemplos aportaron fue la **idea candidata** I.

Lo que falta es demostrar que esa idea funciona para un elemento arbitrario y, por tanto, para todo el dominio.

De la exploración a la reconstrucción. Con esta sección ya disponemos de casi todos los movimientos elementales del espacio de descubrimiento:

avanzar desde las hipótesis,
trabajar hacia atrás desde el objetivo,
hacer coincidir ambos frentes,
cambiar de representación,
introducir objetos auxiliares,
formular lemas,
diagnosticar callejones,
explorar ejemplos para descubrir estructura.

El resultado de este trabajo puede ser un borrador lleno de flechas, notas, cálculos, casos descartados y afirmaciones descubiertas en un orden muy distinto del lógico.

Todavía queda una tarea decisiva:

> **convertir ese borrador de descubrimiento en una prueba final cuya dependencia pueda verificarse de principio a fin.**

Ése será el objetivo de §7.9.

7.9. Del borrador de descubrimiento a la prueba final

Cuando una búsqueda ha tenido éxito, todavía no necesariamente tenemos una demostración lista para presentar.

Podemos tener una hoja de trabajo en la que aparecen:

- consecuencias obtenidas desde las hipótesis;
- metas auxiliares descubiertas trabajando hacia atrás;
- ejemplos que sugirieron una identidad;
- objetos introducidos provisionalmente;
- lemas que finalmente resultaron útiles;
- cálculos abandonados;
- flechas escritas en direcciones distintas;
- observaciones que fueron importantes para descubrir la idea pero que ya no justifican ningún paso.

Ese material puede ser excelente como **borrador de descubrimiento** y, al mismo tiempo, ser inadecuado como prueba final.

La tarea de esta sección es aprender a reconstruirlo.

La pregunta ya no es

> «¿cómo encuentro una idea?»

sino

> **«¿qué partes de lo encontrado forman realmente una cadena justificable desde las hipótesis hasta la conclusión?»**

La prueba final no es la crónica del descubrimiento. Supongamos que durante la búsqueda fuimos encontrando piezas en este orden:

$$Q \rightsquigarrow S_1 \rightsquigarrow S_2,$$

luego, desde las hipótesis,

$$H \Rightarrow R_1 \Rightarrow R_2,$$

y finalmente descubrimos el puente

$$R_2 \Rightarrow S_2.$$

Cronológicamente, quizá pensamos primero en Q, después en S_1, más tarde volvimos a H y sólo al final apareció $R_2 \Rightarrow S_2$.

Pero la dependencia matemática es

$$H \Rightarrow R_1 \Rightarrow R_2 \Rightarrow S_2 \Rightarrow S_1 \Rightarrow Q.$$

Éste es el orden que debe controlar la prueba final.

La demostración no tiene obligación de reproducir el orden en que ocurrió nuestra búsqueda. Debe reproducir el orden en que las afirmaciones **quedan justificadas**.

Un borrador deliberadamente desordenado. Consideremos una función inyectiva

$$f : A \to B$$

y dos subconjuntos

$$X, Y \subseteq A.$$

Supongamos que

$$f(X) \subseteq f(Y).$$

Queremos demostrar

$$X \subseteq Y.$$

Una hoja de descubrimiento podría contener notas como éstas:

OBJETIVO: $X \subseteq Y$.

Tomar $x \in X$.

Necesito $x \in Y$.

Si encontrara $y \in Y$ con $f(x) = f(y)$,

la inyectividad daría $x = y$.

¿Cómo obtener ese y?

$$f(x) \in f(Y) \quad \Longrightarrow \quad \exists y \in Y \; f(y) = f(x).$$

¿Tengo $f(x) \in f(Y)$?

Como $x \in X$,

$$f(x) \in f(X),$$

y como

$$f(X) \subseteq f(Y),$$

entonces sí:

$$f(x) \in f(Y).$$

El borrador ha sido útil. Hemos trabajado desde el objetivo, hemos identificado qué forma de igualdad activaría la inyectividad y después hemos vuelto a las hipótesis para producirla.

Pero no conviene copiar estas notas, en este orden, como demostración.

La prueba debe reconstruirse según las dependencias.

Extraer la cadena que sobrevivirá. Del borrador anterior podemos conservar la secuencia siguiente:

$$x \in X \Longrightarrow f(x) \in f(X) \Longrightarrow f(x) \in f(Y) \Longrightarrow \exists y \in Y\ f(y) = f(x) \Longrightarrow x = y \Longrightarrow x \in Y.$$

Ahora cada paso tiene una justificación disponible:

1. $x \in X$ implica $f(x) \in f(X)$ por definición de imagen;
2. $f(x) \in f(X)$ y $f(X) \subseteq f(Y)$ implican $f(x) \in f(Y)$;
3. $f(x) \in f(Y)$ proporciona un $y \in Y$ con $f(y) = f(x)$;
4. la inyectividad de f implica $y = x$;
5. como $y \in Y$ y $x = y$, obtenemos $x \in Y$.

La prueba final puede escribirse entonces de manera lineal.

Demostración. Sea

$$x \in X.$$

Entonces

$$f(x) \in f(X).$$

Como

$$f(X) \subseteq f(Y),$$

se sigue que

$$f(x) \in f(Y).$$

Por definición de imagen, existe

$$y \in Y$$

tal que

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

Como f es inyectiva,

$$y = x.$$

Por tanto,

$$x \in Y.$$

Como $x \in X$ era arbitrario,

$$X \subseteq Y.$$

□

La idea que permitió descubrir la prueba fue regresiva: preguntamos qué necesitaríamos para obtener $x \in Y$.

La demostración final, en cambio, avanza sólo por dependencias justificadas.

Revertir el trabajo hacia atrás. Éste es uno de los controles más importantes al pasar del borrador a la prueba.

Durante la búsqueda podemos escribir

$$Q \Longleftarrow R \Longleftarrow S,$$

porque hemos descubierto

$$R \Rightarrow Q$$

y

$$S \Rightarrow R.$$

Si después conseguimos demostrar S desde las hipótesis, la prueba final deberá recorrer

$$S \Rightarrow R \Rightarrow Q.$$

No copiamos simplemente las flechas del borrador.

Primero verificamos qué relación expresaba cada una y luego las orientamos según la justificación real.

Esto es especialmente importante cuando durante la exploración hubo pasos equivalentes mezclados con condiciones solamente suficientes.

Una cadena regresiva puede haber sido útil aunque no todas sus transformaciones fueran reversibles.

La prueba final debe utilizar sólo la dirección que verdaderamente está demostrada.

Un candidato descubierto debe reaparecer como una elección justificada. El mismo principio vale para objetos y testigos.

Durante la búsqueda de una afirmación existencial podemos experimentar con varios candidatos antes de encontrar el adecuado.

Supongamos que el borrador contiene

$$c_1 \quad \text{no funciona},$$

$$c_2 \quad \text{parece funcionar},$$

$$c_3 \quad \text{produce la identidad que necesito}.$$

La prueba final no necesita narrar obligatoriamente esos tres intentos.

Puede comenzar directamente:

> Elegimos $c = c_3$.

Pero esa elección debe quedar acompañada por una verificación que muestre por qué c_3 satisface las condiciones exigidas.

La historia del descubrimiento responde

«¿cómo se nos ocurrió c_3?»

La prueba responde

«¿por qué c_3 funciona?»

Son preguntas distintas.

Eliminar ramas que ya no sostienen la conclusión. Un borrador puede contener afirmaciones verdaderas que no pertenecen a la cadena final.

Supongamos que, intentando demostrar Q, encontramos dos ramas:

$$H \Rightarrow R_1 \Rightarrow R_2 \Rightarrow Q$$

y

$$H \Rightarrow T_1 \Rightarrow T_2.$$

Si la segunda rama no interviene en la justificación de Q, no tiene por qué aparecer en la prueba final.

Esto no significa que haya sido inútil durante la búsqueda.

Tal vez permitió descubrir una definición, descartar una conjetura o reconocer qué hipótesis era irrelevante.

Pero una vez reconstruida la dependencia, podemos distinguir

$$\boxed{\text{material de descubrimiento}}$$

de

$$\boxed{\text{material justificativo}}.$$

La prueba final conserva el segundo.

No eliminar una justificación necesaria por parecer obvia después. La limpieza del borrador tiene también un riesgo opuesto.

Después de haber trabajado mucho tiempo en un problema, algunas transiciones pueden parecernos tan familiares que olvidamos que siguen necesitando una razón.

Supongamos que la cadena final contiene

$$R \Rightarrow S \Rightarrow Q.$$

Podemos eliminar cálculos exploratorios que rodeaban el paso $R \Rightarrow S$, pero no podemos eliminar **la justificación de** $R \Rightarrow S$ si no es realmente inmediata en el contexto.

Por eso reconstruir no significa simplemente acortar.

Significa conservar una cadena suficiente para que cada dependencia matemática quede disponible.

El control estilístico fino de cuánto detalle conviene escribir pertenecerá al capítulo 10. Aquí nos interesa una exigencia previa:

> ningún paso esencial puede sobrevivir en la versión final sin una razón identificable.

Recuperar las hipótesis exactas que usa cada paso. Durante la búsqueda es frecuente trabajar mentalmente con todas las hipótesis a la vez.

Al reconstruir la prueba conviene preguntar, para cada paso importante:

> **¿qué información anterior utiliza exactamente?**

Supongamos que disponemos de

$$H_1, \qquad H_2, \qquad H_3$$

y obtenemos

$$R.$$

Quizá en realidad

$$H_1 \wedge H_3 \Rightarrow R,$$

mientras que H_2 no interviene.

Registrar esto puede revelar dos cosas distintas:

- que H_2 será necesaria más adelante;
- o que H_2 es redundante para toda la demostración.

En ambos casos la reconstrucción mejora nuestra comprensión de la prueba.

También permite detectar dependencias ocultas: si un paso usa una condición que nunca fue establecida, el argumento todavía está incompleto aunque el resultado parezca plausible.

El borrador como grafo; la prueba como un orden compatible con sus dependencias. Podemos pensar el borrador de descubrimiento como una red.

Una afirmación puede depender de varias anteriores:

$$R_1, \qquad R_2 \quad \Longrightarrow \quad S.$$

Un lema puede alimentar dos partes distintas de la prueba.

Una rama puede terminar sin llegar al objetivo.

La versión final selecciona la parte de esa red que realmente sostiene la conclusión y la presenta en un orden compatible con las dependencias.

Esquemáticamente, el borrador podría contener

$$\begin{array}{lllll} H_1 & \longrightarrow & R_1 & \longrightarrow & S \\ H_2 & \longrightarrow & R_2 & \nearrow & \\ H_3 & \longrightarrow & T_1 & \longrightarrow & T_2 \end{array}$$

siendo la rama $T_1 \to T_2$ finalmente irrelevante.

La prueba conserva

$$H_1, H_2 \longrightarrow R_1, R_2 \longrightarrow S \longrightarrow Q,$$

si ésa es la dependencia efectiva.

No es necesario conservar la topología visual del borrador; sí es necesario conservar su estructura lógica válida.

Reconstruir un lema en el lugar donde se necesita. Si durante la búsqueda demostramos un lema auxiliar L, la prueba final debe dejar clara su posición en la dependencia.

Podemos tener

$$H \Rightarrow L \Rightarrow Q.$$

Hay dos posibilidades legítimas.

Podemos demostrar primero L y después usarlo.

O, si L ya fue establecido anteriormente como resultado independiente, podemos invocarlo directamente verificando que sus hipótesis se cumplen.

Lo que no podemos hacer es usar L y justificarlo sólo después mediante argumentos que dependen de la propia conclusión Q.

La separación modular introducida en §7.6 debe sobrevivir a la reconstrucción sin crear ciclos.

Auditar la cadena final. Antes de considerar que el borrador ya se ha convertido en prueba conviene recorrer la demostración desde el principio y preguntar, en cada línea significativa:

1. ¿esta afirmación ya estaba dada o acaba de ser demostrada?;
2. si acaba de ser demostrada, ¿qué hechos anteriores la justifican?;
3. ¿la implicación se está utilizando en la dirección correcta?;
4. ¿algún objeto fue introducido sin justificar su existencia?;
5. ¿algún caso particular o cálculo experimental está actuando como si fuera una prueba general?;
6. ¿algún lema depende circularmente del resultado principal?;
7. ¿la conclusión final se sigue realmente de lo anterior?

Este recorrido transforma la prueba en una cadena auditable.

No estamos todavía corrigiendo estilo, puntuación o elegancia. Estamos comprobando **integridad de dependencias**.

Una reconstrucción puede revelar que la prueba todavía no está terminada. A veces el intento parece completo mientras permanece en forma de borrador.

Sólo al ordenarlo descubrimos una laguna.

Por ejemplo, nuestras notas pueden contener

$$H \Rightarrow R,$$

$$S \Rightarrow Q,$$

y una línea informal

$$R \approx S.$$

Mientras explorábamos, la semejanza podía parecernos suficiente para continuar.

Al reconstruir necesitamos reemplazarla por una relación matemática precisa:

$$R \Rightarrow S,$$

$$R \Longleftrightarrow S,$$

o una cadena intermedia explícita.

Si no podemos hacerlo, la prueba no está terminada.

Ésta es una de las ventajas de reconstruir: obliga a transformar intuiciones de conexión en obligaciones verificables.

Qué puede desaparecer y qué debe permanecer. Podemos resumir la transición mediante dos columnas conceptuales.

Puede desaparecer de la prueba final:

- ejemplos que sólo sirvieron para adivinar;
- cálculos exploratorios abandonados;
- candidatos que resultaron inútiles;
- ramas verdaderas pero irrelevantes;
- preguntas que guiaron la búsqueda una vez que ya conocemos la respuesta;

- repeticiones destinadas únicamente a orientarnos durante el trabajo.

Debe permanecer, de una forma u otra:

- las hipótesis efectivamente utilizadas;
- las definiciones necesarias para activar esas hipótesis;
- la existencia y propiedades de los objetos auxiliares relevantes;
- los lemas de los que depende la conclusión;
- las transiciones que conectan cada afirmación con las anteriores;
- la cadena que finalmente alcanza el objetivo.

La diferencia no es entre «largo» y «corto».

Es entre **descubrimiento histórico** y **dependencia justificativa**.

Reconstruir, no transcribir

Una prueba final no es una copia limpia del borrador de descubrimiento.

Es una reconstrucción de sus dependencias válidas.

El procedimiento básico es:

1. identificar la cadena que realmente llega a la conclusión;
2. reorientar correctamente los pasos descubiertos hacia atrás;
3. justificar los candidatos y objetos que sobrevivieron;
4. eliminar ramas exploratorias que ya no sostienen el argumento;
5. comprobar que cada afirmación aparece después de aquello de lo que depende.

Comprobación. Un borrador para demostrar

$$H \Rightarrow Q$$

contiene las notas siguientes, escritas en el orden en que fueron descubiertas:

$$Q \Leftarrow S,$$

$$T \text{ parece relacionado con } Q,$$

$$H \Rightarrow R,$$

$$R \Rightarrow S,$$

$$H \Rightarrow T.$$

¿Qué cadena mínima debería controlar la prueba final y qué información puede quedar fuera si no cumple otra función justificativa?

Respuesta razonada.

La relación

$$Q \Leftarrow S$$

significa que disponemos de

$$S \Rightarrow Q.$$

Además sabemos

$$H \Rightarrow R$$

y

$$R \Rightarrow S.$$

Por tanto, la cadena justificativa completa es

$$H \Rightarrow R \Rightarrow S \Rightarrow Q.$$

Ésa es la dependencia mínima que debe controlar la prueba final.

La rama

$$H \Rightarrow T$$

puede haber sido útil durante la exploración, pero mientras sólo sepamos que T «parece relacionado» con Q y no dispongamos de una inferencia necesaria para la conclusión, esa rama no pertenece a la justificación final.

Puede eliminarse de la prueba sin pérdida lógica.

La reconstrucción no conserva todo lo que descubrimos: conserva todo lo que **demuestra**.

Del argumento reconstruido a la práctica autónoma. Con §7.9 queda cerrado el recorrido conceptual principal del capítulo.

Ya hemos pasado por el ciclo completo:

$$\text{leer el problema} \longrightarrow \text{explorar}$$
$$\longrightarrow \text{localizar el hueco} \longrightarrow \text{construir un puente} \longrightarrow \text{reconstruir la prueba.}$$

La última sección trasladará este ciclo a ejercicios graduados.

No se pedirá únicamente «demuestre». También habrá que identificar obligaciones, comparar estrategias, diagnosticar bloqueos, formular lemas y reconstruir demostraciones a partir de notas de trabajo.

El objetivo será comprobar que las herramientas de descubrimiento pueden transferirse entre distintos microcontextos sin convertirse en recetas ligadas a un único tipo de problema.

7.10. Ejercicios y soluciones desarrolladas

Los ejercicios finales no están organizados por «tipo de prueba», sino por el grado de autonomía que exige la búsqueda. En los primeros se indica con bastante precisión qué parte del proceso conviene observar. En los siguientes hay que diagnosticar el bloqueo o construir una pieza intermedia. En los últimos se retira casi todo el andamiaje y se pide reconstruir una estrategia completa.

La meta no es acumular trucos. Es practicar el ciclo desarrollado a lo largo del capítulo:

$$\text{datos} \longrightarrow \text{exploración} \longrightarrow \text{hueco identificado} \longrightarrow \text{puente} \longrightarrow \text{prueba reconstruida.}$$

Familia I. Búsqueda guiada. Ejercicio 1. Avanzar con dos hipótesis. Sean $a, b, c, d \in \mathbb{R}$ tales que

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

Queremos demostrar

$$a + c \leq b + d.$$

1. Obtén una consecuencia útil de $a \leq b$ manteniendo c fijo.
2. Obtén una consecuencia útil de $c \leq d$ manteniendo b fijo.
3. Haz que ambas consecuencias se encuentren y reconstruye la prueba final.

Ejercicio 2. Elegir una meta regresiva útil. Para $x \in \mathbb{R}$ queremos demostrar

$$x^2 + 9 \geq 6x.$$

Considera las siguientes metas auxiliares:

$$A: \ (x-3)^2 \geq 0,$$

$$B: \ x = 3,$$

$$C: \ x^2 \geq 6x.$$

1. Decide cuáles son condiciones suficientes para el objetivo.
2. Distingue entre «suficiente» y «estratégicamente útil para todo $x \in \mathbb{R}$».
3. Elige la mejor meta auxiliar y escribe una demostración válida en el sentido de las dependencias.

Ejercicio 3. Localizar el punto de encuentro. Sean A, B, C, D conjuntos tales que

$$A \subseteq B, \qquad C \subseteq D, \qquad B \cap D = \varnothing.$$

Demuestra

$$A \cap C = \varnothing.$$

Antes de escribir la prueba final, identifica explícitamente:

1. qué información produces desde $x \in A \cap C$;
2. qué condición sería incompatible con $B \cap D = \varnothing$;
3. cuál es el puente exacto entre ambos frentes.

Ejercicio 4. Desplegar sólo las definiciones necesarias. Sea

$$f : A \to B$$

una función inyectiva, y sean $X, Y \subseteq A$. Supón que

$$f(X) = f(Y).$$

Demuestra

$$X = Y.$$

No expandas todas las definiciones al comienzo. Para una sola de las inclusiones, indica qué definición conviene desplegar primero y en qué momento exacto se vuelve útil la inyectividad.

Ejercicio 5. Construir objetos auxiliares. Sean

$$f : A \to B, \qquad g : B \to C.$$

Supón que f es sobreyectiva y que

$$g \circ f$$

es inyectiva. Demuestra que g es inyectiva.

La prueba deberá explicar por qué, dados $b_1, b_2 \in B$, la sobreyectividad de f sugiere introducir dos objetos auxiliares en A.

Familia II. Diagnóstico y construcción intermedia. Ejercicio 6. Formular un lema que pueda reutilizarse. Sean

$$p : X \to Y, \qquad q : Y \to Z, \qquad r : Z \to W$$

funciones sobreyectivas.

1. Formula y demuestra un lema general sobre la composición de **dos** funciones sobreyectivas.
2. Utiliza ese lema para demostrar que
$$r \circ q \circ p : X \to W$$
es sobreyectiva.
3. Explica por qué el lema es más útil que repetir desde cero una prueba para las tres funciones.

Ejercicio 7. Diagnosticar una cancelación bloqueada. Sea F un cuerpo y supón

$$ac = bc$$

para ciertos $a, b, c \in F$.

Un intento de prueba escribe inmediatamente

$$a = b.$$

1. Localiza la obligación que falta justificar.
2. Identifica una hipótesis adicional suficiente para que esa ruta funcione.
3. Bajo esa hipótesis adicional, completa la demostración.
4. Explica por qué el fracaso de la cancelación, por sí solo, no autoriza a concluir que cualquier enunciado relacionado sea falso.

Ejercicio 8. Un ciclo escondido. Quieres demostrar

$$H \Rightarrow Q.$$

Durante la búsqueda propones un lema L tal que

$$L \Rightarrow Q.$$

Para «demostrar» L, escribes:

> Supongamos Q. A partir de Q obtenemos L.

1. Dibuja el grafo de dependencias que realmente has construido.
2. Explica por qué todavía no existe una prueba de Q desde H.
3. Indica qué clase de enlace independiente tendría que añadirse para reparar la estrategia.

Ejercicio 9. Del experimento a una identidad. Considera

$$p(x) = x^2 - 4x + 5.$$

1. Calcula $p(x)$ para $x = 0, 1, 2, 3, 4$.
2. Formula una conjetura sobre una cota inferior para $p(x)$ y sobre el caso de igualdad.
3. Busca una representación de $p(x)$ que explique el patrón observado.
4. Demuestra la conjetura para todo $x \in \mathbb{R}$.
5. Separa explícitamente qué parte fue exploración y qué parte fue demostración.

Ejercicio 10. Reconstruir una igualdad de preimágenes. Sea

$$f : A \to B$$

y sean $U, V \subseteq B$. Queremos demostrar

$$f^{-1}(U \cap V) = f^{-1}(U) \cap f^{-1}(V).$$

Un borrador contiene, desordenadas, las siguientes notas para un $x \in A$:

$$x \in f^{-1}(U \cap V),$$

$$f(x) \in U \cap V,$$

$$f(x) \in U \text{ y } f(x) \in V,$$

$$x \in f^{-1}(U) \text{ y } x \in f^{-1}(V),$$

$$x \in f^{-1}(U) \cap f^{-1}(V).$$

1. Ordena las notas según sus dependencias.
2. Decide cuáles de los pasos son equivalencias.
3. Reconstruye una demostración de la igualdad para elementos arbitrarios.

Familia III. Transferencia y descubrimiento autónomo limitado. Ejercicio 11. Comparar dos estrategias. Sea $f : A \to B$ y sean $U, V \subseteq B$. Demuestra

$$f^{-1}(U \setminus V) = f^{-1}(U) \setminus f^{-1}(V).$$

Antes de cerrar la prueba, compara estas dos estrategias:

1. demostrar dos inclusiones por separado;
2. tomar $x \in A$ y encadenar equivalencias de pertenencia.

Ambas son legítimas. Explica cuál hace más visible en este caso el puente entre las definiciones y por qué.

Ejercicio 12. Combinar hipótesis para descubrir el candidato. Sean $s, d \in \mathbb{R}$. Demuestra que existe un único par $(x, y) \in \mathbb{R}^2$ que satisface

$$x + y = s$$

y

$$x - y = d.$$

No comiences anunciando las fórmulas de x e y. Explica primero qué combinaciones de las dos ecuaciones permiten descubrirlas. Después reconstruye una prueba de existencia y unicidad.

Ejercicio 13. Reutilizar resultados intermedios. Sean

$$f : A \to B, \qquad g : B \to C.$$

Supón que f es sobreyectiva y que

$$g \circ f : A \to C$$

es biyectiva. Demuestra que tanto f como g son biyectivas.

Antes de escribir la prueba final, separa el problema en cuatro obligaciones:

- inyectividad de f;
- sobreyectividad de f;
- inyectividad de g;
- sobreyectividad de g.

Identifica qué obligaciones ya están dadas, cuáles siguen inmediatamente de la composición y cuál necesita introducir preimágenes mediante la sobreyectividad de f.

Ejercicio 14. Descubrir y verificar un testigo único. Sean $a, b, c \in \mathbb{R}$ con

$$a \neq 0.$$

Demuestra que existe un único $x \in \mathbb{R}$ tal que

$$ax + b = c.$$

Incluye, antes de la prueba final, una línea de descubrimiento que muestre cómo el objetivo sugiere el candidato. Después verifica existencia y unicidad sin utilizar como argumento la mera apariencia de la fórmula.

Ejercicio 15. Reconstruir desde notas no lineales. Sean $x, y, s, t \in \mathbb{R}$ y supón

$$x + y = s$$

y

$$x^2 + y^2 = t.$$

Queremos obtener una fórmula para xy. Un borrador contiene:

$$\text{OBJETIVO: } xy = ?$$

$$(x + y)^2 = x^2 + 2xy + y^2,$$

$$s^2 = t + 2xy,$$

$$(x - y)^2 \geq 0 \quad \text{¿sirve?},$$

$$2xy = s^2 - t,$$

$$xy = \frac{s^2 - t}{2}.$$

1. Reconstruye la cadena justificativa mínima.
2. Identifica qué nota puede eliminarse sin pérdida lógica.
3. Explica por qué elevar al cuadrado la hipótesis $x + y = s$ es una transformación estratégicamente dirigida al objetivo y no una manipulación arbitraria.

Soluciones desarrolladas.

Solución 1. Avanzar con dos hipótesis. De

$$a \leq b$$

podemos sumar c a ambos lados:

$$a + c \leq b + c.$$

Ésta es una consecuencia útil porque introduce el mismo término b que aparecerá en el extremo derecho del objetivo.

De

$$c \leq d$$

podemos sumar b:

$$b + c \leq b + d.$$

Ahora los dos frentes se encuentran en

$$b + c.$$

Por transitividad del orden,

$$a + c \leq b + c \leq b + d.$$

Por tanto,

$$a + c \leq b + d.$$

La prueba no exigió producir todas las consecuencias posibles de las dos desigualdades. Bastó encontrar dos consecuencias compatibles cuya composición alcanzara el objetivo.

Solución 2. Elegir una meta regresiva útil. El objetivo es

$$x^2 + 9 \geq 6x.$$

La condición A es

$$(x - 3)^2 \geq 0.$$

Al expandir,

$$x^2 - 6x + 9 \geq 0,$$

que es equivalente a

$$x^2 + 9 \geq 6x.$$

Por tanto, A no sólo es suficiente: es equivalente al objetivo.

La condición B, $x = 3$, también es suficiente para el objetivo, porque si $x = 3$ entonces ambos lados valen 18. Pero no podemos esperar demostrar $x = 3$ para un real arbitrario. Es una condición demasiado fuerte para esta búsqueda.

La condición C,

$$x^2 \geq 6x,$$

es suficiente, pues al sumar 9 obtenemos

$$x^2 + 9 \geq 6x + 9 \geq 6x.$$

Sin embargo, tampoco es una propiedad válida para todo real. Por ejemplo, no podemos convertirla en una meta universal alcanzable sin información adicional.

La mejor meta auxiliar es A.

Demostración. Para todo $x \in \mathbb{R}$,

$$(x-3)^2 \geq 0.$$

Por tanto,

$$x^2 - 6x + 9 \geq 0,$$

y reordenando,

$$x^2 + 9 \geq 6x.$$

La comparación muestra la diferencia entre una condición meramente suficiente y una condición suficiente que además es alcanzable desde los datos disponibles.

Solución 3. Localizar el punto de encuentro. Para demostrar

$$A \cap C = \varnothing,$$

supongamos que existe

$$x \in A \cap C.$$

Entonces

$$x \in A \qquad \text{y} \qquad x \in C.$$

Como

$$A \subseteq B,$$

obtenemos

$$x \in B.$$

Como

$$C \subseteq D,$$

obtenemos

$$x \in D.$$

Así,

$$x \in B \cap D.$$

Pero

$$B \cap D = \varnothing,$$

lo cual es imposible.

Por tanto,

$$A \cap C = \varnothing.$$

El punto de encuentro es la pertenencia

$$x \in B \cap D.$$

Desde $x \in A \cap C$ podemos producirla; desde la hipótesis $B \cap D = \varnothing$ sabemos que esa pertenencia es imposible. El hueco queda reducido a una contradicción local perfectamente identificada.

Solución 4. Desplegar sólo las definiciones necesarias. Demostraremos primero

$$X \subseteq Y.$$

Sea

$$x \in X.$$

Por definición de imagen,

$$f(x) \in f(X).$$

Como

$$f(X) = f(Y),$$

se sigue que

$$f(x) \in f(Y).$$

Ahora sí conviene desplegar la definición de imagen de Y: existe

$$y \in Y$$

tal que

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

En este punto aparece exactamente la forma que activa la inyectividad. Como f es inyectiva,

$$y = x.$$

Puesto que $y \in Y$,

$$x \in Y.$$

Así,

$$X \subseteq Y.$$

Intercambiando los papeles de X e Y y usando de nuevo

$$f(Y) = f(X),$$

obtenemos

$$Y \subseteq X.$$

Por doble inclusión,

$$X = Y.$$

La inyectividad no era utilizable en la formulación inicial $f(X) = f(Y)$. Se volvió operativa sólo después de producir una igualdad entre dos valores de la función.

Solución 5. Construir objetos auxiliares. Queremos demostrar que g es inyectiva. Sean

$$b_1, b_2 \in B$$

y supongamos

$$g(b_1) = g(b_2).$$

La hipótesis de inyectividad disponible no habla de g sola, sino de

$$g \circ f.$$

Por eso necesitamos representar b_1 y b_2 como valores de f.

Como f es sobreyectiva, existen

$$a_1, a_2 \in A$$

tales que

$$f(a_1) = b_1 \qquad \text{y} \qquad f(a_2) = b_2.$$

Entonces

$$g(f(a_1)) = g(b_1) = g(b_2) = g(f(a_2)).$$

Por tanto,

$$(g \circ f)(a_1) = (g \circ f)(a_2).$$

Como $g \circ f$ es inyectiva,

$$a_1 = a_2.$$

Aplicando f,

$$b_1 = f(a_1) = f(a_2) = b_2.$$

Así g es inyectiva.

Los objetos auxiliares a_1 y a_2 tienen una función precisa: trasladan el problema desde B, donde queremos comparar $g(b_1)$ y $g(b_2)$, hacia A, donde podemos utilizar la inyectividad de la composición.

Solución 6. Formular un lema que pueda reutilizarse. Formulamos el resultado intermedio:

> **Lema.** Si
>
> $$u : M \to N \qquad \text{y} \qquad v : N \to P$$
>
> son sobreyectivas, entonces
>
> $$v \circ u : M \to P$$
>
> es sobreyectiva.

Demostración del lema. Sea

$$p \in P.$$

Como v es sobreyectiva, existe

$$n \in N$$

tal que

$$v(n) = p.$$

Como u es sobreyectiva, existe

$$m \in M$$

tal que

$$u(m) = n.$$

Entonces

$$(v \circ u)(m) = v(u(m)) = v(n) = p.$$

Por tanto, $v \circ u$ es sobreyectiva.

□

Aplicamos ahora el lema a p y q. Como ambas son sobreyectivas,

$$q \circ p : X \to Z$$

es sobreyectiva.

Aplicamos nuevamente el lema a $q \circ p$ y r. Obtenemos que

$$r \circ (q \circ p) : X \to W$$

es sobreyectiva.

Por asociatividad de la composición,

$$r \circ (q \circ p) = r \circ q \circ p.$$

Por tanto,

$$r \circ q \circ p$$

es sobreyectiva.

El lema aísla el mecanismo que se repite. En vez de reconstruir una prueba especial para tres funciones, demostramos una interfaz reutilizable y luego la aplicamos dos veces.

Solución 7. Diagnosticar una cancelación bloqueada. De

$$ac = bc$$

no podemos pasar automáticamente a

$$a = b.$$

La cancelación multiplicativa necesita que el factor cancelado sea no nulo.

El cuello de botella puede escribirse como

$$ac = bc \quad \Big| \quad a = b.$$

Una hipótesis suficiente para cerrar el hueco es

$$c \neq 0.$$

Bajo esa condición, c tiene inverso multiplicativo c^{-1}. Multiplicamos ambos lados de

$$ac = bc$$

por c^{-1}:

$$(ac)c^{-1} = (bc)c^{-1}.$$

Por asociatividad,

$$a(cc^{-1}) = b(cc^{-1}),$$

y por tanto

$$a = b.$$

El intento inicial nos ha enseñado qué propiedad faltaba para esa estrategia: la posibilidad de cancelar c.

Que esa ruta no funcione sin $c \neq 0$ no demuestra, por sí solo, la falsedad de cualquier afirmación que estemos investigando. Sólo muestra que **esta inferencia concreta** necesita una condición que todavía no está disponible.

Solución 8. Un ciclo escondido. Disponemos de

$$L \Rightarrow Q.$$

Pero la propuesta para demostrar L consiste en suponer Q y obtener L:

$$Q \Rightarrow L.$$

El grafo construido es entonces

$$Q \Rightarrow L \Rightarrow Q.$$

No hay ninguna flecha que entre en ese ciclo desde H.

Por tanto, aunque hayamos mostrado una relación entre Q y L, todavía no hemos demostrado

$$H \Rightarrow Q.$$

La estrategia se repararía si pudiéramos obtener L de manera independiente desde las hipótesis, por ejemplo mediante una cadena

$$H \Rightarrow R \Rightarrow L.$$

Entonces sí tendríamos

$$H \Rightarrow R \Rightarrow L \Rightarrow Q.$$

La circularidad no depende de la longitud del argumento. Depende de que la conclusión aparezca como presupuesto necesario para producir aquello que debía demostrarla.

Solución 9. Del experimento a una identidad. Calculamos:

x	0	1	2	3	4
$p(x)$	5	2	1	2	5

Los valores sugieren que

$$p(x) \geq 1$$

y que la igualdad ocurre cuando

$$x = 2.$$

Para explicar el patrón, reescribimos:

$$\begin{aligned} p(x) &= x^2 - 4x + 5 \\ &= x^2 - 4x + 4 + 1 \\ &= (x-2)^2 + 1. \end{aligned}$$

Como

$$(x-2)^2 \geq 0,$$

se sigue que

$$p(x) \geq 1$$

para todo $x \in \mathbb{R}$.
Además,

$$p(x) = 1$$

si y sólo si

$$(x-2)^2 = 0,$$

es decir, si y sólo si

$$x = 2.$$

La tabla fue exploratoria: sugirió la cota y el caso de igualdad. La identidad

$$p(x) = (x-2)^2 + 1$$

es la pieza justificativa que convierte esa sospecha en demostración general.

Solución 10. Reconstruir una igualdad de preimágenes. Sea $x \in A$. La cadena correcta es

$$\begin{aligned} x \in f^{-1}(U \cap V) &\iff f(x) \in U \cap V \\ &\iff (f(x) \in U \text{ y } f(x) \in V) \\ &\iff (x \in f^{-1}(U) \text{ y } x \in f^{-1}(V)) \\ &\iff x \in f^{-1}(U) \cap f^{-1}(V). \end{aligned}$$

Cada paso es una equivalencia:

1. la primera y la tercera usan la definición de preimagen;
2. la segunda usa la definición de intersección;
3. la última vuelve a usar la definición de intersección.

Como para todo $x \in A$ tenemos

$$x \in f^{-1}(U \cap V) \iff x \in f^{-1}(U) \cap f^{-1}(V),$$

los conjuntos tienen exactamente los mismos elementos. Por tanto,

$$f^{-1}(U \cap V) = f^{-1}(U) \cap f^{-1}(V).$$

El borrador podía contener las piezas en cualquier orden. La prueba final las organiza como una cadena de equivalencias cuya dependencia es transparente.

Solución 11. Comparar dos estrategias. Tomemos $x \in A$. Entonces

$$\begin{aligned} x \in f^{-1}(U \setminus V) &\iff f(x) \in U \setminus V \\ &\iff \big(f(x) \in U \text{ y } f(x) \notin V\big) \\ &\iff \big(x \in f^{-1}(U) \text{ y } x \notin f^{-1}(V)\big) \\ &\iff x \in f^{-1}(U) \setminus f^{-1}(V). \end{aligned}$$

Así,

$$f^{-1}(U \setminus V) = f^{-1}(U) \setminus f^{-1}(V).$$

También podríamos demostrar dos inclusiones. Esa estrategia sería correcta: tomaríamos primero un elemento del lado izquierdo y después uno del lado derecho, repitiendo en ambos sentidos las mismas definiciones.

En este problema la cadena de equivalencias hace más visible el puente porque todas las transformaciones relevantes son reversibles y consisten precisamente en desplegar definiciones de pertenencia a preimagen y diferencia de conjuntos.

La comparación no establece una regla universal de estilo. Muestra que la representación elegida puede hacer más o menos visible la dependencia matemática.

Solución 12. Combinar hipótesis para descubrir el candidato. Partimos de

$$x + y = s$$

y

$$x - y = d.$$

Si sumamos ambas ecuaciones, desaparece y:

$$2x = s + d.$$

Esto sugiere

$$x = \frac{s + d}{2}.$$

Si restamos la segunda ecuación de la primera, desaparece x:

$$2y = s - d,$$

y obtenemos

$$y = \frac{s - d}{2}.$$

Estas fórmulas fueron descubiertas combinando las hipótesis con un propósito: aislar una variable cada vez.

Existencia. Definimos

$$x_0 = \frac{s+d}{2}, \qquad y_0 = \frac{s-d}{2}.$$

Entonces

$$x_0 + y_0 = \frac{s+d+s-d}{2} = s,$$

y

$$x_0 - y_0 = \frac{s+d-(s-d)}{2} = d.$$

Por tanto, existe una solución.

Unicidad. Supongamos que (x, y) es cualquier solución. Sumando las dos ecuaciones obtenemos necesariamente

$$2x = s + d,$$

de modo que

$$x = \frac{s+d}{2} = x_0.$$

Restándolas obtenemos

$$2y = s - d,$$

y por tanto

$$y = \frac{s-d}{2} = y_0.$$

Así la solución es única.

Solución 13. Reutilizar resultados intermedios. La sobreyectividad de f ya está dada. Falta demostrar su inyectividad.

Supongamos

$$f(a_1) = f(a_2).$$

Aplicando g,

$$(g \circ f)(a_1) = (g \circ f)(a_2).$$

Como $g \circ f$ es biyectiva, en particular es inyectiva. Luego

$$a_1 = a_2.$$

Así f es inyectiva. Como ya era sobreyectiva, f es biyectiva.

Ahora estudiemos g.

Como $g \circ f$ es biyectiva, es sobreyectiva. Sea

$$c \in C.$$

Existe $a \in A$ tal que

$$g(f(a)) = c.$$

Tomando

$$b = f(a) \in B,$$

obtenemos

$$g(b) = c.$$

Por tanto, g es sobreyectiva.

Para la inyectividad, 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 inyectividad de $g \circ f$ implica

$$a_1 = a_2,$$

y por tanto

$$b_1 = f(a_1) = f(a_2) = b_2.$$

Así g es inyectiva y, junto con su sobreyectividad, biyectiva.

La arquitectura global se volvió manejable al separar cuatro obligaciones y observar que cada una utiliza una parte distinta de la información disponible.

Solución 14. Descubrir y verificar un testigo único. Queremos resolver

$$ax + b = c.$$

Durante la búsqueda podemos trabajar hacia atrás algebraicamente:

$$ax = c - b,$$

y, como $a \neq 0$,

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

Esto descubre el candidato. Todavía debemos verificarlo.

Existencia. Definimos

$$x_0 = \frac{c - b}{a}.$$

Entonces

$$\begin{aligned} ax_0 + b &= a\frac{c - b}{a} + b \\ &= c - b + b \\ &= c. \end{aligned}$$

Así existe al menos una solución.

Unicidad. Supongamos que x e y satisfacen

$$ax + b = c$$

y

$$ay + b = c.$$

Entonces

$$ax + b = ay + b.$$

Cancelando b,

$$ax = ay.$$

Como $a \neq 0$, podemos cancelar a y obtenemos

$$x = y.$$

Por tanto, la solución es única.

La manipulación regresiva produjo el candidato; la prueba final justificó por separado que existe y que ningún otro puede funcionar.

Solución 15. Reconstruir desde notas no lineales. El objetivo es expresar xy usando s y t.

La hipótesis

$$x + y = s$$

contiene una suma, mientras que el objetivo contiene el producto xy. Elevar al cuadrado es útil porque la identidad

$$(x + y)^2 = x^2 + 2xy + y^2$$

hace aparecer precisamente el término xy y, al mismo tiempo, contiene la segunda cantidad conocida

$$x^2 + y^2 = t.$$

La cadena justificativa mínima es:

$$x + y = s \Longrightarrow (x + y)^2 = s^2.$$

Por la identidad algebraica,

$$x^2 + 2xy + y^2 = s^2.$$

Usando

$$x^2 + y^2 = t,$$

obtenemos

$$t + 2xy = s^2.$$

Por tanto,

$$2xy = s^2 - t,$$

y finalmente

$$\boxed{xy = \frac{s^2 - t}{2}}.$$

La nota

$$(x - y)^2 \geq 0$$

puede eliminarse sin pérdida lógica. Podría conducir a una desigualdad relacionada, pero no interviene en la derivación de la fórmula exacta.

El paso de elevar al cuadrado no fue una manipulación arbitraria: fue elegido porque transforma una hipótesis sobre $x + y$ en una expresión que contiene simultáneamente el producto buscado y la suma de cuadrados ya conocida.

Cierre del capítulo. Los ejercicios recorren el mismo ciclo en contextos diferentes: orden, conjuntos, funciones y álgebra elemental. La superficie matemática cambia, pero las preguntas estratégicas permanecen:

- ¿qué puedo obtener legítimamente de los datos?;
- ¿qué me bastaría para alcanzar el objetivo?;
- ¿dónde está el hueco mínimo?;
- ¿qué definición, representación, objeto o lema puede cerrarlo?;
- ¿qué me enseña un intento que se atasca?;
- ¿qué parte del borrador pertenece realmente a la demostración final?

Ésa es la competencia que este capítulo quería hacer visible. Una prueba no se descubre únicamente reconociendo su arquitectura ni se justifica simplemente narrando cómo apareció la idea. Entre ambas tareas hay un trabajo controlado de transformación, diagnóstico y reconstrucción.

El capítulo siguiente introducirá otra familia de herramientas fundamentales —inducción, recursión y buen orden— que ampliará el repertorio disponible para razonar sobre los números naturales y estructuras construidas paso a paso.

Capítulo 8 — Naturales, recursión, inducción y buen orden

En los capítulos anteriores aprendimos a leer afirmaciones cuantificadas, reconocer obligaciones de prueba y construir o descubrir cadenas de justificación. Hasta ahora, sin embargo, la mayor parte de nuestros argumentos podía examinarse localmente: fijábamos objetos arbitrarios, desplegábamos definiciones y buscábamos una cadena que llevara de las hipótesis a la conclusión.

Los números naturales introducen una estructura adicional.

Una afirmación puede pedirnos demostrar una propiedad no de un único objeto ni de un conjunto finito de casos, sino de todos los índices de una sucesión potencialmente infinita:

$$P(0), \quad P(1), \quad P(2), \quad P(3), \quad \dots$$

También podemos querer **definir** un objeto cuyo valor en cada índice dependa de valores anteriores:

$$a_0, \quad a_1, \quad a_2, \quad a_3, \quad \dots$$

Y, cuando una propiedad falla en algún natural, podemos preguntarnos si existe un **primer índice** donde ocurre el fallo.

Estas tres situaciones conducirán respectivamente a la inducción, la recursión y el buen orden.

Conviene mantener desde el comienzo una distinción que gobernará todo el capítulo:

$$\boxed{\text{inducción} = \text{demostrar propiedades}}$$

$$\boxed{\text{recursión} = \text{definir objetos}}$$

$$\boxed{\text{buen orden} = \text{localizar mínimos}}.$$

No son tres nombres para la misma técnica. Son tres maneras diferentes de explotar una misma característica estructural de los números naturales: su organización sucesiva.

8.1. ¿Qué tiene de especial $\mathbb{N}$?

Una convención que debe quedar fijada. En este capítulo utilizaremos

$$\mathbb{N} = \{0, 1, 2, 3, \dots\}.$$

El cero pertenece, por tanto, a nuestros naturales.

Esta convención no es universal. Otros textos utilizan

$$\{1, 2, 3, \dots\}$$

para denotar los números naturales. Ninguna de las dos elecciones crea una diferencia matemática profunda, pero una demostración puede quedar mal formulada si el punto inicial permanece implícito.

Por eso, cuando una afirmación sólo pretenda valer desde un índice determinado

$$n_0 \in \mathbb{N},$$

escribiremos explícitamente el rango correspondiente:

$$n \geq n_0.$$

Por ejemplo,

$$\forall n \in \mathbb{N},\ n \geq 3 \Rightarrow P(n)$$

afirma algo distinto de

$$\forall n \in \mathbb{N}\ P(n).$$

Los valores

$$n = 0, 1, 2$$

no pertenecen a la obligación de la primera afirmación.

Esto será importante cuando aparezca la inducción: el punto donde comienza una afirmación forma parte del teorema que estamos intentando demostrar.

Los naturales avanzan mediante sucesores. Para cada natural

$$n \in \mathbb{N}$$

existe el natural siguiente

$$n + 1.$$

Podemos considerar la función sucesor

$$S : \mathbb{N} \to \mathbb{N}, \qquad S(n) = n + 1.$$

Así,

$$0 \mapsto 1 \mapsto 2 \mapsto 3 \mapsto \cdots .$$

La palabra **sucesor** no significa solamente "un número mayor". Entre dos números naturales consecutivos no hay otro natural:

$$n < k < n + 1$$

no puede ocurrir con

$$k \in \mathbb{N}.$$

En este sentido, $\mathbb{N}$ tiene una estructura discreta.

Esto contrasta con lo que ocurre en $\mathbb{R}$. Si

$$x < y,$$

siempre podemos encontrar otro real estrictamente entre ambos; por ejemplo,

$$\frac{x + y}{2}.$$

Por ello no existe, para cada real x, un "real siguiente" análogo a $n + 1$.

La estructura

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

nos proporciona algo que utilizaremos repetidamente: un punto inicial y una regla natural para pasar de un índice al siguiente.

Todavía no hemos formulado ningún principio de inducción. Sólo estamos identificando la estructura sobre la cual ese principio podrá actuar.

Tramos de naturales. No siempre trabajaremos desde 0.

Si fijamos

$$n_0 \in \mathbb{N},$$

llamaremos tramo desde n_0 al conjunto

$$\mathbb{N}_{\geq n_0} = \{n \in \mathbb{N} : n \geq n_0\}.$$

Por ejemplo,

$$\mathbb{N}_{\geq 4} = \{4, 5, 6, 7, \ldots\}.$$

Este conjunto conserva la misma forma sucesiva:

$$4 \mapsto 5 \mapsto 6 \mapsto 7 \mapsto \cdots .$$

Su primer elemento ya no es 0, sino 4.

Esta observación parece elemental, pero evita un error frecuente: pensar que cualquier argumento sobre naturales debe comenzar obligatoriamente en 0 o en 1.

El dominio de la afirmación decide dónde comienza la obligación.

Si queremos demostrar

$$P(n)$$

para todo

$$n \geq n_0,$$

los casos anteriores a n_0 pueden ser completamente irrelevantes.

Una propiedad indexada no es todavía una afirmación universal. Volvamos ahora al lenguaje de los cuantificadores.

Supongamos que $P(n)$ expresa una propiedad que depende de un natural n.

Por ejemplo, podríamos considerar

$$P(n): \quad n(n+1) \text{ es par.}$$

Mientras n permanece libre, $P(n)$ es una propiedad indexada: su valor de verdad depende del natural que sustituyamos por n.

Podemos formar instancias concretas:

$$P(0), \quad P(1), \quad P(2), \quad \ldots$$

En nuestro ejemplo:

$$P(0): \quad 0 \cdot 1 \text{ es par,}$$

$$P(1): \quad 1 \cdot 2 \text{ es par},$$

$$P(2): \quad 2 \cdot 3 \text{ es par}.$$

Cada una de estas expresiones es ya una proposición.

Pero el enunciado que dice que la propiedad vale para **todos** los naturales es otra proposición:

$$\forall n \in \mathbb{N}\ P(n).$$

En el ejemplo:

$$\forall n \in \mathbb{N}, \quad n(n+1) \text{ es par}.$$

Es importante no confundir

$$P(n)$$

con

$$\forall n \in \mathbb{N}\ P(n).$$

La primera es una propiedad cuyo índice todavía debe fijarse o cuantificarse.

La segunda es una afirmación universal que abarca infinitas instancias.

Esta distinción ya estaba disponible desde C03. Ahora adquirirá una función nueva: la inducción será una forma de justificar ciertas afirmaciones universales sobre familias indexadas por $\mathbb{N}$.

Una universal contiene infinitas obligaciones. Si queremos establecer

$$\forall n \in \mathbb{N}\ P(n),$$

debemos justificar simultáneamente

$$P(0),$$

$$P(1),$$

$$P(2),$$

y así sucesivamente.

No podemos completar esta tarea escribiendo una prueba separada para cada natural. No existe un "último caso" después del cual podamos detenernos.

Tampoco basta verificar una cantidad muy grande de casos.

Incluso si comprobamos

$$P(0), P(1), \ldots, P(10^6),$$

todavía quedan

$$P(10^6+1), \quad P(10^6+2), \quad \ldots$$

fuera de lo comprobado.

Esto marca una diferencia entre

$$\boxed{\text{muchas verificaciones particulares}}$$

y

$$\boxed{\text{una justificación universal}}.$$

La inducción no será una forma elegante de abreviar un millón de comprobaciones. Tendrá que proporcionar algo que ninguna lista finita de ejemplos puede ofrecer: una razón que alcance **todos** los índices del tramo pertinente.

La regularidad que queremos explotar. ¿Qué característica adicional tenemos en $\mathbb{N}$?

Los índices no están dispuestos como puntos aislados sin relación. Cada uno tiene un sucesor.

Podemos representar una propiedad indexada sobre un tramo así:

$$P(n_0), \quad P(n_0+1), \quad P(n_0+2), \quad P(n_0+3), \quad \dots$$

y colocar entre los índices la estructura sucesiva:

$$n_0 \longrightarrow n_0+1 \longrightarrow n_0+2 \longrightarrow n_0+3 \longrightarrow \cdots.$$

Esto sugiere una pregunta distinta de "¿puedo verificar todos los casos uno por uno?":

> **¿Existe una manera de hacer que la verdad de la propiedad se propague a lo largo de la estructura sucesiva de los índices?**

Ésta es todavía una pregunta, no el principio de inducción.

Para responderla necesitaremos distinguir dos tipos de información:

1. una afirmación que nos permita entrar en el tramo;
2. una afirmación que controle el paso de un índice al siguiente.

La sección siguiente formalizará exactamente esa combinación.

El caso inicial depende del enunciado. Supongamos que la propiedad sólo pretende valer para

$$n \geq 5.$$

Entonces la familia relevante es

$$P(5), P(6), P(7), \dots$$

y no

$$P(0), P(1), P(2), \dots.$$

El primer caso pertinente es

$$P(5).$$

Esto puede parecer una cuestión de notación, pero no lo es.

Si una futura prueba consiguiera establecer

$$P(0)$$

y una regla que sólo funcionara desde $n \geq 5$, todavía existiría un hueco entre el caso inicial demostrado y el lugar donde la regla puede comenzar a aplicarse.

De manera inversa, si el teorema sólo afirma la propiedad desde 5, no hay motivo para exigir una demostración de $P(0)$.

El rango cuantificado determina qué casos deben cubrirse.

Por eso conviene escribir explícitamente:

$$\forall n \in \mathbb{N}, \quad n \geq n_0 \Rightarrow P(n),$$

o, de forma abreviada cuando el contexto ya está fijado,

$$\forall n \geq n_0, \quad P(n).$$

Un índice arbitrario sigue siendo un índice concreto dentro de la prueba. Hay otra distinción que necesitaremos más adelante.

Cuando escribimos

$$\forall n \geq n_0, \quad P(n),$$

el símbolo n recorre todos los índices del tramo.

En una demostración, sin embargo, podemos fijar temporalmente un

$$n \geq n_0$$

arbitrario y estudiar la relación entre

$$P(n)$$

y

$$P(n+1).$$

Ese n no representa un valor especial como 3 o 100. Tampoco significa que hayamos supuesto la afirmación universal.

Representa un índice cualquiera dentro del rango donde queremos entender un paso general.

La diferencia entre

$$\boxed{\forall n \geq n_0 \; P(n)}$$

y

$$\boxed{\text{fijar un } n \geq n_0 \text{ y estudiar } P(n)}$$

será esencial para comprender la hipótesis inductiva en §8.4.

El buen orden aparece ya en el horizonte. La organización de $\mathbb{N}$ tiene otra característica que utilizaremos más adelante.

Si tomamos un subconjunto no vacío de naturales, por ejemplo

$$A = \{4, 7, 11, 18, \ldots\},$$

podemos preguntar por su primer elemento.

En $\mathbb{N}$, la idea de "primer índice donde ocurre algo" tiene sentido estructural.

Esto permitirá, por ejemplo, convertir una afirmación del tipo

«la propiedad falla para algún natural»

en el estudio de

$$F = \{n \in \mathbb{N} : \neg P(n)\}$$

y preguntar, si F no es vacío, por su menor elemento.

No utilizaremos todavía este principio para demostrar nada. Su función aquí es mostrar que la estructura discreta de $\mathbb{N}$ permite formular dos preguntas relacionadas:

- ¿cómo se propaga una propiedad hacia los sucesores?;
- si la propiedad falla, ¿dónde aparece el primer fallo?

La relación entre ambas preguntas se desarrollará en §§8.6–8.7.

Demostrar y definir son tareas diferentes. Finalmente, la estructura sucesiva de los naturales no sólo puede organizar demostraciones.

También puede organizar definiciones.

Por ejemplo, podemos imaginar que queremos construir una función

$$a : \mathbb{N} \to X.$$

Una manera natural de hacerlo sería indicar un valor inicial

$$a(0)$$

y una regla que explique cómo obtener

$$a(n+1)$$

a partir de información anterior.

Aquí la pregunta no es

> «¿es verdadera una propiedad para todo n?»

sino

> «¿qué objeto queda definido en cada índice?»

Éste será el territorio de la recursión.

La semejanza externa entre

$$P(n) \Rightarrow P(n+1)$$

y una regla del tipo

$$a(n+1) = F(a(n))$$

puede ser engañosa.

La primera expresión relaciona **proposiciones** dentro de una demostración.

La segunda relaciona **valores** dentro de una definición.

Mantener esta diferencia desde el comienzo evitará que inducción y recursión se confundan más adelante.

Antes de hablar de inducción

Sobre un tramo de naturales debemos distinguir cuatro niveles:

1. el **índice** n;
2. la **propiedad indexada** $P(n)$;
3. una **instancia concreta** como $P(7)$;
4. la **afirmación universal**

$$\forall n \geq n_0, \quad P(n).$$

Comprobar muchas instancias no establece automáticamente la universal.

La estructura sucesiva de $\mathbb{N}$ será lo que permita sustituir una lista infinita de verificaciones por una justificación estructurada.

Comprobación. Considera la propiedad

$$P(n): \quad n^2 + n \text{ es par},$$

con $n \in \mathbb{N}$.

Explica la diferencia entre las tres expresiones siguientes:

$$P(4),$$

$$P(n),$$

y

$$\forall n \in \mathbb{N}\ P(n).$$

¿Bastaría comprobar $P(n)$ para los primeros mil naturales para establecer la tercera?

Respuesta razonada.

La expresión

$$P(4)$$

es una proposición concreta:

$$4^2 + 4 = 20$$

es par.

La expresión

$$P(n)$$

mantiene el índice n libre. Describe una propiedad que puede evaluarse para distintos naturales, pero todavía no afirma por sí sola que todos ellos la satisfagan.

En cambio,

$$\forall n \in \mathbb{N}\ P(n)$$

es una proposición universal: afirma que cada natural satisface la propiedad.

Verificar los primeros mil casos sólo establece

$$P(0), P(1), \ldots, P(999).$$

No proporciona por sí solo información sobre

$$P(1000), P(1001), \ldots.$$

Por tanto, una cantidad finita de comprobaciones favorables no demuestra la afirmación universal.

Lo que necesitaremos es una razón que no dependa de enumerar los casos uno por uno.

La pregunta que queda abierta. Ya tenemos el problema correctamente formulado. Queremos comprender cómo demostrar una afirmación de la forma

$$\forall n \geq n_0, \quad P(n)$$

sin verificar separadamente infinitos casos.

Sabemos además que los índices están conectados mediante

$$n \longmapsto n+1.$$

La pregunta natural es entonces:

> **¿qué combinación de información inicial y regla de paso permite cubrir todo el tramo?**

La respuesta será el principio de inducción matemática.

Ése es el contenido de §8.2.

8.2. El principio de inducción

En la sección anterior dejamos planteada una pregunta precisa. Si queremos demostrar

$$\forall n \in \mathbb{N}, \quad P(n),$$

¿cómo podemos evitar una lista infinita de verificaciones separadas?

La respuesta es aprovechar la estructura sucesiva de los naturales.

No intentaremos demostrar de manera independiente

$$P(0),\ P(1),\ P(2),\ P(3), \ldots$$

En lugar de eso, buscaremos dos tipos de información:

1. una afirmación que nos permita **entrar** en la cadena;
2. una regla que nos permita **propagar** la propiedad de un índice al siguiente.

Ésta es la arquitectura del principio de inducción.

Formulación del principio. Sea $P(n)$ una propiedad definida para $n \in \mathbb{N}$.

Si podemos demostrar:

$$P(0),$$

y además

$$\forall n \in \mathbb{N}, \qquad P(n) \Rightarrow P(n+1),$$

entonces podemos concluir

$$\forall n \in \mathbb{N}, \qquad P(n).$$

Lo escribiremos esquemáticamente como

$$\boxed{P(0)} \quad + \quad \boxed{\forall n \in \mathbb{N}\ \big(P(n) \Rightarrow P(n+1)\big)} \quad \Longrightarrow \quad \boxed{\forall n \in \mathbb{N}\ P(n)}.$$

A la primera obligación la llamamos **caso base** o **base de la inducción**.

A la segunda la llamamos **paso inductivo**.

Dentro del paso inductivo, cuando fijamos un n arbitrario y suponemos temporalmente $P(n)$ para demostrar $P(n+1)$, la afirmación $P(n)$ recibe el nombre de **hipótesis inductiva**.

La conclusión final es la afirmación universal que queríamos establecer.

La base proporciona un punto de entrada. ¿Por qué necesitamos demostrar $P(0)$? Porque la regla

$$P(n) \Rightarrow P(n+1)$$

no produce por sí sola ninguna instancia verdadera.

Supongamos que sabemos

$$P(0) \Rightarrow P(1),$$

$$P(1) \Rightarrow P(2),$$

$$P(2) \Rightarrow P(3),$$

y así sucesivamente.

Si no sabemos que alguna de las afirmaciones del comienzo es verdadera, la cadena no arranca.

Por ejemplo, una implicación

$$P(0) \Rightarrow P(1)$$

puede ser verdadera incluso si tanto $P(0)$ como $P(1)$ son falsas.

Por eso el paso inductivo no sustituye al caso base.

La base afirma que existe un primer punto de la cadena en el que la propiedad está efectivamente establecida.

El paso proporciona propagación. La base tampoco basta por sí sola.

Demostrar

$$P(0)$$

sólo demuestra una instancia.

Para llegar al resto necesitamos una regla que funcione uniformemente.

El paso inductivo exige demostrar

$$\forall n \in \mathbb{N}, \qquad P(n) \Rightarrow P(n+1).$$

La palabra **para todo** es decisiva.

No alcanza con comprobar, por ejemplo,

$$P(0) \Rightarrow P(1),$$

$$P(1) \Rightarrow P(2),$$

$$P(2) \Rightarrow P(3).$$

Eso seguiría siendo una colección finita de casos.

Necesitamos una razón que funcione para un índice natural arbitrario.

Podemos imaginar entonces la combinación:

$$P(0) \Rightarrow P(1) \Rightarrow P(2) \Rightarrow P(3) \Rightarrow \cdots .$$

La base legitima el primer eslabón disponible.

El paso inductivo garantiza que cada eslabón verdadero permite alcanzar el siguiente.

Ninguna de las dos partes puede omitirse. Conviene separar tres situaciones.

Sólo base.

Sabemos

$$P(0),$$

pero no poseemos una regla que conecte $P(n)$ con $P(n+1)$.

Entonces no sabemos nada, en general, sobre $P(1), P(2), \ldots$.

Sólo paso.

Sabemos

$$\forall n \in \mathbb{N}, \qquad P(n) \Rightarrow P(n+1),$$

pero no sabemos que $P(0)$ sea verdadera.

Entonces la regla puede estar perfectamente demostrada y, aun así, no proporcionar ninguna instancia verdadera.

Base y paso.

Cuando ambas piezas están disponibles, podemos avanzar:

$$P(0),$$

por el paso con $n = 0$ obtenemos

$$P(1),$$

por el paso con $n = 1$ obtenemos

$$P(2),$$

y el mismo mecanismo continúa a lo largo de todos los naturales.

La inducción no consiste en escribir efectivamente esta cadena infinita. El principio nos autoriza a concluir de una vez que toda la cadena queda cubierta.

Base + propagación

Una prueba por inducción combina dos obligaciones distintas:

$$\boxed{\text{entrada}} \quad + \quad \boxed{\text{propagación}} \quad \Longrightarrow \quad \boxed{\text{cobertura universal}}.$$

La base establece el primer caso del tramo.

El paso demuestra una regla uniforme que lleva de un caso al siguiente.

Ninguna de las dos partes, tomada aisladamente, produce la conclusión universal.

Un ejemplo completo. Demostremos que para todo $n \in \mathbb{N}$,

$$2^n \geq n + 1.$$

Definimos la propiedad

$$P(n): \quad 2^n \geq n + 1.$$

Queremos establecer

$$\forall n \in \mathbb{N}, \quad P(n).$$

Caso base. Para $n = 0$,

$$2^0 = 1$$

y

$$0 + 1 = 1.$$

Por tanto,

$$2^0 \geq 0 + 1.$$

Así, $P(0)$ es verdadera.
Paso inductivo. Sea $n \in \mathbb{N}$ arbitrario y supongamos

$$P(n),$$

es decir,

$$2^n \geq n + 1.$$

Debemos demostrar

$$P(n + 1),$$

que en este caso significa

$$2^{n+1} \geq n + 2.$$

A partir de la hipótesis inductiva,

$$2^n \geq n + 1,$$

multiplicamos por $2 > 0$:

$$2^{n+1} \geq 2n + 2.$$

Como $n \geq 0$,

$$2n + 2 \geq n + 2.$$

Por transitividad,

$$2^{n+1} \geq n + 2.$$

Así hemos demostrado

$$P(n) \Rightarrow P(n + 1)$$

para un $n \in \mathbb{N}$ arbitrario.
Por el principio de inducción,

$$\boxed{2^n \geq n + 1 \text{ para todo } n \in \mathbb{N}}.$$

La arquitectura y el contenido siguen siendo distintos. El principio de inducción nos dice **qué obligaciones debe cumplir la prueba**:

$$\text{base} \quad + \quad \text{paso} \quad \Longrightarrow \quad \text{conclusión universal.}$$

Pero no nos dice automáticamente cómo demostrar el paso.

En el ejemplo anterior, la parte específicamente matemática fue reconocer que

$$2^{n+1} = 2 \cdot 2^n$$

y que la hipótesis

$$2^n \geq n + 1$$

podía combinarse con

$$2n + 2 \geq n + 2.$$

Otra propiedad $P(n)$ requerirá otra idea.

Esto conecta C08 con la distinción ya desarrollada entre arquitectura y contenido: saber que una prueba debe ser inductiva no significa haber encontrado todavía el argumento inductivo.

Qué significa realmente "suponer $P(n)$". En el paso acabamos de escribir:

> sea $n \in \mathbb{N}$ arbitrario y supongamos $P(n)$.

Esa frase puede parecer sospechosa al principio.

¿No estamos suponiendo precisamente lo que queremos demostrar?

No.

El objetivo global es

$$\forall n \in \mathbb{N}, \quad P(n).$$

Dentro del paso no suponemos esa afirmación universal.

Suponemos temporalmente, para un índice arbitrario pero fijo,

$$P(n),$$

con el único objetivo de demostrar la implicación

$$P(n) \Rightarrow P(n+1).$$

La diferencia lógica es

$$\boxed{\forall n\ P(n)} \qquad \neq \qquad \boxed{\forall n\ \big(P(n) \Rightarrow P(n+1)\big)}.$$

En §8.4 estudiaremos esta distinción con más detalle y veremos varios usos incorrectos de la hipótesis inductiva.

La conclusión no es sólo el caso siguiente. Otro error frecuente consiste en terminar el paso después de demostrar

$$P(n+1)$$

y creer que ésa es la conclusión del teorema.

No lo es.

Demostrar

$$P(n) \Rightarrow P(n+1)$$

establece la regla de propagación.

La conclusión que autoriza el principio, después de haber verificado también la base, es

$$\forall n \in \mathbb{N}, \quad P(n).$$

Por eso una prueba inductiva tiene tres niveles que no deben mezclarse:

$$\begin{array}{rcl} \text{base} & : & P(0), \\ \text{paso} & : & \forall n \in \mathbb{N}\ \big(P(n) \Rightarrow P(n+1)\big), \\ \text{conclusión} & : & \forall n \in \mathbb{N}\ P(n). \end{array}$$

Comprobación. Supón que una propiedad $Q(n)$ está definida para todos los naturales y que has demostrado:

$$Q(0)$$

y

$$\forall n \in \mathbb{N}, \qquad Q(n) \Rightarrow Q(n+1).$$

Un compañero afirma que todavía habría que demostrar por separado

$$Q(1), Q(2), Q(3), \ldots$$

antes de poder concluir que $Q(n)$ vale para todo natural.

¿Tiene razón? Explica exactamente qué papel cumple cada una de las dos premisas.

Respuesta razonada.

No.

La primera premisa,

$$Q(0),$$

proporciona el punto de entrada.

La segunda,

$$\forall n \in \mathbb{N}, \qquad Q(n) \Rightarrow Q(n+1),$$

proporciona una regla de propagación válida para cualquier índice natural.

En particular, como $Q(0)$ es verdadera, el paso con $n = 0$ da $Q(1)$. Después, como $Q(1)$ es verdadera, el paso con $n = 1$ da $Q(2)$, y así sucesivamente.

El principio de inducción afirma precisamente que no necesitamos ejecutar esta cadena como una lista infinita de pruebas separadas. La combinación de base y paso basta para concluir

$$\forall n \in \mathbb{N}, \quad Q(n).$$

La base sin el paso sólo cubriría un caso; el paso sin la base no proporcionaría un caso verdadero desde el cual comenzar.

Lo que todavía no hemos variado. Hasta aquí hemos formulado la inducción con la convención canónica del capítulo:

$$\mathbb{N} = \{0, 1, 2, \ldots\}$$

y hemos usado como base

$$P(0).$$

Pero muchas afirmaciones matemáticas comienzan en otro índice.

Podemos querer demostrar, por ejemplo,

$$\forall n \geq 5, \quad P(n).$$

En ese caso no debemos insertar mecánicamente una base en 0.

La siguiente sección estudiará cómo se desplaza correctamente toda la arquitectura inductiva cuando el tramo empieza en un índice n_0 distinto de 0.

8.3. Empezar en otro índice

La formulación de §8.2 comenzó en 0 porque hemos fijado

$$\mathbb{N} = \{0, 1, 2, \ldots\}.$$

Pero el principio de inducción no está ligado al número 0.

Una afirmación puede pretender valer sólo a partir de cierto índice

$$n_0 \in \mathbb{N}.$$

En ese caso el dominio real del teorema es

$$\mathbb{N}_{\geq n_0} = \{n \in \mathbb{N} : n \geq n_0\},$$

y la arquitectura inductiva debe comenzar precisamente allí.

La forma desplazada del principio. Sea $P(n)$ una propiedad de los naturales. Supongamos que la afirmación que queremos demostrar sólo pretende cubrir el tramo $n \geq n_0$.

Si demostramos

$$P(n_0)$$

y además

$$\forall n \geq n_0, \qquad P(n) \Rightarrow P(n+1),$$

entonces podemos concluir

$$\forall n \geq n_0, \qquad P(n).$$

La estructura es exactamente la misma que antes:

$$\boxed{P(n_0)} \qquad + \qquad \boxed{\forall n \geq n_0 \ (P(n) \Rightarrow P(n+1))} \qquad \Longrightarrow \qquad \boxed{\forall n \geq n_0 \ P(n)}.$$

Sólo ha cambiado el punto de entrada.

Esto significa que la base correcta no se elige por costumbre. Se lee en el propio enunciado.

Si el teorema afirma algo para

$$n \geq 3,$$

la primera obligación es $P(3)$.
Si afirma algo para

$$n \geq 12,$$

la primera obligación es $P(12)$.
El índice inicial es parte de la información matemática del teorema.

Un ejemplo que no comienza en cero. Consideremos la propiedad

$$P(n): \quad n^2 \geq 2n + 1.$$

No es verdadera para todos los naturales.
En efecto,

$$P(0): \quad 0 \geq 1$$

es falsa,

$$P(1): \quad 1 \geq 3$$

es falsa, y

$$P(2): \quad 4 \geq 5$$

es falsa.
En cambio,

$$P(3): \quad 9 \geq 7$$

es verdadera.
El enunciado correcto que estudiaremos es, por tanto,

$$\forall n \geq 3, \qquad n^2 \geq 2n + 1.$$

Caso base. Para $n = 3$,

$$3^2 = 9$$

y

$$2 \cdot 3 + 1 = 7.$$

Por tanto,

$$3^2 \geq 2 \cdot 3 + 1.$$

Así, $P(3)$ es verdadera.

Paso inductivo. Sea $n \geq 3$ y supongamos

$$P(n),$$

es decir,

$$n^2 \geq 2n + 1.$$

Queremos demostrar

$$P(n+1),$$

que significa

$$(n+1)^2 \geq 2(n+1) + 1.$$

Partimos de

$$(n+1)^2 = n^2 + 2n + 1.$$

Usando la hipótesis inductiva,

$$n^2 \geq 2n + 1,$$

obtenemos

$$(n+1)^2 \geq (2n+1) + 2n + 1 = 4n + 2.$$

Como $n \geq 3$, en particular

$$2n - 1 \geq 0,$$

y por tanto

$$4n + 2 \geq 2n + 3.$$

Pero

$$2n + 3 = 2(n+1) + 1.$$

Así,

$$(n+1)^2 \geq 2(n+1) + 1.$$

Hemos probado

$$P(n) \Rightarrow P(n+1)$$

para todo $n \geq 3$.
Por inducción desde 3,

$$\boxed{n^2 \geq 2n + 1 \text{ para todo } n \geq 3}.$$

Este ejemplo muestra por qué no debemos exigir una base en 0: la afirmación que queremos demostrar ni siquiera es verdadera allí.

La base debe pertenecer al dominio real del teorema. Supongamos que nuestro objetivo es

$$\forall n \geq 5, \quad P(n).$$

Demostrar $P(0)$ no cumple la obligación de base de este argumento.

Incluso si $P(0)$ fuese verdadera, el paso podría estar disponible únicamente para

$$n \geq 5.$$

Entonces conoceríamos

$$P(0)$$

y

$$P(5) \Rightarrow P(6), \quad P(6) \Rightarrow P(7), \quad \ldots,$$

pero no tendríamos ninguna cadena que conectara $P(0)$ con $P(5)$.

La base no es simplemente "algún caso verdadero".

Debe ser un caso desde el cual la regla de propagación pueda comenzar a recorrer exactamente el tramo que queremos cubrir.

Base y paso deben conectarse sin huecos. También podemos equivocarnos aunque hayamos elegido un índice cercano al correcto.

Supongamos nuevamente que queremos demostrar

$$\forall n \geq 3, \quad P(n).$$

Consideremos dos conjuntos de información.

Primera situación:

$$P(4)$$

y

$$\forall n \geq 4, \quad P(n) \Rightarrow P(n+1).$$

Esto permite concluir

$$P(4), P(5), P(6), \ldots$$

pero no establece $P(3)$.

Hemos demostrado una afirmación sobre un tramo más corto:

$$\forall n \geq 4, \quad P(n).$$

Segunda situación:

$$P(3)$$

y

$$\forall n \geq 4, \quad P(n) \Rightarrow P(n+1).$$

Ahora sí tenemos la base correcta, pero falta el enlace

$$P(3) \Rightarrow P(4).$$

La regla de propagación comienza demasiado tarde.

Por tanto, una prueba inductiva debe pasar una prueba de continuidad lógica:

$$\boxed{\text{base}} \longrightarrow \boxed{\text{primer paso disponible}} \longrightarrow \boxed{\text{todos los sucesores del tramo}}.$$

Si queda un hueco entre la base y el rango del paso, la inducción no cubre el dominio anunciado.

Saltar dos índices no equivale a recorrer todos. En §8.2 exigimos que la propagación pasara de n a $n+1$. Es posible, sin embargo, que la única implicación que hayamos conseguido demostrar sea

$$\forall n \in \mathbb{N}, \qquad P(n) \Rightarrow P(n+2).$$

Esta información tampoco es inútil, pero **no** constituye el paso inductivo ordinario para una conclusión sobre *todos* los naturales. El salto de dos unidades conserva la paridad del índice: desde 0 permite alcanzar $2, 4, 6, \ldots$, sin llegar a $1, 3, 5, \ldots$.

Contraejemplo a una inducción incompleta. Definamos $P(n)$ como «n es par». La base $P(0)$ es verdadera. Además, si $n = 2k$ para algún $k \in \mathbb{N}$, entonces $n + 2 = 2(k + 1)$, de modo que $P(n) \Rightarrow P(n + 2)$ es verdadera para los índices pares. Para un índice impar, $P(n)$ es falsa y la implicación también es verdadera en lógica clásica. Así, la regla vale para **todos** los naturales. A pesar de ello, la afirmación $\forall n \in \mathbb{N}\ P(n)$ es falsa: 1 no es par. No ha fallado la aritmética del paso; ha fallado su capacidad para cubrir el dominio anunciado.

Proposición (dos bases para un salto de dos). Para una propiedad P de los naturales, supongamos que se han demostrado

$$P(0), \qquad P(1), \qquad \forall n \in \mathbb{N}\ \big(P(n) \Rightarrow P(n+2)\big).$$

Entonces $P(n)$ vale para todo $n \in \mathbb{N}$.

Demostración. Reunamos las dos clases de índices en una sola propiedad:

$$Q(k): \quad P(2k) \wedge P(2k+1).$$

Por las dos bases, $Q(0)$ es verdadera. Supongamos $Q(k)$ para un $k \in \mathbb{N}$ arbitrario. Podemos aplicar el paso dado con $n = 2k$ y con $n = 2k + 1$; obtenemos respectivamente $P(2k + 2)$ y $P(2k+3)$. Ambas constituyen $Q(k+1)$. Por inducción ordinaria, $Q(k)$ vale para todo natural k. Como cualquier natural es de la forma $2k$ o $2k + 1$, todas las instancias de P quedan cubiertas. □

El resultado no rescata la falsa afirmación «todos los naturales son pares»: en ese caso falta precisamente $P(1)$. Muestra, en cambio, cómo una regla que salta posiciones exige suficientes casos iniciales para alcanzar **cada clase de índices**. La prueba agrupó dos casos consecutivos y aplicó el principio ordinario a esa propiedad más rica.

Comprobación. ¿Serían suficientes $P(0)$, $P(2)$ y la regla $P(n) \Rightarrow P(n + 2)$ para todo natural n?

Respuesta razonada. No. Las dos bases pertenecen a la misma clase de índices: son pares. La regla sólo produce otros índices pares y no aporta $P(1)$ ni ninguna instancia impar. Tomando otra vez $P(n)$ como «n es par» se verifican los tres datos, pero no la conclusión universal. Para cubrir todos los naturales mediante ese salto haría falta una base impar, como $P(1)$, además de una base par.

La forma con cuantificadores. La abreviatura

$$\forall n \geq n_0, \quad P(n)$$

puede leerse como una cuantificación sobre el dominio restringido

$$\mathbb{N}_{\geq n_0}.$$

Si $P(n)$ está definida para todo $n \in \mathbb{N}$, podemos expandir esa restricción como

$$\forall n \in \mathbb{N}, \qquad n \geq n_0 \Rightarrow P(n).$$

Del mismo modo, cuando P tiene sentido en todo $\mathbb{N}$, el paso desplazado puede escribirse como

$$\forall n \in \mathbb{N}, \qquad n \geq n_0 \Rightarrow \big(P(n) \Rightarrow P(n+1)\big).$$

Si P sólo está formulada sobre $\mathbb{N}_{\geq n_0}$, conservaremos en cambio la notación restringida y no evaluaremos $P(n)$ fuera de su dominio.

Esto hace visible que la restricción

$$n \geq n_0$$

forma parte de la hipótesis bajo la cual exigimos el paso.

No afirmamos necesariamente nada sobre

$$P(0), P(1), \ldots, P(n_0 - 1).$$

Esos casos están fuera del dominio de la conclusión.

Inducción desde n_0 no es un principio nuevo. La forma desplazada puede reducirse a la inducción ordinaria de §8.2.

Definamos, para $k \in \mathbb{N}$,

$$Q(k) := P(n_0 + k).$$

La base de Q es

$$Q(0) = P(n_0).$$

Ahora supongamos $Q(k)$ para un $k \in \mathbb{N}$. Esto significa

$$P(n_0 + k).$$

Como

$$n_0 + k \geq n_0,$$

podemos aplicar el paso desplazado y obtener

$$P(n_0 + k + 1).$$

Pero esto es exactamente

$$Q(k+1).$$

Por la inducción ordinaria,

$$Q(k)$$

vale para todo $k \in \mathbb{N}$.

Es decir,

$$P(n_0 + k)$$

vale para todo $k \in \mathbb{N}$.

Y todo natural $n \geq n_0$ puede escribirse como

$$n = n_0 + k$$

con

$$k = n - n_0 \in \mathbb{N}.$$

Por tanto,

$$\forall n \geq n_0, \quad P(n).$$

La "inducción desde n_0" es la misma estructura vista después de cambiar el origen de los índices.

El primer índice no es decoración

Para demostrar

$$\forall n \geq n_0, \quad P(n),$$

la arquitectura correcta es

$$P(n_0)$$

y

$$\forall n \geq n_0, \qquad P(n) \Rightarrow P(n+1).$$

La base debe pertenecer al dominio real del teorema y el paso debe comenzar exactamente donde esa base puede activarlo.

Si el paso empieza más tarde o la base pertenece a otro tramo, hay que localizar y justificar explícitamente el hueco restante.

Comprobación. Quieres demostrar

$$\forall n \geq 6, \quad P(n).$$

Has conseguido demostrar

$$P(6)$$

y

$$\forall n \geq 7, \qquad P(n) \Rightarrow P(n+1).$$

¿Basta esta información para aplicar directamente la inducción desde 6? Si no basta, identifica exactamente qué falta y da dos maneras de reparar el argumento.

Respuesta razonada.

No basta.

La base

$$P(6)$$

es correcta, pero la regla disponible sólo empieza en

$$n = 7.$$

Por tanto, conocemos

$$P(7) \Rightarrow P(8), \quad P(8) \Rightarrow P(9), \quad \ldots,$$

pero todavía no hemos justificado

$$P(6) \Rightarrow P(7).$$

El hueco está exactamente entre 6 y 7.

Una primera reparación consiste en fortalecer el paso y demostrar

$$\forall n \geq 6, \qquad P(n) \Rightarrow P(n+1).$$

Entonces $P(6)$ activa inmediatamente la propagación.

Una segunda reparación consiste en demostrar $P(7)$ por separado. Después podemos usar $P(7)$ como base de una inducción que cubra

$$n \geq 7,$$

y combinar ese resultado con el caso ya demostrado $P(6)$.

En ambos casos desaparece el hueco. Lo importante no es que la palabra "base" aparezca en la prueba, sino que la cadena de cobertura alcance efectivamente todos los índices anunciados.

La próxima dificultad no es el índice. Ya sabemos adaptar la base al dominio correcto.

Pero incluso con el índice inicial bien elegido puede persistir una duda más profunda.

En el paso inductivo escribimos

"supongamos $P(n)$ y demostremos $P(n+1)$".

¿Por qué esa suposición es legítima?

¿Qué podemos hacer con ella?

¿Y en qué momento un uso descuidado de $P(n)$ sí produciría circularidad?

Éstas serán las preguntas de §8.4.

8.4. La hipótesis inductiva: uso y límites

En una prueba por inducción hay una frase que merece una atención especial:

supongamos que $P(n)$ es verdadera.

A primera vista puede parecer que estamos suponiendo exactamente lo que queremos demostrar. Si el objetivo global es

$$\forall n \geq n_0, \quad P(n),$$

¿cómo puede ser legítimo escribir $P(n)$ como hipótesis?

La respuesta está en distinguir con precisión **qué afirmación estamos demostrando en cada nivel**.

El paso inductivo es una implicación. Dentro del paso inductivo no intentamos demostrar todavía

$$\forall n \geq n_0, \quad P(n).$$

Intentamos demostrar otra afirmación:

$$\forall n \geq n_0, \qquad P(n) \Rightarrow P(n+1).$$

Para probar una implicación

$$A \Rightarrow B,$$

la arquitectura directa de C06 permite suponer temporalmente A y, bajo esa suposición, demostrar B.

Aquí hacemos exactamente lo mismo.

Fijamos un índice arbitrario

$$n \geq n_0,$$

suponemos

$$P(n),$$

y tratamos de obtener

$$P(n+1).$$

Si lo conseguimos, hemos probado la implicación

$$P(n) \Rightarrow P(n+1)$$

para ese índice arbitrario. Como n no tenía ninguna característica especial, queda establecida la regla universal de propagación.

La hipótesis inductiva es, por tanto, la **hipótesis local de una prueba condicional**.

No es la conclusión universal.

Tres afirmaciones que no deben confundirse. Conviene colocar juntas las tres piezas:

$$\boxed{P(n)}$$

es la hipótesis inductiva para un índice arbitrario pero fijo;

$$\boxed{P(n) \Rightarrow P(n+1)}$$

es la implicación que debemos demostrar en el paso;

$$\boxed{\forall k \geq n_0, \quad P(k)}$$

es la conclusión global que sólo podremos obtener después de combinar base y paso mediante el principio de inducción.

El cambio de letra en la última expresión es deliberado. Ayuda a recordar que el n usado dentro del paso tiene alcance local. La conclusión final habla de **todos** los índices del tramo.

La hipótesis inductiva tiene un alcance limitado. Supongamos que estamos probando por inducción

$$\forall n \geq 2, \quad P(n).$$

En el paso podemos escribir:

> Sea $n \geq 2$ arbitrario. Supongamos $P(n)$. Demostraremos $P(n+1)$.

Durante ese argumento podemos utilizar:

- $P(n)$;
- las hipótesis generales del teorema, si las hay;
- resultados ya demostrados anteriormente;

- hechos algebraicos o lógicos disponibles.

Pero no podemos utilizar gratuitamente

$$P(n+1),$$

porque ésa es justamente la conclusión local del paso.

Tampoco podemos utilizar

$$\forall k \geq 2, \quad P(k),$$

porque ésa es la conclusión global que todavía estamos intentando establecer.

En otras palabras, la hipótesis inductiva concede exactamente

$$P(n),$$

no una licencia para usar cualquier instancia futura de la propiedad.

Un uso legítimo de la hipótesis inductiva. Consideremos la afirmación

$$P(n): \quad 6 \mid 7^n - 1,$$

para $n \in \mathbb{N}$.

El caso base es inmediato:

$$7^0 - 1 = 0,$$

que es divisible por 6.

Para el paso inductivo, fijemos $n \in \mathbb{N}$ y supongamos

$$6 \mid 7^n - 1.$$

Esta hipótesis no es sólo una etiqueta. Debemos convertirla en información utilizable. Por definición de divisibilidad, existe un entero k tal que

$$7^n - 1 = 6k.$$

Ahora observamos que

$$7^{n+1} - 1 = 7 \cdot 7^n - 1.$$

Reescribimos:

$$7^{n+1} - 1 = 7(7^n - 1) + 6.$$

Usando la hipótesis inductiva,

$$7^{n+1} - 1 = 7(6k) + 6 = 6(7k+1).$$

Como $7k+1$ es un entero,

$$6 \mid 7^{n+1} - 1.$$

Así hemos demostrado

$$P(n) \Rightarrow P(n+1).$$

Aquí puede señalarse exactamente dónde se usó la hipótesis inductiva: en la sustitución

$$7^n - 1 = 6k.$$

Una buena prueba inductiva permite localizar ese punto.

La hipótesis debe transformarse, no sólo mencionarse. Un defecto frecuente consiste en escribir algo como:

> Por hipótesis inductiva, $P(n)$ es verdadera. Entonces claramente $P(n+1)$ también es verdadera.

Esto no demuestra el paso.

La hipótesis inductiva sólo tiene valor matemático cuando podemos explicar **qué información concreta aporta** y **cómo esa información contribuye al objetivo** $P(n+1)$.

En el ejemplo anterior, la cadena fue:

$$6 \mid 7^n - 1$$

$$\Downarrow$$

$$7^n - 1 = 6k$$

$$\Downarrow$$

$$7^{n+1} - 1 = 6(7k+1)$$

$$\Downarrow$$

$$6 \mid 7^{n+1} - 1.$$

La hipótesis no funciona por proximidad notacional entre n y $n+1$. Tiene que existir un argumento que conecte ambas instancias.

Cuándo aparece circularidad. Ahora podemos identificar con precisión varios usos ilegítimos.

1. Suponer directamente el objetivo local. Si queremos demostrar

$$P(n+1)$$

y escribimos

> Supongamos que $P(n+1)$ es verdadera.

sin estar dentro de una prueba por contradicción u otra arquitectura que lo justifique, hemos supuesto la conclusión que queríamos obtener.

Eso es circular.

2. Usar la conclusión universal dentro del paso. Si durante el paso escribimos

$$\forall k \geq n_0, \quad P(k),$$

y luego tomamos $k = n+1$, obtenemos inmediatamente $P(n+1)$.

Pero la afirmación universal es precisamente el teorema completo que tratamos de demostrar.

Utilizarla en el paso equivale a dar por demostrado el resultado global.

3. Cambiar el índice de la hipótesis sin justificación. De

$$P(n)$$

no podemos pasar simplemente a

$$P(n+1)$$

sustituyendo n por $n+1$.

La hipótesis inductiva no es una afirmación universal en la variable n. Es una afirmación asumida para el índice arbitrario que hemos fijado dentro del paso.

Precisamente la tarea del paso consiste en justificar el tránsito

$$P(n) \longrightarrow P(n+1).$$

Una distinción sutil: no usar la hipótesis no siempre es un error. Supongamos que durante el paso conseguimos demostrar $P(n+1)$ sin utilizar $P(n)$ en absoluto.

¿La prueba es incorrecta?

No necesariamente.

Si el argumento demuestra $P(n+1)$ a partir de hechos independientes de la hipótesis inductiva, entonces hemos demostrado algo incluso más fuerte que la implicación condicional necesaria.

En efecto, si $P(n+1)$ puede establecerse directamente para un n arbitrario, entonces ciertamente

$$P(n) \Rightarrow P(n+1)$$

es verdadera.

Lo que ocurre es distinto: **la inducción era innecesaria**.

Esto proporciona una buena regla diagnóstica:

- si la hipótesis inductiva no aparece porque hemos encontrado una prueba directa más fuerte, no hay error lógico;
- si la hipótesis inductiva no aparece porque hemos omitido el único puente disponible hacia $P(n+1)$, el paso está incompleto.

La ausencia de uso de la hipótesis es, por tanto, una señal que debe interpretarse, no una falacia automática.

El supuesto se descarga al terminar el paso. La hipótesis inductiva tampoco permanece disponible después del paso como una afirmación demostrada de manera independiente.

El esquema lógico es:

Sea $n \geq n_0$ arbitrario.
  Supongamos $P(n)$.
  $\vdots$
  Demostramos $P(n+1)$.
Por tanto $P(n) \Rightarrow P(n+1)$.

Al llegar a la última línea, la suposición local $P(n)$ ha sido **descargada**: lo que queda demostrado es la implicación.

Después universalizamos sobre el índice arbitrario:

$$\forall n \geq n_0, \qquad P(n) \Rightarrow P(n+1).$$

Sólo entonces, junto con la base, invocamos el principio de inducción para obtener la conclusión global.

Esta secuencia de dependencias puede representarse así:

$$\boxed{P(n_0)} \quad + \quad \boxed{\begin{array}{c} P(n) \text{ [supuesto local]} \\ \Downarrow \\ P(n+1) \end{array}}$$

$$\Downarrow$$

$$\boxed{\forall n \geq n_0 \; (P(n) \Rightarrow P(n+1))}$$

$$\Downarrow$$

$$\boxed{\forall n \geq n_0 \; P(n)}.$$

Alcance de la hipótesis inductiva

La hipótesis inductiva no dice

$$\forall k \geq n_0, \quad P(k).$$

Dentro del paso dice únicamente:

$$\text{para este } n \text{ arbitrario y fijo, supongamos } P(n).$$

Su función es demostrar la implicación

$$P(n) \Rightarrow P(n+1).$$

Es legítima porque estamos probando una condicional. Se vuelve circular si utilizamos sin justificación $P(n+1)$ o la conclusión universal antes de haberlas demostrado.

Comprobación. Quieres demostrar por inducción una propiedad $P(n)$ para todo $n \geq 1$. En el paso inductivo, un borrador contiene estas cuatro líneas:

1. Sea $n \geq 1$ arbitrario.
2. Supongamos $P(n)$.
3. Como el teorema dice que $P(k)$ vale para todo $k \geq 1$, tomando $k = n+1$ tenemos $P(n+1)$.
4. Por tanto, el paso inductivo está demostrado.

Localiza exactamente el error. ¿Por qué la línea 2 sí es legítima pero la línea 3 no lo es?

Respuesta razonada.

La línea 2 es legítima porque el paso inductivo tiene como objetivo demostrar la implicación

$$P(n) \Rightarrow P(n+1).$$

Para probar esa implicación podemos suponer temporalmente su antecedente $P(n)$.

La línea 3, en cambio, utiliza

$$\forall k \geq 1, \quad P(k),$$

que es precisamente la conclusión global del teorema que todavía no ha sido demostrada.

Al tomar $k = n + 1$ se obtiene el objetivo local sólo porque se ha supuesto de antemano el resultado universal completo.

Por tanto, el argumento es circular.

La reparación no consiste en eliminar la hipótesis inductiva, sino en conservar $P(n)$ y encontrar una cadena matemática independiente que, a partir de esa información y de resultados previamente disponibles, conduzca a $P(n+1)$.

De una hipótesis a varias. La inducción ordinaria concede, en el paso, una sola instancia previa:

$$P(n).$$

Hay problemas en los que para demostrar $P(n+1)$ resulta natural necesitar información sobre más de un índice anterior, por ejemplo

$$P(n), \quad P(n-1), \quad P(n-2).$$

Intentar forzar toda esa información dentro de una única hipótesis $P(n)$ puede volver artificial la prueba.

La siguiente sección estudia una formulación en la que el paso puede utilizar todas las instancias anteriores pertinentes.

Ése será el principio de **inducción fuerte**.

8.5. Inducción fuerte

En la inducción ordinaria, el paso tiene la forma

$$P(n) \Rightarrow P(n+1).$$

La información inductiva disponible es, por tanto, una sola instancia previa:

$$P(n).$$

Pero hay situaciones en las que el caso siguiente no depende de manera natural del caso inmediatamente anterior. Para demostrar $P(n+1)$ quizá necesitemos una afirmación sobre

$$P(n-1),$$

o sobre varios índices anteriores a la vez.

La **inducción fuerte** organiza precisamente esa situación.

La forma fuerte del paso. Sea $P(n)$ una propiedad que queremos demostrar para todos los naturales

$$n \geq n_0.$$

En la forma ordinaria intentaríamos probar

$$P(n) \Rightarrow P(n+1).$$

En la forma fuerte podemos suponer, para un índice arbitrario $n \geq n_0$, que todas las instancias anteriores del tramo son verdaderas:

$$P(n_0), P(n_0+1), \ldots, P(n).$$

Bajo esa hipótesis debemos demostrar

$$P(n+1).$$

Esquemáticamente:

$$\boxed{P(n_0), P(n_0+1), \ldots, P(n)} \quad \Longrightarrow \quad \boxed{P(n+1)}.$$

La conclusión final sigue siendo exactamente la misma:

$$\forall n \geq n_0, \quad P(n).$$

Por eso la palabra **fuerte** puede ser engañosa si se interpreta mal.

No estamos intentando demostrar una afirmación más fuerte.

Lo que se fortalece es la **hipótesis local disponible durante el paso**.

Formulación del principio. Una forma básica del principio de inducción fuerte es la siguiente.

Sea $P(n)$ una propiedad de los naturales y supongamos que queremos demostrarla para todo $n \geq n_0$. Supongamos que:

1. $P(n_0)$ es verdadera;
2. para todo $n \geq n_0$, si

$$P(k)$$

es verdadera para cada k con

$$n_0 \leq k \leq n,$$

entonces $P(n+1)$ es verdadera.

Entonces

$$\forall n \geq n_0, \quad P(n).$$

Podemos escribir el paso de manera compacta como

$$[\forall k \in \mathbb{N}, \quad n_0 \leq k \leq n \Rightarrow P(k)] \Rightarrow P(n+1).$$

La expresión entre corchetes es la **hipótesis inductiva fuerte**.

Su alcance sigue siendo local: se supone para demostrar el paso correspondiente a un índice arbitrario n.

No estamos suponiendo de antemano

$$\forall k \geq n_0, \quad P(k).$$

Sólo suponemos las instancias comprendidas entre el inicio del tramo y el índice actual.

Qué información adicional obtenemos. Comparemos las dos formas.

En inducción ordinaria, para demostrar $P(n+1)$ disponemos de

$$\boxed{P(n)}.$$

En inducción fuerte, para demostrar $P(n+1)$ disponemos de

$$\boxed{P(n_0), P(n_0+1), \ldots, P(n)}.$$

La segunda caja contiene la primera, pero también contiene información sobre índices anteriores.

Eso permite, por ejemplo, utilizar legítimamente

$$P(n-2)$$

si

$$n-2 \geq n_0.$$

En una inducción ordinaria, $P(n-2)$ no forma parte automáticamente de la hipótesis inductiva local.

Un ejemplo donde conviene mirar más atrás. Demostremos que todo entero

$$n \geq 8$$

puede escribirse en la forma

$$n = 3a + 5b$$

para ciertos

$$a, b \in \mathbb{N}.$$

Definamos

$$P(n): \quad \exists a, b \in \mathbb{N} \text{ tales que } n = 3a + 5b.$$

Nuestro objetivo es

$$\forall n \geq 8, \quad P(n).$$

Casos iniciales. Tenemos

$$8 = 3 + 5 = 3 \cdot 1 + 5 \cdot 1,$$

$$9 = 3 \cdot 3 + 5 \cdot 0,$$

y

$$10 = 3 \cdot 0 + 5 \cdot 2.$$

Por tanto,

$$P(8), \qquad P(9), \qquad P(10)$$

son verdaderas.

Aquí hemos verificado tres casos iniciales porque el mecanismo que utilizaremos en el paso necesita retroceder tres unidades.

Paso fuerte. Sea

$$n \geq 10$$

arbitrario y supongamos que

$$P(k)$$

es verdadera para todo

$$8 \leq k \leq n.$$

Queremos demostrar

$$P(n+1).$$

Como

$$n \geq 10,$$

tenemos

$$n - 2 \geq 8.$$

Además,

$$n - 2 \leq n.$$

Por la hipótesis inductiva fuerte podemos utilizar específicamente

$$P(n-2).$$

Esto significa que existen

$$a, b \in \mathbb{N}$$

tales que

$$n - 2 = 3a + 5b.$$

Sumando 3 obtenemos

$$n + 1 = 3(a+1) + 5b.$$

Como

$$a + 1 \in \mathbb{N},$$

hemos obtenido una representación de la forma requerida.
Por tanto,

$$P(n+1)$$

es verdadera.
Así, por inducción fuerte,

$$\boxed{\forall n \geq 8, \quad \exists a, b \in \mathbb{N} \text{ tales que } n = 3a + 5b}.$$

Por qué la hipótesis ordinaria sería incómoda aquí. En el paso anterior no utilizamos

$$P(n).$$

Utilizamos

$$P(n-2).$$

La forma fuerte nos permitió hacerlo porque $n-2$ pertenece al intervalo de índices ya cubierto:

$$8 \leq n-2 \leq n.$$

Podríamos intentar reorganizar el argumento para usar sólo inducción ordinaria, pero entonces tendríamos que modificar la propiedad inductiva o empaquetar información adicional.

La inducción fuerte permite expresar directamente la dependencia matemática que el problema ya tiene.

Ésta es su principal ventaja práctica.

Varios casos base no son un ritual. El ejemplo anterior utilizó

$$P(8), P(9), P(10).$$

Esto no significa que toda inducción fuerte requiera tres casos base.

La cantidad de casos iniciales depende de cuándo comienza a ser válida la regla del paso.

En nuestro argumento, para producir $P(n+1)$ utilizamos $P(n-2)$. El primer valor de n para el cual

$$n-2 \geq 8$$

es

$$n = 10.$$

Por eso el paso empieza en 10, y los índices

$$8, 9, 10$$

deben quedar cubiertos antes de activarlo.

La regla general es la misma que encontramos en §8.3:

los casos iniciales y el rango del paso deben conectarse sin huecos.

Fuerte no significa circular. La hipótesis fuerte puede parecer aún más sospechosa que la ordinaria:

> "supongamos que $P(k)$ es verdadera para todos los índices desde n_0 hasta n".

¿Por qué no estamos suponiendo ya el teorema?

Porque el alcance termina en n.

La instancia que debemos demostrar es

$$P(n+1),$$

y no está incluida entre las hipótesis disponibles.

La frontera lógica es

$$\boxed{n_0 \le k \le n} \qquad \text{frente a} \qquad \boxed{k = n+1}.$$

Mientras no utilicemos sin justificación $P(n+1)$ ni una instancia posterior, no hay circularidad.

La inducción fuerte no cambia la conclusión. Comparemos otra vez:

Inducción ordinaria

$$P(n) \Rightarrow P(n+1).$$

Inducción fuerte

$$P(n_0) \wedge P(n_0+1) \wedge \cdots \wedge P(n) \Rightarrow P(n+1).$$

En ambos casos queremos terminar con

$$\forall n \ge n_0, \quad P(n).$$

La diferencia reside en cómo organizamos la información utilizada durante el paso.

En §8.7 demostraremos con detalle que las formas de inducción que estamos estudiando están estrechamente relacionadas. Por ahora basta observar una transformación importante.

Si definimos

$$Q(n): \quad P(n_0), P(n_0+1), \ldots, P(n) \text{ son todas verdaderas,}$$

entonces una prueba por inducción ordinaria sobre $Q(n)$ puede transportar simultáneamente toda la información acumulada hasta n.

Esta idea explica por qué la forma fuerte no introduce una fuente nueva y misteriosa de verdad: puede **empaquetarse** la historia previa en una sola propiedad inductiva.

No desarrollaremos todavía la equivalencia completa; ése será uno de los objetivos de §8.7.

Qué fortalece la inducción fuerte

La inducción fuerte no fortalece la conclusión.

En el paso fortalece la información disponible:

$$P(n)$$

se reemplaza por

$$P(n_0), P(n_0+1), \ldots, P(n).$$

Esto permite usar cualquier caso anterior pertinente, pero no autoriza a usar $P(n+1)$ ni casos futuros.

Comprobación. Quieres demostrar una propiedad $P(n)$ para todo

$$n \ge 4.$$

Has probado

$$P(4), \qquad P(5)$$

y has demostrado que, para todo

$$n \ge 5,$$

si $P(k)$ es verdadera para cada k con

$$4 \leq k \leq n,$$

entonces $P(n+1)$ es verdadera.

Un compañero afirma que también debes demostrar $P(6)$ como caso base antes de aplicar inducción fuerte.

¿Tiene razón?

Respuesta razonada.

No.

Cuando tomamos

$$n = 5,$$

la hipótesis fuerte disponible contiene exactamente

$$P(4) \quad \text{y} \quad P(5),$$

que ya han sido demostradas.

El paso con $n = 5$ produce entonces

$$P(6).$$

Después, el paso con $n = 6$ puede utilizar

$$P(4), P(5), P(6)$$

para obtener $P(7)$, y el mecanismo continúa.

Por tanto, los dos casos iniciales bastan.

Lo que debemos comprobar no es una cantidad prefijada de bases, sino que **el primer paso disponible pueda activarse con los casos iniciales ya establecidos**.

Hacia el primer fallo. La inducción ordinaria y la fuerte razonan hacia adelante: parten de casos establecidos y controlan cómo avanzar.

Existe otra manera de explotar la estructura de los naturales.

Podemos preguntar:

> si una afirmación universal fallara, ¿habría un primer índice donde falla?

Esta pregunta conduce al **principio de buen orden** y al método del menor contraejemplo.

Ése será el contenido de §8.6.

8.6. Buen orden y contraejemplo mínimo

Hasta ahora hemos explotado la estructura de los naturales mirando **hacia adelante**.

En la inducción ordinaria partimos de un caso inicial y demostramos que la propiedad pasa de un índice al siguiente. En la inducción fuerte disponemos, durante el paso, de todas las instancias anteriores pertinentes.

El principio de buen orden permite mirar la misma estructura desde otra dirección.

En lugar de preguntar

> «¿cómo avanzo desde los casos que ya conozco?»,

podemos preguntar

> **«si la afirmación fallara, ¿dónde aparecería el primer fallo?»**

Esa pregunta sólo tiene fuerza porque estamos trabajando sobre los números naturales.

El principio de buen orden. Diremos que un elemento

$$m \in A$$

es un **mínimo** de un conjunto ordenado A si

$$m \leq a$$

para todo

$$a \in A.$$

El principio de buen orden de los naturales afirma:

Principio de buen orden

Todo subconjunto no vacío de $\mathbb{N}$ tiene un elemento mínimo.
Es decir, si

$$A \subseteq \mathbb{N}$$

y

$$A \neq \varnothing,$$

entonces existe

$$m \in A$$

tal que

$$m \leq a$$

para todo $a \in A$.

La condición

$$A \neq \varnothing$$

es esencial. El conjunto vacío no posee ningún elemento que pueda ser su mínimo.

También es esencial que estemos hablando de subconjuntos de los naturales. Por ejemplo, el conjunto real

$$(0, 1)$$

no tiene mínimo: dado cualquier

$$x \in (0, 1),$$

también

$$\frac{x}{2} \in (0, 1)$$

y

$$\frac{x}{2} < x.$$

No existe, por tanto, un "primer" número real positivo del intervalo.

En $\mathbb{N}$, en cambio, todo subconjunto no vacío tiene un primer elemento en el sentido del orden.

Convertir una falla en un conjunto. Supongamos que queremos demostrar

$$\forall n \geq n_0, \quad P(n).$$

Si esta afirmación fuese falsa, entonces existiría al menos un índice del tramo para el cual P falla.

Podemos reunir todos esos índices en el conjunto

$$F = \{n \in \mathbb{N} : n \geq n_0 \text{ y } \neg P(n)\}.$$

Llamaremos a F el **conjunto de índices de fallo**.

La negación de la afirmación universal

$$\forall n \geq n_0, \quad P(n)$$

se traduce precisamente en

$$F \neq \varnothing.$$

Y si F no es vacío, el buen orden garantiza que existe

$$m = \min F.$$

Este m es el **menor contraejemplo** a la afirmación universal.

No significa que hayamos buscado computacionalmente todos los casos hasta encontrarlo.

El argumento es estructural:

$$\boxed{\text{la propiedad falla en algún índice}}$$

$$\Downarrow$$

$$\boxed{F \neq \varnothing}$$

$$\Downarrow$$

$$\boxed{m = \min F}.$$

Qué nos entrega la minimalidad. Elegir el menor contraejemplo proporciona mucha más información que elegir simplemente algún índice de fallo.

Si

$$m = \min F,$$

entonces

$$\neg P(m).$$

Pero además, para todo índice del dominio que sea estrictamente menor que m, la propiedad **no puede fallar**.

En símbolos, si

$$n_0 \leq k < m,$$

entonces

$$k \notin F,$$

y por tanto

$$P(k).$$

Así, la minimalidad convierte automáticamente todos los índices anteriores en casos buenos:

$$\boxed{n_0 \leq k < m \quad \Longrightarrow \quad P(k)}.$$

Ésta es la información decisiva del método.

Comparemos.

Un contraejemplo cualquiera sólo nos dice

$$\neg P(m).$$

El **menor** contraejemplo nos dice simultáneamente

$$\neg P(m)$$

y

$$P(n_0), P(n_0+1), \ldots, P(m-1).$$

El objetivo es utilizar esas instancias anteriores para demostrar que $P(m)$ también tendría que ser verdadera.

Si lo conseguimos, obtenemos una contradicción.

La base reaparece como exclusión del primer índice. El índice inicial sigue importando.

Supongamos que ya hemos demostrado

$$P(n_0).$$

Si m es un contraejemplo, entonces necesariamente

$$m \neq n_0.$$

Como además

$$m \geq n_0,$$

obtenemos

$$m > n_0.$$

Por tanto,

$$m - 1 \geq n_0.$$

Esto permite mirar el índice inmediatamente anterior sin salir del dominio del teorema.

Además, por minimalidad de m,

$$P(m-1)$$

debe ser verdadera.

Aquí reaparece, desde otra perspectiva, la misma continuidad que vimos en §§8.2–8.3:

- la base impide que el primer índice sea un fallo;
- la minimalidad garantiza que todos los índices anteriores al primer fallo son buenos;
- una regla que lleve de esos casos anteriores al índice m destruye la posibilidad de que m sea realmente un contraejemplo.

Un ejemplo completo. Demostremos, mediante menor contraejemplo, que para todo

$$n \geq 1,$$

se cumple

$$1 + 3 + 5 + \cdots + (2n - 1) = n^2.$$

Definamos

$$P(n): \quad 1 + 3 + 5 + \cdots + (2n - 1) = n^2.$$

Supongamos, buscando una contradicción, que la afirmación universal es falsa.
Entonces el conjunto

$$F = \{n \in \mathbb{N} : n \geq 1 \text{ y } \neg P(n)\}$$

es no vacío.
Por buen orden, existe

$$m = \min F.$$

Así,

$$\neg P(m).$$

Pero

$$P(1)$$

es verdadera, porque

$$1 = 1^2.$$

Por tanto,

$$m \neq 1,$$

y como $m \geq 1$,

$$m \geq 2.$$

En consecuencia,

$$m - 1 \geq 1.$$

Como $m - 1 < m$ y m es el menor elemento de F, el índice $m - 1$ no pertenece a F.
Por tanto,

$$P(m - 1)$$

es verdadera:

$$1 + 3 + \cdots + \big(2(m - 1) - 1\big) = (m - 1)^2.$$

Para obtener la suma correspondiente a m, añadimos el siguiente número impar:

$$\begin{aligned} 1+3+\cdots+\big(2(m-1)-1\big)+(2m-1) &= (m-1)^2+(2m-1) \\ &= m^2. \end{aligned}$$

Así,

$$P(m)$$

es verdadera.

Pero habíamos elegido

$$m \in F,$$

lo que significa

$$\neg P(m).$$

Hemos llegado a una contradicción.

Por tanto,

$$F = \varnothing,$$

y concluimos

$$\boxed{1+3+5+\cdots+(2n-1) = n^2 \text{ para todo } n \geq 1}.$$

La contradicción se concentra en un solo índice. La estructura del argumento anterior merece aislarse.

Partimos de una afirmación universal que queremos demostrar:

$$\forall n \geq n_0, \quad P(n).$$

Negamos provisionalmente esa afirmación y construimos

$$F = \{n \geq n_0 : \neg P(n)\}.$$

Si $F \neq \varnothing$, tomamos

$$m = \min F.$$

Entonces disponemos simultáneamente de:

$$\neg P(m)$$

y

$$P(k) \quad \text{para todo } n_0 \leq k < m.$$

La prueba debe encontrar ahora una razón independiente por la cual esos casos anteriores obliguen a que

$$P(m)$$

sea verdadera.

El esquema es:

$$\boxed{m \text{ es el primer fallo}}$$

$$\Downarrow$$

$$\boxed{\text{todo índice anterior pertinente es bueno}}$$

$$\Downarrow$$

$$\boxed{P(m) \text{ debe ser verdadera}}$$

$$\Downarrow$$

$$\boxed{\text{contradicción}}.$$

Una afirmación sobre infinitos índices queda así concentrada en el análisis de un único objeto mínimo.

Menor contraejemplo no significa "probar casos hasta que algo salga mal". La expresión **menor contraejemplo** puede sugerir un procedimiento de búsqueda:

$$n_0, n_0 + 1, n_0 + 2, \ldots$$

hasta encontrar un fallo.

Ésa no es la estructura lógica del método.

En una prueba no necesitamos saber cuál es el valor numérico de m.

Suponemos que existe algún fallo; el buen orden nos garantiza que, entre todos los fallos, hay uno mínimo. Trabajamos entonces con ese m abstracto y utilizamos únicamente las consecuencias de su minimalidad.

Por eso el método no depende de haber calculado muchos casos favorables.

Tampoco demuestra que una afirmación sea falsa. Al contrario: normalmente **supone provisionalmente que hay un contraejemplo** para mostrar que tal contraejemplo no puede existir.

La búsqueda sistemática de afirmaciones falsas y de contraejemplos reales pertenece a C09. Aquí el contraejemplo mínimo es una herramienta para explotar el buen orden de $\mathbb{N}$.

Cuándo el método se atasca. Elegir un menor contraejemplo no produce por sí solo la contradicción.

De

$$m = \min F$$

obtenemos información sobre los índices anteriores, pero todavía necesitamos un mecanismo matemático que conecte esa información con $P(m)$.

Si sabemos únicamente que

$$P(k)$$

es verdadera para todo $k < m$, pero no existe ninguna razón disponible que permita deducir $P(m)$, el argumento se detiene.

El buen orden localiza el punto crítico; no inventa el puente que falta.

Esto refleja una distinción que ya apareció en C07:

$$\boxed{\text{arquitectura de prueba}} \neq \boxed{\text{contenido matemático que cierra la prueba}}.$$

El método del menor contraejemplo decide **dónde** debe concentrarse el argumento. Todavía debemos demostrar **por qué** ese primer fallo sería imposible.

Qué aporta el menor contraejemplo

Si

$$F = \{n \geq n_0 : \neg P(n)\}$$

es no vacío y

$$m = \min F,$$

entonces sabemos dos cosas:

$$\neg P(m)$$

y

$$P(k) \quad \text{para todo } n_0 \leq k < m.$$

La prueba debe utilizar la segunda información para contradecir la primera.

El buen orden garantiza el mínimo; no proporciona automáticamente el argumento que hace imposible ese mínimo.

Comprobación. Quieres demostrar

$$\forall n \geq 4, \quad P(n)$$

mediante menor contraejemplo.

Supón que el conjunto

$$F = \{n \in \mathbb{N} : n \geq 4 \text{ y } \neg P(n)\}$$

es no vacío y sea

$$m = \min F.$$

Sabes además que $P(4)$ es verdadera.

Explica qué puedes concluir sobre m y sobre todos los índices k con

$$4 \leq k < m.$$

¿Qué información adicional necesitarías para obtener una contradicción?

Respuesta razonada.

Como

$$P(4)$$

es verdadera, el índice 4 no pertenece a F.

Por tanto, el mínimo de F no puede ser 4:

$$m > 4.$$

En particular,

$$m - 1 \geq 4.$$

Además, si

$$4 \leq k < m,$$

entonces k no puede pertenecer a F, porque eso contradiciría la minimalidad de m.

Así,

$$P(k)$$

es verdadera para todo

$$4 \leq k < m.$$

En particular disponemos de

$$P(m - 1).$$

Sin embargo, todavía no hay contradicción. Para obtenerla necesitamos una regla, lema o argumento matemático que, a partir de las instancias anteriores disponibles —por ejemplo de $P(m - 1)$, o quizá de varias $P(k)$ con $k < m$—, permita demostrar

$$P(m).$$

Eso contradirá

$$\neg P(m),$$

que viene de $m \in F$.

Tres lenguajes para una misma estructura. Ya hemos conocido tres maneras de razonar sobre una propiedad indexada por los naturales.

La inducción ordinaria pregunta:

$$P(n) \Rightarrow P(n + 1)?$$

La inducción fuerte pregunta:

$$P(n_0), \ldots, P(n) \Rightarrow P(n + 1)?$$

El método del menor contraejemplo pregunta:

si m fuera el primer fallo, ¿cómo podrían los casos anteriores forzar $P(m)$?

Las tres formas utilizan la misma estructura del tramo de naturales, pero organizan la dependencia de maneras distintas.

La siguiente sección hará explícita esa relación y demostrará, en el nivel fundacional apropiado, por qué inducción ordinaria, inducción fuerte y buen orden no son principios desconectados.

8.7. Cómo se relacionan los tres principios

En las tres secciones anteriores hemos presentado tres formas de organizar argumentos sobre los números naturales:

- la **inducción ordinaria**;
- la **inducción fuerte**;
- el **principio de buen orden**, utilizado mediante menor contraejemplo.

Podrían parecer tres recursos independientes.

No lo son.

En el marco elemental de este capítulo, y trabajando con la lógica clásica que hemos utilizado desde C02 y C06, estos principios pueden derivarse unos de otros. No son la misma afirmación escrita con símbolos distintos, pero expresan aspectos equivalentes de la estructura sucesiva y ordenada de $\mathbb{N}$.

Conviene demostrar esta relación porque evita una impresión engañosa: que al pasar de inducción ordinaria a fuerte, o de inducción a menor contraejemplo, estuviéramos incorporando una fuente adicional de verdad.

No ocurre eso.

Lo que cambia es la **organización de la dependencia** dentro de la prueba.

Qué significa aquí que dos principios sean equivalentes. Diremos que dos principios son equivalentes, en el nivel de este capítulo, si podemos demostrar cada uno suponiendo el otro.

Por ejemplo, para justificar

$$\text{inducción ordinaria} \iff \text{inducción fuerte},$$

debemos mostrar dos direcciones:

$$\text{inducción ordinaria} \implies \text{inducción fuerte},$$

y

$$\text{inducción fuerte} \implies \text{inducción ordinaria}.$$

Del mismo modo, relacionaremos inducción y buen orden demostrando

$$\text{buen orden} \implies \text{inducción ordinaria}$$

y

$$\text{inducción ordinaria} \implies \text{buen orden}.$$

Para mantener limpia la notación utilizaremos en las demostraciones la convención

$$\mathbb{N} = \{0, 1, 2, \ldots\}.$$

Las versiones que comienzan en un índice n_0 se recuperan mediante la reindexación estudiada en §8.3.

De inducción ordinaria a inducción fuerte. Supongamos que podemos utilizar el principio de inducción ordinaria.

Queremos deducir de él la inducción fuerte.

Sea $P(n)$ una propiedad de los naturales. Supongamos que hemos demostrado:

$$P(0),$$

y que, para todo $n \in \mathbb{N}$, si

$$P(0), P(1), \ldots, P(n)$$

son todas verdaderas, entonces

$$P(n+1)$$

también es verdadera.

Ésta es exactamente la información que exige una prueba por inducción fuerte.

La dificultad es que la inducción ordinaria sólo permite transportar una propiedad desde n hasta $n+1$.

La solución consiste en **empaquetar toda la historia previa dentro de una sola propiedad**.

Definamos

$$Q(n): \quad P(0), P(1), \ldots, P(n) \text{ son todas verdaderas.}$$

Ahora aplicaremos inducción ordinaria a Q.

Base para Q. La afirmación

$$Q(0)$$

significa simplemente

$$P(0),$$

que ya conocemos.

Paso para Q. Supongamos

$$Q(n).$$

Entonces sabemos simultáneamente

$$P(0), P(1), \ldots, P(n).$$

Por la hipótesis de paso de la inducción fuerte obtenemos

$$P(n+1).$$

Pero ahora conocemos

$$P(0), P(1), \ldots, P(n), P(n+1),$$

que es exactamente

$$Q(n+1).$$

Por inducción ordinaria,

$$Q(n)$$

es verdadera para todo $n \in \mathbb{N}$.

En particular, como $Q(n)$ contiene $P(n)$,

$$P(n)$$

es verdadera para todo $n \in \mathbb{N}$.

Hemos derivado la inducción fuerte a partir de la ordinaria.

La idea conceptual puede resumirse así:

$$\boxed{\text{historia acumulada}} \quad \longrightarrow \quad \boxed{\text{una sola propiedad } Q(n)}.$$

La inducción ordinaria puede transportar un paquete que contiene toda la información previa.

De inducción fuerte a inducción ordinaria. La dirección contraria es más inmediata.

Supongamos ahora que disponemos del principio de inducción fuerte y queremos demostrar la validez de la inducción ordinaria.

Sean dadas las hipótesis ordinarias:

$$P(0)$$

y

$$\forall n \in \mathbb{N}, \qquad P(n) \Rightarrow P(n+1).$$

Para aplicar inducción fuerte, fijemos un $n \in \mathbb{N}$ y supongamos que

$$P(0), P(1), \ldots, P(n)$$

son verdaderas.

Entre todas esas afirmaciones está, en particular,

$$P(n).$$

Entonces el paso ordinario nos da

$$P(n+1).$$

Así hemos verificado el paso fuerte.

Por inducción fuerte concluimos

$$\forall n \in \mathbb{N}, \quad P(n).$$

Por tanto,

$$\boxed{\text{inducción ordinaria} \iff \text{inducción fuerte}}.$$

La forma fuerte puede ser más cómoda cuando el caso siguiente depende de índices anteriores no inmediatos, pero no demuestra una clase misteriosamente mayor de afirmaciones.

Del buen orden a la inducción ordinaria. Supongamos ahora el principio de buen orden:

todo subconjunto no vacío de $\mathbb{N}$ tiene mínimo.

Queremos recuperar el principio de inducción ordinaria.

Sea $P(n)$ una propiedad tal que

$$P(0)$$

y

$$\forall n \in \mathbb{N}, \qquad P(n) \Rightarrow P(n+1).$$

Debemos demostrar

$$\forall n \in \mathbb{N}, \quad P(n).$$

Supongamos, buscando una contradicción, que esto es falso.

Entonces el conjunto de índices de fallo

$$F = \{n \in \mathbb{N} : \neg P(n)\}$$

es no vacío.

Por buen orden, existe

$$m = \min F.$$

Como $P(0)$ es verdadera,

$$m \neq 0.$$

Por tanto,

$$m \geq 1,$$

de modo que

$$m - 1 \in \mathbb{N}.$$

Además,

$$m - 1 < m.$$

Por minimalidad de m, el índice $m - 1$ no pertenece a F. Luego

$$P(m - 1)$$

es verdadera.

Aplicando el paso inductivo con

$$n = m - 1,$$

obtenemos

$$P(m).$$

Pero $m \in F$, por lo que también tenemos

$$\neg P(m).$$

Contradicción.

Por tanto F debe ser vacío, y concluimos

$$\forall n \in \mathbb{N}, \quad P(n).$$

Así, el principio de buen orden implica la inducción ordinaria.

Esta demostración hace visible una idea importante: **la inducción puede verse como la imposibilidad de un primer fallo**.

De inducción ordinaria al buen orden. Falta la dirección inversa.
Supongamos ahora que disponemos del principio de inducción ordinaria.
Sea

$$A \subseteq \mathbb{N}$$

un conjunto no vacío.
Queremos demostrar que A tiene mínimo.
Supongamos, buscando una contradicción, que A **no** tiene mínimo.
Definamos la propiedad

$$Q(n): \quad A \cap \{0, 1, \ldots, n\} = \varnothing.$$

En palabras, $Q(n)$ afirma que ninguno de los naturales desde 0 hasta n pertenece a A.
Base. Si

$$0 \in A,$$

entonces 0 sería automáticamente el mínimo de A.
Eso contradice nuestra suposición de que A no tiene mínimo.
Por tanto,

$$0 \notin A,$$

y así

$$Q(0)$$

es verdadera.
Paso inductivo. Supongamos

$$Q(n).$$

Entonces ningún natural

$$0, 1, \ldots, n$$

pertenece a A.
Si ocurriera que

$$n + 1 \in A,$$

entonces no habría ningún elemento de A menor que $n + 1$.
Por tanto, $n + 1$ sería el mínimo de A.
Eso contradice otra vez la suposición de que A no tiene mínimo.
Luego

$$n + 1 \notin A.$$

Junto con $Q(n)$, esto implica

$$Q(n + 1).$$

Por inducción ordinaria,

$$Q(n)$$

es verdadera para todo $n \in \mathbb{N}$.

Pero A es no vacío. Podemos elegir

$$a \in A.$$

La afirmación $Q(a)$ dice, en particular, que

$$a \notin A,$$

contradicción.

Por tanto nuestra suposición era imposible: todo subconjunto no vacío de $\mathbb{N}$ tiene mínimo.

Hemos obtenido

$$\boxed{\text{inducción ordinaria} \iff \text{buen orden}}.$$

Las tres formas quedan conectadas. Ya tenemos las relaciones necesarias:

$$\text{inducción ordinaria} \iff \text{inducción fuerte}$$

y

$$\text{inducción ordinaria} \iff \text{buen orden}.$$

Por tanto, en el marco de este capítulo,

$$\boxed{\text{inducción ordinaria} \iff \text{inducción fuerte} \iff \text{buen orden}}.$$

Esto no significa que las tres demostraciones tengan la misma forma práctica.

Cada una organiza la información de manera distinta.

Podemos resumirlas así:

$$\begin{aligned} \text{inducción ordinaria} &: P(n) \Rightarrow P(n+1), \\ \text{inducción fuerte} &: P(0), \ldots, P(n) \Rightarrow P(n+1), \\ \text{menor contraejemplo} &: P(0), \ldots, P(m-1) \text{ y } \neg P(m) \Rightarrow \text{ contradicción}. \end{aligned}$$

La estructura subyacente es la misma: ningún índice puede aparecer como un punto aislado que rompa para siempre la cadena de los naturales.

Inducción fuerte y menor contraejemplo son especialmente cercanos. La relación entre estas dos formas merece verse de manera directa.

En un paso de inducción fuerte para demostrar $P(n+1)$ podemos utilizar

$$P(0), P(1), \ldots, P(n).$$

En una prueba por menor contraejemplo, si

$$m$$

fuera el primer fallo, la minimalidad nos concede precisamente

$$P(0), P(1), \ldots, P(m-1).$$

Entonces intentamos demostrar

$$P(m).$$

Si escribimos

$$n = m - 1,$$

la información disponible es formalmente la misma que en un paso fuerte.

La diferencia es narrativa y lógica:

- la inducción fuerte formula directamente una regla de avance;
- el menor contraejemplo supone provisionalmente un primer fallo y demuestra que los casos anteriores lo hacen imposible.

Una mira hacia adelante.

La otra concentra la atención en el supuesto primer punto de ruptura.

Equivalentes no significa intercambiables en toda prueba. Que los principios sean equivalentes no significa que toda demostración deba escribirse en cualquiera de las tres formas con la misma comodidad.

Una propiedad con dependencia inmediata puede quedar transparente mediante inducción ordinaria.

Una propiedad cuyo caso siguiente depende de varios índices anteriores puede expresarse naturalmente mediante inducción fuerte.

Una afirmación cuya negación produce un objeto mínimo con buenas propiedades puede volverse especialmente clara mediante buen orden o menor contraejemplo.

La equivalencia nos dice que no cambia la fuerza lógica disponible en este nivel.

La elección de la forma cambia la **geometría de la prueba**.

Esta distinción enlaza con C07: dos arquitecturas pueden ser lógicamente suficientes, pero una de ellas puede hacer visible el puente matemático con mucha mayor claridad.

Una precisión fundacional. En esta sección hemos demostrado equivalencias **dentro del marco elemental adoptado por el libro**: los números naturales, su orden y la lógica clásica ordinaria ya forman parte del contexto.

No estamos afirmando que las expresiones “inducción”, “inducción fuerte” y “buen orden” sean sintácticamente idénticas, ni que ocupen el mismo lugar primitivo en toda axiomatización posible de los naturales.

Una teoría fundacional puede tomar uno de estos principios como básico y derivar otros.

Lo que hemos establecido aquí es exactamente lo necesario para la práctica matemática del libro:

> una vez aceptada cualquiera de estas formulaciones en el marco considerado,
> las otras pueden recuperarse mediante argumentos explícitos.

Tres formulaciones, una estructura

En el nivel de C08:

$$\boxed{\text{inducción ordinaria} \iff \text{inducción fuerte} \iff \text{buen orden}}.$$

La equivalencia no elimina las diferencias de uso.

- La inducción ordinaria transporta una instancia.
- La inducción fuerte transporta un bloque de instancias anteriores.
- El buen orden convierte un supuesto fallo en un primer fallo.

Las tres explotan la estructura de $\mathbb{N}$ desde perspectivas distintas.

Comprobación. Supón que sólo aceptas como principio la inducción ordinaria y quieres demostrar que todo subconjunto no vacío de $\mathbb{N}$ tiene mínimo.

Sea

$$A \subseteq \mathbb{N}, \qquad A \neq \varnothing,$$

y supón, buscando una contradicción, que A no tiene mínimo.

¿Por qué es útil definir

$$Q(n): \quad A \cap \{0, 1, \ldots, n\} = \varnothing?$$

Explica la base, el paso inductivo y la contradicción final.

Respuesta razonada.

La propiedad $Q(n)$ traduce la ausencia de un mínimo en una afirmación que puede propagarse por inducción.

Para la base, si

$$0 \in A,$$

entonces 0 sería el mínimo de A. Como hemos supuesto que A no tiene mínimo,

$$0 \notin A,$$

y por tanto $Q(0)$ es verdadera.

Para el paso, supongamos $Q(n)$. Entonces ninguno de los números

$$0, 1, \ldots, n$$

pertenece a A.

Si

$$n + 1 \in A,$$

entonces $n + 1$ sería menor o igual que cualquier elemento de A, porque no existe ningún elemento de A entre 0 y n.

Así, $n + 1$ sería el mínimo de A, contradicción.

Por tanto,

$$n + 1 \notin A,$$

y obtenemos $Q(n + 1)$.

La inducción ordinaria da entonces $Q(n)$ para todo $n \in \mathbb{N}$.

Pero como A es no vacío, existe algún

$$a \in A.$$

La afirmación $Q(a)$ implica que $a \notin A$, contradicción.

Luego nuestra suposición era falsa y A debe tener mínimo.

Lo que cambia a partir de aquí. Hasta §8.7 hemos estudiado principios que sirven para **demostrar** propiedades de los naturales.

El siguiente paso cambia de tarea.

Querremos construir objetos indexados por $\mathbb{N}$ mediante reglas como

$$a_0 = c, \qquad a_{n+1} = F(a_n).$$

Ahora la pregunta ya no será

«¿cómo pruebo que una propiedad vale para todos los índices?»

sino

«¿cómo queda definido legítimamente un objeto en cada índice?»

Ése es el problema de la **recursión**, y será el contenido de §8.8.

8.8. Definir por recursión

Hasta aquí hemos utilizado la estructura de los naturales para **demostrar** afirmaciones. La inducción ordinaria, la inducción fuerte y el buen orden nos ayudaban a justificar que una propiedad debía valer en todos los índices pertinentes.

La recursión cambia la tarea.

Ahora no partimos de una propiedad $P(n)$ que queremos demostrar. Queremos **definir** un objeto en cada índice natural.

La forma más simple es ésta:

$$a_0 = c,$$

y, para cada $n \in \mathbb{N}$,

$$a_{n+1} = F(a_n).$$

El primer dato nos dice dónde comienza la construcción. La segunda regla nos dice cómo obtener un valor nuevo a partir del anterior.

La semejanza con una prueba inductiva es visible, pero la función lógica es distinta:

$$\boxed{P(n) \Rightarrow P(n+1)}$$

relaciona **proposiciones** dentro de una demostración, mientras que

$$\boxed{a_{n+1} = F(a_n)}$$

relaciona **valores** dentro de una definición.

Por eso conviene conservar la distinción rectora del capítulo:

$$\boxed{\text{inducción} = \text{demostrar}} \qquad \boxed{\text{recursión} = \text{definir}}.$$

Una definición recursiva necesita un punto de partida. Consideremos solamente la regla

$$a_{n+1} = a_n + 3.$$

Esta ecuación no determina por sí sola una única sucesión.

Si elegimos

$$a_0 = 0,$$

obtenemos

$$0, 3, 6, 9, \ldots$$

Pero si elegimos

$$a_0 = 5,$$

obtenemos

$$5, 8, 11, 14, \ldots$$

La misma regla de transición es compatible con muchas sucesiones distintas.

Por tanto, una definición recursiva elemental necesita al menos dos componentes:

1. **dato inicial**;
2. **regla de generación**.

En la forma

$$a_0 = c, \qquad a_{n+1} = F(a_n),$$

el dato inicial fija el primer valor y la regla determina cómo continuar.

La regla debe producir valores legítimos. Tampoco basta escribir cualquier expresión que contenga a_n.

Supongamos que pretendemos definir una función

$$a : \mathbb{N} \to \mathbb{R}$$

mediante

$$a_0 = 1, \qquad a_{n+1} = \frac{1}{a_n - 1}.$$

Al intentar calcular el siguiente valor obtenemos

$$a_1 = \frac{1}{1-1},$$

que no es un número real definido.

La construcción se detiene inmediatamente.

El problema no es una dificultad de cálculo posterior. La propia regla no produce un valor legítimo a partir del dato inicial elegido.

Por eso, si queremos construir valores dentro de un conjunto X, la regla debe estar formulada de manera que, cada vez que recibe datos permitidos, produzca nuevamente un elemento de X.

En la forma más simple esto significa disponer de una función

$$F : X \to X.$$

Si

$$a_n \in X,$$

entonces automáticamente

$$F(a_n) \in X,$$

y el siguiente valor

$$a_{n+1} = F(a_n)$$

permanece en el dominio declarado de la construcción.

El teorema elemental de recursión. La intuición de "seguir aplicando la regla para siempre" necesita una justificación matemática.

En este capítulo utilizaremos la siguiente forma elemental del **teorema de recursión**.

Teorema elemental de recursión

Sea X un conjunto no vacío, sea

$$c \in X,$$

y sea

$$G : \mathbb{N} \times X \to X$$

una función.
Entonces existe una única función

$$a : \mathbb{N} \to X$$

tal que

$$a(0) = c$$

y, para todo $n \in \mathbb{N}$,

$$a(n+1) = G(n, a(n)).$$

La versión

$$a_{n+1} = F(a_n)$$

se obtiene como caso particular tomando

$$G(n, x) = F(x).$$

La forma con $G(n, x)$ es útil porque muchas reglas dependen tanto del índice actual como del valor ya construido.

El teorema afirma dos cosas distintas:

$$\boxed{\text{existencia}} \qquad \text{y} \qquad \boxed{\text{unicidad}}.$$

Existencia: hay efectivamente una función definida en todos los naturales que satisface los datos recursivos.

Unicidad: no hay dos funciones diferentes que satisfagan exactamente el mismo dato inicial y la misma regla.

En FND utilizaremos este teorema en esta forma elemental. Una construcción set-teórica completa de la existencia exigiría infraestructura adicional sobre familias de conjuntos y no será introducida aquí de manera encubierta; esa cuestión pertenece a una capa fundacional posterior. Sí podemos, en cambio, comprender y demostrar ahora por qué la **unicidad** es natural.

Por qué el objeto recursivo es único. Supongamos que dos funciones

$$a, b : \mathbb{N} \to X$$

satisfacen el mismo dato inicial

$$a(0) = c = b(0)$$

y la misma regla

$$a(n+1) = G(n, a(n)),$$

$$b(n+1) = G(n, b(n)).$$

Queremos demostrar que

$$a = b.$$

Por la igualdad de funciones estudiada en C04, basta demostrar

$$a(n) = b(n)$$

para todo $n \in \mathbb{N}$.
Definamos

$$P(n): \quad a(n) = b(n).$$

Base. Tenemos

$$a(0) = c = b(0),$$

así que $P(0)$ es verdadera.
Paso. Supongamos

$$a(n) = b(n).$$

Entonces, usando la misma regla G,

$$\begin{aligned} a(n+1) &= G(n, a(n)) \\ &= G(n, b(n)) \\ &= b(n+1). \end{aligned}$$

Por inducción,

$$a(n) = b(n)$$

para todo $n \in \mathbb{N}$.
Por tanto,

$$\boxed{a = b}.$$

Aquí aparece por primera vez una interacción importante:

- la **recursión** especifica qué función queremos construir;
- la **inducción** demuestra que dos funciones que satisfacen esa especificación no pueden diferir.

Un ejemplo propio: construir una sucesión. Definamos una sucesión

$$a : \mathbb{N} \to \mathbb{N}$$

mediante

$$a_0 = 2$$

y

$$a_{n+1} = a_n + 2n + 1.$$

La regla depende del índice y del valor anterior. Podemos escribirla mediante

$$G : \mathbb{N} \times \mathbb{N} \to \mathbb{N},$$

$$G(n, x) = x + 2n + 1.$$

Como

$$x + 2n + 1 \in \mathbb{N}$$

siempre que

$$n, x \in \mathbb{N},$$

la regla produce valores legítimos en el codominio declarado.

El teorema de recursión garantiza entonces una única función

$$a : \mathbb{N} \to \mathbb{N}$$

con esos datos.

Podemos calcular sus primeros valores:

$$a_0 = 2,$$

$$a_1 = a_0 + 1 = 3,$$

$$a_2 = a_1 + 3 = 6,$$

$$a_3 = a_2 + 5 = 11,$$

$$a_4 = a_3 + 7 = 18.$$

Así obtenemos

$$2, 3, 6, 11, 18, \ldots$$

Estos cálculos ilustran la definición, pero no la sustituyen.

La sucesión no está definida porque hayamos calculado cinco términos. Está definida por el dato inicial, la regla y el teorema que garantiza que ambos determinan una única función sobre todo $\mathbb{N}$.

Calcular términos no es demostrar una fórmula general. Al observar

$$2, 3, 6, 11, 18, \ldots$$

podríamos sospechar que existe una fórmula cerrada sencilla para a_n.

Esa conjetura pertenece a otra tarea.

La definición recursiva ya nos dice **qué es** la sucesión:

$$a_0 = 2, \qquad a_{n+1} = a_n + 2n + 1.$$

Si luego proponemos una fórmula explícita

$$a_n = H(n),$$

tendremos que **demostrar** que esa fórmula describe realmente los valores definidos recursivamente.

En otras palabras:

$$\boxed{\text{definir por recursión}} \neq \boxed{\text{demostrar una fórmula para el objeto definido}}.$$

La segunda tarea será el tema de §8.9.

Una regla sin dato inicial está incompleta. Consideremos

$$b_{n+1} = 2b_n + 1.$$

No hay todavía una sucesión única.

Por ejemplo, si

$$b_0 = 0,$$

obtenemos

$$0, 1, 3, 7, 15, \ldots$$

mientras que si

$$b_0 = 1,$$

obtenemos

$$1, 3, 7, 15, 31, \ldots$$

La regla explica cómo continuar **una vez conocido un valor inicial**, pero no elige ese valor por nosotros.

Por tanto, la especificación recursiva está incompleta mientras no fijemos el punto de partida.

Un dato inicial sin regla también está incompleto. De manera simétrica, escribir solamente

$$c_0 = 4$$

no define una función sobre $\mathbb{N}$.

Existen infinitas sucesiones que empiezan en 4.

Por ejemplo,

$$4, 4, 4, 4, \ldots$$

y

$$4, 5, 6, 7, \ldots$$

comparten el mismo valor inicial.

El dato inicial determina dónde empieza la construcción; la regla determina cómo continúa.

Recursión que comienza en otro índice. También una definición recursiva puede empezar en un índice distinto de 0.

Si queremos definir valores sólo para

$$n \geq n_0,$$

podemos especificar

$$a_{n_0} = c$$

y una regla

$$a_{n+1} = G(n, a_n)$$

para

$$n \geq n_0.$$

Como ocurrió con la inducción desplazada, esto no requiere un principio nuevo. Podemos reindexar mediante

$$b_k = a_{n_0+k}, \qquad k \in \mathbb{N},$$

y convertir la construcción en una recursión que comienza en $k = 0$.

El índice inicial forma parte, otra vez, de la especificación matemática del objeto.

La recursión puede necesitar más memoria. La forma

$$a_{n+1} = G(n, a_n)$$

usa sólo el valor inmediatamente anterior.

Pero también aparecen definiciones donde el siguiente término depende de varios términos previos. Por ejemplo, una regla podría tener la forma

$$a_{n+2} = H(n, a_n, a_{n+1}).$$

Entonces harán falta suficientes datos iniciales para que la primera aplicación de la regla sea legítima.

El principio conceptual sigue siendo el mismo:

1. proporcionar todos los datos iniciales necesarios;
2. dar una regla determinista;
3. asegurar que la regla produce valores dentro del dominio declarado;
4. justificar que la especificación determina un único objeto.

No desarrollaremos aquí una teoría general de recurrencias de orden superior. En C08 sólo necesitamos reconocer la arquitectura recursiva y controlar que esté bien formada.

Qué debe contener una definición recursiva

Para construir un objeto indexado por los naturales debemos controlar:

1. **dominio de índices**;
2. **codominio de los valores**;
3. **dato o datos iniciales**;
4. **regla que produce el valor siguiente**;
5. **legitimidad de la regla dentro del codominio**;
6. **existencia y unicidad** del objeto determinado por esos datos.

Una relación recursiva no es todavía una definición completa si alguno de estos componentes necesarios queda indeterminado.

Comprobación. Se pretende definir una sucesión

$$c : \mathbb{N} \to \mathbb{R}$$

mediante

$$c_0 = 2, \qquad c_{n+1} = \frac{1}{c_n - 2}.$$

¿Esta especificación produce legítimamente una función $c : \mathbb{N} \to \mathbb{R}$?

Respuesta razonada.

No.

El dato inicial pertenece al codominio declarado:

$$c_0 = 2 \in \mathbb{R}.$$

Pero al intentar aplicar la regla por primera vez obtenemos

$$c_1 = \frac{1}{c_0 - 2} = \frac{1}{0},$$

que no es un número real definido.

Por tanto, la expresión

$$G(n, x) = \frac{1}{x - 2}$$

no determina una función

$$G : \mathbb{N} \times \mathbb{R} \to \mathbb{R}$$

sobre todos los pares permitidos: falla, en particular, cuando $x = 2$.

No podemos invocar el teorema elemental de recursión con ese codominio y ese dato inicial.

Para reparar la definición habría que modificar el dato inicial, la regla o el dominio de valores de modo que cada aplicación requerida sea legítima; además habría que comprobar que la construcción nunca sale del dominio elegido.

Definir primero, demostrar después. Ya disponemos de la sucesión

$$a_0 = 2, \qquad a_{n+1} = a_n + 2n + 1.$$

La recursión nos garantiza qué objeto estamos estudiando.

Ahora podemos formular preguntas sobre él:

- ¿existe una expresión cerrada para a_n?;
- ¿todos sus valores satisfacen cierta desigualdad?;
- ¿qué paridad tienen?;
- ¿cómo crecen?

Esas preguntas ya no son definiciones. Son **afirmaciones acerca de un objeto definido**.

La herramienta natural para muchas de ellas será precisamente la inducción.

§8.9 estudiará esta combinación:

$$\boxed{\text{recursión para definir}} \quad + \quad \boxed{\text{inducción para demostrar}}.$$

8.9. Probar propiedades de objetos recursivos

En §8.8 construimos, mediante recursión, una sucesión concreta:

$$a_0 = 2, \qquad a_{n+1} = a_n + 2n + 1.$$

La definición ya resolvió una pregunta:

¿qué valor tiene la sucesión en cada índice natural?

Pero dejó abiertas muchas otras.

Podemos preguntar, por ejemplo, si existe una fórmula explícita para a_n, si la sucesión es creciente, si sus términos satisfacen cierta paridad o si alguna desigualdad vale para todos los índices.

Estas preguntas no forman parte de la definición.

Son **propiedades del objeto ya definido**.

Aquí aparece la interacción central de esta sección:

$$\boxed{\text{recursión para definir}} \quad + \quad \boxed{\text{inducción para demostrar}}.$$

La proximidad entre ambas técnicas no debe borrar su diferencia lógica.

La ecuación recursiva

$$a_{n+1} = a_n + 2n + 1$$

es un hecho disponible porque forma parte de la definición de a.

Una afirmación como

$$a_n = n^2 + 2$$

no forma parte de esa definición. Si queremos utilizarla para todo n, debemos demostrarla.

Los primeros términos sugieren, pero no prueban. Calculemos de nuevo algunos valores:

$$a_0 = 2,$$

$$a_1 = 3,$$

$$a_2 = 6,$$

$$a_3 = 11,$$

$$a_4 = 18.$$

La sucesión comienza, por tanto,

$$2, 3, 6, 11, 18, \ldots$$

Si restamos 2 obtenemos

$$0, 1, 4, 9, 16, \ldots,$$

lo que sugiere la conjetura

$$\boxed{a_n = n^2 + 2}.$$

Esta observación es útil para el **descubrimiento** de la fórmula, pero todavía no la demuestra.

Comprobar que

$$a_0 = 0^2 + 2, \quad a_1 = 1^2 + 2, \quad \ldots, \quad a_4 = 4^2 + 2$$

sólo verifica cinco instancias.

El objetivo real es universal:

$$\forall n \in \mathbb{N}, \qquad a_n = n^2 + 2.$$

Necesitamos una razón que cubra todos los índices.

La definición recursiva proporciona el puente del paso. Definamos la propiedad

$$P(n): \quad a_n = n^2 + 2.$$

Demostraremos $P(n)$ para todo $n \in \mathbb{N}$ por inducción.

Caso base. Para $n = 0$, la definición recursiva da

$$a_0 = 2.$$

Por otra parte,

$$0^2 + 2 = 2.$$

Luego

$$a_0 = 0^2 + 2,$$

y $P(0)$ es verdadera.

Paso inductivo. Sea $n \in \mathbb{N}$ arbitrario y supongamos

$$P(n),$$

es decir,

$$a_n = n^2 + 2.$$

Debemos demostrar

$$P(n+1),$$

que significa

$$a_{n+1} = (n+1)^2 + 2.$$

Aquí entran simultáneamente dos tipos de información.

La **definición recursiva** nos dice

$$a_{n+1} = a_n + 2n + 1.$$

La **hipótesis inductiva** nos dice

$$a_n = n^2 + 2.$$

Sustituyendo la segunda información en la primera obtenemos

$$\begin{aligned} a_{n+1} &= a_n + 2n + 1 \\ &= (n^2 + 2) + 2n + 1 \\ &= n^2 + 2n + 3 \\ &= (n+1)^2 + 2. \end{aligned}$$

Así hemos demostrado

$$P(n) \Rightarrow P(n+1).$$

Por inducción,

$$\boxed{a_n = n^2 + 2 \text{ para todo } n \in \mathbb{N}}.$$

Qué hizo cada herramienta. La prueba anterior permite localizar con precisión la función de cada componente.

La recursión proporcionó

$$\boxed{a_{n+1} = a_n + 2n + 1}.$$

Esa igualdad relaciona **valores consecutivos del objeto definido**.

La inducción proporcionó la arquitectura para demostrar

$$\boxed{a_n = n^2 + 2 \text{ para todo } n}.$$

Dentro del paso, la hipótesis inductiva permitió reemplazar el valor anterior por la descripción explícita que queríamos propagar:

$$a_n \quad \rightsquigarrow \quad n^2 + 2.$$

Podemos resumir el mecanismo así:

$$\boxed{\text{regla recursiva para } a_{n+1}} \quad + \quad \boxed{\text{hipótesis inductiva sobre } a_n}$$

$$\Longrightarrow \boxed{\text{propiedad en } n+1}.$$

La definición y la prueba cooperan, pero no se identifican.

La hipótesis inductiva habla de una propiedad, no de la existencia del término. Conviene precisar un punto lógico.

En el paso anterior no supusimos que a_n **existiera**.

La existencia de todos los valores

$$a_0, a_1, a_2, \ldots$$

ya estaba garantizada por la definición recursiva legitimada en §8.8.

Lo que supusimos inductivamente fue una propiedad adicional del valor ya definido:

$$a_n = n^2 + 2.$$

Por eso las dos capas deben permanecer separadas:

$$\begin{aligned} \text{recursión} &: \text{construye } a_n, \\ \text{inducción} &: \text{demuestra algo acerca de } a_n. \end{aligned}$$

Confundirlas conduce a frases engañosas como

> "suponemos que a_n está definido y entonces definimos a_{n+1} por inducción".

En nuestro marco, no hacemos eso. La sucesión ya está definida globalmente por el teorema de recursión; la inducción entra después para demostrar propiedades.

Una segunda ruta: verificar la misma recursión. La fórmula

$$a_n = n^2 + 2$$

puede justificarse también explotando la **unicidad** de la definición recursiva.

Definamos explícitamente una función

$$b : \mathbb{N} \to \mathbb{N}, \qquad b(n) = n^2 + 2.$$

Comprobemos primero el dato inicial:

$$b(0) = 0^2 + 2 = 2.$$

Por tanto,

$$b(0) = a_0.$$

Ahora calculemos el valor siguiente de b:

$$\begin{aligned} b(n+1) &= (n+1)^2 + 2 \\ &= n^2 + 2n + 3 \\ &= (n^2 + 2) + 2n + 1 \\ &= b(n) + 2n + 1. \end{aligned}$$

Así, b satisface exactamente la misma especificación recursiva que a:

$$b(0) = 2, \qquad b(n+1) = b(n) + 2n + 1.$$

Pero §8.8 estableció que esa especificación determina una **única** función sobre $\mathbb{N}$.

Por tanto,

$$a = b.$$

En consecuencia,

$$a_n = b(n) = n^2 + 2$$

para todo $n \in \mathbb{N}$.

Esta segunda prueba no repite el paso inductivo sobre la fórmula. Utiliza un resultado estructural previo: la unicidad del objeto recursivo.

Sin embargo, la conexión con la inducción no ha desaparecido por completo: la unicidad del teorema de recursión fue demostrada en §8.8 precisamente mediante inducción.

Una fórmula candidata debe cumplir dos pruebas distintas. Cuando proponemos una expresión explícita

$$H(n)$$

para una sucesión definida recursivamente, hay dos preguntas diferentes.

La primera es experimental o heurística:

¿parece que $H(n)$ coincide con los primeros valores?

La segunda es matemática:

¿qué argumento demuestra que coincide con **todos** los valores?

Comprobar sólo

$$H(0) = a_0, \quad H(1) = a_1, \quad \ldots, \quad H(10) = a_{10}$$

no resuelve la segunda pregunta.

Dos caminos legítimos, cuando están disponibles, son:

1. demostrar por inducción
$$a_n = H(n);$$
2. demostrar que H satisface el mismo dato inicial y la misma regla recursiva, y usar unicidad.

Ambas estrategias convierten una conjetura obtenida de casos particulares en una afirmación universal justificada.

No toda propiedad de un objeto recursivo necesita inducción. La presencia de una definición recursiva tampoco obliga a utilizar inducción en toda afirmación posterior.

Para nuestra sucesión,

$$a_{n+1} = a_n + 2n + 1.$$

Como

$$2n + 1 > 0$$

para todo $n \in \mathbb{N}$, obtenemos directamente

$$a_{n+1} > a_n.$$

Por tanto, la sucesión es estrictamente creciente.

Aquí no fue necesario suponer una propiedad en n para demostrarla en $n + 1$. La regla recursiva ya contenía toda la información requerida para comparar términos consecutivos.

Éste es un criterio diagnóstico importante:

> **que un objeto esté definido recursivamente no implica que toda propiedad suya deba probarse por inducción.**

La arquitectura de la prueba debe responder a la propiedad concreta que queremos demostrar.

Cuándo la forma de la recursión sugiere la forma de la inducción. Si una definición utiliza sólo el término anterior,

$$a_{n+1} = G(n, a_n),$$

es frecuente que una propiedad de a_{n+1} pueda relacionarse con una propiedad de a_n mediante inducción ordinaria.

Pero si la definición tiene memoria mayor, por ejemplo

$$a_{n+2} = H(n, a_n, a_{n+1}),$$

entonces una propiedad del término nuevo puede depender de información sobre dos términos anteriores.

En tal situación podemos necesitar:

- varios casos iniciales;
- una propiedad inductiva que empaquete varias instancias;
- o una formulación mediante inducción fuerte.

La regla recursiva no determina mecánicamente qué forma de inducción debemos usar, pero sí revela las **dependencias** que una prueba tendrá que respetar.

Un error típico: usar la fórmula que todavía se está demostrando. Supongamos que queremos probar

$$a_n = n^2 + 2$$

para nuestra sucesión.

Un borrador incorrecto podría decir:

Como $a_n = n^2 + 2$, entonces

$$a_{n+1} = a_n + 2n + 1 = (n+1)^2 + 2.$$

Luego la fórmula es verdadera para todo n.

El cálculo algebraico es correcto, pero la justificación está incompleta si la igualdad

$$a_n = n^2 + 2$$

se utiliza sin declarar que es la **hipótesis inductiva local** dentro de un paso inductivo y sin haber verificado la base.

La reparación consiste en reconstruir las dependencias:

1. demostrar $P(0)$;
2. fijar n arbitrario;
3. suponer temporalmente $P(n)$;
4. usar la regla recursiva para demostrar $P(n+1)$;
5. invocar el principio de inducción.

La misma ecuación puede ser legítima o circular según el lugar que ocupe en la arquitectura de la prueba.

Definir no es demostrar

Si un objeto está definido por

$$a_0 = c, \qquad a_{n+1} = G(n, a_n),$$

la recursión garantiza **qué objeto** tenemos.

Para demostrar una propiedad universal

$$P(n, a_n)$$

podemos necesitar inducción:

$$P(0, a_0) \qquad + \qquad P(n, a_n) \Rightarrow P(n+1, a_{n+1}).$$

La regla recursiva suele ser el puente que permite reescribir a_{n+1} en términos de información ya disponible en el índice n.

Comprobación. Sea una sucesión

$$b : \mathbb{N} \to \mathbb{N}$$

definida recursivamente por

$$b_0 = 4, \qquad b_{n+1} = b_n + 4n + 2.$$

Los primeros términos son

$$4, 6, 12, 22, \ldots$$

y sugieren la fórmula

$$b_n = 2n^2 + 4.$$

Explica cómo demostrarla por inducción e identifica, en el paso, qué información procede de la definición recursiva y cuál procede de la hipótesis inductiva.

Respuesta razonada.

Definimos

$$P(n): \quad b_n = 2n^2 + 4.$$

Para la base,

$$b_0 = 4 = 2 \cdot 0^2 + 4,$$

así que $P(0)$ es verdadera.
Para el paso, fijamos $n \in \mathbb{N}$ y suponemos la hipótesis inductiva

$$b_n = 2n^2 + 4.$$

La **definición recursiva** proporciona

$$b_{n+1} = b_n + 4n + 2.$$

Sustituyendo la **hipótesis inductiva** obtenemos

$$\begin{aligned} b_{n+1} &= b_n + 4n + 2 \\ &= (2n^2 + 4) + 4n + 2 \\ &= 2n^2 + 4n + 6 \\ &= 2(n+1)^2 + 4. \end{aligned}$$

Por tanto,

$$P(n) \Rightarrow P(n+1).$$

El principio de inducción da

$$\boxed{b_n = 2n^2 + 4 \text{ para todo } n \in \mathbb{N}}.$$

La separación de funciones es exacta:

- la recursión aporta la relación entre b_{n+1} y b_n;
- la hipótesis inductiva aporta la fórmula explícita para b_n en el índice actual;
- la inducción convierte ese mecanismo local en una conclusión universal.

Hacia la integración del capítulo. Ya tenemos las cuatro piezas centrales de C08:

$$\boxed{\text{inducción ordinaria}}$$

$$\boxed{\text{inducción fuerte}}$$

$$\boxed{\text{buen orden}}$$

$$\boxed{\text{recursión}}.$$

También sabemos cómo interactúan sin confundirse:

- inducción ordinaria, inducción fuerte y buen orden son formulaciones equivalentes en el marco elemental adoptado;
- recursión cumple otra función: **define** objetos sobre los naturales;
- una vez definido un objeto recursivamente, la inducción puede utilizar su regla de definición para demostrar propiedades universales sobre él.

La sección final integrará estas ideas mediante ejercicios que exigirán no sólo ejecutar pruebas, sino también diagnosticar bases insuficientes, elegir entre inducción ordinaria y fuerte, reconstruir argumentos por menor contraejemplo y decidir si una especificación recursiva está realmente bien formada.

8.10. Ejercicios y soluciones desarrolladas

Esta sección integra los principios del capítulo sin convertirlos en una colección de trucos separados. Los ejercicios están organizados en tres familias:

1. **diagnóstico y arquitectura**: reconocer qué obligación falta, qué supuesto es legítimo y qué tipo de argumento está realmente disponible;
2. **ejecución guiada**: construir pruebas completas por inducción ordinaria, inducción fuerte y buen orden, y trabajar con definiciones recursivas;
3. **integración y comparación**: relacionar los principios, justificar unicidad, reparar definiciones y distinguir cuándo la inducción es natural y cuándo no es necesaria.

La meta no es acumular identidades. Es aprender a leer la estructura de cada problema.

Ejercicios.

E1. La base pertenece al dominio del teorema. Considera la afirmación

$$P(n): \quad n^2 \geq n + 6.$$

Se quiere demostrar

$$\forall n \geq 3, \quad P(n).$$

1. Identifica la base correcta.
2. Formula el paso inductivo con su rango exacto.
3. Demuestra la afirmación por inducción.
4. Explica por qué demostrar $P(0)$ no cumpliría la obligación de base de este teorema.

E2. Una base correcta y un paso demasiado tardío. Supón que quieres demostrar

$$\forall n \geq 4, \quad P(n).$$

Has demostrado

$$P(4)$$

y también

$$\forall n \geq 5, \qquad P(n) \Rightarrow P(n+1).$$

1. Explica por qué esta información no basta todavía.
2. Localiza exactamente el hueco.
3. Da dos reparaciones distintas que permitan concluir el teorema deseado.

E3. Diagnóstico de circularidad. Un borrador de prueba por inducción de

$$\forall n \geq 1, \quad P(n)$$

contiene el siguiente paso:

> Sea $n \geq 1$ arbitrario y supongamos $P(n)$. Como el teorema afirma que $P(k)$ vale para todo $k \geq 1$, tomamos $k = n + 1$ y obtenemos $P(n+1)$. Por tanto, el paso está demostrado.

1. ¿Qué línea es legítima?
2. ¿Dónde aparece la circularidad?
3. ¿Qué clase de información tendría que reemplazar el paso circular para que la inducción fuese válida?

E4. ¿Ordinaria o fuerte?. Se quiere demostrar una propiedad $P(n)$ para todo $n \geq 6$. El único mecanismo disponible para producir el caso siguiente tiene esta forma:

$$P(n-2) \Longrightarrow P(n+1),$$

siempre que $n \geq 8$.

1. ¿Qué forma de inducción expresa de manera más natural esta dependencia?
2. ¿Qué casos iniciales deben quedar demostrados antes de activar el primer paso?
3. Explica cómo podría, en principio, reorganizarse el argumento usando inducción ordinaria sobre una propiedad más rica.

E5. Una regla que sale del codominio. Se pretende definir

$$a : \mathbb{N} \to \mathbb{N}$$

mediante

$$a_0 = 1, \qquad a_{n+1} = a_n - 2.$$

1. ¿Es ésta una definición recursiva legítima de una función $\mathbb{N} \to \mathbb{N}$?
2. Identifica exactamente el fallo.
3. Propón una reparación mínima que conserve la misma regla algebraica.

E6. Inducción ordinaria. Demuestra por inducción que

$$3^n \geq 2n + 1$$

para todo

$$n \in \mathbb{N}.$$

Indica explícitamente la base, la hipótesis inductiva, el objetivo del paso y la conclusión universal.

E7. Inducción desde un índice no nulo. Demuestra que

$$n^3 \geq 4n$$

para todo

$$n \geq 2.$$

Explica por qué el índice 2 es un punto inicial natural para el enunciado, aunque $n = 0$ también satisfaga accidentalmente la desigualdad.

E8. Una aplicación de inducción fuerte. Demuestra que todo entero

$$n \geq 18$$

puede escribirse en la forma

$$n = 4a + 7b$$

para ciertos

$$a, b \in \mathbb{N}.$$

Tu prueba debe hacer explícitos los casos iniciales necesarios y el índice anterior que se utiliza en el paso fuerte.

E9. No existe una cadena infinita estrictamente decreciente en $\mathbb{N}$. Demuestra, utilizando el principio de buen orden, que no existe una función

$$a : \mathbb{N} \to \mathbb{N}$$

tal que

$$a_{n+1} < a_n$$

para todo $n \in \mathbb{N}$.

E10. Definir recursivamente y luego demostrar. Sea

$$\mathbb{Q}_{>0} = \{q \in \mathbb{Q} : q > 0\}.$$

Considera la especificación

$$c_0 = 1, \qquad c_{n+1} = \frac{c_n}{1 + c_n}.$$

1. Justifica que esta regla define legítimamente una única función
$$c : \mathbb{N} \to \mathbb{Q}_{>0}.$$
2. Demuestra por inducción que
$$c_n = \frac{1}{n+1}$$
para todo $n \in \mathbb{N}$.
3. Señala qué parte de tu argumento pertenece a la recursión y qué parte pertenece a la inducción.

E11. Unicidad de un objeto recursivo. Sean

$$x, y : \mathbb{N} \to \mathbb{R}$$

dos funciones tales que

$$x_0 = y_0 = 2$$

y, para todo $n \in \mathbb{N}$,

$$x_{n+1} = 3x_n + 1, \qquad y_{n+1} = 3y_n + 1.$$

Demuestra por inducción que

$$x = y.$$

Explica por qué ésta es una prueba de **unicidad** y no la definición de las funciones.

E12. Una recursión de segundo orden. Se propone la regla

$$r_{n+2} = r_{n+1} + 2r_n$$

junto con el único dato

$$r_0 = 1.$$

1. Explica por qué la especificación está incompleta.
2. Añade el dato
$$r_1 = 2.$$
Calcula r_2, r_3, r_4.

3. Demuestra, utilizando inducción fuerte, que
$$r_n = 2^n$$
para todo $n \in \mathbb{N}$.
4. Explica por qué una inducción ordinaria también sería posible si se empaquetara suficiente información en la propiedad inductiva.

E13. Reconstruir la inducción como imposibilidad de un primer fallo. Supón que una propiedad $P(n)$ satisface

$$P(0)$$

y

$$\forall n \in \mathbb{N}, \qquad P(n) \Rightarrow P(n+1).$$

Sin invocar directamente el principio de inducción, demuestra mediante buen orden que

$$\forall n \in \mathbb{N}, \quad P(n).$$

Debes construir el conjunto de índices de fallo y explicar exactamente cómo se contradice la minimalidad de su supuesto mínimo.

E14. Elegir la arquitectura natural. Para cada situación, indica qué herramienta expresa de manera más natural la estructura del problema y justifica tu elección.

1. Para demostrar $P(n+1)$ sólo necesitas $P(n)$.
2. Para demostrar $P(n+1)$ necesitas una instancia como $P(n-3)$, además de saber que todos los índices anteriores pertinentes ya están cubiertos.
3. La negación del teorema produce un conjunto no vacío de índices de fallo, y la minimalidad del primero permite demostrar que también debería ser un caso bueno.
4. Quieres construir una sucesión indicando un valor inicial y una regla que produce el siguiente término.

Compara, en particular, **inducción ordinaria**, **inducción fuerte**, **menor contraejemplo** y **recursión**.

E15. Integración final: definir, conjeturar, demostrar y comparar. Define una sucesión

$$s : \mathbb{N} \to \mathbb{N}$$

por

$$s_0 = 3, \qquad s_{n+1} = s_n + 6n + 4.$$

1. Justifica que la especificación es una definición recursiva legítima sobre $\mathbb{N}$.
2. Calcula s_1, s_2, s_3 y formula una conjetura para s_n.
3. Demuestra por inducción que
$$s_n = 3n^2 + n + 3$$
para todo $n \in \mathbb{N}$.
4. Demuestra que la sucesión es estrictamente creciente.
5. Explica por qué la parte 4 admite una prueba directa muy breve y no requiere inducción, aunque el objeto haya sido definido recursivamente.

Soluciones desarrolladas.

Solución E1. La conclusión que queremos demostrar es

$$\forall n \geq 3, \qquad n^2 \geq n + 6.$$

Por tanto, el primer índice del tramo es 3.
La base correcta es

$$P(3): \quad 3^2 \geq 3 + 6.$$

En efecto,

$$9 = 9.$$

Para el paso, fijamos un $n \geq 3$ arbitrario y suponemos

$$P(n): \quad n^2 \geq n + 6.$$

Debemos demostrar

$$P(n+1): \quad (n+1)^2 \geq (n+1) + 6 = n + 7.$$

A partir de la hipótesis inductiva,

$$\begin{aligned}(n+1)^2 &= n^2 + 2n + 1 \\ &\geq (n+6) + 2n + 1 \\ &= 3n + 7.\end{aligned}$$

Como $n \geq 3$, en particular $2n \geq 0$, y por tanto

$$3n + 7 \geq n + 7.$$

Así,

$$(n+1)^2 \geq n + 7,$$

y hemos probado

$$P(n) \Rightarrow P(n+1)$$

para todo $n \geq 3$.
Por inducción desde 3,

$$\boxed{n^2 \geq n + 6 \text{ para todo } n \geq 3}.$$

Demostrar $P(0)$ no cumpliría la obligación de base porque el teorema no empieza en 0. Más aún,

$$P(0): \quad 0 \geq 6$$

es falsa. La base debe pertenecer al dominio real de la conclusión y conectar con el primer paso disponible.

Solución E2. La base

$$P(4)$$

es correcta, pero el paso disponible comienza en

$$n = 5.$$

Por tanto sabemos

$$P(5) \Rightarrow P(6), \quad P(6) \Rightarrow P(7), \quad \ldots,$$

pero no hemos justificado

$$P(4) \Rightarrow P(5).$$

El hueco está exactamente entre los índices 4 y 5.

Hay al menos dos reparaciones naturales.

Primera reparación. Fortalecer el paso y demostrar

$$\forall n \geq 4, \qquad P(n) \Rightarrow P(n+1).$$

Entonces $P(4)$ activa inmediatamente la propagación.

Segunda reparación. Demostrar $P(5)$ por separado. Después podemos usar $P(5)$ como base de una inducción que cubra todos los índices $n \geq 5$, y combinar ese resultado con el caso $P(4)$ ya conocido.

Lo que importa no es la apariencia formal de la prueba, sino que no quede ningún índice del dominio sin cubrir.

Solución E3. La frase

> sea $n \geq 1$ arbitrario y supongamos $P(n)$

es legítima. Dentro del paso inductivo estamos intentando demostrar la implicación

$$P(n) \Rightarrow P(n+1),$$

y para probar una condicional podemos suponer temporalmente su antecedente.

La circularidad aparece después, cuando el borrador utiliza

$$\forall k \geq 1, \quad P(k).$$

Ésa es precisamente la conclusión universal que todavía queremos demostrar. Tomar $k = n + 1$ sólo produce el objetivo porque el teorema completo ya fue supuesto de antemano.

La reparación exige un puente matemático independiente:

$$P(n) \quad + \quad \text{definiciones, hipótesis o resultados previos} \quad \Longrightarrow \quad P(n+1).$$

La hipótesis inductiva debe transformarse en información útil; no puede sustituirse por la conclusión universal.

Solución E4. La dependencia disponible es

$$P(n-2) \Rightarrow P(n+1).$$

Por tanto, para producir el caso $n + 1$ no basta necesariamente con la instancia inmediata $P(n)$. La **inducción fuerte** expresa de manera natural la situación, porque permite utilizar cualquier caso anterior ya cubierto.

El paso sólo está disponible para

$$n \geq 8.$$

El primer paso produce $P(9)$ y utiliza

$$P(6).$$

Por tanto, antes de activar el mecanismo deben estar cubiertos

$$P(6), \quad P(7), \quad P(8).$$

Con esos tres casos iniciales, la hipótesis fuerte para $n = 8$ contiene $P(6)$ y permite obtener $P(9)$. Después la hipótesis para $n = 9$ contiene todos los casos desde 6 hasta 9, y el proceso continúa.

También sería posible usar inducción ordinaria sobre una propiedad acumulativa, por ejemplo una afirmación que transporte simultáneamente tres casos consecutivos. Esquemáticamente podríamos definir

$$Q(n): \quad P(n) \wedge P(n+1) \wedge P(n+2).$$

Si el argumento permite demostrar

$$Q(n) \Rightarrow Q(n+1),$$

la inducción ordinaria puede transportar el bloque completo. La forma fuerte, sin embargo, refleja más directamente la dependencia original.

Solución E5. La especificación no define legítimamente una función

$$a : \mathbb{N} \to \mathbb{N}.$$

El dato inicial es válido:

$$a_0 = 1 \in \mathbb{N}.$$

Pero la primera aplicación de la regla produce

$$a_1 = a_0 - 2 = -1,$$

y

$$-1 \notin \mathbb{N}.$$

Por tanto, la regla

$$G(n, x) = x - 2$$

no determina una función

$$G : \mathbb{N} \times \mathbb{N} \to \mathbb{N}.$$

Una reparación mínima consiste en cambiar el codominio y trabajar con

$$a : \mathbb{N} \to \mathbb{Z}.$$

Entonces

$$G : \mathbb{N} \times \mathbb{Z} \to \mathbb{Z}, \qquad G(n, x) = x - 2$$

sí es una función total, y el teorema de recursión determina una única sucesión entera. De hecho, sus valores satisfacen

$$a_n = 1 - 2n,$$

aunque esa fórmula explícita sería una propiedad que todavía habría que demostrar si quisiéramos utilizarla universalmente.

Solución E6. Definimos

$$P(n): \quad 3^n \geq 2n + 1.$$

Base. Para $n = 0$,

$$3^0 = 1 = 2 \cdot 0 + 1.$$

Así, $P(0)$ es verdadera.

Paso. Sea $n \in \mathbb{N}$ arbitrario y supongamos

$$3^n \geq 2n + 1.$$

Debemos demostrar

$$3^{n+1} \geq 2(n + 1) + 1 = 2n + 3.$$

Multiplicando la hipótesis inductiva por $3 > 0$,

$$3^{n+1} \geq 6n + 3.$$

Como $n \geq 0$,

$$6n + 3 \geq 2n + 3.$$

Por tanto,

$$3^{n+1} \geq 2n + 3.$$

Hemos demostrado

$$P(n) \Rightarrow P(n + 1)$$

para un n arbitrario. Por inducción,

$$\boxed{3^n \geq 2n + 1 \text{ para todo } n \in \mathbb{N}}.$$

La hipótesis inductiva es la desigualdad en el índice n; el objetivo local es la desigualdad en $n + 1$; la conclusión universal sólo aparece después de combinar base y paso.

Solución E7. Definimos

$$P(n): \quad n^3 \geq 4n,$$

para $n \geq 2$.

Base. Para $n = 2$,

$$2^3 = 8 = 4 \cdot 2.$$

Paso. Sea $n \geq 2$ y supongamos

$$n^3 \geq 4n.$$

Entonces

$$\begin{aligned}(n+1)^3 &= n^3 + 3n^2 + 3n + 1 \\ &\geq 4n + 3n^2 + 3n + 1.\end{aligned}$$

Como $n \geq 2$, tenemos en particular

$$3n^2 + 3n + 1 \geq 4.$$

Luego

$$(n+1)^3 \geq 4n + 4 = 4(n+1).$$

Por inducción desde 2,

$$\boxed{n^3 \geq 4n \text{ para todo } n \geq 2}.$$

El valor $n = 0$ satisface también $0^3 \geq 4 \cdot 0$, pero eso no cambia el dominio natural del enunciado propuesto. El caso $n = 1$ falla:

$$1^3 < 4.$$

Por eso no existe una cadena inductiva continua desde 0 que cubra el teorema sin tratar el hueco en 1. El tramo $n \geq 2$ expresa exactamente la región continua donde la afirmación es verdadera y el paso puede propagarse.

Solución E8. Sea

$$P(n): \quad \exists a, b \in \mathbb{N} \text{ tales que } n = 4a + 7b.$$

Queremos demostrar $P(n)$ para todo $n \geq 18$.

Los primeros cuatro casos son:

$$18 = 4 \cdot 1 + 7 \cdot 2,$$

$$19 = 4 \cdot 3 + 7 \cdot 1,$$

$$20 = 4 \cdot 5 + 7 \cdot 0,$$

y

$$21 = 4 \cdot 0 + 7 \cdot 3.$$

Por tanto,

$$P(18), P(19), P(20), P(21)$$

son verdaderas.

Ahora sea $n \geq 21$ y supongamos, como hipótesis inductiva fuerte, que

$$P(k)$$

es verdadera para todo

$$18 \leq k \leq n.$$

Queremos demostrar $P(n+1)$.
Como $n \geq 21$,

$$n - 3 \geq 18.$$

Además,

$$n - 3 \leq n.$$

Por la hipótesis fuerte podemos utilizar

$$P(n-3).$$

Así existen $a, b \in \mathbb{N}$ tales que

$$n - 3 = 4a + 7b.$$

Sumando 4,

$$n + 1 = 4(a+1) + 7b.$$

Como $a + 1 \in \mathbb{N}$, queda demostrada $P(n+1)$.
Por inducción fuerte,

$$\boxed{\forall n \geq 18,\ \exists a, b \in \mathbb{N} \text{ tales que } n = 4a + 7b}.$$

Los cuatro casos iniciales no son un ritual: son exactamente los índices que deben quedar cubiertos antes de que el mecanismo "retroceder cuatro unidades y sumar 4" pueda activarse sin huecos.

Solución E9. Supongamos, buscando una contradicción, que existe una función

$$a : \mathbb{N} \to \mathbb{N}$$

tal que

$$a_{n+1} < a_n$$

para todo $n \in \mathbb{N}$.
Consideremos el conjunto de todos sus valores:

$$S = \{a_n : n \in \mathbb{N}\}.$$

El conjunto S es un subconjunto no vacío de $\mathbb{N}$, porque contiene al menos a_0.
Por el principio de buen orden, S tiene un mínimo. Por tanto, existe algún índice $m \in \mathbb{N}$ tal que

$$a_m = \min S.$$

Pero la condición de decrecimiento estricto da

$$a_{m+1} < a_m.$$

Además,

$$a_{m+1} \in S.$$

Esto contradice que a_m sea el mínimo de S.
Por tanto, no existe una cadena infinita estrictamente decreciente de números naturales:

$$\boxed{\nexists a : \mathbb{N} \to \mathbb{N} \text{ con } a_{n+1} < a_n \text{ para todo } n}.$$

El buen orden concentra el argumento en un supuesto valor mínimo y hace imposible que la cadena siga descendiendo después de alcanzarlo.

Solución E10. Trabajemos en

$$X = \mathbb{Q}_{>0}.$$

Definamos

$$G : X \to X, \qquad G(x) = \frac{x}{1+x}.$$

Si $x \in \mathbb{Q}_{>0}$, entonces $1 + x > 0$, el cociente es racional y

$$\frac{x}{1+x} > 0.$$

Por tanto,

$$G(x) \in \mathbb{Q}_{>0}.$$

El dato inicial

$$c_0 = 1$$

pertenece también a X. La forma elemental del teorema de recursión garantiza entonces una única función

$$c : \mathbb{N} \to \mathbb{Q}_{>0}$$

tal que

$$c_0 = 1, \qquad c_{n+1} = \frac{c_n}{1+c_n}.$$

Ahora demostremos la fórmula explícita. Definimos

$$P(n) : \quad c_n = \frac{1}{n+1}.$$

Base. Para $n = 0$,

$$c_0 = 1 = \frac{1}{0+1}.$$

Paso. Supongamos

$$c_n = \frac{1}{n+1}.$$

Usando la definición recursiva,

$$\begin{aligned} c_{n+1} &= \frac{c_n}{1+c_n} \\ &= \frac{\frac{1}{n+1}}{1+\frac{1}{n+1}} \\ &= \frac{\frac{1}{n+1}}{\frac{n+2}{n+1}} \\ &= \frac{1}{n+2}. \end{aligned}$$

Esto es exactamente $P(n+1)$.

Por inducción,

$$\boxed{c_n = \frac{1}{n+1} \text{ para todo } n \in \mathbb{N}}.$$

La separación conceptual es ésta:

- la **recursión** garantiza la existencia y unicidad de la función c;
- la ecuación recursiva proporciona el puente entre c_{n+1} y c_n;
- la **inducción** demuestra que la fórmula $1/(n+1)$ describe todos los valores de ese objeto ya definido.

Solución E11. Queremos demostrar que las funciones x e y coinciden en todos los naturales.

Definamos

$$P(n): \quad x_n = y_n.$$

Base. Por hipótesis,

$$x_0 = 2 = y_0.$$

Así, $P(0)$ es verdadera.

Paso. Supongamos

$$x_n = y_n.$$

Entonces

$$\begin{aligned} x_{n+1} &= 3x_n + 1 \\ &= 3y_n + 1 \\ &= y_{n+1}. \end{aligned}$$

Por tanto,

$$P(n) \Rightarrow P(n+1).$$

La inducción da

$$x_n = y_n$$

para todo $n \in \mathbb{N}$. Por igualdad de funciones,

$$\boxed{x = y}.$$

La recursión especifica cómo debe avanzar cada función desde el dato inicial. La inducción demuestra que dos funciones que obedecen exactamente la misma especificación no pueden separarse en ningún índice. Por eso el argumento establece **unicidad**; no es el acto de definir x o y.

Solución E12. La regla

$$r_{n+2} = r_{n+1} + 2r_n$$

necesita dos valores anteriores para producir el siguiente término.

Conocer sólo

$$r_0 = 1$$

no permite calcular r_2, porque antes necesitamos r_1.

Por eso la especificación original está incompleta.

Añadamos

$$r_1 = 2.$$

Entonces

$$r_2 = r_1 + 2r_0 = 2 + 2 = 4,$$

$$r_3 = r_2 + 2r_1 = 4 + 4 = 8,$$

y

$$r_4 = r_3 + 2r_2 = 8 + 8 = 16.$$

Esto sugiere

$$r_n = 2^n.$$

Demostremos la fórmula mediante inducción fuerte.

Los dos casos iniciales son

$$r_0 = 1 = 2^0$$

y

$$r_1 = 2 = 2^1.$$

Ahora supongamos que la fórmula es verdadera para todos los índices hasta $n + 1$. En particular podemos utilizar

$$r_n = 2^n$$

y

$$r_{n+1} = 2^{n+1}.$$

Entonces

$$\begin{aligned} r_{n+2} &= r_{n+1} + 2r_n \\ &= 2^{n+1} + 2 \cdot 2^n \\ &= 2^{n+1} + 2^{n+1} \\ &= 2^{n+2}. \end{aligned}$$

Por inducción fuerte,

$$\boxed{r_n = 2^n \text{ para todo } n \in \mathbb{N}}.$$

También podríamos usar inducción ordinaria si empaquetamos las dos instancias necesarias en una sola propiedad. Por ejemplo,

$$Q(n): \quad r_n = 2^n \text{ y } r_{n+1} = 2^{n+1}.$$

El paso $Q(n) \Rightarrow Q(n+1)$ transportaría simultáneamente la información de dos índices consecutivos. La forma fuerte evita tener que introducir artificialmente ese paquete.

Solución E13. Queremos demostrar

$$\forall n \in \mathbb{N}, \quad P(n)$$

a partir de

$$P(0)$$

y

$$\forall n \in \mathbb{N}, \qquad P(n) \Rightarrow P(n+1).$$

Supongamos, buscando una contradicción, que la conclusión universal es falsa.

Definamos el conjunto de índices de fallo:

$$F = \{n \in \mathbb{N} : \neg P(n)\}.$$

Bajo nuestra suposición,

$$F \neq \varnothing.$$

Por buen orden existe

$$m = \min F.$$

Como $P(0)$ es verdadera,

$$0 \notin F,$$

y por tanto

$$m \geq 1.$$

Así,

$$m - 1 \in \mathbb{N}.$$

Además,

$$m - 1 < m.$$

Por minimalidad de m,

$$m - 1 \notin F,$$

de modo que

$$P(m-1)$$

es verdadera.

Aplicando la regla de paso al índice $m-1$ obtenemos

$$P(m).$$

Pero $m \in F$, así que también

$$\neg P(m).$$

Contradicción.

Por tanto,

$$F = \varnothing,$$

y concluimos

$$\boxed{\forall n \in \mathbb{N},\ P(n)}.$$

El argumento muestra de manera concreta que la inducción ordinaria puede interpretarse como la imposibilidad de un **primer fallo**.

Solución E14. **Situación 1.** Si para demostrar $P(n+1)$ sólo necesitamos $P(n)$, la arquitectura más natural es la **inducción ordinaria**:

$$P(n) \Rightarrow P(n+1).$$

Situación 2. Si el caso siguiente depende de una instancia anterior no inmediata, como $P(n-3)$, y queremos disponer de todos los casos anteriores pertinentes, la forma más natural es la **inducción fuerte**. Ésta organiza la hipótesis local como un bloque

$$P(n_0), \ldots, P(n).$$

Situación 3. Si suponemos que el teorema falla, obtenemos un conjunto no vacío de índices de fallo y la minimalidad del primero convierte todos los casos anteriores en casos buenos. Entonces la arquitectura natural es el **menor contraejemplo**, apoyado en buen orden.

Situación 4. Si la tarea consiste en construir una sucesión mediante un valor inicial y una regla para producir valores posteriores, no estamos todavía demostrando una propiedad universal. La herramienta es la **recursión**.

Podemos resumir:

inducción ordinaria	:	transportar una instancia,
inducción fuerte	:	transportar historia previa,
menor contraejemplo	:	hacer imposible el primer fallo,
recursión	:	definir el objeto índice por índice.

Las tres primeras son arquitecturas para demostrar propiedades. La cuarta cumple otra función lógica: construir un objeto.

Solución E15. La especificación es

$$s_0 = 3, \qquad s_{n+1} = s_n + 6n + 4.$$

Podemos escribir la regla como

$$G : \mathbb{N} \times \mathbb{N} \to \mathbb{N}, \qquad G(n, x) = x + 6n + 4.$$

Si $n, x \in \mathbb{N}$, entonces

$$x + 6n + 4 \in \mathbb{N}.$$

Además,

$$s_0 = 3 \in \mathbb{N}.$$

Por el teorema elemental de recursión existe una única función

$$s : \mathbb{N} \to \mathbb{N}$$

que satisface la especificación.
Calculemos algunos términos:

$$s_1 = s_0 + 4 = 7,$$

$$s_2 = s_1 + 10 = 17,$$

y

$$s_3 = s_2 + 16 = 33.$$

Los valores

$$3, 7, 17, 33, \ldots$$

sugieren la fórmula

$$s_n = 3n^2 + n + 3.$$

Demostrémosla.
Definimos

$$P(n) : \quad s_n = 3n^2 + n + 3.$$

Base. Para $n = 0$,

$$s_0 = 3 = 3 \cdot 0^2 + 0 + 3.$$

Paso. Supongamos

$$s_n = 3n^2 + n + 3.$$

La definición recursiva da

$$s_{n+1} = s_n + 6n + 4.$$

Sustituyendo la hipótesis inductiva,

$$\begin{aligned} s_{n+1} &= (3n^2 + n + 3) + 6n + 4 \\ &= 3n^2 + 7n + 7. \end{aligned}$$

Por otra parte,

$$\begin{aligned} 3(n+1)^2 + (n+1) + 3 &= 3(n^2 + 2n + 1) + n + 1 + 3 \\ &= 3n^2 + 7n + 7. \end{aligned}$$

Por tanto,

$$s_{n+1} = 3(n+1)^2 + (n+1) + 3.$$

Así hemos demostrado el paso inductivo. Por inducción,

$$\boxed{s_n = 3n^2 + n + 3 \text{ para todo } n \in \mathbb{N}}.$$

Ahora estudiemos el crecimiento. De la propia definición,

$$s_{n+1} - s_n = 6n + 4.$$

Como

$$6n + 4 > 0$$

para todo $n \in \mathbb{N}$, se sigue directamente que

$$s_{n+1} > s_n$$

para todo n.

Por tanto, la sucesión es estrictamente creciente.

No hizo falta inducción para esta última afirmación porque la regla recursiva proporciona directamente, para un índice arbitrario, una diferencia positiva entre términos consecutivos. La inducción fue natural para demostrar una fórmula explícita universal; la monotonía, en cambio, se deduce localmente de la definición.

Cierre del capítulo. A lo largo de C08 hemos utilizado la estructura de los naturales de cuatro maneras que ahora pueden distinguirse con precisión:

$$\boxed{\text{inducción ordinaria}} \qquad \boxed{\text{inducción fuerte}} \qquad \boxed{\text{buen orden}} \qquad \boxed{\text{recursión}}.$$

Las tres primeras permiten organizar demostraciones universales sobre $\mathbb{N}$ y, en el marco elemental del capítulo, quedaron conectadas por equivalencias explícitas. La recursión cumple una función distinta: permite construir objetos indexados por los naturales a partir de datos iniciales y reglas de generación.

La relación más útil puede resumirse así:

$$\boxed{\text{recursión construye}} \qquad \text{y} \qquad \boxed{\text{inducción verifica}}.$$

Pero tampoco debemos convertir esta fórmula en una receta mecánica. Un objeto recursivo puede poseer propiedades demostrables directamente; una prueba inductiva puede requerir una idea matemática no suministrada por la arquitectura; y el buen orden sólo localiza un mínimo, no inventa la contradicción que debe excluirlo.

El capítulo siguiente cambiará de pregunta. Ya no preguntará principalmente cómo propagar una propiedad verdadera, sino cómo reconocer que una afirmación universal **es falsa**, cómo encontrar un contraejemplo y cómo diagnosticar exactamente dónde se rompe un argumento.

Ése será el problema de C09.

Capítulo 9 — ¿Cómo sabemos que una afirmación universal es falsa?

En los capítulos anteriores aprendimos a reconocer qué exige una afirmación matemática y a construir argumentos capaces de satisfacer esa exigencia. C06 organizó varias arquitecturas de demostración. C07 distinguió entre conocer la forma de una prueba y descubrir el contenido matemático que permite completarla. C08 añadió principios específicos para razonar sobre los números naturales.

Pero hay otra posibilidad que no debe quedar en segundo plano:

la afirmación que estamos intentando demostrar puede ser falsa.

En ese caso, insistir en encontrar una demostración no sólo será difícil. Será imposible.

La dificultad consiste entonces en cambiar correctamente de tarea.

No debemos pasar de

«todavía no encontré una prueba»

al juicio

«por tanto, el enunciado es falso».

Un intento fallido sólo demuestra que **ese intento** no funcionó. No demuestra todavía nada sobre el valor de verdad del enunciado.

Para establecer falsedad necesitamos un argumento.

Cuando la afirmación es universal, ese argumento puede ser extraordinariamente pequeño: un solo objeto puede bastar. Pero ese objeto debe cumplir exactamente las condiciones que la afirmación promete cubrir.

Ésta será la tensión central del capítulo:

muchos casos favorables no prueban una universal

mientras que

un contraejemplo válido basta para refutarla.

La palabra **válido** será decisiva.

Un número elegido fuera del dominio no es un contraejemplo. Un objeto que no cumple las hipótesis tampoco lo es. Y un caso que sólo parece extraño, pero no hace falsa la conclusión pertinente, tampoco refuta nada.

C09 estudiará por eso tres tareas relacionadas pero distintas:

refutar diagnosticar reparar.

Refutar será demostrar que un enunciado es falso.

Diagnosticar será localizar la primera inferencia que deja de estar justificada dentro de un argumento.

Reparar será modificar el enunciado o el argumento sólo en la medida necesaria para eliminar el fallo.

Comenzaremos por la forma más elemental y más importante de refutación: la de una afirmación universal.

9.1. ¿Qué significa refutar una afirmación universal?

Supongamos que tenemos una afirmación de la forma

$$\forall x \in D, \quad P(x).$$

Esta expresión promete algo muy fuerte.
Afirma que **cada** objeto del dominio D satisface la propiedad P.
No afirma que la mayoría lo haga.
No afirma que los primeros cien casos que revisemos lo hagan.
No afirma que no hayamos encontrado todavía una excepción.
Afirma que no existe ninguna excepción dentro de D.
Por eso, para refutarla, debemos demostrar exactamente lo contrario.

La negación de una universal es existencial. Desde C03 sabemos que

$$\neg\,(\forall x \in D, \quad P(x))$$

es equivalente a

$$\exists x \in D, \quad \neg P(x).$$

Esta equivalencia determina la arquitectura completa de la refutación.
Para demostrar que

$$\forall x \in D, \quad P(x)$$

es falsa, basta encontrar un objeto

$$a \in D$$

para el cual

$$\neg P(a)$$

sea verdadera.
Ese objeto a recibe el nombre de **contraejemplo**.
Así, un contraejemplo no es simplemente un caso que "no encaja bien" con nuestra intuición.
Es un **testigo de una negación existencial**.
La estructura lógica es

$$\boxed{\forall x \in D\ P(x)}$$

$$\Downarrow \text{ negamos}$$

$$\boxed{\exists x \in D\ \neg P(x)}.$$

Y la refutación consiste en exhibir y verificar ese testigo.

Primer ejemplo: una universal falsa. Consideremos la afirmación

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

Para refutarla buscamos un real a tal que

$$a^2 < a.$$

Tomemos

$$a = \frac{1}{2}.$$

Primero comprobamos el dominio:

$$\frac{1}{2} \in \mathbb{R}.$$

Después evaluamos la propiedad:

$$\left(\frac{1}{2}\right)^2 = \frac{1}{4}.$$

Y tenemos

$$\frac{1}{4} < \frac{1}{2}.$$

Por tanto,

$$\neg\left(\left(\frac{1}{2}\right)^2 \geq \frac{1}{2}\right).$$

Así,

$$\boxed{\frac{1}{2} \text{ es un contraejemplo}}.$$

La afirmación universal queda refutada.

Obsérvese que no hemos demostrado que la desigualdad falle “muchas veces”.

No lo necesitamos.

La afirmación prometía que funcionaría para **todo** real. Un solo real que rompa la promesa es suficiente.

Nombrar un objeto no basta. Una refutación completa no termina con una frase como

> «un contraejemplo es 1/2».

Todavía debemos mostrar por qué ese objeto constituye efectivamente un contraejemplo.

En el ejemplo anterior hubo dos verificaciones explícitas:

1. $1/2$ pertenece al dominio anunciado;
2. la propiedad universal falla cuando $x = 1/2$.

Esta segunda comprobación fue

$$\frac{1}{4} < \frac{1}{2}.$$

Sin ella sólo habríamos propuesto un candidato.

Un contraejemplo debe ser **verificado**, no únicamente declarado.

El dominio forma parte de la refutación. Consideremos ahora la afirmación

$$\forall n \in \mathbb{Z}, \quad n^2 \geq n.$$

Alguien podría observar que

$$x = \frac{1}{2}$$

satisface

$$\left(\frac{1}{2}\right)^2 < \frac{1}{2}$$

y concluir que ha refutado el enunciado.

Pero no lo ha hecho.

El dominio anunciado es

$$\mathbb{Z},$$

y

$$\frac{1}{2} \notin \mathbb{Z}.$$

Por tanto, $1/2$ no pertenece al conjunto de objetos sobre los que la afirmación hizo una promesa.

El cálculo puede ser correcto y, sin embargo, ser irrelevante para la refutación.

Esto produce una regla que utilizaremos durante todo C09:

$$\boxed{\text{un objeto fuera del dominio no puede ser contraejemplo}}.$$

La pertenencia al dominio no es una formalidad previa al verdadero argumento.

Es parte del argumento.

Cuando la universal contiene una implicación. Muchas afirmaciones universales tienen una estructura ligeramente más rica:

$$\forall x \in D, \qquad H(x) \Rightarrow C(x).$$

Aquí $H(x)$ representa las hipótesis que debe satisfacer un objeto y $C(x)$ la conclusión prometida para esos objetos.

La negación es

$$\exists x \in D, \qquad H(x) \wedge \neg C(x).$$

Por tanto, un contraejemplo a una implicación universal debe satisfacer simultáneamente tres condiciones:

$$\boxed{x \in D}$$

$$\boxed{H(x)}$$

$$\boxed{\neg C(x)}.$$

No basta con hacer falsa la conclusión.

También debemos mantener verdaderas las hipótesis bajo las cuales esa conclusión fue prometida.

Ejemplo: refutar una implicación. Consideremos

$$\forall x \in \mathbb{R}, \qquad x^2 = 1 \Rightarrow x = 1.$$

Para construir un contraejemplo necesitamos un real a tal que

$$a^2 = 1$$

y, al mismo tiempo,

$$a \neq 1.$$

Tomemos

$$a = -1.$$

Verificamos el dominio:

$$-1 \in \mathbb{R}.$$

Verificamos la hipótesis:

$$(-1)^2 = 1.$$

Y verificamos el fallo de la conclusión:

$$-1 \neq 1.$$

Por tanto,

$$\boxed{-1 \text{ es un contraejemplo}}.$$

Hemos demostrado que la implicación universal es falsa.

Una conclusión falsa no basta si la hipótesis también falla. Consideremos ahora el enunciado verdadero

$$\forall x \in \mathbb{R}, \qquad x > 0 \Rightarrow x + 1 > 1.$$

Supongamos que alguien propone

$$x = -2$$

como contraejemplo porque

$$-2 + 1 = -1$$

y

$$-1 \not> 1.$$

La conclusión efectivamente es falsa para $x = -2$.

Pero la hipótesis también es falsa:

$$-2 \not> 0.$$

Por tanto, ese objeto no refuta la implicación.

El enunciado sólo prometía

$$x + 1 > 1$$

para los reales que satisfacen

$$x > 0.$$

Como -2 no pertenece a esa subclase, el caso no ataca la promesa realizada.
Esta distinción puede resumirse así:

$$\boxed{\neg C(a)}$$

no basta para refutar

$$H(x) \Rightarrow C(x).$$

Necesitamos

$$\boxed{H(a) \wedge \neg C(a)}.$$

Tres controles para validar un contraejemplo. A partir de los ejemplos anteriores podemos formular un protocolo mínimo.

Ante una afirmación universal, un supuesto contraejemplo a debe pasar tres controles.
Control 1 — dominio.
¿Se cumple

$$a \in D?$$

Control 2 — hipótesis.
Si el enunciado contiene condiciones adicionales, ¿las satisface a?
Control 3 — fallo.
¿La conclusión que el teorema promete es efectivamente falsa para a?
Sólo después de estas tres verificaciones podemos afirmar que el objeto refuta el enunciado.

Un contraejemplo es un testigo, no una sospecha

Para refutar

$$\forall x \in D, \quad P(x),$$

hay que verificar un objeto a tal que

$$a \in D \qquad \text{y} \qquad \neg P(a).$$

Si la universal tiene la forma

$$\forall x \in D, \qquad H(x) \Rightarrow C(x),$$

el contraejemplo debe satisfacer

$$\boxed{a \in D, \qquad H(a), \qquad \neg C(a)}.$$

Un candidato fuera del dominio, un objeto que viola las hipótesis o un caso cuya conclusión no falla no refuta el enunciado.

La asimetría entre casos favorables y desfavorables. La lógica de la cuantificación universal produce una asimetría importante.

Supongamos que consideramos

$$P(n): \quad n < 100$$

para $n \in \mathbb{N}$.

Podemos verificar

$$P(0), P(1), \ldots, P(99).$$

Tenemos cien casos favorables.

Sin embargo, eso no demuestra

$$\forall n \in \mathbb{N}, \quad P(n).$$

El caso

$$n = 100$$

satisface

$$100 \in \mathbb{N}$$

pero

$$100 \not< 100.$$

Por tanto, 100 es un contraejemplo y la universal es falsa.

La diferencia lógica es profunda:

$$\boxed{P(a_1), \ldots, P(a_{100})}$$

sólo habla de cien objetos concretos.

En cambio,

$$\boxed{\forall n \in \mathbb{N}\ P(n)}$$

habla de todos los naturales.

Pero para negar esa universal basta demostrar

$$\boxed{\exists n \in \mathbb{N}\ \neg P(n)}.$$

Por eso una sola excepción decide la cuestión.

Refutar no es lo mismo que fracasar al demostrar. Hay otra distinción que deberá permanecer visible durante todo el capítulo.

Supongamos que intentamos demostrar

$$\forall x \in D, \quad P(x)$$

y nuestro argumento se atasca.

De ese fracaso no podemos concluir

$$\exists x \in D, \quad \neg P(x).$$

Sólo hemos aprendido que una cierta estrategia no produjo una prueba.

Puede ocurrir que:

- la afirmación sea verdadera y necesitemos otra idea;
- la afirmación sea verdadera pero falte una hipótesis intermedia en nuestro argumento;
- la afirmación sea falsa y todavía no hayamos encontrado el contraejemplo;
- el enunciado necesite reformulación.

C07 estudió el diagnóstico de estrategias que no cierran una prueba. C09 añadirá una exigencia nueva: si sospechamos falsedad, debemos **certificarla** mediante un argumento apropiado.

Contraejemplo real y "menor contraejemplo" no son lo mismo. En C08 utilizamos la expresión **menor contraejemplo** dentro de una prueba por buen orden.

Allí suponíamos provisionalmente que existía algún fallo y elegíamos el primero para demostrar que tal fallo era imposible.

El objetivo era concluir que el conjunto de fallos era vacío.

En C09 la situación es distinta.

Un **contraejemplo real** es un objeto que hemos encontrado y verificado y que demuestra que el conjunto de fallos es no vacío.

Podemos representar el contraste así:

$$\begin{aligned} \text{C08} &: \text{“supongamos que existe un primer fallo”} \to \text{contradicción}, \\ \text{C09} &: \text{“aquí hay un fallo concreto”} \to \text{refutación}. \end{aligned}$$

La semejanza terminológica no debe ocultar que las funciones lógicas son opuestas.

El contraejemplo debe atacar exactamente el enunciado. Una buena refutación puede ser muy breve, pero no es imprecisa.

Debe quedar claro:

1. **qué afirmación** se está refutando;
2. **qué objeto** se propone;
3. **por qué pertenece al dominio**;
4. **qué hipótesis satisface**;
5. **qué conclusión hace falsa**.

Una refutación de una línea puede ser completa si contiene todo eso.

Una página entera de cálculos puede ser irrelevante si el objeto elegido no pertenece al dominio o no satisface el antecedente de la implicación.

La longitud no decide la validez.

La correspondencia exacta entre el objeto y la negación lógica sí.

Comprobación. Examina los tres candidatos siguientes.

A. Se quiere refutar

$$\forall x \in \mathbb{R}, \qquad x^2 + 2x + 2 > 1.$$

Se propone

$$x = -1.$$

B. Se quiere refutar

$$\forall n \in \mathbb{Z}, \qquad n^2 \geq n.$$

Se propone

$$n = \frac{1}{2}.$$

C. Se quiere refutar

$$\forall x \in \mathbb{R}, \qquad x > 0 \Rightarrow x + 1 > 1.$$

Se propone

$$x = -2.$$

¿Cuáles de los tres son contraejemplos válidos? Justifica cada respuesta utilizando explícitamente dominio, hipótesis y conclusión.

Respuesta razonada.

En **A**, el candidato sí es un contraejemplo.

Primero,

$$-1 \in \mathbb{R}.$$

Después,

$$(-1)^2 + 2(-1) + 2 = 1 - 2 + 2 = 1.$$

La afirmación exigía

$$1 > 1,$$

lo cual es falso.

Por tanto,

$$\boxed{x = -1 \text{ refuta A}}.$$

En **B**, el cálculo podría sugerir una falla porque

$$\left(\frac{1}{2}\right)^2 = \frac{1}{4} < \frac{1}{2}.$$

Pero el dominio es

$$\mathbb{Z},$$

y

$$\frac{1}{2} \notin \mathbb{Z}.$$

Así, el candidato no pertenece al dominio anunciado y no puede refutar la afirmación.

Por tanto,

$$\boxed{n = \frac{1}{2} \text{ no es contraejemplo a B}}.$$

En **C**, el candidato pertenece al dominio:

$$-2 \in \mathbb{R}.$$

Además, la conclusión

$$x + 1 > 1$$

es falsa en $x = -2$, pues

$$-1 \not> 1.$$

Sin embargo, la hipótesis

$$x > 0$$

es también falsa.

Tenemos

$$-2 \not> 0.$$

Por tanto, el caso no satisface las condiciones bajo las cuales la implicación promete su conclusión.

Así,

$$\boxed{x = -2 \text{ no es contraejemplo a C}}.$$

Sólo el candidato A supera los controles necesarios.

La pregunta que queda abierta. Ya sabemos **qué forma lógica** debe tener un contraejemplo.

Para una universal simple buscamos

$$a \in D \qquad \text{con} \qquad \neg P(a).$$

Para una implicación universal buscamos

$$a \in D, \qquad H(a), \qquad \neg C(a).$$

Pero saber qué necesitamos encontrar no nos dice todavía **dónde buscarlo**.

En problemas sencillos el caso puede saltar a la vista. En otros, el dominio es grande y hay demasiadas posibilidades.

La siguiente sección estudia precisamente esa tarea:

> **¿cómo elegir casos de prueba que tengan una probabilidad razonable de revelar el fallo, sin confundir la búsqueda heurística con la refutación ya demostrada?**

Ése será el contenido de §9.2.

9.2. Cómo buscar un contraejemplo

En §9.1 resolvimos una cuestión lógica: ya sabemos **qué debe verificar** un contraejemplo.

Si la afirmación es

$$\forall x \in D, \quad P(x),$$

buscamos un elemento

$$a \in D$$

para el cual

$$\neg P(a).$$

Y si la afirmación tiene forma condicional,

$$\forall x \in D, \quad H(x) \Rightarrow C(x),$$

necesitamos un caso que satisfaga simultáneamente

$$H(a)$$

y

$$\neg C(a).$$

Pero esta descripción todavía deja abierto un problema práctico.

Si el dominio contiene miles, millones o infinitos objetos, ¿por dónde conviene empezar?

No existe un algoritmo universal que produzca un contraejemplo cada vez que una afirmación es falsa. Si existiera una receta mecánica de ese tipo para toda situación matemática, gran parte del trabajo de descubrimiento desaparecería.

Sí existen, en cambio, **sondas de búsqueda** que suelen revelar rápidamente dónde una afirmación es demasiado fuerte.

La regla será:

buscar de manera dirigida, pero validar de manera lógica.

Una sonda puede sugerir un candidato. Sólo las verificaciones de §9.1 convierten ese candidato en un contraejemplo.

Primero escribimos qué tendría que fallar. Antes de probar valores conviene traducir el enunciado a la forma exacta de su negación.

Supongamos que queremos examinar la afirmación

$$\forall x \in D, \quad H(x) \Rightarrow C(x).$$

La búsqueda no consiste simplemente en preguntar:

«¿para qué x es falsa $C(x)$?»

La pregunta correcta es más exigente:

«¿para qué $x \in D$ se cumple $H(x)$ y falla $C(x)$?»

Podemos escribir una pequeña ficha de búsqueda:

$$\begin{aligned} \text{dominio} &: a \in D, \\ \text{debe cumplirse} &: H(a), \\ \text{debe fallar} &: C(a). \end{aligned}$$

Esta ficha evita perder tiempo con candidatos que jamás podrían refutar el enunciado.

Un valor para el cual $C(a)$ sea falsa pero también lo sea $H(a)$ no sirve. Un valor que satisfaga todo lo deseado pero quede fuera de D tampoco sirve.

La búsqueda comienza, por tanto, **antes del cálculo**: comienza leyendo la obligación lógica.

Casos simples: cero, uno e identidades. Los objetos más simples del dominio suelen ser buenos lugares donde mirar.

En contextos numéricos aparecen con frecuencia

$$0, \quad 1, \quad -1.$$

En contextos conjuntistas pueden aparecer

$$\varnothing$$

o conjuntos con uno o dos elementos.

En contextos funcionales puede ser útil probar funciones constantes, la identidad o transformaciones especialmente sencillas.

¿Por qué?

Porque una afirmación universal debe sobrevivir también a los casos de baja complejidad. Si su formulación ha olvidado una excepción elemental, esos casos tienden a exponerla con poco cálculo.

Pero no debemos convertir esto en una superstición.

Que 0 sea un buen primer lugar donde mirar no significa que toda afirmación falsa falle en 0.

El procedimiento correcto es:

$$\boxed{\text{elegir una sonda simple}} \longrightarrow \boxed{\text{comprobar las condiciones}} \longrightarrow \boxed{\text{decidir si realmente refuta}}.$$

Casos de frontera. Cuando el dominio posee una restricción como

$$x \geq 0,$$

$$|x| \leq 1,$$

o

$$a \leq x \leq b,$$

los puntos de frontera merecen atención especial.

Las afirmaciones con desigualdades estrictas son particularmente sensibles a ellos.

Consideremos, por ejemplo, la afirmación

$$\forall x \in [0,1], \quad x^2 < x.$$

Los valores interiores como

$$x = \frac{1}{2}$$

satisfacen

$$\left(\frac{1}{2}\right)^2 = \frac{1}{4} < \frac{1}{2}.$$

Pero eso no demuestra la universal.

El dominio incluye también sus extremos. Probemos

$$x = 0.$$

Tenemos

$$0 \in [0,1]$$

y

$$0^2 = 0.$$

La conclusión exigiría

$$0 < 0,$$

que es falsa.
Por tanto,

$$\boxed{x = 0 \text{ es un contraejemplo}}.$$

La búsqueda en la frontera no fue una prueba por sí misma. Fue una buena elección de candidato que después quedó certificada.

Signos y regiones cualitativamente distintas. Otra sonda poderosa consiste en preguntar si el comportamiento cambia cuando cambia el **tipo cualitativo** del objeto.
En números reales puede ser relevante separar:

$$x < 0, \qquad x = 0, \qquad x > 0.$$

Consideremos la afirmación

$$\forall x, y \in \mathbb{R}, \qquad x < y \Rightarrow x^2 < y^2.$$

Si sólo ensayamos números positivos, la afirmación parece plausible:

$$1 < 2 \quad \text{y} \quad 1^2 < 2^2.$$

Pero el cuadrado no conserva de la misma manera el orden en la región negativa.
Probemos

$$x = -2, \qquad y = -1.$$

Primero,

$$-2 < -1,$$

así que la hipótesis se cumple.
Sin embargo,

$$(-2)^2 = 4$$

y

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

de modo que

$$4 < 1$$

es falso.
Así,

$$\boxed{(x, y) = (-2, -1) \text{ refuta la afirmación}}.$$

El punto estratégico no fue «probar números negativos porque sí».
Fue advertir que la operación central,

$$x \longmapsto x^2,$$

se comporta de manera diferente en regiones de signo distintas.

Casos mínimos y primeros casos no triviales. En dominios discretos conviene examinar a menudo los primeros objetos que realmente satisfacen las hipótesis.

La palabra **mínimo** se usa aquí en un sentido heurístico: buscamos un caso pequeño o el primer caso no trivial que pueda poner a prueba el enunciado.

No estamos aplicando todavía el argumento de **menor contraejemplo** de C08.

Por ejemplo, supongamos que queremos examinar

$$\forall n \in \mathbb{Z}, \qquad 4 \mid n^2 \Rightarrow 4 \mid n.$$

La ficha lógica dice que un contraejemplo tendría que cumplir

$$4 \mid n^2$$

y

$$4 \nmid n.$$

Probemos valores pequeños.

El caso

$$n = 0$$

satisface la hipótesis, pero también la conclusión, de modo que no refuta nada.

Los valores

$$n = 1, -1$$

ni siquiera satisfacen la hipótesis.

El siguiente candidato natural es

$$n = 2.$$

Tenemos

$$2^2 = 4,$$

por lo que

$$4 \mid 2^2.$$

Pero

$$4 \nmid 2.$$

Entonces

$$\boxed{n = 2 \text{ es un contraejemplo}}.$$

Aquí el examen de casos pequeños fue eficiente porque estuvo filtrado por la hipótesis. No enumeramos enteros sin criterio: buscamos los primeros que podían satisfacer la condición relevante.

Casos degenerados. Un caso **degenerado** es un objeto permitido por las definiciones pero que se encuentra en una situación límite de complejidad: un conjunto vacío, un conjunto unitario, una función constante, dos objetos que coinciden, una longitud nula cuando el dominio la permite, entre otros ejemplos.

Estos casos son útiles porque una afirmación puede haber sido formulada pensando sólo en situaciones «genéricas» y olvidar que sus cuantificadores abarcan también configuraciones límite.

El criterio sigue siendo el mismo:

1. confirmar que el caso degenerado está realmente permitido;
2. comprobar todas las hipótesis;
3. verificar si la conclusión falla.

Si la definición excluye el caso degenerado, utilizarlo no produce un contraejemplo: produce un objeto fuera del problema.

Romper una simetría aparente. Algunas afirmaciones parecen plausibles porque dos objetos desempeñan papeles visualmente parecidos.

Una buena sonda consiste en construir un caso donde esa simetría se rompa.

Consideremos la afirmación:

> Si dos conjuntos A y B tienen intersección no vacía, entonces uno de ellos está contenido en el otro.

En símbolos:

$$A \cap B \neq \varnothing \quad \Longrightarrow \quad A \subseteq B \text{ o } B \subseteq A.$$

En lugar de probar conjuntos al azar, traduzcamos directamente qué necesitaríamos para refutarla.

Queremos simultáneamente:

$$A \cap B \neq \varnothing,$$

$$A \nsubseteq B,$$

y

$$B \nsubseteq A.$$

Estas tres condiciones prácticamente construyen el ejemplo por nosotros.

Para lograr una intersección no vacía necesitamos un elemento común; llamémoslo c.

Para conseguir

$$A \nsubseteq B,$$

necesitamos un elemento a que esté en A pero no en B.

Para conseguir

$$B \nsubseteq A,$$

necesitamos un elemento b que esté en B pero no en A.

Tomemos entonces

$$A = \{a, c\}, \qquad B = \{b, c\},$$

con a, b, c distintos.

Ahora verificamos.

Como

$$c \in A \cap B,$$

tenemos

$$A \cap B \neq \varnothing.$$

Pero

$$a \in A \qquad \text{y} \qquad a \notin B,$$

así que

$$A \nsubseteq B.$$

Del mismo modo,

$$b \in B \qquad \text{y} \qquad b \notin A,$$

por lo que

$$B \nsubseteq A.$$

Por tanto la conclusión es falsa y hemos construido un contraejemplo.

Este ejemplo muestra una estrategia más potente que la enumeración:

$$\boxed{\text{descomponer la negación en requisitos}} \longrightarrow \boxed{\text{construir un objeto que los satisfaga}}.$$

Buscar no siempre significa enumerar. La palabra «buscar» puede sugerir una lista:

$$a_1, a_2, a_3, \ldots$$

hasta que algo falle.

A veces esa exploración es útil, especialmente en dominios pequeños. Pero en matemáticas la búsqueda de un contraejemplo suele ser más conceptual.

Podemos preguntarnos:

- ¿qué condición de la conclusión parece más frágil?;
- ¿qué operación cambia de comportamiento según el signo?;
- ¿hay un extremo incluido en el dominio?;
- ¿qué ocurre en el primer caso no trivial?;
- ¿la afirmación presupone implícitamente que dos objetos son distintos?;
- ¿podemos forzar directamente la negación de la conclusión?;
- ¿qué información exige la hipótesis y cómo podemos conservarla mientras rompemos la conclusión?

Estas preguntas convierten la búsqueda en una exploración dirigida por la estructura del enunciado.

Una sonda fallida no demuestra el teorema. Supongamos que probamos

$$0, 1, -1, 2, -2$$

y todos los casos satisfacen la afirmación.

¿Qué hemos demostrado?

Sólo esas cinco instancias.

No podemos concluir automáticamente que la universal es verdadera.

Tampoco podemos concluir que nuestra estrategia de búsqueda fue mala. Quizá el primer fallo esté lejos, quizá necesite una estructura distinta o quizá la afirmación sí sea verdadera.

Ésta es una disciplina importante:

$$\boxed{\text{no encontrar un contraejemplo}} \neq \boxed{\text{demostrar que no existe}}.$$

El fracaso de una búsqueda modifica nuestro estado de información, pero no decide por sí solo el valor de verdad de la universal.

En §9.8 volveremos sobre esta asimetría al estudiar tablas numéricas y figuras.

Sondas de búsqueda, no certificados de verdad

Para buscar un contraejemplo puede ser útil examinar:

- casos simples y elementos neutros;
- puntos de frontera;
- regiones de signo o comportamiento diferente;
- primeros casos no triviales;
- configuraciones degeneradas permitidas;
- situaciones que rompen una simetría aparente;
- objetos construidos directamente desde la negación.

Estas son **heurísticas de búsqueda**.

Después de encontrar un candidato todavía debemos certificar:

$$\boxed{\text{dominio}} \quad + \quad \boxed{\text{hipótesis}} \quad + \quad \boxed{\text{fallo de la conclusión}}.$$

Si no encontramos ningún candidato, no hemos demostrado todavía la afirmación universal.

Comprobación. Considera la afirmación

$$\forall n \in \mathbb{Z}, \qquad 12 \mid n^2 \Rightarrow 12 \mid n.$$

Un estudiante decide probar, en orden,

$$0, 1, 2, 3, 4, 5, 6, 7, \ldots$$

sin distinguir qué valores satisfacen la hipótesis.

1. Reescribe primero la condición exacta que debe cumplir un contraejemplo.
2. Explica cómo esa condición permite descartar candidatos sin hacer todos los cálculos posibles.
3. Encuentra un contraejemplo pequeño y verifícalo completamente.
4. Explica por qué el hecho de haber encontrado un contraejemplo pequeño no convierte «probar valores pequeños» en un método infalible para toda afirmación universal falsa.

Respuesta razonada.

La afirmación tiene forma condicional. Un contraejemplo debe satisfacer simultáneamente

$$12 \mid n^2$$

y

$$12 \nmid n.$$

Por tanto, cualquier valor para el cual

$$12 \nmid n^2$$

puede descartarse de inmediato: no satisface la hipótesis y no puede refutar la implicación.

Los valores

$$n = 1, 2, 3, 4, 5$$

no proporcionan todavía un contraejemplo. Pero para

$$n = 6$$

tenemos

$$6^2 = 36$$

y

$$12 \mid 36.$$

La hipótesis se cumple.

En cambio,

$$12 \nmid 6.$$

La conclusión es falsa.

Así,

$$\boxed{n = 6 \text{ es un contraejemplo}}.$$

La búsqueda fue eficiente porque no tratamos cada número como igualmente informativo: nos interesaban especialmente aquellos cuyo cuadrado podía ser múltiplo de 12 sin que el propio número lo fuera.

Pero nada de esto convierte la inspección de valores pequeños en una regla universal. Una afirmación falsa puede tener su primer contraejemplo muy lejos de los casos iniciales, o requerir una estructura cualitativamente distinta de la que estamos probando.

Lo que decide la refutación no es que el testigo sea pequeño, sino que satisfaga exactamente

$$12 \mid n^2 \qquad \text{y} \qquad 12 \nmid n.$$

Los casos pequeños fueron una **estrategia de descubrimiento**; la verificación de esas dos condiciones es la **refutación**.

Hacia los condicionales. Hasta aquí hemos tratado la forma condicional como parte de la validación de candidatos, pero todavía no hemos estudiado un error especialmente frecuente.

De

$$P \Rightarrow Q$$

no se sigue en general

$$Q \Rightarrow P.$$

El recíproco puede parecer muy cercano al teorema original y, precisamente por eso, resultar engañoso.

La siguiente sección estudiará cómo refutar recíprocos falsos y cómo distinguirlos de la contraposición legítima.

Ése será el contenido de §9.3.

9.3. Recíprocos falsos y condicionales

En §9.1 vimos que una implicación universal

$$\forall x \in D, \qquad P(x) \Rightarrow Q(x)$$

se refuta mediante un objeto $a \in D$ para el cual

$$P(a)$$

sea verdadera y

$$Q(a)$$

sea falsa.

Ahora necesitamos afinar esa lectura, porque alrededor de una misma implicación aparecen varias afirmaciones formalmente parecidas que no cumplen la misma función lógica.

Partiremos de

$$P \Rightarrow Q.$$

A su alrededor aparecen, entre otras, estas tres expresiones:

$$Q \Rightarrow P,$$

$$\neg Q \Rightarrow \neg P,$$

y

$$P \wedge \neg Q.$$

Conviene aprender a identificarlas sin depender de la semejanza visual.

El recíproco cambia la dirección. El **recíproco** de

$$P \Rightarrow Q$$

es

$$Q \Rightarrow P.$$

La operación parece pequeña: sólo intercambiamos antecedente y consecuente.

Pero ese intercambio puede cambiar el valor de verdad.

Consideremos, para $n \in \mathbb{Z}$, la afirmación

$$8 \mid n \Rightarrow 4 \mid n.$$

Esta implicación es verdadera. Si

$$8 \mid n,$$

entonces existe $k \in \mathbb{Z}$ tal que

$$n = 8k.$$

Pero

$$n = 8k = 4(2k),$$

y como $2k \in \mathbb{Z}$,

$$4 \mid n.$$

Ahora escribamos el recíproco:

$$4 \mid n \Rightarrow 8 \mid n.$$

Para decidir si también es verdadero no podemos invocar automáticamente el resultado anterior. Debemos examinar esta nueva implicación.

Su negación exige un entero n tal que

$$4 \mid n$$

y

$$8 \nmid n.$$

El caso

$$n = 4$$

sirve:

$$4 \mid 4$$

pero

$$8 \nmid 4.$$

Por tanto,

$$\boxed{4 \mid n \Rightarrow 8 \mid n \text{ es falso}}.$$

Hemos demostrado simultáneamente dos cosas distintas:

$$\boxed{8 \mid n \Rightarrow 4 \mid n}$$

es verdadera, mientras que

$$\boxed{4 \mid n \Rightarrow 8 \mid n}$$

es falsa.

El recíproco no viene incluido en el teorema original.

Refutar el recíproco no refuta la implicación original. El ejemplo anterior permite eliminar otra confusión.

El número

$$n = 4$$

es un contraejemplo al recíproco porque satisface

$$4 \mid 4$$

y

$$8 \nmid 4.$$

Pero no es un contraejemplo a la implicación original.

Para refutar

$$8 \mid n \Rightarrow 4 \mid n$$

necesitaríamos exactamente lo contrario:

$$8 \mid n$$

y

$$4 \nmid n.$$

El valor 4 ni siquiera satisface el antecedente original.

Por tanto,

$$\boxed{\text{contraejemplo al recíproco}} \neq \boxed{\text{contraejemplo al enunciado original}}.$$

Cada condicional debe auditarse con sus propias posiciones de antecedente y consecuente.

La contraposición no es el recíproco. La **contraposición** de

$$P \Rightarrow Q$$

es

$$\neg Q \Rightarrow \neg P.$$

A diferencia del recíproco, la contraposición es lógicamente equivalente a la implicación original:

$$\boxed{P \Rightarrow Q \Longleftrightarrow \neg Q \Rightarrow \neg P}.$$

Volvamos al ejemplo:

$$8 \mid n \Rightarrow 4 \mid n.$$

Su contraposición es

$$4 \nmid n \Rightarrow 8 \nmid n.$$

Esto tiene exactamente el mismo contenido lógico.

Si un número no es divisible por 4, entonces no puede ser divisible por 8, porque toda divisibilidad por 8 produciría automáticamente divisibilidad por 4.

Compárese con el recíproco:

$$4 \mid n \Rightarrow 8 \mid n.$$

Las dos afirmaciones comienzan hablando de 4, pero una dice

$$4 \nmid n$$

y la otra

$$4 \mid n.$$

No son variantes intercambiables.

La negación tampoco es el recíproco. Para refutar

$$P \Rightarrow Q$$

no demostramos su recíproco.
Demostramos su **negación**.
Sabemos que

$$\neg(P \Rightarrow Q)$$

es equivalente a

$$P \wedge \neg Q.$$

En una afirmación universal,

$$\forall x \in D, \qquad P(x) \Rightarrow Q(x),$$

la negación completa es

$$\exists x \in D, \qquad P(x) \wedge \neg Q(x).$$

Por tanto, alrededor de un condicional conviene mantener separadas cuatro formas:

$$\begin{array}{rcl} \text{original} & : & P \Rightarrow Q, \\ \text{recíproco} & : & Q \Rightarrow P, \\ \text{contraposición} & : & \neg Q \Rightarrow \neg P, \\ \text{negación} & : & P \wedge \neg Q. \end{array}$$

Sólo la tercera es equivalente al original.
La cuarta describe exactamente el tipo de caso que lo refuta.

Un mismo objeto puede refutar una dirección y ser irrelevante para otra. Consideremos ahora la afirmación sobre números reales

$$x = 3 \Rightarrow x^2 = 9.$$

Es verdadera.
Su recíproco es

$$x^2 = 9 \Rightarrow x = 3.$$

Probemos

$$x = -3.$$

Tenemos

$$(-3)^2 = 9,$$

así que el antecedente del recíproco es verdadero.
Pero

$$-3 \neq 3.$$

Por tanto,

$$\boxed{x = -3 \text{ refuta el recíproco}}.$$

¿Qué ocurre con la contraposición del enunciado original?

Es

$$x^2 \neq 9 \Rightarrow x \neq 3.$$

El valor $x = -3$ no la refuta, porque su antecedente

$$(-3)^2 \neq 9$$

es falso.

Y tampoco refuta el enunciado original, porque allí el antecedente

$$x = 3$$

es falso.

Este ejemplo muestra por qué no basta decir que un objeto «hace algo raro» respecto de una familia de fórmulas.

Hay que preguntar siempre:

> **¿qué condicional concreto estoy tratando de refutar, y cuál es su antecedente?**

El recíproco puede ser verdadero, pero debe demostrarse aparte. Que el recíproco no se siga automáticamente del original no significa que siempre sea falso.

Consideremos

$$x = 0 \Rightarrow x^2 = 0.$$

Su recíproco es

$$x^2 = 0 \Rightarrow x = 0.$$

Sobre $\mathbb{R}$, ambas afirmaciones son verdaderas.

En ese caso podemos reunirlas en un bicondicional:

$$x = 0 \Longleftrightarrow x^2 = 0.$$

Pero la segunda dirección no es verdadera **porque** la primera lo sea.

Necesita su propia justificación.

Ésta es precisamente la estructura de una equivalencia:

$$\boxed{P \Longleftrightarrow Q}$$

significa que debemos tener

$$P \Rightarrow Q$$

y también

$$Q \Rightarrow P.$$

Un contraejemplo a cualquiera de las dos direcciones basta para refutar el bicondicional completo.

Una forma práctica de auditar direcciones. Cuando aparezcan varias versiones cercanas de un condicional, conviene escribir una tabla conceptual antes de probar casos:

afirmación	debe ser verdadero	debe ser falso para refutar
$P \Rightarrow Q$	P	Q
$Q \Rightarrow P$	Q	P
$\neg Q \Rightarrow \neg P$	$\neg Q$	$\neg P$

La tabla no demuestra ninguna de las afirmaciones.

Su función es impedir que busquemos el tipo equivocado de testigo.

Por ejemplo, para refutar el recíproco

$$Q \Rightarrow P$$

debemos buscar

$$Q \wedge \neg P.$$

No debemos buscar

$$P \wedge \neg Q,$$

porque eso refutaría el condicional original.

Un error de razonamiento relacionado. Supongamos que sabemos que

$$P \Rightarrow Q$$

es verdadera y, en un caso concreto, observamos que

$$Q$$

también es verdadera.

¿Podemos concluir P?

No.

Eso equivaldría a utilizar el recíproco

$$Q \Rightarrow P$$

sin haberlo demostrado.

En §9.7 estudiaremos este tipo de inferencia defectuosa dentro de argumentos completos. Por ahora basta identificar la dependencia:

$$P \Rightarrow Q$$

autoriza avanzar de P a Q, pero no regresar de Q a P sin una razón adicional.

No intercambiar direcciones

A partir de

$$P \Rightarrow Q$$

podemos obtener legítimamente su contraposición:

$$\neg Q \Rightarrow \neg P.$$

No obtenemos automáticamente su recíproco:

$$Q \Rightarrow P.$$

Y para refutar el original buscamos

$$P \wedge \neg Q,$$

mientras que para refutar el recíproco buscamos

$$Q \wedge \neg P.$$

Cambiar la dirección cambia también el tipo exacto de contraejemplo que necesitamos.

Comprobación. Considera el enunciado

$$\forall n \in \mathbb{Z}, \qquad 10 \mid n \Rightarrow 5 \mid n.$$

Responde:

1. escribe su recíproco;
2. escribe su contraposición;
3. escribe la condición exacta que debería satisfacer un contraejemplo al enunciado original;
4. escribe la condición exacta que debería satisfacer un contraejemplo al recíproco;
5. decide qué afirmación refuta $n = 5$, si refuta alguna.

Respuesta razonada.

El enunciado original es

$$10 \mid n \Rightarrow 5 \mid n.$$

Su recíproco es

$$5 \mid n \Rightarrow 10 \mid n.$$

Su contraposición es

$$5 \nmid n \Rightarrow 10 \nmid n.$$

Para refutar el enunciado original necesitaríamos un entero n tal que

$$10 \mid n$$

y

$$5 \nmid n.$$

El valor $n = 5$ no cumple la primera condición:

$$10 \nmid 5.$$

Por tanto, no refuta el enunciado original.

Para refutar el recíproco, en cambio, necesitamos

$$5 \mid n$$

y

$$10 \nmid n.$$

El valor $n = 5$ satisface exactamente ambas condiciones:

$$5 \mid 5$$

y

$$10 \nmid 5.$$

Así,

$$\boxed{n = 5 \text{ refuta el recíproco}}.$$

No refuta la contraposición, que es equivalente al enunciado original.

La diferencia no depende de que hayamos cambiado poco los símbolos. Depende de qué condición funciona como antecedente y cuál como conclusión en cada afirmación.

Hacia los cuantificadores. Los condicionales no son el único lugar donde dos enunciados muy parecidos pueden tener fuerzas lógicas distintas.

También ocurre con

$$\forall x \, \exists y \, P(x, y)$$

y

$$\exists y \, \forall x \, P(x, y).$$

El orden de los cuantificadores cambia qué puede depender de qué.

La siguiente sección estudiará cómo detectar y refutar intercambios ilegítimos de cuantificadores y cómo controlar la dependencia de los testigos.

Ése será el contenido de §9.4.

9.4. Cuantificadores intercambiados y testigos

En §9.3 vimos que cambiar la dirección de un condicional cambia la obligación lógica. Con los cuantificadores ocurre algo semejante: cambiar su orden puede transformar completamente lo que una afirmación exige.

Comparemos

$$\forall x \in D \, \exists y \in E, \quad P(x, y)$$

y

$$\exists y \in E \, \forall x \in D, \quad P(x, y).$$

Las dos fórmulas contienen las mismas variables, los mismos dominios, los mismos cuantificadores y el mismo predicado. Sin embargo, en general no dicen lo mismo.

La diferencia está en **cuándo** debe elegirse el testigo existencial.

Testigo dependiente y testigo uniforme. En

$$\forall x \in D \, \exists y \in E, \quad P(x, y),$$

primero se fija un objeto arbitrario x. Después podemos elegir un y adecuado para ese x.

Por eso el testigo puede depender de la entrada:

$$y = y(x).$$

En cambio, en

$$\exists y \in E\, \forall x \in D, \quad P(x, y),$$

el testigo y debe elegirse **antes** de recorrer todos los x.

Ese mismo valor debe funcionar simultáneamente para cada objeto del dominio.

Podemos resumir la diferencia así:

$$\boxed{\forall x\, \exists y : \text{ puede variar el testigo con } x}$$

mientras que

$$\boxed{\exists y\, \forall x : \text{ un solo testigo debe servir para todos los } x}.$$

Esta distinción es una cuestión de dependencia lógica, no de notación superficial.

Un ejemplo donde el orden cambia la verdad. Consideremos enteros n y m y la propiedad

$$P(n, m) : \quad n + m \text{ es par.}$$

Primero estudiemos

$$\forall n \in \mathbb{Z}\, \exists m \in \mathbb{Z}, \qquad n + m \text{ es par.}$$

Esta afirmación es verdadera.

Dado cualquier entero n, podemos elegir

$$m = n.$$

Entonces

$$n + m = n + n = 2n,$$

que es par.

El testigo utilizado fue

$$m = n,$$

y por tanto dependió del valor de n.

Ahora intercambiemos los cuantificadores:

$$\exists m \in \mathbb{Z}\, \forall n \in \mathbb{Z}, \qquad n + m \text{ es par.}$$

Esta afirmación exige mucho más.

Afirma que existe un entero fijo m que, una vez elegido, hace par a $n + m$ para **todo** entero n.

No existe tal m.

En efecto, tomemos cualquier entero m. Si elegimos

$$n = m + 1,$$

entonces

$$n + m = (m + 1) + m = 2m + 1,$$

que es impar.

Por tanto, para cada candidato m podemos construir un entero n que lo derrota.

Así,

$$\boxed{\forall n\, \exists m\ P(n,m)}$$

es verdadera, mientras que

$$\boxed{\exists m\, \forall n\ P(n,m)}$$

es falsa.

El testigo uniforme es una exigencia más fuerte. Si sabemos que

$$\exists y \in E\ \forall x \in D, \quad P(x,y),$$

entonces existe un mismo $y_0 \in E$ tal que

$$P(x, y_0)$$

para todo $x \in D$.

En particular, cuando se fija cualquier x, podemos reutilizar ese mismo y_0 y concluir

$$\exists y \in E, \quad P(x,y).$$

Por tanto,

$$\boxed{\exists y \in E\ \forall x \in D\ P(x,y) \Longrightarrow \forall x \in D\ \exists y \in E\ P(x,y)}.$$

La implicación inversa no vale en general, como mostró el ejemplo de la paridad.

La razón es conceptual: saber resolver el problema **por separado para cada** x no produce automáticamente una única solución que funcione simultáneamente para todos.

Cómo se niega una alternancia de cuantificadores. Para refutar correctamente estas afirmaciones necesitamos negar la fórmula completa.

Partamos de

$$\forall x \in D\ \exists y \in E, \quad P(x,y).$$

Su negación es

$$\exists x \in D\ \forall y \in E, \quad \neg P(x,y).$$

Por tanto, para refutar una afirmación de tipo

$$\forall x \exists y,$$

no basta encontrar un par concreto (a,b) para el cual $P(a,b)$ falle.

Necesitamos un a tal que **ningún** y permitido funcione:

$$\boxed{\exists a \in D\ \forall y \in E, \quad \neg P(a,y)}.$$

Ahora partamos de

$$\exists y \in E\ \forall x \in D, \quad P(x,y).$$

Su negación es

$$\forall y \in E\ \exists x \in D, \quad \neg P(x,y).$$

Por tanto, para refutar una afirmación de tipo

$$\exists y \forall x,$$

no alcanza con encontrar un único par fallido.

Debemos mostrar que **todo** candidato a testigo uniforme puede ser derrotado por algún x:

$$\boxed{\forall y \in E\ \exists x \in D, \quad \neg P(x,y)}.$$

Ésta fue exactamente la estructura usada en el ejemplo anterior: dado cualquier m, elegimos $n = m + 1$ y obtenemos una suma impar.

Un par que falla puede no refutar nada. Volvamos a la afirmación verdadera

$$\forall n \in \mathbb{Z}\ \exists m \in \mathbb{Z}, \qquad n + m \text{ es par.}$$

Consideremos el par

$$(n, m) = (1, 0).$$

Tenemos

$$1 + 0 = 1,$$

que es impar.

Así,

$$P(1, 0)$$

es falsa.

¿Hemos refutado la afirmación universal-existencial?

No.

Para $n = 1$ sólo se prometía que **existiría algún** entero m que funcionara.

Y efectivamente podemos elegir

$$m = 1,$$

pues

$$1 + 1 = 2$$

es par.

El fracaso de un testigo concreto no demuestra la inexistencia de todos los testigos.

Por tanto:

$$\boxed{\neg P(a, b)}$$

no basta para refutar

$$\forall x \exists y\ P(x, y).$$

Necesitamos

$$\boxed{\forall y\ \neg P(a, y)}$$

para algún a.

Refutar una universal-existencial. Consideremos ahora la afirmación

$$\forall n \in \mathbb{Z}\ \exists m \in \mathbb{Z}, \qquad n = 2m.$$

Esta fórmula dice que todo entero es el doble de algún entero.
Para refutarla negamos:

$$\exists n \in \mathbb{Z}\ \forall m \in \mathbb{Z}, \qquad n \neq 2m.$$

Elijamos

$$n = 1.$$

Debemos demostrar ahora que **ningún** entero m satisface

$$1 = 2m.$$

Pero $2m$ es par para todo $m \in \mathbb{Z}$, mientras que 1 es impar.
Así,

$$\forall m \in \mathbb{Z}, \qquad 1 \neq 2m.$$

Por tanto,

$$\boxed{n = 1 \text{ refuta la afirmación}}$$

en el sentido correcto: hemos encontrado un valor del cuantificador universal para el cual todos los posibles testigos existenciales fallan.

Refutar una existencial-universal. Consideremos la afirmación

$$\exists m \in \mathbb{Z}\ \forall n \in \mathbb{Z}, \qquad n - m \text{ es múltiplo de } 3.$$

La fórmula afirma que existe un entero fijo m congruente con **todos** los enteros módulo 3.
Su negación es

$$\forall m \in \mathbb{Z}\ \exists n \in \mathbb{Z}, \qquad n - m \text{ no es múltiplo de } 3.$$

Tomemos un candidato arbitrario m.
Elegimos

$$n = m + 1.$$

Entonces

$$n - m = 1,$$

que no es múltiplo de 3.
Como esta construcción funciona para cualquier m, hemos demostrado

$$\forall m \in \mathbb{Z}\ \exists n \in \mathbb{Z}, \qquad 3 \nmid (n - m).$$

Por tanto, no existe ningún testigo uniforme y la afirmación original queda refutada.

Aquí no hubo un único contraejemplo numérico aislado. El certificado de falsedad fue una **regla que derrota a cada candidato existencial**.

Los dominios también participan en la dependencia. En una fórmula como

$$\forall x \in D\ \exists y \in E, \quad P(x, y),$$

el testigo debe pertenecer a E.

No basta encontrar un objeto fuera de ese conjunto que haga verdadera la relación.

Del mismo modo, al refutar

$$\exists y \in E\ \forall x \in D, \quad P(x, y),$$

cada candidato y que debemos derrotar es un elemento de E, y el objeto x utilizado para derrotarlo debe pertenecer a D.

El control de dominio de §9.1 no desaparece cuando aparecen varios cuantificadores. Se vuelve más importante, porque cada variable puede tener restricciones diferentes.

No todos los intercambios son problemáticos. La advertencia sobre el orden se refiere especialmente a cuantificadores de distinto tipo.

En condiciones ordinarias,

$$\forall x\ \forall y\ P(x, y)$$

y

$$\forall y\ \forall x\ P(x, y)$$

expresan la misma exigencia.

Del mismo modo,

$$\exists x\ \exists y\ P(x, y)$$

y

$$\exists y\ \exists x\ P(x, y)$$

son equivalentes.

La dificultad aparece cuando alternamos

$$\forall \quad \text{y} \quad \exists.$$

Entonces el orden controla la dependencia del testigo.

El orden fija quién puede depender de quién

En

$$\forall x \in D\ \exists y \in E\ P(x, y),$$

el testigo y puede depender de x.

En

$$\exists y \in E\ \forall x \in D\ P(x, y),$$

el mismo y debe funcionar para todos los x.

Para refutar la primera forma buscamos

$$\exists x \in D\ \forall y \in E\ \neg P(x, y).$$

Para refutar la segunda buscamos

$$\forall y \in E\ \exists x \in D\ \neg P(x, y).$$

Un único par (x, y) donde $P(x, y)$ falle puede ser insuficiente en ambos casos.

Comprobación. Considera las dos afirmaciones siguientes:

$$A: \qquad \forall n \in \mathbb{Z}\ \exists m \in \mathbb{Z}, \qquad n + m \text{ es impar},$$

y

$$B: \qquad \exists m \in \mathbb{Z}\ \forall n \in \mathbb{Z}, \qquad n + m \text{ es impar}.$$

Responde:

1. decide si A es verdadera y, si lo es, exhibe una regla para escoger m a partir de n;
2. decide si B es verdadera;
3. si B es falsa, escribe primero su negación y luego demuestra esa negación;
4. explica por qué el par $(n, m) = (0, 0)$, aunque produce una suma par, no refuta por sí solo A.

Respuesta razonada.

La afirmación A es verdadera.

Dado cualquier $n \in \mathbb{Z}$, elegimos

$$m = n + 1.$$

Entonces

$$n + m = n + (n + 1) = 2n + 1,$$

que es impar.

Por tanto,

$$\boxed{A \text{ es verdadera}}.$$

El testigo depende de n.

La afirmación B exige, en cambio, un único entero m que produzca una suma impar con todo entero n.

Su negación es

$$\forall m \in \mathbb{Z}\ \exists n \in \mathbb{Z}, \qquad n + m \text{ es par}.$$

Sea m arbitrario. Elegimos

$$n = m.$$

Entonces

$$n + m = 2m,$$

que es par.

Así hemos demostrado la negación de B, y por tanto

$$\boxed{B \text{ es falsa}}.$$

Finalmente, el par $(0, 0)$ produce

$$0 + 0 = 0,$$

que es par y por ello no satisface la propiedad pedida en A.

Sin embargo, A no afirmaba que **todo** m funcionara para $n = 0$; afirmaba que existía **algún** m.

Para $n = 0$ podemos elegir, por ejemplo,

$$m = 1,$$

pues

$$0 + 1 = 1$$

es impar.

Por eso un testigo fallido no refuta una afirmación existencial.

Hacia el dominio de las operaciones. Hasta aquí los errores que hemos diagnosticado eran lógicos: cambiar una dirección o alterar la dependencia de un testigo.

La siguiente familia de fallos aparece cuando una transformación algebraica se usa fuera de las condiciones que la hacen legítima.

Dividir por una cantidad que puede ser cero, cancelar sin conocer que el factor es no nulo o evaluar una expresión fuera de su dominio no son simples descuidos de notación: cambian qué inferencias están autorizadas.

Ése será el contenido de §9.5.

9.5. Dominio, operaciones ilegítimas y casos excluidos

Hasta aquí hemos diagnosticado errores producidos por la forma lógica de los enunciados: invertir una implicación, confundir su negación o alterar la dependencia de un testigo.

Ahora aparece otra fuente de fallos.

Una cadena algebraica puede parecer impecable línea por línea y, sin embargo, contener una operación que sólo era válida bajo una condición que nunca fue comprobada.

El ejemplo más conocido es dividir por cero, pero el principio es más general.

Toda operación tiene un dominio de legitimidad.

No basta con reconocer el símbolo de una transformación. Debemos comprobar que las condiciones que autorizan esa transformación se cumplen en el caso que estamos estudiando.

El dominio no es información decorativa. Consideremos la expresión

$$\frac{x^2 - 1}{x - 1}.$$

Factorizando el numerador obtenemos

$$\frac{(x - 1)(x + 1)}{x - 1}.$$

Cuando

$$x \neq 1,$$

podemos cancelar el factor $x - 1$ y concluir

$$\frac{x^2 - 1}{x - 1} = x + 1.$$

Pero la igualdad correcta es

$$\boxed{\frac{x^2-1}{x-1} = x+1 \qquad \text{para } x \neq 1}.$$

En $x = 1$ la expresión original contiene

$$\frac{0}{0},$$

que no está definida.

En cambio, la expresión simplificada

$$x + 1$$

sí está definida en $x = 1$.

Por tanto, la simplificación no autoriza a olvidar la restricción original.

Hemos obtenido dos expresiones que coinciden en el dominio donde la primera está definida, pero no dos funciones con el mismo dominio si una se considera sobre todos los reales.

Ésta es una forma elemental de un fenómeno que reaparecerá muchas veces:

$$\boxed{\text{simplificar una expresión no autoriza a ampliar silenciosamente su dominio}}.$$

Dividir exige un divisor no nulo. En un cuerpo como $\mathbb{R}$, de

$$a = b$$

podemos obtener, para $c \neq 0$,

$$\frac{a}{c} = \frac{b}{c}.$$

Pero si $c = 0$, la expresión

$$\frac{a}{c}$$

no está definida.

Así, una línea como

$$ac = bc \quad \Longrightarrow \quad a = b$$

no es válida sin una condición adicional.

La inferencia correcta es

$$\boxed{c \neq 0 \ \text{ y } \ ac = bc \quad \Longrightarrow \quad a = b}.$$

La hipótesis $c \neq 0$ no es un detalle técnico agregado después del razonamiento. Es exactamente lo que permite utilizar la cancelación.

Una cancelación que destruye un caso decisivo. Consideremos la afirmación

$$\forall x \in \mathbb{R}, \qquad (x-2)(x+3) = (x-2)(2x+1) \Rightarrow x+3 = 2x+1.$$

Un argumento apresurado podría decir:

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

por lo tanto cancelamos $x - 2$ y obtenemos

$$x + 3 = 2x + 1.$$

El problema está exactamente en la palabra **cancelamos**.

No sabemos todavía que

$$x - 2 \neq 0.$$

Probemos el caso excluido por esa cancelación:

$$x = 2.$$

Entonces la hipótesis se convierte en

$$0 \cdot 5 = 0 \cdot 5,$$

que es verdadera.

Pero la conclusión exige

$$5 = 5,$$

que en este caso también es verdadera.

Este valor, por tanto, no refuta el enunciado; sólo muestra que el paso de cancelación no estaba justificado en ese punto.

Esto es importante: una operación ilegítima puede invalidar un argumento sin que el enunciado que intentábamos probar sea necesariamente falso.

Para decidir el valor de verdad del enunciado debemos continuar el análisis por separado.

En este caso, si $x \neq 2$ la cancelación es legítima y obtenemos

$$x + 3 = 2x + 1,$$

es decir,

$$x = 2,$$

lo que contradice $x \neq 2$.

Por tanto, ningún valor con $x \neq 2$ satisface la hipótesis. El único caso que puede satisfacerla es $x = 2$, y allí la conclusión también es verdadera.

Así, la afirmación es verdadera, pero la supuesta demostración por cancelación era incompleta.

Este ejemplo separa dos diagnósticos:

$$\boxed{\text{la prueba contiene un paso inválido}}$$

no implica automáticamente

$$\boxed{\text{el enunciado es falso}}.$$

Un paso inválido puede ocultar una afirmación falsa. Modifiquemos ligeramente el ejemplo:

$$\forall x \in \mathbb{R}, \qquad (x-2)(x+3) = (x-2)(2x+4) \Rightarrow x + 3 = 2x + 4.$$

De nuevo, cancelar $x - 2$ sin comprobar que sea no nulo es ilegítimo.

Ahora probemos precisamente el caso problemático:

$$x = 2.$$

La hipótesis es

$$0 \cdot 5 = 0 \cdot 8,$$

por tanto

$$0 = 0,$$

que es verdadera.
La conclusión exige

$$5 = 8,$$

que es falsa.
Así,

$$\boxed{x = 2 \text{ es un contraejemplo}}.$$

Aquí el caso excluido por la cancelación no sólo revela un fallo en el argumento: revela además que el propio enunciado es falso.

La moraleja no es «si una cancelación es dudosa, el teorema es falso».
La moraleja es:

> **cuando una operación excluye un caso, ese caso debe quedar controlado por separado.**

Puede resultar irrelevante, puede necesitar una prueba independiente o puede contener exactamente el contraejemplo que buscamos.

Un mal argumento no vuelve falso un teorema verdadero. Consideremos el resultado elemental

$$xy = 0 \quad \Longrightarrow \quad x = 0 \text{ o } y = 0,$$

para $x, y \in \mathbb{R}$.
Alguien podría escribir:

$$xy = 0.$$

Dividimos por x y obtenemos

$$y = 0.$$

Por tanto,

$$x = 0 \text{ o } y = 0.$$

La conclusión es correcta, pero el argumento no lo es.
La división por x sólo está autorizada si

$$x \neq 0.$$

La prueba correcta separa los casos.
Si

$$x = 0,$$

la disyunción ya es verdadera.

Si

$$x \neq 0,$$

entonces sí podemos dividir

$$xy = 0$$

por x y obtener

$$y = 0.$$

Por tanto, en ambos casos,

$$x = 0 \text{ o } y = 0.$$

Este ejemplo refuerza una distinción que será esencial en §9.7:

$$\boxed{\text{argumento defectuoso}} \neq \boxed{\text{enunciado falso}}.$$

Para demostrar que el enunciado es falso necesitamos un contraejemplo o alguna otra demostración de su negación. Localizar un error sólo establece que **esa prueba** no funciona.

Cancelar es una inferencia condicional. Podemos formular la regla de cancelación de manera que su hipótesis quede visible.

En $\mathbb{R}$,

$$ca = cb$$

y

$$c \neq 0$$

permiten concluir

$$a = b.$$

Si omitimos $c \neq 0$, la inferencia falla.

Por ejemplo,

$$0 \cdot 7 = 0 \cdot 13$$

es verdadera, pero

$$7 = 13$$

es falsa.

Por eso, cada vez que veamos una cancelación conviene preguntar:

¿qué garantiza que el factor cancelado no es cero?

Si la respuesta no aparece en las hipótesis ni se ha demostrado antes, el paso necesita reparación.

Expresiones parcialmente definidas. El mismo problema aparece sin ninguna división explícita dentro de la cadena de prueba.

Consideremos la función real

$$f(x) = \sqrt{x-4}.$$

Para que $f(x)$ sea un número real necesitamos

$$x - 4 \geq 0,$$

es decir,

$$x \geq 4.$$

Por tanto, si trabajamos en los reales, la expresión

$$\sqrt{x-4}$$

no puede evaluarse en $x = 1$.

Un cálculo que sustituya $x = 1$ y continúe manipulando

$$\sqrt{-3}$$

como si fuera un número real ha abandonado el dominio del problema.

Esto no produce un contraejemplo a una afirmación sobre valores permitidos. Produce una expresión fuera del marco declarado.

De forma semejante, si

$$g(x) = \frac{1}{x+5},$$

entonces g no está definida en

$$x = -5.$$

No podemos utilizar $x = -5$ como si fuera una entrada ordinaria de esa función real.

Las hipótesis excluyen casos de manera deliberada. Supongamos que se afirma:

> Para todos $x, y, z \in \mathbb{R}$, si $x \neq 0$ y $xy = xz$, entonces $y = z$.

La forma simbólica es

$$\forall x, y, z \in \mathbb{R}, \qquad (x \neq 0 \land xy = xz) \Rightarrow y = z.$$

El caso

$$x = 0, \qquad y = 2, \qquad z = 7$$

satisface

$$xy = xz,$$

pues ambos lados valen 0.

Además,

$$y \neq z.$$

Sin embargo, no es un contraejemplo al teorema, porque viola la hipótesis

$$x \neq 0.$$

El teorema nunca prometió cancelación cuando $x = 0$.

Esto conecta directamente con el protocolo de §9.1:

$$\boxed{\text{un caso excluido por las hipótesis no puede refutar la conclusión condicional}}.$$

Los casos excluidos sí son útiles para comprender **por qué** la hipótesis está allí. A menudo muestran qué saldría mal si intentáramos eliminarla.

Cambiar una fórmula puede cambiar el problema. Volvamos a

$$\frac{x^2 - 1}{x - 1} = x + 1.$$

La factorización demuestra que ambas expresiones tienen el mismo valor cuando

$$x \neq 1.$$

Pero si reemplazamos la expresión racional por $x + 1$ y después tratamos $x = 1$ como si siempre hubiera estado permitido, hemos cambiado el problema.

Esto puede ocurrir de varias maneras:

- cancelar un factor y olvidar el cero que anulaba ese factor;
- simplificar una fracción y olvidar los ceros del denominador original;
- introducir una raíz sin controlar dónde existe en el dominio elegido;
- usar una función fuera de su dominio declarado;
- aplicar una regla de cancelación sin la hipótesis de no nulidad.

En todos los casos, el error comparte la misma estructura:

$$\boxed{\text{se aplica una transformación más allá del conjunto donde está justificada}}.$$

Un protocolo para auditar operaciones. Cuando una línea de cálculo parezca sospechosa, conviene detenerse antes de continuar y responder cuatro preguntas.

1. ¿Qué operación se realizó?

¿División, cancelación, raíz, evaluación de una función, inversión, sustitución?

2. ¿Qué condición hace válida esa operación?

Por ejemplo:

$$\text{dividir por } c \quad \Longrightarrow \quad c \neq 0.$$

3. ¿Esa condición está en las hipótesis o fue demostrada?

Si no, el paso todavía no está autorizado.

4. ¿Qué casos quedaron excluidos por la operación?

Esos casos deben analizarse por separado. Pueden ser inofensivos o contener precisamente el fallo del enunciado.

Este protocolo transforma una sospecha vaga —«aquí algo no me convence»— en un diagnóstico matemático concreto.

Las operaciones también tienen hipótesis

Una transformación algebraica no se justifica sólo porque sea familiar.

Antes de dividir o cancelar por una expresión E, hay que saber que

$$E \neq 0.$$

Antes de evaluar una función, hay que comprobar que la entrada pertenece a su dominio.

Antes de simplificar una expresión racional, hay que conservar las restricciones heredadas de los denominadores originales.

Cuando una operación excluye ciertos casos, esos casos deben quedar controlados por separado.

Y recordar siempre:

$$\boxed{\text{fallo de una prueba} \neq \text{falsedad del enunciado}}.$$

Comprobación. Analiza los tres razonamientos siguientes.
A. Se afirma, para $x \in \mathbb{R}$:

$$(x-3)(x+2) = (x-3)(2x+1) \Rightarrow x+2 = 2x+1.$$

Un estudiante cancela $x-3$ en ambos lados.
B. Se escribe

$$\frac{x^2-4}{x-2} = x+2$$

«para todo $x \in \mathbb{R}$».
C. Se supone

$$x \neq 0$$

y

$$xy = xz,$$

y se cancela x para concluir

$$y = z.$$

Para cada caso:

1. identifica la operación relevante;
2. indica la condición que la hace legítima;
3. decide si esa condición está disponible;
4. determina qué puede concluirse realmente.

Respuesta razonada.
En **A**, la operación es cancelar el factor

$$x-3.$$

Para que la cancelación sea válida necesitamos

$$x-3 \neq 0,$$

es decir,

$$x \neq 3.$$

Pero esa hipótesis no aparece en el enunciado.
Por tanto, la cancelación no puede aplicarse uniformemente a todos los casos.
Además, el caso excluido debe revisarse.
Si

$$x = 3,$$

la hipótesis vale porque

$$0 \cdot 5 = 0 \cdot 7 = 0.$$

La conclusión, en cambio, exige

$$5 = 7,$$

que es falsa.

Así,

$$\boxed{x = 3 \text{ es un contraejemplo a A}}.$$

El mismo valor revela tanto la operación ilegítima como la falsedad del enunciado.

En **B**, la cancelación de $x - 2$ es válida sólo cuando

$$x \neq 2.$$

En efecto,

$$\frac{x^2 - 4}{x - 2} = \frac{(x - 2)(x + 2)}{x - 2} = x + 2$$

para

$$x \neq 2.$$

Pero en $x = 2$ la expresión original no está definida.

Por tanto, la formulación «para todo $x \in \mathbb{R}$» no conserva el dominio correcto. La igualdad debe entenderse con la restricción

$$\boxed{x \in \mathbb{R} \setminus \{2\}}.$$

En **C**, la condición necesaria para cancelar x es precisamente

$$x \neq 0,$$

y esa hipótesis está disponible desde el comienzo.

Por tanto, la cancelación es legítima y podemos concluir

$$\boxed{y = z}.$$

Los tres ejemplos contienen una operación visualmente parecida. Lo que cambia es si la condición que la autoriza está o no demostrada.

Hacia los pasos irreversibles. Hasta ahora el problema fue que una operación se usó fuera de su dominio de legitimidad.

Hay otra dificultad distinta.

Una transformación puede estar perfectamente definida y ser válida en una dirección, pero no conservar equivalencia en la dirección inversa.

Elevar al cuadrado, por ejemplo, puede transformar una ecuación en otra que admite candidatos adicionales.

Entonces no hay división por cero ni expresión fuera de dominio: el problema es que una flecha lógica fue tratada como si fuera una equivalencia.

Ése será el contenido de §9.6.

9.6. Pasos irreversibles y soluciones espurias

En §9.5 estudiamos operaciones que se utilizaban **fuera de las condiciones que las autorizaban**. Dividir por una cantidad que podía ser cero o evaluar una expresión fuera de su dominio hacía que el paso mismo fuese ilegítimo.

Ahora estudiaremos un problema diferente.

Una transformación puede ser completamente válida y, sin embargo, conservar sólo una dirección lógica.

Es decir, podemos tener

$$A(x) \Rightarrow B(x)$$

sin tener

$$B(x) \Rightarrow A(x).$$

En ese caso, pasar de A a B es correcto. El error aparece cuando después tratamos la transformación como si hubiese sido una equivalencia.

Éste es el mecanismo básico por el que surgen muchas **soluciones espurias**.

Implicación no significa equivalencia. Supongamos que

$$a = b.$$

Entonces podemos elevar ambos miembros al cuadrado y obtener

$$a^2 = b^2.$$

Por tanto,

$$\boxed{a = b \Rightarrow a^2 = b^2}.$$

El paso es perfectamente legítimo.

Pero el recíproco no vale en general.

Por ejemplo,

$$2^2 = (-2)^2,$$

aunque

$$2 \neq -2.$$

Así,

$$a^2 = b^2$$

no permite concluir automáticamente

$$a = b.$$

La información del signo se ha perdido.

Por eso no debemos reemplazar sin justificación

$$a = b \Rightarrow a^2 = b^2$$

por

$$a = b \iff a^2 = b^2.$$

La primera fórmula expresa una transformación válida.

La segunda afirma además que podemos regresar desde el resultado cuadrado hasta la igualdad original.

Eso exige información adicional.

Aplicar una función puede perder información. El fenómeno anterior puede formularse de manera más general.

Si

$$a = b,$$

entonces, para cualquier función f definida en ambos valores,

$$f(a) = f(b).$$

Por tanto,

$$a = b \Rightarrow f(a) = f(b).$$

Pero para invertir esa inferencia necesitamos que f distinga los valores relevantes.

En términos de C04, necesitamos una condición de **inyectividad** sobre el dominio que estamos utilizando.

Si f no es inyectiva, distintos valores pueden producir la misma salida.

El ejemplo

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

sobre $\mathbb{R}$ lo muestra inmediatamente:

$$f(2) = f(-2) = 4.$$

Así, cuando una transformación aplica una función no inyectiva a ambos lados de una ecuación, puede comprimir información.

La igualdad transformada puede tener más soluciones que la original.

Ver las soluciones como conjuntos. Sea

$$S_A = \{x \in D : A(x)\}$$

el conjunto de soluciones de una condición $A(x)$, y sea

$$S_B = \{x \in D : B(x)\}$$

el conjunto de soluciones de una condición $B(x)$.

Si hemos demostrado

$$A(x) \Rightarrow B(x)$$

para todo $x \in D$, entonces

$$\boxed{S_A \subseteq S_B}.$$

Toda solución original sigue siendo solución del problema transformado.

Pero pueden aparecer elementos nuevos en

$$S_B \setminus S_A.$$

Esos elementos son exactamente los **candidatos espurios** creados por la pérdida de equivalencia.

En cambio, si tenemos

$$A(x) \iff B(x),$$

entonces

$$\boxed{S_A = S_B}.$$

Esta interpretación conjuntista ofrece una regla muy útil:

> **una implicación hacia adelante puede ampliar el conjunto de candidatos; una equivalencia conserva exactamente el conjunto de soluciones.**

Primer ejemplo: elevar al cuadrado. Consideremos la ecuación real

$$x + 2 = \sqrt{4x + 13}.$$

Si x satisface la ecuación, podemos elevar ambos miembros al cuadrado:

$$(x + 2)^2 = 4x + 13.$$

Desarrollando,

$$x^2 + 4x + 4 = 4x + 13,$$

de donde

$$x^2 = 9.$$

Por tanto,

$$x = 3 \qquad \text{o} \qquad x = -3.$$

Hasta aquí hemos obtenido **candidatos**.

No hemos demostrado todavía que ambos resuelvan la ecuación original.

Comprobemos.

Para

$$x = 3,$$

tenemos

$$x + 2 = 5$$

y

$$\sqrt{4x + 13} = \sqrt{25} = 5.$$

Por tanto,

$$x = 3$$

sí es solución.

Ahora probemos

$$x = -3.$$

Entonces

$$x + 2 = -1,$$

mientras que

$$\sqrt{4x + 13} = \sqrt{1} = 1.$$

Así,

$$-1 \neq 1.$$

Por tanto,

$$\boxed{x = -3 \text{ es una solución espuria}}.$$

La ecuación cuadrada admitió un candidato que no satisface la ecuación original.
El paso correcto fue

$$x + 2 = \sqrt{4x + 13} \Rightarrow (x + 2)^2 = 4x + 13.$$

No teníamos todavía derecho a escribir

$$x + 2 = \sqrt{4x + 13} \Longleftrightarrow (x + 2)^2 = 4x + 13.$$

Por qué apareció la solución espuria. La raíz cuadrada real satisface

$$\sqrt{4x + 13} \geq 0.$$

Por tanto, cualquier solución de la ecuación original debe cumplir también

$$x + 2 \geq 0.$$

Es decir,

$$x \geq -2.$$

Al elevar al cuadrado, esa información de signo dejó de estar visible.
La ecuación

$$(x + 2)^2 = 4x + 13$$

acepta tanto

$$x = 3$$

como

$$x = -3,$$

porque sólo compara cuadrados.
Pero la ecuación original comparaba una raíz cuadrada —que es no negativa— con $x + 2$.
El candidato -3 viola precisamente esa información perdida.

Esto muestra que una solución espuria no aparece «por magia». Suele indicar **qué información dejó de conservar la transformación**.

Multiplicar también puede ser irreversible. En §9.5 vimos que dividir por una expresión exige comprobar que no sea cero.

Multiplicar, en cambio, siempre está definido en $\mathbb{R}$.

Pero eso no significa que la multiplicación sea siempre reversible.

Consideremos la ecuación

$$x - 4 = 0.$$

Si multiplicamos ambos lados por $x - 2$, obtenemos legítimamente

$$(x - 2)(x - 4) = 0.$$

Así,

$$x - 4 = 0 \Rightarrow (x - 2)(x - 4) = 0.$$

Pero la ecuación transformada tiene dos soluciones:

$$x = 2 \qquad \text{o} \qquad x = 4.$$

El valor

$$x = 2$$

no satisface la ecuación original.

Es otro candidato espurio.

¿Por qué?

Porque multiplicar por $x - 2$ puede convertir una expresión no nula en cero cuando

$$x - 2 = 0.$$

La transformación es válida hacia adelante, pero no puede invertirse uniformemente.

Para recuperar la dirección inversa necesitaríamos saber que

$$x - 2 \neq 0.$$

Entonces sí podríamos cancelar el factor y volver a

$$x - 4 = 0.$$

Así, una misma condición de no nulidad puede aparecer en dos papeles distintos:

- en §9.5, para **autorizar una división o cancelación**;
- en §9.6, para **hacer reversible** una multiplicación que de otro modo puede añadir soluciones.

No toda transformación que cambia la forma pierde equivalencia. También debemos evitar el error opuesto.

Muchas transformaciones sí son reversibles.

Por ejemplo,

$$x + 7 = 12$$

es equivalente a

$$x = 5,$$

porque sumar o restar la misma cantidad en ambos lados puede invertirse sin restricciones adicionales.

Podemos escribir legítimamente

$$x + 7 = 12 \iff x = 5.$$

Del mismo modo,

$$3x = 12 \iff x = 4,$$

porque dividir por el número fijo 3 es legítimo y reversible.

La pregunta no es, por tanto,

> «¿estoy haciendo álgebra?»

sino

> **«¿esta transformación conserva exactamente el mismo conjunto de soluciones?»**

Si la respuesta es sí, podemos usar una equivalencia.

Si sólo sabemos que toda solución anterior produce una solución nueva, debemos usar una implicación.

La notación de las flechas importa. Consideremos una cadena

$$A_0, \quad A_1, \quad A_2, \quad A_3.$$

Si cada paso es reversible, podemos escribir

$$A_0 \iff A_1 \iff A_2 \iff A_3.$$

Entonces cualquiera de las cuatro condiciones tiene exactamente las mismas soluciones.

Pero si sólo sabemos

$$A_0 \Rightarrow A_1$$

y después

$$A_1 \iff A_2 \iff A_3,$$

lo único garantizado es

$$A_0 \Rightarrow A_3.$$

Resolver A_3 produce un conjunto que **contiene** todas las soluciones de A_0, pero puede contener más.

No podemos recorrer la cadena de regreso sin revisar el primer paso.

Por eso escribir símbolos de equivalencia como decoración es peligroso.

Cada

$$\iff$$

afirma dos inferencias.

Resolver después de un paso irreversible. Un paso irreversible no vuelve inútil una resolución.

Puede utilizarse de forma perfectamente rigurosa si controlamos la dirección.

Supongamos que queremos resolver

$$A(x).$$

Realizamos una transformación válida

$$A(x) \Rightarrow B(x),$$

y conseguimos resolver completamente $B(x)$.

Obtenemos una lista de candidatos

$$c_1, \ldots, c_k.$$

Como toda solución de A debe satisfacer B, no hemos perdido soluciones originales.

Ahora verificamos cada candidato en A.

Los que satisfacen A son soluciones reales.

Los que no satisfacen A son espurios.

El procedimiento es:

$$\boxed{\text{ecuación original}} \Rightarrow \boxed{\text{ecuación más débil}} \longrightarrow \boxed{\text{candidatos}} \longrightarrow \boxed{\text{recontrol en la original}}.$$

El último paso no es opcional.

Es lo que restaura la información que la transformación pudo haber perdido.

Una condición adicional puede recuperar la equivalencia. Volvamos a los cuadrados.

En general,

$$a^2 = b^2$$

no implica

$$a = b.$$

Pero si sabemos además que

$$a \geq 0 \qquad \text{y} \qquad b \geq 0,$$

entonces sí podemos concluir

$$a = b.$$

Sobre los reales no negativos, la función

$$x \longmapsto x^2$$

es inyectiva.

Por tanto, bajo esas restricciones,

$$\boxed{a, b \geq 0 \quad \Longrightarrow \quad (a = b \Longleftrightarrow a^2 = b^2)}.$$

Esto sugiere una forma de reparar cadenas que parecen casi equivalentes:

identificar qué información adicional hace reversible el paso.

A veces será una condición de signo.

Otras veces será una condición de no nulidad o una restricción del dominio.

En §9.9 convertiremos esta observación en una estrategia general de reparación.

Solución espuria no significa error de cálculo. Es importante distinguir dos diagnósticos.

Podemos hacer todos los cálculos correctamente y obtener una solución espuria.

En el ejemplo

$$x + 2 = \sqrt{4x + 13},$$

la expansión

$$(x + 2)^2 = x^2 + 4x + 4$$

fue correcta.

La resolución

$$x^2 = 9 \Rightarrow x = \pm 3$$

también fue correcta.

El problema no fue un error aritmético.

El problema fue lógico:

$$\text{ecuación original} \Rightarrow \text{ecuación cuadrada}$$

no era una equivalencia sin condiciones adicionales.

Por eso conviene separar:

$$\boxed{\text{cálculo correcto}}$$

de

$$\boxed{\text{dirección lógica correcta}}.$$

Una cadena puede estar algebraicamente bien ejecutada y, sin embargo, justificar menos de lo que afirma.

Perder soluciones y añadir soluciones son problemas distintos. Si realmente tenemos

$$A(x) \Rightarrow B(x),$$

entonces toda solución de A pertenece al conjunto de soluciones de B.

Por tanto, esa transformación no pierde soluciones originales: puede **añadir candidatos**.

Pero si la supuesta transformación hacia adelante no era válida —por ejemplo, porque dividimos por cero—, ya no podemos asegurar ni siquiera esa inclusión.

Ésa es la frontera entre §§9.5 y 9.6:

§9.5 : el paso puede no estar autorizado,

§9.6 : el paso está autorizado, pero puede no ser reversible.

La diferencia es pequeña en apariencia y decisiva en el diagnóstico.

Un protocolo para auditar una cadena de equivalencias. Cuando una resolución usa varias transformaciones, conviene auditar cada flecha.

1. ¿El paso hacia adelante es válido?

Si no lo es, estamos ante un problema del tipo estudiado en §9.5.

2. ¿El paso inverso también es válido?

Si sí, podemos escribir

$$\Longleftrightarrow .$$

Si no, debemos conservar sólo

$$\Rightarrow .$$

3. ¿Qué información se pierde?

¿Signo, no nulidad, orden, pertenencia a un dominio, distinción entre valores?

4. ¿Puede esa pérdida generar candidatos adicionales?

Si puede, las soluciones de la expresión transformada son sólo candidatos.

5. ¿Se verificaron esos candidatos en la condición original?

Sin esta comprobación, la resolución queda incompleta.

Una flecha no se convierte sola en equivalencia

De

$$A(x) \Rightarrow B(x)$$

sólo podemos concluir

$$S_A \subseteq S_B.$$

Resolver B puede producir candidatos que no pertenecen a S_A.

Para escribir

$$A(x) \Longleftrightarrow B(x)$$

debemos justificar también

$$B(x) \Rightarrow A(x).$$

Si esa dirección no está disponible, hay que recontrolar cada candidato en el problema original.

Una **solución espuria** es precisamente un elemento de

$$S_B \setminus S_A.$$

Comprobación. Analiza las tres situaciones siguientes.

A. Se quiere resolver, en $\mathbb{R}$,

$$x + 2 = \sqrt{4x + 13}.$$

Se elevan ambos lados al cuadrado y se obtienen los candidatos

$$x = 3, \qquad x = -3.$$

B. A partir de

$$x - 4 = 0$$

se multiplica por $x - 2$ y se obtiene

$$(x - 2)(x - 4) = 0.$$

C. Sean $a, b \in \mathbb{R}$ con

$$a \geq 0, \qquad b \geq 0.$$

Se sabe que

$$a^2 = b^2.$$

Para cada caso:

1. indica si el paso utilizado es válido hacia adelante;
2. decide si es reversible bajo las hipótesis disponibles;
3. identifica los candidatos espurios, si existen;
4. escribe la relación lógica correcta usando $\Rightarrow$ o $\Longleftrightarrow$.

Respuesta razonada.

En **A**, elevar al cuadrado es válido hacia adelante:

$$x + 2 = \sqrt{4x + 13} \Rightarrow (x + 2)^2 = 4x + 13.$$

La ecuación transformada produce

$$x^2 = 9,$$

y por tanto los candidatos

$$x = 3, \qquad x = -3.$$

Al recontrolarlos en la ecuación original, $x = 3$ funciona y $x = -3$ no.

Por tanto,

$$\boxed{x = -3 \text{ es espurio}},$$

y la solución original es

$$\boxed{x = 3}.$$

No podemos reemplazar la flecha por una equivalencia sin añadir la información de signo necesaria.

En **B**, multiplicar por $x - 2$ es siempre una operación válida:

$$x - 4 = 0 \Rightarrow (x - 2)(x - 4) = 0.$$

Pero el paso inverso no es válido para todos los reales.

La ecuación transformada tiene soluciones

$$x = 2 \qquad \text{y} \qquad x = 4.$$

El valor $x = 2$ no satisface

$$x - 4 = 0.$$

Así,

$$\boxed{x = 2 \text{ es un candidato espurio}}.$$

La relación correcta sigue siendo una implicación.

En **C**, la situación cambia porque las hipótesis

$$a \geq 0, \qquad b \geq 0$$

hacen inyectiva a la operación de elevar al cuadrado sobre los valores considerados.

De

$$a^2 = b^2$$

podemos concluir

$$a = b.$$

Y, naturalmente,

$$a = b \Rightarrow a^2 = b^2.$$

Por tanto, bajo esas hipótesis,

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

No aparecen soluciones espurias.

Los tres casos muestran que la pregunta correcta no es sólo si una transformación puede ejecutarse, sino si **preserva exactamente la información necesaria para regresar**.

Hacia los argumentos que se prueban a sí mismos. Ya podemos diagnosticar dos grandes familias de fallos.

En §9.5 vimos operaciones utilizadas sin cumplir sus hipótesis.

En §9.6 vimos operaciones legítimas tratadas como si fueran reversibles cuando no lo eran.

Queda una tercera familia: argumentos cuyo problema no está en una operación algebraica particular, sino en su estructura lógica global.

Puede ocurrir que la conclusión se use, directa o indirectamente, para justificar un paso del que luego se pretende derivar esa misma conclusión.

También pueden aparecer inferencias como afirmar el consecuente o negar el antecedente dentro de una cadena aparentemente plausible.

Ése será el contenido de §9.7.

9.7. Circularidad y "nonproofs"

En §§9.5 y 9.6 aprendimos a auditar transformaciones algebraicas.

Primero vimos operaciones que sólo eran legítimas bajo ciertas condiciones. Después vimos transformaciones perfectamente válidas hacia adelante que no podían recorrerse automáticamente en sentido inverso.

Ahora el foco cambia.

Un argumento puede no contener ninguna división por cero, ninguna raíz fuera de dominio y ningún cálculo incorrecto y, sin embargo, no constituir una demostración.

El fallo puede estar en la **estructura de dependencia** entre sus afirmaciones.

Una línea puede apoyarse en algo que todavía no ha sido demostrado. Una conclusión puede reaparecer disfrazada como premisa intermedia. O una regla de inferencia puede utilizarse en una dirección que no autoriza.

Para estos casos utilizaremos de manera amplia la expresión **“nonproof”**: un texto que tiene apariencia de demostración, quizá contiene pasos correctos e incluso puede terminar en una conclusión verdadera, pero no establece todavía esa conclusión mediante una cadena justificada de inferencias.

La tarea no será simplemente ponerle una etiqueta.

Será localizar el primer punto donde se rompe la dependencia lógica.

Un teorema verdadero puede tener una prueba inválida. Ésta es la primera distinción que debemos conservar.

Supongamos que una afirmación T es verdadera.

De eso no se sigue que cualquier argumento que termine escribiendo T sea una prueba de T.

Del mismo modo, si encontramos un error dentro de una demostración propuesta, sólo hemos establecido que **ese argumento** no demuestra el resultado.

No hemos demostrado todavía que el resultado sea falso.

Podemos resumirlo así:

$$\boxed{\text{conclusión verdadera}} \not\Rightarrow \boxed{\text{argumento válido}}$$

y

$$\boxed{\text{argumento inválido}} \not\Rightarrow \boxed{\text{conclusión falsa}}.$$

Esta separación ya apareció en §9.5, pero ahora se vuelve el centro del diagnóstico.

Para refutar un **enunciado** necesitamos demostrar su negación: por ejemplo, mediante un contraejemplo válido cuando se trata de una universal.

Para refutar una **prueba propuesta** basta encontrar una inferencia que no esté justificada.

Son tareas distintas.

Circularidad: usar la conclusión antes de haberla obtenido. La forma más directa de circularidad aparece cuando una prueba utiliza precisamente aquello que pretende demostrar.

Supongamos que queremos probar una afirmación C a partir de ciertas hipótesis H.

Una demostración debe construir una dependencia del tipo

$$H \Longrightarrow A_1 \Longrightarrow A_2 \Longrightarrow \cdots \Longrightarrow C.$$

Pero si en algún momento utilizamos C para justificar uno de los pasos intermedios, la estructura cambia:

$$H \quad + \boxed{C} \Longrightarrow A_1 \Longrightarrow \cdots \Longrightarrow C.$$

El argumento puede terminar correctamente en C, pero sólo porque C ya había sido introducida como información disponible.

No hemos construido una prueba de la conclusión.

La hemos supuesto.

Un ejemplo de circularidad directa. Consideremos el enunciado verdadero

$$\forall n \in \mathbb{Z}, \qquad 8 \mid n \Rightarrow 4 \mid n.$$

Una demostración defectuosa podría escribirse así:

> Supongamos que $8 \mid n$. Como todo múltiplo de 8 es múltiplo de 4, podemos escribir $n = 4m$ para algún $m \in \mathbb{Z}$. Por tanto, $4 \mid n$.

La conclusión es verdadera.

Sin embargo, examinemos la frase

«como todo múltiplo de 8 es múltiplo de 4».

Eso es exactamente lo que se quería demostrar.

El argumento ha utilizado la conclusión universal para justificar la representación

$$n = 4m.$$

Por tanto, la primera dependencia defectuosa aparece **antes** del último renglón.

El problema no es que

$$n = 4m$$

sea una expresión incorrecta para los múltiplos de 4.

El problema es que todavía no teníamos derecho a afirmar que el n considerado era un múltiplo de 4.

La prueba correcta debe extraer esa información de la hipótesis disponible:

si

$$8 \mid n,$$

entonces existe $k \in \mathbb{Z}$ tal que

$$n = 8k.$$

Pero

$$n = 8k = 4(2k),$$

y como $2k \in \mathbb{Z}$, concluimos

$$4 \mid n.$$

La conclusión no cambió.

Cambió la **fuente** de la que dependía.

La circularidad puede quedar escondida en un lema. No siempre aparece una frase tan evidente como «por lo que queremos demostrar».

Supongamos que queremos demostrar un teorema T y escribimos:

$$H \Rightarrow L \Rightarrow T,$$

donde L es un lema auxiliar.

A primera vista la dependencia parece correcta.

Pero si la demostración de L utiliza a su vez el teorema T, tenemos en realidad

$$T \Rightarrow L \Rightarrow T.$$

El círculo puede quedar repartido entre páginas, secciones o resultados con nombres diferentes.

Por eso una auditoría seria no pregunta sólo

«¿cada línea parece razonable?»

sino también

«¿de qué resultados depende esta línea, y estaban ya disponibles sin utilizar el teorema actual?»

Una prueba matemática no es sólo una lista de proposiciones verdaderas.

Es una red de dependencias que debe estar orientada desde lo conocido hacia lo que se quiere establecer.

Afirmar el consecuente. Otra inferencia defectuosa frecuente aparece alrededor de un condicional.

Supongamos que sabemos

$$P \Rightarrow Q.$$

Y en un caso concreto sabemos además que

$$Q$$

es verdadera.

De estas dos informaciones **no** podemos concluir en general

$$P.$$

El patrón inválido es

$$\boxed{P \Rightarrow Q, \qquad Q, \qquad \therefore P}.$$

Esta inferencia se llama **afirmación del consecuente**.

El problema puede verse inmediatamente con divisibilidad.

Sabemos que, para todo $n \in \mathbb{Z}$,

$$6 \mid n \Rightarrow 3 \mid n.$$

Tomemos

$$n = 3.$$

El consecuente es verdadero:

$$3 \mid 3.$$

Pero el antecedente es falso:

$$6 \nmid 3.$$

Por tanto, de

$$6 \mid n \Rightarrow 3 \mid n$$

y

$$3 \mid n$$

no podemos regresar automáticamente a

$$6 \mid n.$$

Eso equivaldría a utilizar el recíproco

$$3 \mid n \Rightarrow 6 \mid n,$$

que §9.3 nos enseñó a distinguir del condicional original.

Una conclusión accidentalmente verdadera no valida la inferencia. La afirmación del consecuente puede resultar especialmente engañosa cuando, para el objeto concreto que estamos estudiando, la conclusión obtenida por la inferencia inválida es casualmente verdadera.

Tomemos

$$n = 12.$$

Tenemos

$$6 \mid 12 \Rightarrow 3 \mid 12,$$

y también

$$3 \mid 12.$$

Si alguien concluye

$$6 \mid 12,$$

ha llegado a una afirmación verdadera.

Pero la regla utilizada sigue siendo inválida.

El caso $n = 3$ muestra que el patrón no preserva verdad en general.

Esto ilustra una idea fundamental:

> **la validez de una inferencia no se decide comprobando que acertó una vez.**

Una inferencia debe estar autorizada por su forma y sus hipótesis, no por el hecho de que en un ejemplo particular haya producido una conclusión verdadera.

Negar el antecedente. Existe un error simétrico.

De

$$P \Rightarrow Q$$

y

$$\neg P$$

no se sigue en general

$$\neg Q.$$

El patrón inválido es

$$\boxed{P \Rightarrow Q, \qquad \neg P, \qquad \therefore \neg Q}.$$

Esta inferencia se llama **negación del antecedente**.

Volvamos a

$$6 \mid n \Rightarrow 3 \mid n.$$

Tomemos ahora

$$n = 9.$$

Tenemos

$$6 \nmid 9,$$

pero también

$$3 \mid 9.$$

Por tanto,

$$6 \nmid n$$

no autoriza a concluir

$$3 \nmid n.$$

El condicional sólo dice qué ocurre **si** se cumple el antecedente.

No afirma que el consecuente pueda ocurrir únicamente de esa manera.

No confundir negación del antecedente con contraposición. La cercanía simbólica puede ocultar la diferencia.

De

$$P \Rightarrow Q$$

sí podemos obtener legítimamente

$$\neg Q \Rightarrow \neg P.$$

Ésta es la contraposición.

Pero no podemos sustituirla por

$$\neg P \Rightarrow \neg Q.$$

Compárese:

$$\begin{aligned} \text{contraposición válida} &: \neg Q \Rightarrow \neg P, \\ \text{negación del antecedente} &: \neg P \Rightarrow \neg Q. \end{aligned}$$

El orden de las negaciones importa tanto como el orden del condicional original.

Demostrar algo verdadero pero demasiado débil. No todos los "nonproofs" contienen una inferencia formalmente inválida del tipo anterior.

A veces cada paso escrito es correcto, pero el argumento termina demostrando **menos** de lo que el objetivo exigía.

Supongamos que queremos demostrar

$$12 \mid n \Rightarrow 6 \mid n.$$

Partimos de

$$12 \mid n.$$

Entonces existe $k \in \mathbb{Z}$ tal que

$$n = 12k.$$

Podemos escribir correctamente

$$n = 12k = 2(6k),$$

y concluir

$$2 \mid n.$$

Todo lo escrito hasta aquí es verdadero.

Pero el objetivo era

$$6 \mid n.$$

Hemos obtenido una consecuencia más débil.

El argumento todavía no termina la demostración.

Para cerrar el objetivo necesitamos, por ejemplo, observar que

$$n = 12k = 6(2k)$$

y que

$$2k \in \mathbb{Z}.$$

Entonces sí obtenemos

$$6 \mid n.$$

La falla del primer argumento no era una falsedad intermedia.

Era un **desajuste entre lo demostrado y lo requerido**.

Podemos expresarlo así:

$$H \Rightarrow R$$

no demuestra automáticamente

$$H \Rightarrow C$$

cuando R es sólo una consecuencia más débil que C.

Debemos justificar además

$$R \Rightarrow C$$

o encontrar otra ruta hasta C.

Una línea puede ser verdadera y aun así estar injustificada. En una demostración no basta preguntar si cada oración es verdadera en algún sentido general.

Debemos preguntar si se sigue de lo que estaba disponible en ese punto.

Supongamos que durante una prueba aparece una afirmación R que sabemos, por conocimiento externo, que es verdadera.

Si la demostración actual no ha establecido R, ni lo tiene como hipótesis, ni puede invocar legítimamente un resultado anterior que lo garantice, entonces R constituye una laguna.

La pregunta correcta es

$$\boxed{\text{¿por qué puedo afirmar esto aquí?}}$$

no simplemente

$$\boxed{\text{¿esto es verdadero?}}.$$

Esta distinción es especialmente importante en argumentos largos, donde una proposición conocida puede introducirse sin registrar la dependencia que la hace utilizable.

El primer paso inválido importa más que el último. Cuando una conclusión es absurda o claramente falsa, existe la tentación de revisar sólo las últimas líneas.

Pero el error puede haber ocurrido mucho antes.

Supongamos una cadena

$$A_0 \Rightarrow A_1 \Rightarrow A_2 \Rightarrow A_3 \Rightarrow A_4.$$

Si el paso

$$A_1 \Rightarrow A_2$$

no está justificado, todo lo que venga después puede ser algebraicamente impecable y seguir sin reparar la prueba.

Por eso el diagnóstico debe localizar **la primera ruptura**.

Esa localización tiene dos ventajas.

Primero, evita atribuir el problema a pasos posteriores que quizá sean perfectamente correctos bajo las nuevas premisas.

Segundo, revela qué información faltaba exactamente cuando se produjo el salto.

Protocolo de diagnóstico de un "nonproof". Ante una demostración sospechosa, podemos trabajar línea por línea.

1. ¿Qué estaba disponible antes del paso?

Registrar las hipótesis, definiciones desplegadas y resultados ya demostrados.

2. ¿Qué se afirmó después?

Escribir con precisión la nueva conclusión introducida por esa línea.

3. ¿Qué regla permitiría pasar de una a otra?

¿Es una equivalencia, una implicación conocida, una definición, una operación algebraica, una regla lógica?

4. ¿Están presentes todas las hipótesis de esa regla?

Si falta alguna, la inferencia todavía no está autorizada.

5. ¿La línea utiliza directa o indirectamente lo que se quiere demostrar?

Si lo hace, aparece circularidad.

6. ¿Se está utilizando un condicional en la dirección correcta?

Comprobar que no aparezca afirmación del consecuente, negación del antecedente o un recíproco no demostrado.

7. ¿Lo obtenido es realmente el objetivo?

Una consecuencia relacionada pero más débil no cierra la prueba.

Este protocolo convierte la lectura crítica en una tarea localizable.

No buscamos una impresión general de que «algo está raro».

Buscamos una dependencia concreta que pueda escribirse y auditarse.

Un mapa de dependencias puede revelar el círculo. En argumentos breves basta seguir las líneas.

En argumentos más largos puede ser útil representar las dependencias como un pequeño grafo.

Supongamos que el objetivo es T y el argumento usa dos resultados auxiliares:

$$H \Rightarrow L_1,$$

$$L_1 \Rightarrow L_2,$$

$$L_2 \Rightarrow T.$$

Esta estructura es legítima si L_1 y L_2 se obtienen sin depender de T.

Pero si descubrimos que

$$T \Rightarrow L_1,$$

entonces el grafo contiene un ciclo:

$$T \Rightarrow L_1 \Rightarrow L_2 \Rightarrow T.$$

El nombre de «lema» no elimina la circularidad.

Lo que importa es la dirección real de dependencia.

Un "nonproof" puede enseñar dónde está la hipótesis ausente. Diagnosticar un argumento defectuoso no es sólo una actividad negativa.

El punto exacto donde la inferencia falla suele indicar qué falta.

Si intentamos pasar de

$$Q$$

a

$$P$$

usando sólo

$$P \Rightarrow Q,$$

el diagnóstico muestra que necesitaríamos justificar además el recíproco

$$Q \Rightarrow P$$

o sustituir la estrategia.

Si una cadena termina en una propiedad demasiado débil, el diagnóstico muestra qué tramo adicional falta hasta el objetivo.

Si un lema depende del teorema que debía sostener, necesitamos una prueba independiente del lema o una reorganización de dependencias.

En §9.9 estudiaremos sistemáticamente esas reparaciones.

Aquí nos detenemos un paso antes:

primero hay que saber exactamente qué se rompió.

Fallo local y falsedad global. Conviene reunir ahora las distinciones de las secciones anteriores.

Un argumento puede fracasar porque:

- una operación no estaba autorizada (§9.5);
- una transformación válida fue tratada como reversible (§9.6);
- una dependencia lógica es circular o utiliza un condicional incorrectamente (§9.7);
- se demuestra una consecuencia que no alcanza el objetivo.

Ninguno de estos diagnósticos, por sí solo, demuestra que el teorema sea falso.

Para concluir falsedad necesitamos una prueba de la negación correspondiente.

En una universal condicional, por ejemplo, debemos volver a la forma

$$\exists x \in D, \qquad H(x) \wedge \neg C(x).$$

Así mantenemos separadas dos preguntas:

$$\boxed{\text{¿esta prueba funciona?}}$$

y

$$\boxed{\text{¿este enunciado es verdadero?}}$$

La respuesta a la primera puede ser «no» mientras la respuesta a la segunda sea «sí».

Un error lógico no se corrige con más cálculo. Cuando el problema es estructural, continuar manipulando expresiones rara vez lo resuelve.

Si hemos afirmado el consecuente, diez líneas adicionales de álgebra no vuelven válida esa inferencia.

Si el argumento es circular, desarrollar más consecuencias de la conclusión supuesta no rompe el círculo.

Si sólo demostramos una propiedad más débil, repetir esa propiedad con otra notación no alcanza el objetivo.

La reparación debe actuar precisamente sobre la dependencia defectuosa.

Esto explica por qué el diagnóstico precede a la reparación.

Sin identificar qué tipo de obligación falló, podemos modificar muchas líneas sin tocar el problema real.

Diagnosticar no es etiquetar

Ante un argumento defectuoso, no basta escribir «circular», «falacia» o «no demuestra nada».

Hay que localizar la primera inferencia no justificada y describirla:

$$\boxed{\text{información disponible}} \not\Rightarrow \boxed{\text{afirmación introducida}}.$$

Preguntamos:

- ¿se utilizó la conclusión antes de demostrarla?;
- ¿se afirmó el consecuente?;
- ¿se negó el antecedente?;
- ¿se invocó un recíproco no demostrado?;
- ¿se obtuvo sólo una consecuencia más débil que el objetivo?;
- ¿un lema depende circularmente del resultado que debía sostener?

Y mantenemos separadas dos tareas:

$$\boxed{\text{invalidar una prueba}} \neq \boxed{\text{refutar un enunciado}}.$$

Comprobación. Analiza los tres argumentos siguientes.

A. Se quiere demostrar, para $n \in \mathbb{Z}$,

$$8 \mid n \Rightarrow 4 \mid n.$$

Se escribe:

> Como todo múltiplo de 8 es múltiplo de 4, tenemos $n = 4m$ para algún entero m. Por tanto, $4 \mid n$.

B. Se sabe que

$$6 \mid n \Rightarrow 3 \mid n.$$

Para cierto entero n se verifica

$$3 \mid n.$$

Se concluye

$$6 \mid n.$$

C. Se quiere demostrar

$$12 \mid n \Rightarrow 6 \mid n.$$

A partir de $12 \mid n$ se escribe

$$n = 12k = 2(6k),$$

por lo que

$$2 \mid n.$$

El argumento termina allí.

Para cada caso:

1. identifica la primera obligación que no ha sido satisfecha;
2. decide si el problema es circularidad, uso inválido de un condicional o conclusión insuficiente;
3. explica si el fallo demuestra que el enunciado es falso;
4. indica qué información faltaría para convertir el argumento en una prueba.

Respuesta razonada.

En **A**, la dificultad aparece en la primera frase.

El argumento afirma

$$8 \mid n \Rightarrow 4 \mid n$$

como justificación para escribir

$$n = 4m.$$

Pero ésa es precisamente la implicación que se quería demostrar.

Por tanto, el argumento es circular.

El enunciado no queda refutado; de hecho, puede demostrarse directamente.

De

$$8 \mid n$$

obtenemos

$$n = 8k = 4(2k)$$

para algún $k \in \mathbb{Z}$, y por tanto

$$4 \mid n.$$

En **B**, la forma disponible es

$$P \Rightarrow Q$$

con

$$P: \ 6 \mid n$$

y

$$Q: \ 3 \mid n.$$

Saber Q y concluir P es afirmar el consecuente.
La inferencia no es válida.
El valor

$$n = 3$$

lo muestra:

$$3 \mid 3$$

pero

$$6 \nmid 3.$$

Este caso refuta el recíproco

$$3 \mid n \Rightarrow 6 \mid n,$$

no el enunciado original.
En **C**, todos los pasos escritos son correctos.
De

$$12 \mid n$$

sí se sigue

$$2 \mid n.$$

Pero el objetivo era

$$6 \mid n.$$

La prueba termina con una consecuencia demasiado débil.
No hay todavía una inferencia que conecte lo demostrado con la conclusión requerida.
Podemos completar el argumento observando directamente que

$$n = 12k = 6(2k),$$

y, como $2k \in \mathbb{Z}$,

$$6 \mid n.$$

Así, los tres argumentos fallan por razones distintas:

$$\begin{aligned} A &: \text{circularidad}, \\ B &: \text{afirmación del consecuente}, \\ C &: \text{conclusión insuficiente}. \end{aligned}$$

En ninguno de los tres casos el mero fallo de la prueba demuestra por sí solo que el enunciado correspondiente sea falso.

Hacia la evidencia numérica y visual. Hasta ahora hemos analizado argumentos que pretenden ser deductivos.

Pero en la práctica matemática también encontramos tablas, cálculos experimentales, diagramas y figuras que sugieren fuertemente que una afirmación es verdadera o falsa.

Esa evidencia puede ser extremadamente útil para descubrir patrones y para localizar contraejemplos.

La dificultad consiste en no atribuirle automáticamente una fuerza lógica que no posee.

La siguiente sección distinguirá entre

exploración, evidencia, contraejemplo y demostración.

Ése será el contenido de §9.8.

9.8. Evidencia numérica y figuras

En §9.7 analizamos argumentos que pretendían ser demostraciones y aprendimos a localizar la primera inferencia que dejaba de estar justificada.

Pero no todo lo que hacemos al trabajar en matemática pretende ser una demostración.

Antes de encontrar una prueba podemos calcular ejemplos, construir tablas, dibujar figuras, representar funciones, probar casos con una computadora o examinar patrones que parecen repetirse.

Todo eso puede ser matemáticamente valioso.

La dificultad aparece cuando confundimos el **valor heurístico** de esa información con su **fuerza deductiva**.

Una tabla puede sugerir una regularidad.

Una figura puede hacer visible una relación que antes no habíamos notado.

Una colección grande de casos favorables puede aumentar nuestra confianza en una conjetura.

Pero una afirmación universal exige algo más: debemos justificar que la propiedad vale para **todos** los objetos cubiertos por el enunciado.

La pregunta de esta sección será, por tanto:

> **¿qué nos autoriza realmente a concluir la evidencia que tenemos delante?**

Cinco funciones que conviene separar. En la práctica aparecen cinco niveles que a veces se mezclan en una sola frase.

Exploración.

Probamos objetos, hacemos cálculos, dibujamos, buscamos regularidades o sometemos una afirmación a casos difíciles.

La exploración tiene una función de descubrimiento.

Todavía no exige que hayamos formulado una afirmación definitiva.

Evidencia.

Llamaremos evidencia a los resultados observados durante esa exploración: valores de una tabla, casos comprobados, configuraciones representadas o patrones repetidos.

La evidencia puede ser exacta o aproximada, abundante o escasa, favorable o desfavorable.

Conjetura.

Una conjetura es una afirmación que proponemos porque la evidencia y la estructura del problema la hacen plausible, pero que todavía no hemos demostrado.

Por ejemplo, después de observar muchos casos en los que se cumple una propiedad $P(n)$, podemos conjeturar

$$\forall n \in \mathbb{N}, \quad P(n).$$

Contraejemplo.

Si la conjetura es universal y encontramos un objeto válido para el cual la propiedad falla, ese objeto no es simplemente nueva evidencia: es una refutación completa.

Un contraejemplo verificado demuestra

$$\exists x \in D, \quad \neg P(x),$$

y por tanto decide la falsedad de

$$\forall x \in D, \quad P(x).$$

Demostración.

Una demostración proporciona una cadena de inferencias que cubre exactamente la obligación lógica del enunciado.

Si queremos establecer una universal, la prueba debe justificar por qué un objeto arbitrario del dominio satisface la propiedad, o utilizar alguna arquitectura equivalente que cubra todos los casos.

Podemos resumir las funciones así:

$$\boxed{\text{explorar}} \longrightarrow \boxed{\text{obtener evidencia}} \longrightarrow \boxed{\text{formular una conjetura}}$$

pero después la conjetura debe someterse a una de dos pruebas decisivas:

$$\boxed{\text{demostración}} \qquad \text{o} \qquad \boxed{\text{contraejemplo}}.$$

La primera establece verdad.

El segundo establece falsedad cuando la afirmación es universal.

Una tabla puede hacer una conjetura muy plausible. Consideremos la expresión

$$p(n) = n^2 + n + 17$$

para $n \in \mathbb{N}$.

Calculemos algunos valores:

n	$p(n)$
0	17
1	19
2	23
3	29
4	37
5	47
6	59
7	73
8	89
9	107
10	127
11	149
12	173
13	199
14	227
15	257

n	$p(n)$

Todos esos números son primos.
La tabla proporciona una regularidad notable.
Sería razonable formular la conjetura

$$\forall n \in \mathbb{N}, \qquad n^2 + n + 17 \text{ es primo.}$$

Pero la tabla no demuestra esa universal.
Sólo ha establecido las dieciséis instancias

$$P(0), P(1), \ldots, P(15).$$

La afirmación universal exige infinitas instancias.
El salto desde

$$P(0) \wedge P(1) \wedge \cdots \wedge P(15)$$

hasta

$$\forall n \in \mathbb{N}, \quad P(n)$$

no está justificado por la cantidad de casos favorables.

Un solo caso desfavorable cambia completamente la situación. Probemos ahora el siguiente valor:

$$n = 16.$$

Tenemos

$$16^2 + 16 + 17 = 256 + 16 + 17 = 289 = 17^2.$$

Por tanto,

$$289$$

no es primo.
Así,

$$\boxed{n = 16 \text{ es un contraejemplo}}.$$

Los dieciséis casos favorables anteriores eran evidencia a favor de una conjetura.
El caso $n = 16$, en cambio, es lógicamente decisivo.
Esto recupera la asimetría central de C09:

$$\boxed{\text{muchos casos favorables} \not\Rightarrow \text{universal verdadera}}$$

mientras que

$$\boxed{\text{un contraejemplo válido} \Rightarrow \text{universal falsa}}.$$

No se trata de que el contraejemplo tenga «más peso estadístico» que los demás casos.
Su fuerza procede de la forma lógica de la negación de una universal.

La cantidad de evidencia no cambia su tipo lógico. Supongamos que una computadora verifica una propiedad para

$$10, \quad 10^3, \quad 10^6$$

casos distintos.

Si el dominio continúa siendo infinito y no hemos demostrado que esos casos agotan alguna clasificación completa, seguimos teniendo evidencia finita.

Aumentar el número de comprobaciones puede ser muy útil para:

- descubrir regularidades;
- localizar excepciones;
- descartar conjeturas falsas;
- comparar formulaciones posibles;
- orientar la búsqueda de una demostración.

Pero no aparece automáticamente una inferencia nueva del tipo

$$\text{“muchos casos”} \Rightarrow \text{“todos los casos”}.$$

La dificultad no está en decidir cuántos casos son «suficientes».

Para una universal infinita, **ningún número fijo de casos favorables es suficiente por sí solo**.

La pregunta correcta es otra:

> **¿qué argumento conecta los casos observados con todos los objetos del dominio?**

Si no hay tal argumento, todavía no hay demostración universal.

Cuando una tabla sí puede ser una demostración. La advertencia anterior no significa que una tabla nunca pueda formar parte de una prueba.

La diferencia está en si los casos mostrados constituyen una **muestra** o una **clasificación exhaustiva**.

Consideremos la afirmación

> para todo entero n, el resto de n^2 al dividir por 4 es 0 o 1.

Todo entero pertenece exactamente a una de las cuatro clases módulo 4:

$$n \equiv 0, 1, 2, 3 \pmod 4.$$

Podemos examinar esas cuatro posibilidades:

$n \pmod 4$	$n^2 \pmod 4$
0	0
1	1
2	0
3	1

Esta tabla tiene una función diferente de la anterior.

No hemos elegido cuatro enteros representativos entre infinitos.

Hemos demostrado que **todo** entero cae en una de cuatro clases y hemos verificado la propiedad en cada clase.

Por tanto, los cuatro casos son exhaustivos respecto de la información relevante.

La estructura lógica es

$$\boxed{\text{todos los objetos caen en una lista finita de casos}}$$

seguida de

$$\boxed{\text{la propiedad se verifica en cada caso}}.$$

Eso es una prueba por casos, no una extrapolación experimental.

Así, debemos distinguir:

$$\boxed{\text{tabla de muestra}} \neq \boxed{\text{tabla exhaustiva}}.$$

La primera produce evidencia.

La segunda puede contener una demostración si la exhaustividad y cada cálculo están justificados.

Una gráfica también depende de qué puntos estamos mirando. La misma dificultad aparece al representar funciones.

Consideremos

$$f(x) = (x-1)(x-2)$$

sobre $\mathbb{R}$.

Supongamos que sólo evaluamos valores enteros cercanos:

x	$f(x)$
-1	6
0	2
1	0
2	0
3	2
4	6

Todos los valores de la tabla satisfacen

$$f(x) \geq 0.$$

Si representáramos únicamente esos puntos, podríamos sentirnos tentados a conjeturar

$$\forall x \in \mathbb{R}, \qquad f(x) \geq 0.$$

Pero el dominio no son los enteros muestreados.

Incluye también los reales situados entre ellos.

Tomemos

$$x = \frac{3}{2}.$$

Entonces

$$f\left(\frac{3}{2}\right) = \left(\frac{1}{2}\right)\left(-\frac{1}{2}\right) = -\frac{1}{4} < 0.$$

Por tanto,

$$\boxed{x = \frac{3}{2} \text{ refuta la conjetura}}.$$

El problema no fue que los cálculos de la tabla estuvieran equivocados.

Todos eran correctos.

El problema fue confundir una colección de puntos del dominio con el dominio completo.

Una representación gráfica con resolución insuficiente puede producir exactamente el mismo efecto: un rasgo importante puede quedar oculto entre los puntos dibujados o por debajo de la escala utilizada.

Una figura puede sugerir un teorema sin demostrarlo. Los diagramas geométricos cumplen una función especialmente poderosa en descubrimiento.

Supongamos que dibujamos un triángulo y trazamos sus tres medianas.

En una figura bien construida parece que las tres rectas concurren en un mismo punto.

La imagen puede sugerir la conjetura:

las tres medianas de todo triángulo son concurrentes.

La conjetura es verdadera.

Pero la apariencia de una figura particular no demuestra todavía la palabra **todo**.

El dibujo representa un triángulo concreto, o a lo sumo una familia limitada por la forma en que lo hemos construido.

Una demostración deberá mostrar que la concurrencia se sigue de las definiciones y propiedades pertinentes para un triángulo arbitrario.

La figura puede ayudar a descubrir esa demostración:

- muestra qué objetos auxiliares conviene considerar;
- permite reconocer simetrías;
- sugiere relaciones entre segmentos;
- hace visible qué punto debería caracterizarse.

Pero la frase

«se ve en el dibujo»

no sustituye la inferencia que falta.

Una figura especial puede introducir hipótesis que nunca fueron dadas. Existe otro peligro.

Queremos estudiar una afirmación sobre un cuadrilátero arbitrario, pero dibujamos sin querer una figura casi rectangular o casi simétrica.

El dibujo puede sugerir propiedades que proceden de esa configuración especial, no de la definición de cuadrilátero.

Por ejemplo, en un paralelogramo las diagonales se bisecan mutuamente.

Si nuestro dibujo de un «cuadrilátero cualquiera» parece un paralelogramo, podríamos concluir visualmente que las diagonales siempre se bisecan.

Pero esa universal es falsa.

La figura ha introducido, de manera silenciosa, una estructura adicional.

Por eso, al usar un diagrama debemos preguntar:

1. ¿qué propiedades pertenecen realmente a las hipótesis?;
2. ¿qué propiedades proceden sólo de cómo quedó dibujada la figura?;
3. ¿estamos usando una relación demostrada o sólo una apariencia de igualdad, perpendicularidad, colinealidad o simetría?

Éste es el análogo visual de un error de dominio: la representación puede hacernos trabajar con un caso más especial que el problema original.

Medir un dibujo no es lo mismo que demostrar una relación. Supongamos que una figura parece mostrar que dos segmentos tienen la misma longitud.

Podemos medirlos en la pantalla y obtener, por ejemplo,

$$4.02 \text{ cm}$$

y

$$4.01 \text{ cm}.$$

Ese resultado puede sugerir que la diferencia se debe al redondeo o a la precisión del dibujo.

Pero no demuestra igualdad exacta.

Del mismo modo, si una medición produce

$$89.9^\circ,$$

no hemos demostrado que el ángulo sea recto.

La medición es información aproximada sobre una representación.

Una afirmación matemática exacta como

$$AB = CD$$

o

$$\angle ABC = 90^\circ$$

requiere una justificación que no dependa sólo de la resolución del instrumento.

Esto no vuelve inútil la medición.

Puede revelar qué relación conviene intentar demostrar.

La función cambia de

$$\boxed{\text{certificar}}$$

a

$$\boxed{\text{orientar}}.$$

Una figura puede participar en una prueba si sus relaciones están justificadas. Tampoco debemos adoptar la regla opuesta de que «las figuras nunca prueban nada».

Una figura puede ser parte legítima de una demostración cuando cada relación utilizada está respaldada por las hipótesis, por una construcción definida o por resultados ya demostrados.

Por ejemplo, si construimos un punto como intersección de dos rectas y después usamos esa incidencia, no estamos concluyendo que las rectas se cortan porque «parece» que se cortan.

La intersección forma parte de la construcción.

Del mismo modo, un diagrama puede organizar una demostración mostrando qué segmentos, ángulos o regiones corresponden a cada paso deductivo.

La frontera correcta no es

$$\text{figura} \quad \text{contra} \quad \text{prueba}.$$

La frontera es

$$\boxed{\text{relación justificada}} \qquad \text{frente a} \qquad \boxed{\text{relación inferida sólo por apariencia}}.$$

Evidencia favorable y evidencia refutatoria no son simétricas. Volvamos a una universal

$$\forall x \in D, \quad P(x).$$

Si una tabla, una figura o un cálculo produce un caso $a \in D$ para el cual

$$P(a)$$

es verdadera, hemos confirmado una instancia.

Si produce cien casos favorables, hemos confirmado cien instancias.

En cambio, si produce un candidato $b \in D$ para el cual parece ocurrir

$$\neg P(b),$$

la situación es diferente.

Ese candidato puede convertirse en una refutación completa si verificamos rigurosamente:

$$\boxed{b \in D} \qquad \text{y} \qquad \boxed{\neg P(b)}.$$

Así, una exploración puede tener dos desenlaces muy distintos:

$$\begin{aligned} \text{caso favorable} &\longrightarrow \text{evidencia favorable,} \\ \text{caso desfavorable verificado} &\longrightarrow \text{contraejemplo.} \end{aligned}$$

La primera línea no cierra una universal.

La segunda sí la cierra, pero negativamente.

La computadora no cambia la lógica. Las herramientas computacionales pueden ampliar enormemente nuestra capacidad de exploración.

Podemos probar millones de casos, dibujar gráficos con gran resolución o buscar configuraciones que sería tedioso examinar a mano.

Pero el hecho de que una verificación sea automática no cambia por sí solo qué se ha demostrado.

Si el programa examina una muestra finita de un dominio infinito, sigue siendo una muestra finita.

Si encuentra un posible contraejemplo, todavía debemos comprobar que:

- corresponde realmente al dominio anunciado;
- las hipótesis se satisfacen;
- el supuesto fallo no procede de redondeo, tolerancias o errores de representación;
- el cálculo relevante es suficientemente exacto para la conclusión que queremos extraer.

Hay contextos en los que una verificación computacional exhaustiva de un dominio finito puede constituir parte de una demostración.

Pero entonces debemos justificar también la exhaustividad y la corrección del procedimiento utilizado.

C09 no desarrollará teoría de pruebas asistidas por computadora. Aquí basta conservar la disciplina lógica:

> **la herramienta cambia cuánto podemos explorar; no cambia qué significa cubrir todos los casos.**

Un protocolo para leer evidencia sin sobreinterpretarla. Cuando una tabla, figura, gráfica o cálculo parezca apoyar una afirmación, conviene responder cinco preguntas.

1. ¿Cuál es exactamente el enunciado que estamos considerando?

¿Es universal, existencial, condicional, finito, infinito?

2. ¿Qué objetos fueron realmente examinados?

¿Todo el dominio, una clasificación exhaustiva o sólo una muestra?

3. ¿La información es exacta o aproximada?

¿Estamos usando igualdades demostradas o mediciones y redondeos?

4. ¿La evidencia favorable cubre la obligación lógica completa?

Si no, produce apoyo para una conjetura, no una demostración.

5. ¿Apareció un caso desfavorable?

Si apareció, debemos volver al protocolo de §9.1 y verificar dominio, hipótesis y fallo de la conclusión.

Este protocolo permite utilizar evidencia de manera fuerte sin atribuirle más de lo que realmente contiene.

Evidencia no es demostración, pero puede decidir falsedad

Una tabla, una figura o una gráfica pueden:

- sugerir una conjetura;
- revelar una estructura útil;
- orientar la búsqueda de una prueba;
- producir candidatos a contraejemplo.

Muchos casos favorables no demuestran por sí solos una universal infinita.

Pero un caso desfavorable, después de ser verificado rigurosamente, sí puede convertirse en un contraejemplo y refutarla.

Una tabla finita puede constituir una prueba sólo cuando se ha demostrado que sus casos son exhaustivos respecto del dominio o de una clasificación completa.

Comprobación. Considera las tres situaciones siguientes.

A. Para $n = 0, 1, \ldots, 15$ se verifica que

$$n^2 + n + 17$$

es primo. Se concluye que la expresión es prima para todo $n \in \mathbb{N}$.

B. Para demostrar que todo cuadrado entero deja resto 0 o 1 módulo 4, se examinan las cuatro clases

$$0, 1, 2, 3 \pmod 4$$

y se calcula el cuadrado de cada una.

C. En un dibujo de un cuadrilátero, sus diagonales parecen cortarse en sus puntos medios. Se concluye que las diagonales de todo cuadrilátero se bisecan mutuamente.

Para cada caso:

1. decide si estamos ante exploración, evidencia, conjetura o demostración;
2. explica qué obligación lógica está o no cubierta;
3. si la conclusión no está justificada, indica cómo podría ponerse a prueba;
4. identifica cualquier diferencia entre una muestra favorable y una clasificación exhaustiva.

Respuesta razonada.

En **A**, tenemos evidencia favorable muy fuerte, pero no una demostración universal.

Los cálculos sólo cubren

$$n = 0, 1, \ldots, 15.$$

La conclusión habla de todos los naturales.

Por tanto, el paso

$$\text{“dieciséis casos favorables”} \Rightarrow \text{“todos los casos”}$$

no está justificado.

De hecho, el siguiente caso produce

$$16^2 + 16 + 17 = 289 = 17^2,$$

por lo que

$$\boxed{n = 16 \text{ es un contraejemplo}}.$$

En **B**, la situación es distinta.

Todo entero tiene exactamente uno de los cuatro restos

$$0, 1, 2, 3$$

módulo 4.

La tabla no es una muestra elegida entre infinitos enteros: representa una clasificación exhaustiva de los casos relevantes.

Al comprobar los cuatro cuadrados obtenemos únicamente los restos

$$0 \quad \text{o} \quad 1.$$

Por tanto, el argumento sí constituye una demostración por casos.

En **C**, la figura proporciona evidencia visual sobre un cuadrilátero concreto, y quizá ni siquiera exacta si sólo estamos juzgando por apariencia.

No hay ninguna razón para que ese dibujo represente todas las formas posibles de cuadrilátero.

Además, puede haber introducido accidentalmente una estructura especial, como la de un paralelogramo.

La conclusión universal no está justificada.

Para ponerla a prueba conviene intentar romper la simetría del dibujo y construir un cuadrilátero que no sea paralelogramo.

Si encontramos uno cuyas diagonales no se bisecan mutuamente, tendremos un contraejemplo a la universal.

Los tres casos muestran tres funciones diferentes de una colección de ejemplos:

$$\begin{aligned} A &: \text{muestra favorable que sugiere una conjetura,} \\ B &: \text{clasificación exhaustiva que permite una prueba,} \\ C &: \text{figura especial cuya apariencia no cubre la universal.} \end{aligned}$$

Hacia la reparación. Después de refutar una afirmación o detectar que un argumento no demuestra lo que pretendía, queda una pregunta constructiva.

¿Qué hacemos con el fallo encontrado?

A veces el contraejemplo revela que el dominio era demasiado amplio.

Otras veces muestra que falta una hipótesis, que la conclusión es demasiado fuerte o que un paso necesita reemplazarse por una inferencia legítima.

La reparación no consiste en cambiar arbitrariamente el problema hasta volverlo verdadero.

Debe responder exactamente al defecto que hemos diagnosticado.

La siguiente sección estudiará cómo restringir un dominio, fortalecer una hipótesis, debilitar una conclusión o sustituir un paso inválido sin modificar más de lo necesario.

Ése será el contenido de §9.9.

9.9. Reparar una afirmación o un argumento falso

Hasta ahora hemos trabajado principalmente en modo diagnóstico.

Un contraejemplo nos permitió demostrar que una afirmación era falsa. Una auditoría de dominio mostró cuándo una operación no estaba autorizada. El control de las flechas reveló pasos irreversibles. Los "nonproofs" nos obligaron a localizar la primera dependencia no justificada. Y la evidencia numérica o visual nos enseñó a separar exploración de demostración.

Pero el trabajo matemático no siempre termina cuando encontramos el fallo.

A menudo comienza una segunda tarea:

> **¿qué modificación exacta convierte el enunciado o el argumento defectuoso en uno correcto?**

Ésta es la tarea de **reparación**.

Reparar no significa sustituir el problema por otro cualquiera que sepamos resolver. Tampoco significa añadir hipótesis hasta que la conclusión resulte trivial.

Una buena reparación debe responder al defecto que hemos identificado y modificar sólo lo necesario para eliminarlo.

Diagnóstico antes de reparación. La regla fundamental es simple:

$$\boxed{\text{no reparar antes de saber qué se rompió}}.$$

Supongamos que tenemos una afirmación condicional

$$\forall x \in D, \qquad H(x) \Rightarrow C(x).$$

Puede ocurrir que el enunciado sea falso. Entonces existe algún

$$a \in D$$

para el cual

$$H(a)$$

es verdadera y

$$C(a)$$

es falsa.

Ese contraejemplo no nos dice automáticamente cuál es la reparación correcta, pero sí impone una condición sobre cualquier versión reparada.

Si sustituimos el dominio D por D', la hipótesis H por H' o la conclusión C por C', el mismo objeto a no puede seguir satisfaciendo simultáneamente

$$a \in D', \qquad H'(a), \qquad \neg C'(a).$$

De lo contrario, seguiría siendo un contraejemplo.

Por tanto, una reparación del enunciado debe hacer al menos una de estas tres cosas:

$$\boxed{a \notin D'}$$

$$\boxed{\neg H'(a)}$$

$$\boxed{C'(a)}.$$

Esto conduce a tres mecanismos naturales:

1. **restringir el dominio**;
2. **fortalecer las hipótesis**;
3. **debilitar la conclusión**.

Cuando el enunciado es verdadero pero la prueba es defectuosa, aparece un cuarto mecanismo:

4. **sustituir el paso inválido por una inferencia legítima**.

Estas cuatro operaciones organizan la sección.

Restringir el dominio. Volvamos a una afirmación que apareció al comienzo del capítulo:

$$\forall x \in \mathbb{R}, \qquad x^2 \geq x.$$

Sabemos que es falsa. Por ejemplo,

$$x = \frac{1}{2}$$

satisface

$$\left(\frac{1}{2}\right)^2 = \frac{1}{4} < \frac{1}{2}.$$

El contraejemplo nos obliga a abandonar la afirmación sobre **todos** los reales si queremos conservar exactamente la misma conclusión.

Preguntemos entonces para qué reales sí se cumple

$$x^2 \geq x.$$

Reescribimos:

$$x^2 - x \geq 0,$$

es decir,

$$x(x-1) \geq 0.$$

El producto es no negativo exactamente cuando

$$x \leq 0$$

o

$$x \geq 1.$$

Por tanto podemos reemplazar el dominio por

$$D' = (-\infty, 0] \cup [1, \infty).$$

La afirmación reparada es

$$\boxed{\forall x \in (-\infty, 0] \cup [1, \infty), \qquad x^2 \geq x}.$$

Aquí no hemos añadido una condición arbitraria.

Hemos identificado exactamente la región donde la desigualdad es verdadera.

El contraejemplo 1/2 queda excluido porque

$$\frac{1}{2} \notin D'.$$

Y la nueva afirmación se demuestra mediante la factorización anterior.

Una restricción puede ser correcta y, sin embargo, excesiva. Podríamos haber reparado la afirmación escribiendo simplemente

$$\forall x \geq 10, \qquad x^2 \geq x.$$

Este enunciado es verdadero.

Pero como reparación del problema original es pobre: elimina una parte enorme de los casos donde la conclusión también era válida.

Esto muestra que la palabra **mínima** debe usarse con cuidado.

No siempre existe una única reparación mínima en un sentido absoluto. Distintas modificaciones pueden ser incomparables.

Pero sí podemos exigir una disciplina:

> **no imponer una restricción más fuerte de lo que nuestra justificación necesita.**

En el ejemplo anterior, el conjunto

$$(-\infty, 0] \cup [1, \infty)$$

no sólo produce un teorema verdadero: describe exactamente el conjunto de reales para los que la desigualdad original se cumple.

Fortalecer una hipótesis. Consideremos ahora el enunciado falso

$$\forall x \in \mathbb{R}, \qquad x^2 = 1 \Rightarrow x = 1.$$

El contraejemplo es

$$x = -1.$$

Tenemos

$$(-1)^2 = 1$$

pero

$$-1 \neq 1.$$

Una posible reparación consiste en conservar el dominio y la conclusión, pero añadir información que elimine la rama negativa:

$$\boxed{\forall x \in \mathbb{R}, \qquad (x^2 = 1 \wedge x \geq 0) \Rightarrow x = 1}.$$

La nueva hipótesis es más fuerte que la anterior.

En efecto,

$$x^2 = 1 \land x \geq 0$$

implica naturalmente

$$x^2 = 1,$$

pero añade una restricción de signo.

Ahora el antiguo contraejemplo queda excluido porque

$$-1 \not\geq 0.$$

Y la reparación puede demostrarse: de

$$x^2 = 1$$

obtenemos

$$x = 1 \qquad \text{o} \qquad x = -1,$$

y la hipótesis $x \geq 0$ elimina la segunda posibilidad.

Fortalecer no significa esconder la conclusión dentro de la hipótesis. Hay una forma trivial de "reparar" casi cualquier condicional falso.

Podríamos escribir

$$\big(H(x) \land C(x)\big) \Rightarrow C(x).$$

Este enunciado es verdadero, pero la conclusión ya fue introducida como hipótesis.

No hemos aprendido nada sobre la relación que queríamos estudiar.

Por eso una reparación mediante hipótesis debe responder al mecanismo del fallo, no limitarse a insertar la conclusión en el antecedente.

En el ejemplo anterior, $x \geq 0$ tiene una función matemática clara: recupera la información de signo que faltaba para distinguir 1 de -1.

Debilitar la conclusión. El mismo enunciado falso admite otra reparación.

De

$$x^2 = 1$$

no podemos concluir siempre

$$x = 1,$$

pero sí podemos concluir

$$|x| = 1.$$

Por tanto otra versión correcta es

$$\boxed{\forall x \in \mathbb{R}, \qquad x^2 = 1 \Rightarrow |x| = 1}.$$

Aquí no hemos excluido al contraejemplo $x = -1$.

Lo hemos conservado dentro del dominio y de la hipótesis, pero hemos reemplazado la conclusión por una afirmación que también es verdadera para él.

En términos lógicos, hemos pasado de una conclusión C a una conclusión C' más débil, de modo que

$$C \Rightarrow C'.$$

En este caso,

$$x = 1 \Rightarrow |x| = 1,$$

pero el recíproco no vale para todos los reales.

La conclusión nueva exige menos.

Una reparación puede no ser única. Las dos modificaciones anteriores reparan la misma afirmación falsa de maneras diferentes:

$$\begin{aligned} \text{fortalecer hipótesis} &: (x^2 = 1 \wedge x \geq 0) \Rightarrow x = 1, \\ \text{debilitar conclusión} &: x^2 = 1 \Rightarrow |x| = 1. \end{aligned}$$

Ambas son verdaderas.

Pero responden a preguntas distintas.

La primera conserva la conclusión original y reduce los casos a los que se aplica.

La segunda conserva todos los casos de la hipótesis original y afirma menos sobre ellos.

Por tanto, la reparación correcta depende de qué parte del problema queramos preservar.

No debemos preguntar simplemente

> «¿cómo hago que esto sea verdadero?»

sino

> **«¿qué parte del enunciado quiero conservar y qué modificación explica mejor el contraejemplo?»**

Reparar una conclusión usando el contraejemplo. Volvamos a otro ejemplo del capítulo:

$$\forall n \in \mathbb{Z}, \qquad 12 \mid n^2 \Rightarrow 12 \mid n.$$

Sabemos que es falso porque

$$n = 6$$

satisface

$$12 \mid 36$$

pero

$$12 \nmid 6.$$

¿Qué propiedad sí está forzada por la hipótesis?

Si

$$12 \mid n^2,$$

entonces en particular

$$4 \mid n^2$$

y

$$3 \mid n^2.$$

De $4 \mid n^2$ se sigue que n es par. En efecto, si n fuera impar, entonces n^2 sería impar.

De $3 \mid n^2$ se sigue que $3 \mid n$. Para comprobarlo elementalmente, basta separar los tres restos posibles módulo 3: si $n \equiv 1$ o $2 \pmod 3$, entonces

$$n^2 \equiv 1 \pmod 3,$$

por lo que 3 no dividiría a n^2.

Así,

$$2 \mid n \qquad \text{y} \qquad 3 \mid n.$$

Como $3 \mid n$, existe $k \in \mathbb{Z}$ tal que

$$n = 3k.$$

Además, $2 \mid n$, de modo que $3k$ es par. Como 3 es impar, k debe ser par; por tanto existe $r \in \mathbb{Z}$ con

$$k = 2r.$$

Así,

$$n = 3k = 6r,$$

y concluimos

$$6 \mid n.$$

Obtenemos la reparación

$$\boxed{\forall n \in \mathbb{Z}, \qquad 12 \mid n^2 \Rightarrow 6 \mid n}.$$

El antiguo contraejemplo $n = 6$ ya no rompe la conclusión:

$$6 \mid 6.$$

Pero la reparación no se acepta sólo porque el contraejemplo haya desaparecido.

La demostración anterior debe cubrir todos los enteros que satisfacen la hipótesis.

El contraejemplo guía, pero no prueba la reparación. Ésta es una precaución esencial.

Supongamos que una afirmación falsa tiene un contraejemplo a.

Modificamos el enunciado de manera que a deje de ser contraejemplo.

¿Qué hemos conseguido?

Sólo hemos eliminado **ese** fallo.

Todavía puede existir otro objeto

$$b \neq a$$

que refute la nueva versión.

Por tanto,

$$\boxed{\text{eliminar el contraejemplo conocido}} \neq \boxed{\text{demostrar el enunciado reparado}}.$$

Una reparación necesita dos fases:

$$\boxed{\text{modificar a partir del diagnóstico}} \longrightarrow \boxed{\text{demostrar la nueva afirmación}}.$$

Reparar una prueba sin cambiar el teorema. Hasta ahora hemos modificado enunciados falsos.

Pero a veces el teorema es verdadero y lo defectuoso es sólo el argumento.

Consideremos

$$\forall n \in \mathbb{Z}, \qquad 15 \mid n \Rightarrow 5 \mid n.$$

Supongamos que alguien escribe:

> Como todo múltiplo de 15 es múltiplo de 5, podemos concluir que $5 \mid n$.

La frase utilizada como justificación es exactamente el resultado que se quiere probar.

El problema es circularidad.

No necesitamos cambiar el teorema.

Necesitamos reemplazar la dependencia defectuosa.

Si

$$15 \mid n,$$

entonces existe $k \in \mathbb{Z}$ tal que

$$n = 15k.$$

Pero

$$n = 15k = 5(3k),$$

y

$$3k \in \mathbb{Z}.$$

Por tanto,

$$5 \mid n.$$

La reparación fue local:

$$\boxed{\text{conclusión supuesta}} \quad \rightsquigarrow \quad \boxed{\text{definición de divisibilidad aplicada a la hipótesis}}.$$

El enunciado permaneció intacto.

Reparar una prueba incompleta por casos. En §9.5 vimos que, a partir de

$$xy = 0,$$

no podemos dividir inmediatamente por x, porque quizá

$$x = 0.$$

El teorema

$$xy = 0 \Rightarrow x = 0 \text{ o } y = 0$$

sobre $\mathbb{R}$ es verdadero.

La reparación correcta no consiste en añadir la hipótesis $x \neq 0$, porque eso cambiaría innecesariamente el teorema.

La estructura adecuada es separar los casos que la operación problemática confundía.
Si

$$x = 0,$$

la conclusión ya se cumple.
Si

$$x \neq 0,$$

entonces la división por x sí está autorizada y de

$$xy = 0$$

obtenemos

$$y = 0.$$

Así, en ambos casos,

$$x = 0 \text{ o } y = 0.$$

Este ejemplo muestra una idea importante:

> **si el fallo de una prueba proviene de una operación válida sólo bajo cierta condición, puede ser mejor separar casos que modificar el enunciado.**

Reparar un paso irreversible. Supongamos que una resolución contiene

$$A(x) \Rightarrow B(x)$$

pero luego trata la flecha como si fuera

$$A(x) \Longleftrightarrow B(x).$$

Hay al menos dos maneras de reparar el argumento.
Primera opción: conservar la implicación y verificar candidatos.
Resolvemos $B(x)$, obtenemos candidatos y los recontrolamos en $A(x)$.
Segunda opción: añadir una condición que haga reversible el paso.
Por ejemplo, para reales no negativos,

$$a^2 = b^2 \Rightarrow a = b,$$

porque la función $x \mapsto x^2$ es inyectiva en $[0, \infty)$.
La elección entre ambas reparaciones depende del problema.
Si podemos verificar fácilmente los candidatos, no siempre hace falta modificar las hipótesis.
Si la condición de reversibilidad es parte natural del contexto, explicitarla puede producir una cadena de equivalencias limpia.

Reparar una dependencia circular. En argumentos largos, la reparación puede exigir reorganizar resultados auxiliares.

Supongamos que pretendíamos demostrar

$$H \Rightarrow L \Rightarrow T,$$

pero descubrimos que la prueba del lema L utilizaba a su vez T.

La dependencia real era

$$T \Rightarrow L \Rightarrow T.$$

Tenemos varias posibilidades:

- demostrar L directamente a partir de resultados anteriores;
- sustituir L por un lema más débil que sí pueda probarse sin T;
- abandonar esa ruta y demostrar T por otra arquitectura;
- si T y L son realmente equivalentes, demostrar ambas direcciones desde una base independiente en lugar de utilizar una para justificar circularmente la otra.

La reparación no consiste en ocultar la referencia a T.

Debe romper la dependencia circular.

Reparar demasiado puede destruir el problema. Consideremos una afirmación falsa

$$\forall x \in D, \qquad H(x) \Rightarrow C(x).$$

Podemos hacerla verdadera mediante una hipótesis exagerada, por ejemplo

$$H(x) \wedge C(x).$$

También podríamos restringir el dominio al conjunto de objetos que ya sabemos que satisfacen la conclusión:

$$D' = \{x \in D : C(x)\}.$$

Formalmente, estas modificaciones producen afirmaciones verdaderas.

Pero no constituyen necesariamente reparaciones matemáticamente informativas.

Han incorporado en la formulación aquello que queríamos explicar.

Por eso conviene distinguir:

$$\boxed{\text{hacer verdadero un enunciado}}$$

no es siempre lo mismo que

$$\boxed{\text{reparar el problema de forma informativa}}.$$

Una buena reparación conserva tanta estructura del problema original como sea razonable y hace visible qué hipótesis, dominio o conclusión era responsable del fallo.

Comparar reparaciones. Cuando varias reparaciones parecen posibles, podemos evaluarlas con cuatro preguntas.

1. ¿Elimina exactamente el fallo diagnosticado?

El contraejemplo conocido debe dejar de refutar la nueva afirmación, o el paso inválido debe desaparecer de la prueba.

2. ¿La nueva afirmación está realmente demostrada?

No basta con que parezca plausible después de excluir el caso problemático.

3. ¿Preserva el mayor contenido relevante del problema original?

Una hipótesis innecesariamente fuerte o una conclusión demasiado débil pueden volver correcto el enunciado a costa de vaciarlo.

4. ¿La modificación puede explicarse matemáticamente?

Debemos poder decir por qué esa condición aparece: signo, no nulidad, pertenencia a un dominio, dependencia de testigos, dirección de una implicación, entre otras posibilidades.

Estas preguntas no producen siempre una única respuesta, pero permiten comparar reparaciones con criterio.

Un protocolo completo de reparación. Podemos reunir el trabajo de C09 en un procedimiento de seis pasos.

1. Localizar el fallo.

¿Tenemos un contraejemplo al enunciado o sólo una inferencia defectuosa en la prueba?

2. Clasificar el defecto.

¿Dominio demasiado amplio? ¿Hipótesis insuficiente? ¿Conclusión excesiva? ¿Operación ilegítima? ¿Paso irreversible? ¿Circularidad? ¿Consecuencia demasiado débil?

3. Elegir la intervención correspondiente.

- restringir dominio;
- fortalecer hipótesis;
- debilitar conclusión;
- sustituir o completar el paso inválido.

4. Recontrolar el caso que reveló el problema.

El contraejemplo o la inferencia defectuosa deben quedar efectivamente neutralizados.

5. Demostrar la versión reparada.

La reparación no se certifica por intuición ni por ausencia del fallo conocido.

6. Auditar la fuerza de la modificación.

Preguntar si hemos cambiado más de lo necesario o si hemos introducido la conclusión dentro de las hipótesis.

Podemos representar el flujo así:

$$\boxed{\text{diagnóstico}} \longrightarrow \boxed{\text{intervención local}}$$

$$\longrightarrow \boxed{\text{nueva demostración}} \longrightarrow \boxed{\text{control de minimalidad}}.$$

Reparar es modificar con una razón

Ante un enunciado falso de la forma

$$\forall x \in D, \qquad H(x) \Rightarrow C(x),$$

una reparación puede:

- reemplazar D por un dominio más apropiado;
- fortalecer H con una hipótesis que elimine el mecanismo del contraejemplo;
- debilitar C hasta una conclusión realmente forzada por las hipótesis.

Ante un teorema verdadero con una prueba defectuosa, la primera opción debe ser reparar la **inferencia**, no cambiar innecesariamente el enunciado.

Y en todos los casos:

$$\boxed{\text{eliminar un fallo conocido}} \neq \boxed{\text{haber demostrado la reparación}}.$$

Comprobación. Analiza las cuatro situaciones siguientes.

A. Se afirma

$$\forall x \in \mathbb{R}, \qquad x^2 \geq x.$$

Se conoce el contraejemplo $x = 1/2$.

B. Se afirma

$$\forall x \in \mathbb{R}, \qquad x^2 = 1 \Rightarrow x = 1.$$

Se conoce el contraejemplo $x = -1$.

C. Se quiere demostrar

$$15 \mid n \Rightarrow 5 \mid n,$$

pero la prueba dice «todo múltiplo de 15 es múltiplo de 5» como justificación principal.

D. Se quiere demostrar

$$xy = 0 \Rightarrow x = 0 \text{ o } y = 0$$

para reales x, y, y el argumento divide inmediatamente por x.

Para cada caso:

1. decide si debe repararse el enunciado, la prueba o ambos;
2. identifica el defecto exacto;
3. propone una reparación controlada;
4. justifica que la reparación elimina el fallo;
5. explica por qué la modificación no es meramente arbitraria.

Respuesta razonada.

En **A**, el problema está en el enunciado.

La universal sobre todos los reales es falsa.

Resolver

$$x^2 \geq x$$

equivale a resolver

$$x(x-1) \geq 0,$$

lo que produce

$$x \leq 0 \qquad \text{o} \qquad x \geq 1.$$

Por tanto una reparación exacta del dominio es

$$\boxed{\forall x \in (-\infty, 0] \cup [1, \infty), \qquad x^2 \geq x}.$$

El valor $1/2$ queda excluido y la factorización demuestra la nueva versión.

En **B**, el enunciado también es falso, pero admite más de una reparación razonable.

Si queremos conservar la conclusión $x = 1$, podemos fortalecer la hipótesis:

$$\boxed{(x^2 = 1 \wedge x \geq 0) \Rightarrow x = 1}.$$

La condición $x \geq 0$ elimina exactamente la rama $x = -1$.

Si queremos conservar todos los casos de $x^2 = 1$, podemos debilitar la conclusión:

$$\boxed{x^2 = 1 \Rightarrow |x| = 1}.$$

Las dos reparaciones son correctas, pero preservan partes distintas del enunciado original.

En **C**, el teorema es verdadero y no necesita modificarse.

El fallo es circular: la justificación utilizada coincide con lo que se quiere demostrar.

La reparación consiste en volver a la hipótesis

$$15 \mid n.$$

Existe $k \in \mathbb{Z}$ tal que

$$n = 15k = 5(3k),$$

y por tanto

$$5 \mid n.$$

Hemos sustituido la dependencia circular por una aplicación directa de la definición de divisibilidad.

En **D**, el teorema también es verdadero.

La división por x sólo está autorizada cuando

$$x \neq 0.$$

No conviene añadir $x \neq 0$ al teorema, porque perderíamos innecesariamente el caso $x = 0$.

La reparación natural es separar casos.

Si $x = 0$, la conclusión ya se cumple.

Si $x \neq 0$, dividimos legítimamente $xy = 0$ por x y obtenemos $y = 0$.

Por tanto,

$$x = 0 \text{ o } y = 0.$$

Los cuatro ejemplos resumen las cuatro intervenciones principales:

A : restringir dominio,
B : fortalecer hipótesis o debilitar conclusión,
C : sustituir una dependencia circular,
D : completar la prueba mediante una separación de casos.

Hacia los ejercicios integrados. Ya disponemos de las herramientas conceptuales centrales de C09.

Podemos:

- construir y verificar contraejemplos;
- buscar casos de manera dirigida;
- distinguir recíproco, contraposición y negación;
- controlar cuantificadores y dependencia de testigos;
- auditar dominios y operaciones;
- reconocer pasos irreversibles;
- localizar circularidad e inferencias no autorizadas;
- separar evidencia de demostración;
- reparar enunciados y argumentos sin cambiar más de lo necesario.

La última sección del capítulo reunirá estas habilidades en problemas donde no siempre estará anunciado de antemano si el enunciado es verdadero, falso o simplemente está acompañado por una prueba defectuosa.

Ése será el contenido de §9.10.

9.10. Ejercicios y soluciones desarrolladas

Esta sección reúne las habilidades del capítulo en una secuencia donde la estrategia no siempre viene anunciada de antemano. Los ejercicios están organizados en tres familias:

1. **reconocimiento y búsqueda guiada**: leer la obligación lógica exacta, validar contraejemplos, controlar recíprocos, cuantificadores, dominios y evidencia;
2. **diagnóstico y reparación**: localizar la primera dependencia defectuosa, distinguir enunciado falso de prueba inválida y producir reparaciones matemáticamente justificadas;
3. **integración y transferencia**: decidir si conviene demostrar, refutar, reformular o reparar, y explicar por qué la intervención elegida resuelve exactamente el problema detectado.

En varios ejercicios no se indica de antemano si la afirmación es verdadera o falsa. Parte de la tarea consiste precisamente en decidir qué clase de trabajo exige el enunciado.

Ejercicios.

E1. Un contraejemplo debe satisfacer la negación completa. Considera la afirmación

$$\forall n \in \mathbb{Z}, \qquad 18 \mid n^2 \Rightarrow 18 \mid n.$$

1. Escribe la negación completa del enunciado.
2. Decide si $n = 6$ es un contraejemplo válido.
3. Verifica por separado dominio, hipótesis y fallo de la conclusión.
4. Explica por qué $n = 3$ no sirve, aunque $18 \nmid 3$.

E2. Original, recíproco, contraposición y negación. Considera

$$\forall n \in \mathbb{Z}, \qquad 14 \mid n \Rightarrow 7 \mid n.$$

1. Escribe el recíproco.
2. Escribe la contraposición.
3. Escribe la negación del enunciado original.
4. Decide si el recíproco es verdadero o falso.
5. Explica qué afirmación refuta $n = 7$ y cuál no refuta.

E3. Testigo dependiente y testigo uniforme. Compara las afirmaciones

$$A: \qquad \forall n \in \mathbb{Z}\ \exists m \in \mathbb{Z}, \qquad n + m = 0,$$

y

$$B: \qquad \exists m \in \mathbb{Z}\ \forall n \in \mathbb{Z}, \qquad n + m = 0.$$

1. Decide si A es verdadera y exhibe una regla para escoger m.
2. Decide si B es verdadera.
3. Si B es falsa, escribe y demuestra su negación.
4. Explica por qué el par $(n, m) = (1, 0)$ no refuta por sí solo A.

E4. Cancelación ilegítima y caso decisivo. Se afirma, para $x \in \mathbb{R}$,

$$(x - 2)(x + 1) = 0 \Rightarrow x = -1.$$

Un estudiante escribe:

> Como $(x - 2)(x + 1) = 0$, cancelamos $x - 2$ y obtenemos $x + 1 = 0$. Por tanto, $x = -1$.

1. Localiza la primera operación no justificada.
2. Indica qué condición permitiría realizar esa cancelación.
3. Decide si el enunciado original es verdadero o falso.
4. Si es falso, exhibe un contraejemplo.

5. Propón una reparación mínima del enunciado y demuéstrala.

E5. Muestra finita frente a clasificación exhaustiva. Una computadora verifica para

$$n = 0, 1, 2, \ldots, 100$$

que

$$n^2 + n$$

es par.

1. Explica por qué esos 101 casos no constituyen por sí solos una demostración de
$$\forall n \in \mathbb{Z}, \qquad n^2 + n \text{ es par.}$$
2. Demuestra la afirmación para todo entero.
3. Reescribe tu prueba como una clasificación exhaustiva en los dos casos n par / n impar.
4. Explica por qué esa clasificación sí cubre la obligación universal completa.

E6. Un paso válido que introduce una solución espuria. Resuelve en $\mathbb{R}$ la ecuación

$$\sqrt{x+6} = x.$$

1. Determina las restricciones que debe satisfacer toda solución antes de elevar al cuadrado.
2. Eleva al cuadrado y encuentra todos los candidatos.
3. Recontrola los candidatos en la ecuación original.
4. Identifica cualquier solución espuria.
5. Escribe correctamente la relación lógica entre la ecuación original y la ecuación cuadrada usando $\Rightarrow$ o $\Longleftrightarrow$.

E7. Afirmar el consecuente y reparar la inferencia. Sabemos que, para todo $n \in \mathbb{Z}$,

$$10 \mid n \Rightarrow 5 \mid n.$$

Un argumento dice:

> Como $5 \mid n$, concluimos que $10 \mid n$.

1. Identifica la inferencia inválida.
2. Da un contraejemplo a esa inferencia.
3. Explica por qué el contraejemplo no refuta el enunciado original.
4. Repara el enunciado añadiendo una hipótesis elemental sobre paridad que haga verdadera la conclusión $10 \mid n$.
5. Demuestra la versión reparada.

E8. Una prueba que se detiene demasiado pronto. Se quiere demostrar

$$\forall n \in \mathbb{Z}, \qquad 20 \mid n \Rightarrow 5 \mid n.$$

Un estudiante parte de $20 \mid n$ y concluye correctamente

$$2 \mid n.$$

Después da la prueba por terminada.

1. Explica por qué todos los pasos escritos pueden ser correctos y, sin embargo, la prueba estar incompleta.
2. Identifica el desajuste entre lo demostrado y lo requerido.
3. Completa la demostración de la forma más directa posible.

E9. Circularidad escondida en una igualdad de conjuntos. Se quiere demostrar:

Si $A \subseteq B$, entonces $A \cap B = A$.

Se propone el argumento:

> Como $A \cap B = A$, todo elemento de A pertenece a $A \cap B$. Además, todo elemento de $A \cap B$ pertenece a A. Por tanto, $A \cap B = A$.

1. Localiza la circularidad.
2. Explica por qué la segunda inclusión del argumento sí puede justificarse sin circularidad.
3. Reescribe una prueba completa mediante doble inclusión.
4. Indica exactamente dónde se utiliza la hipótesis $A \subseteq B$.

E10. Dos reparaciones distintas de una misma afirmación falsa. Considera

$$\forall x \in \mathbb{R}, \qquad x^2 + 1 > 2x.$$

1. Decide si la afirmación es verdadera o falsa.
2. Si es falsa, encuentra un contraejemplo.
3. Produce una reparación debilitando la conclusión.
4. Produce otra reparación restringiendo el dominio sin cambiar el signo estricto $>$.
5. Demuestra ambas versiones reparadas y compara qué cambió en cada una.

E11. Elegir entre demostración, contraejemplo y clasificación exhaustiva. Decide si cada afirmación es verdadera o falsa. Si es verdadera, demuéstrala. Si es falsa, refútala con un contraejemplo.

A.

$$\forall n \in \mathbb{Z}, \qquad 10 \mid n^2 \Rightarrow 10 \mid n.$$

B.

$$\forall n \in \mathbb{Z}, \qquad 8 \mid n^2 \Rightarrow 8 \mid n.$$

C.

$$\forall n \in \mathbb{Z}, \qquad 9 \mid n^2 \Rightarrow 3 \mid n.$$

En las afirmaciones verdaderas puedes utilizar clases de residuos finitas, siempre que justifiques que la clasificación es exhaustiva.

E12. Una afirmación $\forall\exists$ y otra $\exists\forall$. Considera

$$A: \qquad \forall x \in \mathbb{R}\ \exists y \in \mathbb{R}, \qquad y > x,$$

y

$$B: \qquad \exists y \in \mathbb{R}\ \forall x \in \mathbb{R}, \qquad y > x.$$

1. Decide el valor de verdad de A y B.
2. Para la afirmación verdadera, da una regla explícita para el testigo.
3. Para la afirmación falsa, escribe su negación y demuéstrala.
4. Explica qué cambia en la dependencia del testigo cuando se intercambian los cuantificadores.

E13. Una gráfica muestreada puede ocultar el contraejemplo. Sea

$$f(x) = \left(x - \frac{1}{2}\right)^2 - \frac{1}{16}, \qquad x \in \mathbb{R}.$$

Un programa evalúa $f(k)$ para muchos enteros k y encuentra siempre valores positivos. A partir de ello se conjetura

$$\forall x \in \mathbb{R}, \qquad f(x) \geq 0.$$

1. Explica qué tipo de evidencia proporcionan los valores enteros.
2. Busca un contraejemplo real que no sea entero.
3. Verifica rigurosamente el contraejemplo.
4. Explica por qué una gráfica dibujada sólo a partir de puntos enteros podría ocultar el fallo.
5. Determina exactamente para qué valores de x se cumple $f(x) < 0$.

E14. Del diagnóstico a la reparación mínima. Considera la afirmación

$$\forall a, b, c \in \mathbb{R}, \qquad ca = cb \Rightarrow a = b.$$

1. Decide si es verdadera o falsa.
2. Si es falsa, da un contraejemplo que haga visible la condición ausente.
3. Explica por qué cancelar c no está autorizado en general.
4. Formula una reparación mínima mediante una hipótesis adicional.
5. Demuestra la versión reparada.
6. Explica por qué una reparación como «supongamos además que $a = b$» sería matemáticamente inútil aunque volviera verdadero el enunciado.

E15. Integración final: refutar, diagnosticar y reparar. Se propone la afirmación

$$\forall x, y \in \mathbb{R}, \qquad xy = 1 \Rightarrow x + y \geq 2.$$

1. Examina algunos casos simples y formula una sospecha razonable sobre su valor de verdad.
2. Decide finalmente si la afirmación es verdadera o falsa y justifica tu respuesta.
3. Si es falsa, identifica qué rasgo del contraejemplo revela el fallo de la formulación.
4. Repara la afirmación fortaleciendo las hipótesis, sin cambiar la conclusión $x + y \geq 2$, y demuéstrala.
5. Produce una segunda reparación que conserve el dominio completo $\mathbb{R}$ y la hipótesis $xy = 1$, pero modifique la conclusión.
6. Demuestra también esta segunda reparación.
7. Compara las dos reparaciones y explica por qué ninguna de ellas es simplemente «borrar el contraejemplo».

Soluciones desarrolladas.

Solución E1. El enunciado es

$$\forall n \in \mathbb{Z}, \qquad 18 \mid n^2 \Rightarrow 18 \mid n.$$

La negación de una universal condicional es

$$\exists n \in \mathbb{Z}, \qquad 18 \mid n^2 \quad \text{y} \quad 18 \nmid n.$$

Probemos $n = 6$.

Primero,

$$6 \in \mathbb{Z}.$$

Después,

$$6^2 = 36$$

y

$$18 \mid 36.$$

Por tanto, la hipótesis del condicional se cumple.
En cambio,

$$18 \nmid 6.$$

Así,

$$\boxed{n = 6 \text{ es un contraejemplo válido}}.$$

El valor $n = 3$ no sirve. Es cierto que

$$18 \nmid 3,$$

pero también

$$3^2 = 9$$

y

$$18 \nmid 9.$$

Por tanto, $n = 3$ no satisface el antecedente del condicional. Un contraejemplo debe hacer verdadero el antecedente y falsa la conclusión.

Solución E2. El enunciado original es

$$14 \mid n \Rightarrow 7 \mid n.$$

Su **recíproco** es

$$7 \mid n \Rightarrow 14 \mid n.$$

Su **contraposición** es

$$7 \nmid n \Rightarrow 14 \nmid n.$$

Y la negación del enunciado original es

$$14 \mid n \quad \text{y} \quad 7 \nmid n$$

para algún $n \in \mathbb{Z}$.
El recíproco es falso. Tomemos

$$n = 7.$$

Entonces

$$7 \mid 7,$$

pero

$$14 \nmid 7.$$

Así,

$$\boxed{n = 7 \text{ refuta el recíproco}}.$$

No refuta el enunciado original, porque allí el antecedente exige

$$14 \mid n,$$

y ese antecedente es falso para $n = 7$.
El original sí es verdadero: si $14 \mid n$, existe $k \in \mathbb{Z}$ con

$$n = 14k = 7(2k),$$

y por tanto $7 \mid n$.
La contraposición tiene el mismo valor de verdad que el original.
Solución E3. La afirmación

$$A: \quad \forall n \in \mathbb{Z}\ \exists m \in \mathbb{Z}, \quad n + m = 0$$

es verdadera.
Dado un entero arbitrario n, elegimos

$$m = -n.$$

Entonces $m \in \mathbb{Z}$ y

$$n + m = n - n = 0.$$

El testigo depende de n.
La afirmación

$$B: \quad \exists m \in \mathbb{Z}\ \forall n \in \mathbb{Z}, \quad n + m = 0$$

es falsa.
Su negación es

$$\forall m \in \mathbb{Z}\ \exists n \in \mathbb{Z}, \qquad n + m \neq 0.$$

Sea $m \in \mathbb{Z}$ arbitrario. Elegimos

$$n = 1 - m.$$

Entonces

$$n + m = (1 - m) + m = 1 \neq 0.$$

Por tanto, todo candidato a testigo uniforme m puede ser derrotado por algún n.
Finalmente, el par $(1, 0)$ satisface

$$1 + 0 \neq 0,$$

pero eso no refuta A. Para $n = 1$ sólo se exigía que existiera **algún** m adecuado. El valor

$$m = -1$$

sí funciona.

Solución E4. La operación problemática es la cancelación del factor

$$x - 2.$$

De

$$(x-2)(x+1) = 0$$

podemos escribir

$$(x-2)(x+1) = (x-2) \cdot 0.$$

Cancelar $x - 2$ requiere

$$x - 2 \neq 0,$$

es decir,

$$x \neq 2.$$

Esa hipótesis no aparece en el enunciado original.

Probemos precisamente el caso excluido:

$$x = 2.$$

Entonces

$$(x-2)(x+1) = 0 \cdot 3 = 0,$$

así que la hipótesis se cumple.

Pero la conclusión exige

$$2 = -1,$$

que es falsa.

Por tanto,

$$\boxed{x = 2 \text{ es un contraejemplo}},$$

y el enunciado original es falso.

Una reparación mínima consiste en afirmar:

$$\forall x \in \mathbb{R}, \qquad \big(x \neq 2 \land (x-2)(x+1) = 0\big) \Rightarrow x = -1.$$

Ahora sí, como $x \neq 2$, tenemos $x - 2 \neq 0$ y podemos cancelar:

$$(x-2)(x+1) = (x-2) \cdot 0 \Rightarrow x + 1 = 0.$$

Por tanto,

$$\boxed{x = -1}.$$

La reparación añade exactamente la hipótesis que faltaba para autorizar la inferencia.

Solución E5. Verificar los valores

$$0, 1, \ldots, 100$$

sólo demuestra 101 instancias.

La afirmación

$$\forall n \in \mathbb{Z}, \qquad n^2 + n \text{ es par}$$

habla de infinitos enteros. No existe una inferencia automática

$$\text{“101 casos favorables”} \Rightarrow \text{“todos los enteros”}.$$

La afirmación sí es verdadera. Observemos que

$$n^2 + n = n(n + 1).$$

Los enteros n y $n + 1$ son consecutivos, y entre dos enteros consecutivos uno es par. Por tanto, el producto

$$n(n + 1)$$

es par.

Así,

$$\boxed{n^2 + n \text{ es par para todo } n \in \mathbb{Z}}.$$

Podemos expresar la misma prueba mediante una clasificación exhaustiva.

Si n es par, existe $k \in \mathbb{Z}$ tal que

$$n = 2k.$$

Entonces

$$n^2 + n = 2k(2k + 1),$$

que es par.

Si n es impar, existe $k \in \mathbb{Z}$ tal que

$$n = 2k + 1.$$

Entonces

$$\begin{aligned} n^2 + n &= (2k + 1)^2 + (2k + 1) \\ &= 4k^2 + 6k + 2 \\ &= 2(2k^2 + 3k + 1), \end{aligned}$$

que también es par.

La diferencia con la muestra computacional es que **todo entero** es par o impar. Los dos casos agotan el dominio y, por tanto, la clasificación es exhaustiva.

Solución E6. Partimos de

$$\sqrt{x+6} = x.$$

Para que la raíz esté definida en $\mathbb{R}$ necesitamos

$$x + 6 \geq 0.$$

Además, como una raíz cuadrada real es no negativa, toda solución debe satisfacer

$$x \geq 0.$$

Esta segunda condición es más fuerte que $x \geq -6$.
Ahora elevamos al cuadrado. Toda solución original satisface

$$x + 6 = x^2.$$

Así,

$$x^2 - x - 6 = 0,$$

y factorizando,

$$(x-3)(x+2) = 0.$$

Los candidatos son

$$x = 3 \qquad \text{o} \qquad x = -2.$$

Revisemos la ecuación original.
Para $x = 3$,

$$\sqrt{3+6} = \sqrt{9} = 3,$$

por lo que 3 sí es solución.
Para $x = -2$,

$$\sqrt{-2+6} = \sqrt{4} = 2 \neq -2.$$

Por tanto,

$$\boxed{x = -2 \text{ es una solución espuria}}.$$

La única solución es

$$\boxed{x = 3}.$$

La relación lógica correcta es

$$\sqrt{x+6} = x \Rightarrow x + 6 = x^2.$$

Sin incorporar la condición $x \geq 0$, no debemos reemplazar esa flecha por una equivalencia.

Solución E7. La forma disponible es

$$10 \mid n \Rightarrow 5 \mid n.$$

El argumento toma el consecuente

$$5 \mid n$$

y concluye el antecedente

$$10 \mid n.$$

Eso es **afirmar el consecuente**, o equivalentemente utilizar el recíproco sin haberlo demostrado.

El valor

$$n = 5$$

es un contraejemplo a esa inferencia:

$$5 \mid 5,$$

pero

$$10 \nmid 5.$$

No refuta el enunciado original porque $10 \nmid 5$ hace falso el antecedente original.

Una reparación natural es

$$\forall n \in \mathbb{Z}, \qquad (5 \mid n \wedge 2 \mid n) \Rightarrow 10 \mid n.$$

Demostrémosla. Supongamos

$$5 \mid n$$

y

$$2 \mid n.$$

De $5 \mid n$ existe $k \in \mathbb{Z}$ tal que

$$n = 5k.$$

Como n es par y 5 es impar, k debe ser par: si k fuese impar, el producto $5k$ sería impar.

Por tanto existe $r \in \mathbb{Z}$ tal que

$$k = 2r.$$

Entonces

$$n = 5k = 5(2r) = 10r.$$

Así,

$$\boxed{10 \mid n}.$$

La reparación no declara verdadero el recíproco original; añade exactamente la información de paridad que faltaba.

Solución E8. De

$$20 \mid n$$

se sigue ciertamente que

$$2 \mid n.$$

Por ejemplo, si

$$n = 20k,$$

entonces

$$n = 2(10k).$$

Así, el razonamiento puede ser correcto hasta ese punto.

Pero el objetivo era

$$5 \mid n,$$

no

$$2 \mid n.$$

El problema es una **conclusión insuficiente**: se demostró una consecuencia verdadera de la hipótesis, pero no la consecuencia pedida.

La demostración directa se completa inmediatamente. De

$$20 \mid n$$

existe $k \in \mathbb{Z}$ tal que

$$n = 20k.$$

Como

$$20k = 5(4k)$$

y $4k \in \mathbb{Z}$, obtenemos

$$\boxed{5 \mid n}.$$

No había que corregir un cálculo; había que completar la dependencia hasta el objetivo correcto.

Solución E9. El argumento comienza diciendo

como $A \cap B = A$.

Pero ésa es precisamente la igualdad que se quiere demostrar.

Por tanto, la prueba usa la conclusión como premisa y es circular.

La inclusión

$$A \cap B \subseteq A$$

sí puede demostrarse sin circularidad. Si

$$x \in A \cap B,$$

por definición de intersección tenemos

$$x \in A$$

y

$$x \in B.$$

En particular, $x \in A$.
Para la inclusión inversa, sea ahora

$$x \in A.$$

Como

$$A \subseteq B,$$

se sigue que

$$x \in B.$$

Así,

$$x \in A \quad \text{y} \quad x \in B,$$

por lo que

$$x \in A \cap B.$$

Hemos demostrado

$$A \subseteq A \cap B.$$

Combinando ambas inclusiones,

$$\boxed{A \cap B = A}.$$

La hipótesis $A \subseteq B$ sólo es necesaria en la segunda dirección, cuando partimos de $x \in A$ y necesitamos obtener además $x \in B$.

Solución E10. La afirmación

$$\forall x \in \mathbb{R}, \qquad x^2 + 1 > 2x$$

es falsa.
Tomemos

$$x = 1.$$

Entonces

$$1^2 + 1 = 2 = 2 \cdot 1.$$

Por tanto, la desigualdad estricta falla y

$$\boxed{x = 1 \text{ es un contraejemplo}}.$$

Una primera reparación consiste en **debilitar la conclusión**:

$$\forall x \in \mathbb{R}, \qquad x^2 + 1 \geq 2x.$$

En efecto,

$$(x-1)^2 \geq 0$$

para todo real x. Al desarrollar,

$$x^2 - 2x + 1 \geq 0,$$

es decir,

$$\boxed{x^2 + 1 \geq 2x}.$$

Una segunda reparación conserva la desigualdad estricta pero **restringe el dominio**:

$$\forall x \in \mathbb{R} \setminus \{1\}, \qquad x^2 + 1 > 2x.$$

Si $x \neq 1$, entonces

$$(x-1)^2 > 0.$$

Por tanto,

$$x^2 - 2x + 1 > 0,$$

y así

$$\boxed{x^2 + 1 > 2x}.$$

Las dos reparaciones responden al mismo diagnóstico, pero intervienen en lugares distintos: una cambia la conclusión de $>$ a $\geq$; la otra conserva la conclusión y excluye exactamente el caso de igualdad.

Solución E11. **A.** La afirmación

$$10 \mid n^2 \Rightarrow 10 \mid n$$

es verdadera.

Podemos demostrarla mediante una clasificación exhaustiva módulo 10.

Todo entero satisface exactamente una de las congruencias

$$n \equiv 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 \pmod{10}.$$

Los cuadrados correspondientes módulo 10 son

$$0, 1, 4, 9, 6, 5, 6, 9, 4, 1.$$

El único resto cuyo cuadrado es congruente con 0 módulo 10 es 0.

Por tanto, si

$$10 \mid n^2,$$

entonces

$$n \equiv 0 \pmod{10},$$

y por consiguiente

$$\boxed{10 \mid n}.$$

La lista de diez casos no es una muestra: es exhaustiva respecto de las clases módulo 10.

B. La afirmación

$$8 \mid n^2 \Rightarrow 8 \mid n$$

es falsa.

Tomemos

$$n = 4.$$

Entonces

$$n^2 = 16$$

y

$$8 \mid 16,$$

pero

$$8 \nmid 4.$$

Así,

$$\boxed{n = 4 \text{ es un contraejemplo}}.$$

C. La afirmación

$$9 \mid n^2 \Rightarrow 3 \mid n$$

es verdadera.

Todo entero cae en una de las tres clases módulo 3:

$$n \equiv 0, 1, 2 \pmod 3.$$

Sus cuadrados son congruentes con

$$0, 1, 1 \pmod 3.$$

Si

$$9 \mid n^2,$$

entonces en particular

$$3 \mid n^2,$$

por lo que

$$n^2 \equiv 0 \pmod 3.$$

La clasificación anterior obliga a

$$n \equiv 0 \pmod 3.$$

Así,

$$\boxed{3 \mid n}.$$

Los tres apartados muestran tres usos distintos: prueba exhaustiva finita, contraejemplo y prueba exhaustiva sobre una clasificación más pequeña.

Solución E12. La afirmación

$$A: \quad \forall x \in \mathbb{R}\ \exists y \in \mathbb{R}, \quad y > x$$

es verdadera.

Dado cualquier real x, elegimos

$$y = x + 1.$$

Entonces

$$y > x.$$

El testigo depende de x.

La afirmación

$$B: \quad \exists y \in \mathbb{R}\ \forall x \in \mathbb{R}, \quad y > x$$

es falsa.

Su negación es

$$\forall y \in \mathbb{R}\ \exists x \in \mathbb{R}, \qquad y \leq x.$$

Sea y arbitrario. Elegimos simplemente

$$x = y.$$

Entonces

$$y \leq x$$

porque $y = x$.

Por tanto, ningún real puede ser mayor que todos los reales.

La diferencia lógica es ésta:

$$\forall x \exists y$$

permite que y dependa de x, mientras que

$$\exists y \forall x$$

exige elegir un único y antes de recorrer todos los valores de x.

Solución E13. Los valores enteros proporcionan evidencia sobre una **muestra discreta** del dominio real.

No cubren los puntos situados entre enteros.

Consideremos

$$x = \frac{1}{2}.$$

Entonces

$$\begin{aligned} f\left(\frac{1}{2}\right) &= \left(\frac{1}{2} - \frac{1}{2}\right)^2 - \frac{1}{16} \\ &= 0 - \frac{1}{16} \\ &= -\frac{1}{16} < 0. \end{aligned}$$

Por tanto,

$$\boxed{x = \frac{1}{2} \text{ es un contraejemplo}}.$$

La conjetura universal es falsa.

Una gráfica construida sólo a partir de valores enteros puede unir visualmente puntos positivos o simplemente no muestrear la región donde la función desciende bajo cero. El problema no está en los valores calculados, sino en confundir los puntos observados con todo el dominio.

Para determinar exactamente dónde $f(x) < 0$, resolvemos

$$\left(x - \frac{1}{2}\right)^2 - \frac{1}{16} < 0.$$

Esto equivale a

$$\left(x - \frac{1}{2}\right)^2 < \frac{1}{16}.$$

Por tanto,

$$\left|x - \frac{1}{2}\right| < \frac{1}{4}.$$

Así,

$$-\frac{1}{4} < x - \frac{1}{2} < \frac{1}{4},$$

y sumando 1/2,

$$\boxed{\frac{1}{4} < x < \frac{3}{4}}.$$

Precisamente entre 1/4 y 3/4 la función es negativa.

Solución E14. La afirmación

$$ca = cb \Rightarrow a = b$$

para todos los reales a, b, c es falsa.

Tomemos

$$c = 0, \qquad a = 1, \qquad b = 2.$$

Entonces

$$ca = 0 = cb,$$

pero

$$a \neq b.$$

Por tanto,

$$\boxed{(a, b, c) = (1, 2, 0) \text{ es un contraejemplo}}.$$

El contraejemplo revela exactamente la condición ausente: no podemos cancelar c cuando

$$c = 0.$$

La reparación mínima es

$$\forall a, b, c \in \mathbb{R}, \qquad (c \neq 0 \land ca = cb) \Rightarrow a = b.$$

Ahora, como $c \neq 0$, existe $1/c$ y podemos multiplicar ambos lados de

$$ca = cb$$

por $1/c$:

$$\frac{1}{c}(ca) = \frac{1}{c}(cb).$$

Así,

$$\boxed{a = b}.$$

Añadir como hipótesis directamente

$$a = b$$

haría trivialmente verdadera la conclusión, pero destruiría el problema: la supuesta reparación ya contendría exactamente lo que se pretendía deducir.

Una reparación controlada debe aportar la condición que hace válida la inferencia faltante, no asumir la conclusión.

Solución E15. La afirmación es

$$\forall x, y \in \mathbb{R}, \qquad xy = 1 \Rightarrow x + y \geq 2.$$

Los casos positivos pueden hacerla parecer plausible. Por ejemplo,

$$1 \cdot 1 = 1$$

y

$$1 + 1 = 2,$$

mientras que

$$2 \cdot \frac{1}{2} = 1$$

y

$$2 + \frac{1}{2} > 2.$$

Pero la hipótesis

$$xy = 1$$

no obliga a x e y a ser positivos. También puede cumplirse con ambos negativos.
Tomemos

$$x = y = -1.$$

Entonces

$$xy = (-1)(-1) = 1,$$

pero

$$x + y = -2 < 2.$$

Por tanto,

$$\boxed{(x, y) = (-1, -1) \text{ es un contraejemplo}}.$$

El fallo de la formulación es una hipótesis de **signo** ausente.
Una primera reparación consiste en fortalecer las hipótesis:

$$\forall x, y \in \mathbb{R}, \qquad (xy = 1 \land x > 0) \Rightarrow x + y \geq 2.$$

Si $xy = 1$ y $x > 0$, entonces

$$y = \frac{1}{x} > 0.$$

Además,

$$(x - 1)^2 \geq 0.$$

Desarrollando,

$$x^2 - 2x + 1 \geq 0.$$

Como $x > 0$, podemos dividir por x sin cambiar el sentido de la desigualdad:

$$x - 2 + \frac{1}{x} \geq 0.$$

Por tanto,

$$x + \frac{1}{x} \geq 2.$$

Como $y = 1/x$,

$$\boxed{x + y \geq 2}.$$

Una segunda reparación conserva el dominio completo y la única hipótesis $xy = 1$, pero cambia la conclusión:

$$\forall x, y \in \mathbb{R}, \qquad xy = 1 \Rightarrow |x + y| \geq 2.$$

Para demostrarla, partimos de

$$(x - y)^2 \geq 0.$$

Entonces

$$x^2 - 2xy + y^2 \geq 0.$$

Sumando $4xy$ a ambos lados,

$$x^2 + 2xy + y^2 \geq 4xy.$$

Como $xy = 1$,

$$(x + y)^2 \geq 4.$$

Por tanto,

$$\boxed{|x + y| \geq 2}.$$

Las dos reparaciones actúan en lugares diferentes.

La primera conserva la conclusión original pero restringe los casos permitidos mediante una hipótesis de signo.

La segunda conserva todos los pares reales con $xy = 1$, pero debilita la conclusión de una desigualdad orientada

$$x + y \geq 2$$

a una desigualdad simétrica en signo

$$|x + y| \geq 2.$$

En ambos casos el contraejemplo no fue simplemente eliminado por decreto: reveló qué información faltaba, y la nueva afirmación fue después demostrada para todos los casos que conserva.

Cierre del capítulo. C09 comenzó con una pregunta aparentemente sencilla: ¿cómo sabemos que una afirmación universal es falsa?

La respuesta completa resultó exigir una disciplina más amplia.

Aprendimos a traducir la negación de una universal y a verificar contraejemplos; a buscar casos de manera dirigida sin confundir exploración con prueba; a distinguir recíprocos, contraposiciones y negaciones; a controlar la dependencia de testigos; a tratar el dominio como parte de la inferencia; a separar operaciones ilegítimas de pasos legítimos pero irreversibles; a localizar circularidad y otras dependencias defectuosas; a valorar tablas, gráficas y figuras sin atribuirles una fuerza lógica que no tienen; y a reparar enunciados o argumentos sin modificar más de lo necesario.

Estas tareas pueden resumirse en tres movimientos:

$$\boxed{\text{refutar}} \qquad \boxed{\text{diagnosticar}} \qquad \boxed{\text{reparar}}.$$

Refutar exige una prueba de falsedad.

Diagnosticar exige localizar exactamente qué obligación dejó de cumplirse.

Reparar exige producir una nueva afirmación o una nueva dependencia que pueda demostrarse, no sólo una formulación que sobreviva al ejemplo que reveló el fallo.

El capítulo siguiente cambiará de foco una vez más. Su problema no será descubrir si una prueba existe o si un argumento falla, sino **cómo escribir una demostración válida de manera que otro matemático pueda leerla, reconstruirla y verificarla sin tener que adivinar sus dependencias**.

Ése será el problema de C10.

Capítulo 10 — ¿Cómo se escribe una prueba para que otro matemático pueda verificarla?

En los capítulos anteriores hemos trabajado casi siempre con una pregunta matemática: ¿qué debe demostrarse, cómo puede encontrarse una prueba, qué ocurre si el enunciado es falso o en qué punto falla un argumento?

C06 organizó arquitecturas de demostración. C07 separó el trabajo de descubrimiento de la prueba final. C09 entrenó el diagnóstico de argumentos defectuosos y la reparación de dependencias inválidas.

Pero incluso después de haber encontrado una demostración correcta queda una tarea distinta.

Tenemos que **escribirla**.

No basta con que el autor sepa por qué funciona.

La prueba debe quedar construida de manera que otro lector pueda reconstruirla.

Ésa es una exigencia más fuerte que «anotar las cuentas que hice» y distinta de «contar todo lo que pensé mientras buscaba la solución».

Una prueba final no es un diario del descubrimiento.

Es una exposición organizada de dependencias.

El lector debe poder reconocer:

- qué objetos están en juego;
- qué hipótesis están disponibles;
- qué variables se introducen y qué representan;
- qué resultado autoriza cada paso no rutinario;
- cuándo se abre una obligación nueva;
- cuándo esa obligación queda cerrada;
- y por qué la conclusión final es exactamente la que se quería demostrar.

Esto conduce al principio rector de C10:

una prueba está lista cuando otro lector puede auditarla sin adivinar sus dependencias.

La palabra **auditar** será importante.

No significa formalizar cada línea en un asistente de pruebas ni justificar operaciones elementales hasta volver ilegible la idea principal.

Significa algo más básico y más práctico: que un lector del nivel previsto pueda seguir el argumento, comprobarlo y localizar de dónde proviene cada paso importante.

La escritura matemática tendrá entonces una función lógica.

Una notación inestable puede ocultar qué objeto estamos usando. Un pronombre ambiguo puede dejar incierto el referente de una afirmación. Una cadena de símbolos puede borrar la diferencia entre igualdad, implicación y equivalencia. Una referencia vaga como «por lo anterior» puede obligar al lector a buscar por ensayo qué resultado se está invocando.

Estos problemas no son idénticos a una inferencia falsa.

Una prueba puede ser correcta y estar mal escrita.

También puede estar escrita con gran claridad y ser matemáticamente incorrecta.

C10 conservará siempre esa distinción.

Nuestro objetivo no será confundir estilo con verdad, sino aprender a presentar una demostración correcta de modo que su corrección pueda **verificarse**.

Durante el capítulo trabajaremos con una pregunta recurrente:

> **¿qué información necesita realmente el lector para reconstruir este paso sin inventar nada que el autor no haya dicho?**

Comenzaremos precisamente por aclarar qué significa que una prueba sea auditable.

10.1. Una prueba correcta también debe poder auditarse

Consideremos una situación elemental.

Queremos demostrar:

> Para todo entero n, si $18 \mid n$, entonces $6 \mid n$.

La idea matemática es muy simple.

Si n es múltiplo de 18, entonces puede escribirse como 18 por un entero; por tanto también puede escribirse como 6 por un entero.

Supongamos que alguien redacta:

> Como $18 \mid n$, tenemos $n = 18k = 6(3k)$. Luego $6 \mid n$.

Un lector experimentado probablemente entiende lo que quiso decir el autor.

Puede reconstruir mentalmente que k debe ser un entero y que la primera igualdad proviene de la definición de divisibilidad.

Pero observe qué trabajo tuvo que hacer el lector.

Tuvo que completar información que no estaba escrita.

La prueba puede arreglarse con muy poco:

> Sea $n \in \mathbb{Z}$ y supongamos que $18 \mid n$. Por definición de divisibilidad, existe $k \in \mathbb{Z}$ tal que $n = 18k$. Como
>
> $$n = 18k = 6(3k)$$
>
> y $3k \in \mathbb{Z}$, se sigue que $6 \mid n$.

La matemática esencial no cambió.

No apareció una idea nueva.

Lo que cambió fue la **trazabilidad**.

Ahora podemos señalar con precisión:

1. cuál es el objeto arbitrario: n;
2. cuál es su dominio: $\mathbb{Z}$;
3. cuál es la hipótesis: $18 \mid n$;
4. de dónde surge k: de la definición de divisibilidad;
5. qué sabemos sobre k: $k \in \mathbb{Z}$;
6. por qué $3k$ sirve como nuevo factor entero;
7. qué conclusión queda establecida: $6 \mid n$.

La segunda versión no es mejor porque sea más larga.

Es mejor porque deja visibles las dependencias necesarias.

Corrección, legibilidad y auditabilidad. Conviene separar tres propiedades.

Corrección matemática.

Cada afirmación utilizada es verdadera bajo las hipótesis disponibles y cada inferencia está autorizada.

Legibilidad.

El texto puede seguirse sin fricción innecesaria: los símbolos son estables, las oraciones son comprensibles y la organización permite reconocer la estructura general.

Auditabilidad.

El lector puede reconstruir qué información estaba disponible en cada punto y por qué el paso siguiente se sigue de ella.

Estas propiedades están relacionadas, pero no son equivalentes.

Una prueba puede ser matemáticamente correcta y aun así obligar al lector a reconstruir demasiado.

También puede ser muy legible y, sin embargo, contener una inferencia inválida.

Podemos resumir las posibilidades así:

	Auditable	No auditable
Correcta	objetivo de C10	necesita edición
Incorrecta	el error puede localizarse con claridad	falla matemática y expositivamente

La escritura no convierte una prueba inválida en válida.

Pero una buena escritura sí permite **ver** mejor por qué un argumento funciona o dónde deja de funcionar.

El lector no debe leer la mente del autor. Cuando el autor conoce la solución, muchas dependencias le parecen invisibles por exceso de familiaridad.

Puede pensar:

> «es obvio de dónde salió k»

porque él mismo acaba de usar la definición de divisibilidad.

Pero el lector sólo dispone de lo que el texto le ofrece.

Ésta es una asimetría fundamental entre escribir y leer una prueba.

El autor conoce:

- el objetivo;
- la estrategia;
- los intentos descartados;
- el significado que quiso dar a cada símbolo;
- el resultado al que quiere llegar.

El lector, en cambio, recibe una secuencia de frases y fórmulas.

Por eso una prueba no debe evaluarse preguntando únicamente

> «¿yo sé lo que quise decir?»

sino también

> **«¿puede reconstruirlo un lector que conoce la materia pero no conoce esta solución?»**

Ése será nuestro lector de referencia durante C10.

No será un principiante absoluto.

Podrá utilizar lo ya establecido en FND: definiciones, cuantificadores, métodos de prueba, conjuntos, funciones, inducción y diagnóstico de argumentos.

Pero no deberá adivinar la solución concreta que el autor tiene en mente.

Auditar es reconstruir dependencias. Podemos imaginar una prueba como una red orientada.

Tenemos ciertos datos iniciales:

$$H_1, H_2, \ldots, H_r.$$

A partir de ellos obtenemos afirmaciones intermedias:

$$A_1, A_2, \ldots, A_s,$$

hasta alcanzar la conclusión

$$C.$$

La estructura ideal no necesita ser literalmente una cadena lineal, pero cada nodo importante debe tener una procedencia reconocible.

Podemos representar una versión simple como

$$\boxed{H_1, \ldots, H_r} \longrightarrow \boxed{A_1} \longrightarrow \boxed{A_2} \longrightarrow \cdots \longrightarrow \boxed{C}.$$

Una auditoría pregunta, en cada transición:

¿qué autoriza esta flecha?

La respuesta puede ser:

- una definición;
- una hipótesis;
- un resultado demostrado antes;
- una operación algebraica legítima;
- una regla lógica;
- una separación exhaustiva de casos;
- una construcción realizada en la propia prueba.

Si el lector debe inventar la respuesta, la dependencia no está suficientemente visible.

No todas las razones deben escribirse con la misma amplitud. De aquí podría surgir una reacción excesiva.

Tal vez pensemos que toda línea necesita una explicación completa.

No es así.

Supongamos que en una prueba ya sabemos que

$$a = 3$$

y necesitamos escribir

$$a + 2 = 5.$$

Para el lector previsto no hace falta insertar una teoría de la suma de enteros.

El paso es rutinario.

En cambio, si usamos un resultado anterior cuyas hipótesis no son evidentes, sí necesitamos dejar visible por qué podemos aplicarlo.

Por tanto, auditabilidad no significa máxima longitud.

Significa **detalle suficiente en los puntos donde aparece nueva dependencia matemática**.

Más adelante, en §10.8, estudiaremos directamente cómo decidir cuánto detalle escribir.

Por ahora basta con una regla provisional:

explicar lo que el lector no puede recuperar de manera segura a partir del contexto disponible

.

Una prueba puede ser demasiado comprimida sin ser falsa. Volvamos al ejemplo de divisibilidad.

La versión

> Como $18 \mid n$, tenemos $n = 18k = 6(3k)$. Luego $6 \mid n$.

no contiene necesariamente una idea matemática equivocada.

El problema es que deja implícitas varias piezas:

- que n es entero;
- que k es entero;
- que k existe por la definición de $18 \mid n$;
- que $3k$ sigue siendo entero.

En un contexto muy avanzado, algunas de estas piezas podrían comprimirse legítimamente.

En un capítulo fundacional, donde precisamente se está aprendiendo a escribir pruebas verificables, conviene hacerlas visibles.

Esto muestra que la auditabilidad depende en parte del **lector previsto**.

No existe una longitud universalmente correcta para toda demostración.

Una misma frase puede ser suficiente en un artículo especializado y demasiado comprimida en una primera introducción a las pruebas.

Una prueba también puede ser demasiado larga. El defecto contrario existe.

Podríamos expandir la prueba anterior así:

> Como $k \in \mathbb{Z}$, entonces $3 \in \mathbb{Z}$. Como el producto de enteros es entero, $3k \in \mathbb{Z}$. Como $6 \in \mathbb{Z}$ y $3k \in \mathbb{Z}$, el producto $6(3k)$ es entero. Además, $18 = 6 \cdot 3$. Por tanto $18k = 6 \cdot 3 \cdot k = 6(3k)\ldots$

Nada de esto es necesariamente falso.

Pero ahora la idea principal comienza a desaparecer bajo información que el lector ya domina.

La prueba deja de ser eficiente.

Por eso C10 buscará un equilibrio:

$$\boxed{\text{ningún hueco decisivo}} \quad \text{y} \quad \boxed{\text{ningún detalle rutinario que oculte la idea}}.$$

La economía es parte de la claridad, pero sólo después de asegurar la trazabilidad.

Auditabilidad no significa formalización completa. También debemos evitar otra confusión.

Una demostración ordinaria y una prueba formalizada en un sistema lógico no tienen el mismo nivel de explicitación.

En una formalización completa pueden necesitarse pasos que, en prosa matemática, se consideran parte del conocimiento compartido del lector.

C10 no convertirá cada demostración en una derivación formal simbólica.

Nuestro criterio será más práctico:

> **¿puede un lector competente reconstruir el argumento sin introducir hipótesis, significados o resultados que el texto nunca hizo disponibles?**

Si la respuesta es sí, la prueba puede ser suficientemente auditable aunque omita operaciones rutinarias.

Si la respuesta es no, falta información.

El test de reconstrucción. Podemos formular un primer protocolo de lectura.

Ante una prueba breve, intenta responder sin consultar la intención del autor:

1. ¿Qué objetos están fijados?

¿Sabemos qué representan y a qué dominios pertenecen?

2. ¿Qué hipótesis están activas?

¿Se distinguen de las conclusiones que todavía deben demostrarse?

3. ¿Qué símbolos nuevos aparecen?

¿Fueron introducidos antes de usarse?

4. ¿De dónde proviene cada paso no rutinario?

¿Definición, hipótesis, teorema previo, construcción, cálculo?

5. ¿Cuándo se cierra la obligación?

¿La conclusión obtenida coincide exactamente con el objetivo?

Si podemos responder estas preguntas, tenemos una ruta de auditoría.

Si alguna respuesta exige adivinar, hemos localizado un punto de edición.

Editar no significa cambiar la matemática. Una de las habilidades centrales de C10 será transformar versiones opacas en versiones auditables **sin alterar el contenido matemático**.

Esto importa porque editar una prueba no es reparar un teorema falso.

En C09 podíamos necesitar cambiar una hipótesis, restringir un dominio o sustituir una inferencia inválida.

En C10 muchas veces la matemática ya está bien.

La tarea consiste en volver visible algo que estaba implícito:

- introducir una variable;
- nombrar su dominio;
- separar dos casos en párrafos;
- indicar qué definición se usa;
- reemplazar «por lo anterior» por una referencia identificable;
- añadir una frase que explique por qué una igualdad produce la conclusión.

La edición correcta conserva la prueba y mejora su trazabilidad.

Una dependencia visible también ayuda a detectar errores. Aunque C10 no repite el catálogo de fallos de C09, una consecuencia importante de la auditabilidad es que los errores quedan más expuestos.

Si escribimos

«por el Teorema 6.4, como se cumplen sus hipótesis. . . »

el lector puede comprobar esas hipótesis.

Si escribimos simplemente

«por un teorema conocido. . . »

la dependencia queda oculta.

Si declaramos

«sea $x \in A \cap B$»

queda claro qué sabemos sobre x.

Si aparece de pronto una letra x sin introducción, el lector debe reconstruir su función.

La buena exposición no garantiza que la prueba sea correcta.

Pero crea mejores condiciones para verificarla.

Podemos condensar esta idea así:

hacer visible una dependencia no la hace verdadera, pero permite comprobarla.

La prueba final es un objeto público. Durante el descubrimiento podemos escribir abreviaturas personales:

- flechas;
- palabras sueltas;
- símbolos temporales;
- diagramas incompletos;
- pasos hacia atrás;
- conjeturas que después descartaremos.

Ese trabajo puede ser excelente para pensar.

Pero la versión final tiene otra función.

Debe poder salir de la libreta del autor y ser leída por otra persona.

En ese momento deja de bastar con que el texto funcione como recordatorio privado.

Debe funcionar como argumento público.

Ésta es la transición que organiza todo C10:

$$\boxed{\text{yo sé por qué funciona}} \quad \longrightarrow \quad \boxed{\text{otro lector puede verificar por qué funciona}}.$$

La prueba no debe exigir telepatía

Una prueba auditable permite identificar:

- los objetos y sus dominios;
- las hipótesis activas;
- los símbolos introducidos;
- la procedencia de cada paso no rutinario;
- la obligación que se está resolviendo;
- el punto donde esa obligación queda cerrada.

No hace falta escribir cada operación elemental.

Sí hace falta evitar que el lector deba inventar una hipótesis, un referente, un significado o una dependencia esencial que el texto no declaró.

Comprobación. Compara las dos versiones siguientes del mismo argumento.

Se quiere demostrar:

Para todo $n \in \mathbb{Z}$, si $15 \mid n$, entonces $5 \mid n$.

Versión A.

Como $15 \mid n$, $n = 15k = 5(3k)$. Luego $5 \mid n$.

Versión B.

Sea $n \in \mathbb{Z}$ y supongamos que $15 \mid n$. Por definición de divisibilidad, existe $k \in \mathbb{Z}$ tal que $n = 15k$. Entonces

$$n = 15k = 5(3k).$$

Como $3k \in \mathbb{Z}$, concluimos que $5 \mid n$.

Responde:

1. ¿Utilizan las dos versiones la misma idea matemática?
2. ¿Qué información debe reconstruir el lector en la versión A?
3. ¿Hay una nueva idea matemática en la versión B?
4. ¿Por qué la versión B es más auditable?
5. ¿Sería útil añadir una demostración completa de que $3k \in \mathbb{Z}$? Justifica según el lector previsto.

Respuesta razonada.

Las dos versiones utilizan la misma idea central: desplegar la definición de divisibilidad y reescribir un múltiplo de 15 como múltiplo de 5.

En la versión A, el lector debe inferir que la variable k es un entero y que aparece porque $15 \mid n$ significa que existe algún entero k con

$$n = 15k.$$

También debe reconstruir que el factor

$$3k$$

es entero y, por tanto, puede funcionar como testigo de que $5 \mid n$.

La versión B no introduce una nueva estrategia ni una nueva propiedad profunda.

Hace explícitas las dependencias que la versión A había comprimido.

Por eso es más auditable: un lector puede identificar el dominio de n, la hipótesis activa, el origen y dominio de k, la transformación utilizada y el punto exacto donde se satisface la definición de $5 \mid n$.

No hace falta demostrar desde cero que el producto de dos enteros es un entero.

Para el lector previsto de C10, saber que $3k \in \mathbb{Z}$ cuando $k \in \mathbb{Z}$ es un paso rutinario. Basta con registrarlo cuando cumple una función lógica importante, porque necesitamos saber que el factor de 5 pertenece al dominio correcto.

La diferencia entre ambas versiones no es, por tanto, «larga frente a corta».

Es:

$$\boxed{\text{dependencias implícitas que el lector debe reconstruir}}$$

$$\text{frente a } \boxed{\text{dependencias visibles en el texto}}.$$

La siguiente obligación aparece antes del primer cálculo. En los ejemplos anteriores la auditabilidad dependió, antes que nada, de saber qué representan n y k.

Esto señala el siguiente problema del capítulo.

Antes de escribir una sola inferencia debemos controlar la entrada de cada objeto al argumento:

- ¿qué variable estamos fijando?;
- ¿a qué dominio pertenece?;
- ¿es arbitraria o elegida?;
- ¿qué hipótesis suponemos sobre ella?;
- ¿qué símbolos auxiliares introduciremos después?

Una gran cantidad de ambigüedades puede eliminarse si la apertura de la prueba está bien construida.

Ése será el contenido de §10.2.

10.2. Introducir objetos, variables, dominios e hipótesis

En §10.1 formulamos una prueba sencilla de auditabilidad: otro lector debe poder reconstruir qué se supone, qué se introduce, por qué cada paso está autorizado y cuándo se alcanza la conclusión.

La primera oportunidad de cumplir —o de romper— ese contrato aparece antes de la primera inferencia.

Aparece en la **entrada de los objetos al argumento**.

Cuando una prueba utiliza una letra, el lector necesita saber qué representa. Cuando fija una variable, necesita saber de qué dominio proviene. Cuando introduce una hipótesis, necesita reconocer desde qué punto está disponible. Cuando aparece un símbolo auxiliar, debe

quedar claro si fue dado, elegido, construido o garantizado por una definición o un resultado previo.

Una gran cantidad de pruebas difíciles de leer no falla porque contenga matemática sofisticada.

Falla porque empieza demasiado tarde.

El texto comienza manipulando símbolos cuya historia el lector todavía no conoce.

La prueba comienza antes del primer cálculo. Consideremos la afirmación

$$\forall n \in \mathbb{Z}, \qquad 6 \mid n \Rightarrow 12 \mid n^2.$$

Una versión opaca podría comenzar así:

> $n = 6k$, luego $n^2 = 36k^2 = 12(3k^2)$, por tanto $12 \mid n^2$.

La idea matemática es correcta.

Pero el lector debe reconstruir varias piezas:

- ¿qué es n?;
- ¿por qué pertenece a $\mathbb{Z}$?;
- ¿se está suponiendo $6 \mid n$?;
- ¿de dónde apareció k?;
- ¿por qué $k \in \mathbb{Z}$?;
- ¿qué permite concluir finalmente $12 \mid n^2$?

Una apertura auditable hace visible esa información antes de utilizarla:

> Sea $n \in \mathbb{Z}$ y supongamos que $6 \mid n$. Por definición de divisibilidad, existe $k \in \mathbb{Z}$ tal que $n = 6k$. Entonces
>
> $$n^2 = 36k^2 = 12(3k^2).$$
>
> Como $3k^2 \in \mathbb{Z}$, concluimos que $12 \mid n^2$.

La segunda versión no contiene una idea matemática adicional.

Contiene una **entrada ordenada de la información**.

El lector puede distinguir ahora:

$$\boxed{n \in \mathbb{Z}} \qquad \boxed{6 \mid n} \qquad \boxed{k \in \mathbb{Z} \text{ con } n = 6k}.$$

Cada dato tiene un origen identificable.

Un símbolo nuevo necesita una historia. No hace falta escribir una biografía de cada letra.

Sí hace falta que el símbolo tenga una introducción suficiente para responder, cuando corresponda, tres preguntas:

1. **¿qué objeto representa?**
2. **¿a qué dominio pertenece?**
3. **¿por qué tenemos derecho a utilizarlo?**

Por ejemplo, si durante una prueba aparece

$$x \in A \cap B,$$

puede ser suficiente escribir:

> Sea $x \in A \cap B$.

Con una sola frase sabemos que x es un elemento, que está siendo fijado en ese momento y que pertenece al conjunto indicado.

En cambio, comenzar con

> $x \in A \cap B$, entonces...

puede ser comprensible en unas notas privadas, pero deja sin marcar el acto lógico de fijar el objeto con el que trabajaremos.

Del mismo modo, si escribimos

$$y = f(x),$$

debemos haber dejado disponible qué es f y qué es x.

El orden de aparición importa:

$$\boxed{\text{introducir}} \longrightarrow \boxed{\text{usar}}.$$

No al revés.

Arbitrario, fijo, elegido y construido no significan lo mismo. Las frases con las que introducimos objetos registran papeles lógicos diferentes.

Supongamos que queremos demostrar una afirmación universal

$$\forall x \in D, \qquad P(x).$$

Una apertura natural es:

> Sea $x \in D$ arbitrario.

La palabra **arbitrario** indica que no hemos escogido un caso favorable. El argumento debe funcionar para cualquier elemento del dominio.

En muchas pruebas también basta escribir

> Sea $x \in D$.

si el contexto deja inequívoco que se está demostrando una universal y que no se impondrá sobre x ninguna propiedad adicional no autorizada.

Pero la situación cambia cuando necesitamos un testigo existencial.

Para demostrar, por ejemplo, que para cada $x \in \mathbb{R}$ existe un real mayor que x, podemos comenzar:

> Sea $x \in \mathbb{R}$. Elijamos $y = x + 1$.

Aquí x es el objeto universal fijado y y es un objeto **construido a partir de él**.

La dependencia queda visible:

$$y = y(x) = x + 1.$$

Si, en cambio, un teorema previo garantiza la existencia de algún objeto sin darnos una fórmula, la redacción debe reflejarlo:

> Como se cumplen las hipótesis del teorema, existe $y \in E$ tal que $P(x, y)$.

En ese caso y no fue inventado por el autor ni elegido libremente entre todos los objetos de E.

Fue garantizado por una dependencia previa.

Estas diferencias de lenguaje ayudan a registrar la arquitectura real del argumento:

- **sea**: fija un objeto para trabajar con él;
- **sea ... arbitrario**: subraya que no se eligió un caso especial;
- **elijamos**: selecciona un objeto entre opciones permitidas;
- **definamos**: introduce un objeto mediante una regla explícita;
- **existe ... tal que**: registra una existencia ya justificada;
- **tomemos el objeto garantizado por...**: hace visible la procedencia de un testigo.

No son fórmulas rígidas.

Son maneras de hacer legible el papel de cada objeto.

El dominio no debe quedar a cargo de la memoria del lector. C03 enseñó que el dominio forma parte del significado de una afirmación. C10 no reabre esa teoría.

La pregunta ahora es editorial:

> **¿dónde debe aparecer el dominio para que la prueba pueda seguirse sin ambigüedad?**

Consideremos una prueba que empieza:

> Sea x. Como $x^2 = 1$, tenemos $x = \pm 1$.

El texto puede ser correcto o incorrecto según qué objetos se permitan.

Si trabajamos en $\mathbb{R}$, la conclusión tiene el sentido esperado.

Si el contexto fuera otro, podría cambiar qué operaciones, raíces o soluciones están disponibles.

Una versión auditable no obliga al lector a recuperar el dominio desde varias páginas atrás cuando éste es relevante para la inferencia:

> Sea $x \in \mathbb{R}$ y supongamos que $x^2 = 1$.

No hace falta repetir $x \in \mathbb{R}$ en cada línea posterior.

Hace falta **establecerlo una vez y mantenerlo estable**.

La repetición constante produce ruido.

La omisión inicial produce incertidumbre.

Las hipótesis deben entrar antes de ser utilizadas. Una demostración condicional necesita distinguir con claridad entre:

- el objeto que estamos considerando;
- la hipótesis adicional que suponemos sobre ese objeto;
- la conclusión que pretendemos obtener.

Para demostrar

$$\forall x \in \mathbb{R}, \qquad x > 2 \Rightarrow x^2 > 4,$$

una apertura suficiente es:

> Sea $x \in \mathbb{R}$ y supongamos que $x > 2$.

Desde ese momento están disponibles dos datos diferentes:

$$x \in \mathbb{R}$$

y

$$x > 2.$$

La primera afirmación fija el dominio.

La segunda es la hipótesis del condicional.

Si el texto comienza directamente con

> Como $x > 2$, entonces...

el lector quizá pueda reconstruir la intención, pero no sabe todavía si x fue fijado, si es arbitrario, si pertenece al dominio correcto o si la hipótesis procede del enunciado o de un paso anterior.

Una prueba auditable no debería depender de esa reconstrucción cuando una frase breve puede evitarla.

Las hipótesis locales tienen duración limitada. No toda hipótesis permanece activa durante toda la demostración.

En una prueba por casos podemos escribir:

Caso 1. Supongamos que $x = 0$.

La hipótesis

$$x = 0$$

pertenece a ese caso.

Cuando el caso termina y comenzamos

Caso 2. Supongamos que $x \neq 0$,

la primera hipótesis ya no está disponible.

Algo semejante ocurre en una prueba por contradicción: la negación provisional del objetivo está activa dentro de esa subdemostración y desaparece cuando la contradicción cierra el argumento.

C10 no reenseñará estas arquitecturas de C06.

Sí exigirá que su **alcance textual** quede reconocible.

Los recursos más sencillos suelen bastar:

- abrir un párrafo nuevo;
- rotular los casos;
- anunciar «supongamos, para llegar a una contradicción, que. . . »;
- cerrar explícitamente el caso o la contradicción antes de volver al argumento principal.

La disposición textual ayuda a responder:

¿qué hipótesis están activas en esta línea?

Un testigo existencial debe entrar con sus propiedades pertinentes. Supongamos que sabemos

$$6 \mid n.$$

Por definición de divisibilidad obtenemos

$$\exists k \in \mathbb{Z}, \qquad n = 6k.$$

Si la prueba escribe simplemente

Entonces $n = 6k$,

ha ocultado una parte esencial de la información: que el testigo pertenece a $\mathbb{Z}$.

Esto puede parecer pequeño, pero esa pertenencia suele ser exactamente lo que permitirá cerrar la conclusión más adelante.

Por eso es preferible:

Existe $k \in \mathbb{Z}$ tal que $n = 6k$.

A partir de ahí no necesitamos repetir en cada línea que k es entero.

Ya ha quedado incorporado al contexto activo.

La introducción correcta funciona como una especie de registro:

$$\boxed{k \in \mathbb{Z}} \qquad + \qquad \boxed{n = 6k}.$$

Ambas piezas pueden utilizarse después.

Varias variables exigen conservar sus dependencias. Cuando aparecen varios objetos, la apertura también debe mostrar quién depende de quién.

Consideremos

$$\forall x \in \mathbb{R}\ \exists y \in \mathbb{R}, \qquad y > x.$$

Una prueba clara puede escribir:

Sea $x \in \mathbb{R}$ arbitrario. Definamos $y = x + 1$.

Aquí la elección de y ocurre **después** de fijar x.

El texto refleja por tanto la dependencia lógica

$$x \longmapsto y = x + 1.$$

En cambio, una apertura como

Sea $y = x + 1$ y sea $x \in \mathbb{R}$,

puede acabar describiendo la misma fórmula, pero presenta los objetos en un orden que oscurece la dependencia que la prueba debe certificar.

La regla editorial es sencilla:

> **introducir los objetos en un orden compatible con la dependencia lógica que la demostración utiliza.**

Esto será especialmente importante cuando una misma prueba combine parámetros, objetos arbitrarios, testigos y construcciones auxiliares.

Declarar demasiado también puede entorpecer. La solución no consiste en comenzar cada prueba con un inventario enciclopédico.

Una apertura como

Sean $n, k, m, r, s, t \in \mathbb{Z}$, sea A un conjunto, sea f una función...

es peor si la mayoría de esos objetos todavía no tiene ninguna función.

Conviene introducir un símbolo **cuando entra en la dependencia**, no páginas antes por si acaso aparece después.

Podemos pensar en un principio de proximidad:

$$\boxed{\text{declarar antes de usar}}$$

pero también

$$\boxed{\text{declarar cerca del primer uso}}.$$

Eso reduce la carga de memoria del lector y permite que cada objeto aparezca acompañado por la razón que lo introduce.

Un protocolo para auditar la apertura de una prueba. Antes de revisar cálculos o inferencias internas, podemos auditar sólo las primeras líneas.

1. ¿Qué objetos vienen dados por el enunciado?

¿Conjuntos, números, funciones, relaciones, índices?

2. ¿Qué variable se fija primero?

¿Es arbitraria, particular, un parámetro o un candidato elegido?

3. ¿Su dominio quedó declarado?

¿El dominio es relevante para las operaciones posteriores?

4. ¿Qué hipótesis quedan activas?

¿Proceden del enunciado, de un caso, de una contradicción o de un resultado anterior?

5. ¿Qué símbolos auxiliares aparecen?

¿Cada uno fue introducido antes de usarse?

6. ¿Qué justifica su existencia?

¿Una definición, una construcción explícita, una elección o un teorema previo?

7. ¿El orden de introducción refleja las dependencias?

Si y depende de x, ¿se fijó primero x?

Este control no demuestra todavía que la prueba sea correcta.

Pero permite saber si el lector dispone de un contexto suficientemente bien formado para empezar a verificarla.

Antes de usar un símbolo, el lector debe saber de dónde salió

Una apertura auditable hace visibles, cuando son relevantes:

- el objeto que se fija;
- su dominio;
- las hipótesis activas;
- el origen de cada símbolo auxiliar;
- la dependencia entre objetos.

No es necesario repetir esos datos en cada línea.

Sí es necesario declararlos antes de que una inferencia dependa de ellos.

Una regla práctica es:

$$\boxed{\text{introducir } \rightarrow \text{ establecer dominio/papel } \rightarrow \text{ usar}}.$$

Comprobación. Examina las tres aperturas siguientes.

A. Se quiere demostrar:

$$\forall n \in \mathbb{Z}, \qquad 10 \mid n \Rightarrow 5 \mid n.$$

La prueba comienza:

$n = 10k = 5(2k)$, luego $5 \mid n$.

B. Se quiere demostrar:

$$\forall x \in \mathbb{R}\ \exists y \in \mathbb{R}, \qquad y > x.$$

La prueba comienza:

Sea $x \in \mathbb{R}$ arbitrario. Definamos $y = x + 1$.

C. Se quiere demostrar, para conjuntos A y B,

$$A \cap B \subseteq A.$$

La prueba comienza:

Sea $x \in A \cap B$. Entonces $x \in A$ y $x \in B$.

Para cada caso:

1. identifica los objetos que el lector necesita tener registrados;
2. decide si la apertura ya los introduce de manera suficiente;
3. si falta información, reescribe sólo lo necesario;
4. explica qué parte de la mejora es matemática y qué parte es expositiva.

Respuesta razonada.

En **A**, la idea matemática es correcta, pero la apertura es insuficiente como texto final.

El lector necesita saber que estamos trabajando con un entero arbitrario n que satisface la hipótesis

$$10 \mid n.$$

También necesita saber por qué existe el símbolo k y cuál es su dominio.

Una versión suficiente es:

> Sea $n \in \mathbb{Z}$ y supongamos que $10 \mid n$. Por definición de divisibilidad, existe $k \in \mathbb{Z}$ tal que $n = 10k$. Entonces
>
> $$n = 10k = 5(2k).$$

Como $2k \in \mathbb{Z}$, concluimos que $5 \mid n$.

No hemos cambiado la demostración.

Hemos hecho explícitos el dominio de n, la hipótesis activa y el origen de k.

En **B**, la apertura ya es adecuada.

El objeto universal x aparece primero, con dominio y carácter arbitrario. Después se introduce el testigo

$$y = x + 1,$$

cuya dependencia respecto de x queda visible en la propia definición.

Todavía habrá que verificar que $y \in \mathbb{R}$ y que $y > x$, pero no falta información sobre el origen de los objetos.

En **C**, la apertura también es suficiente.

Al escribir

Sea $x \in A \cap B$,

fijamos el elemento con el que se demostrará la inclusión. La definición de intersección permite después obtener

$$x \in A \qquad \text{y} \qquad x \in B.$$

No necesitamos añadir «sea x arbitrario» si el contexto de una prueba de inclusión deja claro que el elemento fijado no posee ninguna propiedad adicional aparte de pertenecer a $A \cap B$.

Tampoco hace falta volver a definir A y B dentro de la prueba si el enunciado ya los introdujo como conjuntos.

Los tres casos muestran el criterio buscado:

$$\boxed{\text{hacer explícita la información necesaria}} \neq \boxed{\text{repetir toda la información disponible}}.$$

De introducir símbolos a conservar su significado. Una apertura correcta resuelve sólo el primer problema.

Después de introducir un símbolo todavía debemos conservar estable su significado.

Si una letra representa un entero, no debería reaparecer unas líneas más abajo nombrando un conjunto. Si f representa una función, debemos distinguirla de un valor $f(x)$. Si creamos una notación local, el lector debe poder recordar qué comprime y por qué resulta útil.

La siguiente sección estudia precisamente esa responsabilidad:

> **¿cómo escoger y mantener una notación que reduzca trabajo al lector en lugar de crear nuevas ambigüedades?**

Ése será el contenido de §10.3.

10.3. Notación coherente y nombres estables

En §10.2 vimos cómo entra un símbolo en una prueba.

Ahora aparece una responsabilidad distinta:

> **una vez que un símbolo ha entrado, su significado debe mantenerse estable mientras siga activo.**

Ésta parece una regla menor.

No lo es.

En una demostración breve, una sola letra puede transportar mucha información:

- qué tipo de objeto representa;
- de qué conjunto proviene;

- si es arbitrario o elegido;
- qué propiedades sabemos de él;
- y qué papel cumple en la arquitectura del argumento.

Si una misma letra cambia de objeto sin aviso, el lector ya no puede saber qué hipótesis siguen disponibles.

La notación deja entonces de comprimir información y empieza a destruirla.

Podemos expresar el principio así:

un símbolo útil conserva una identidad matemática reconocible durante su alcance.

Esto no significa que cada letra deba tener un significado universal para toda la matemática.

La letra x puede representar un número en una prueba y un elemento de un conjunto en otra.

La exigencia es **local**:

> dentro de un mismo tramo de razonamiento, el lector debe poder seguir la historia de cada símbolo sin preguntarse si el autor cambió silenciosamente su significado.

Un símbolo no es sólo una marca gráfica. Consideremos una prueba elemental.

Queremos demostrar que si n es múltiplo de 12, entonces es múltiplo de 4.

Una redacción razonable comienza:

> Sea $n \in \mathbb{Z}$ y supongamos que $12 \mid n$. Entonces existe $k \in \mathbb{Z}$ tal que $n = 12k$.

Hasta aquí tenemos dos símbolos con papeles claros:

- n: entero arbitrario sometido a la hipótesis principal;
- k: entero cuya existencia está garantizada por la definición de divisibilidad.

Supongamos ahora que la prueba continúa:

> Como $n = 12k = 4(3k)$, definamos $n = 3k$. Entonces $n \in \mathbb{Z}$ y se sigue que $4 \mid n$.

La dificultad no es sólo estética.

La letra n ya tenía un significado: era el entero original del teorema.

La frase «definamos $n = 3k$» intenta reutilizarla para nombrar otro entero.

A partir de ese momento, preguntas muy simples quedan sin respuesta inmediata:

- ¿el símbolo n sigue nombrando al entero original?;
- ¿ahora nombra al factor $3k$?;
- ¿la conclusión $4 \mid n$ se refiere al objeto que queríamos estudiar o al objeto redefinido?

La prueba puede repararse sin cambiar la matemática:

> Como $n = 12k = 4(3k)$ y $3k \in \mathbb{Z}$, concluimos que $4 \mid n$.

O, si realmente necesitamos nombrar el factor:

> Definamos $m = 3k$. Entonces $m \in \mathbb{Z}$ y $n = 4m$, por lo que $4 \mid n$.

La mejora consiste en preservar una identidad ya asignada y utilizar un nombre nuevo para un objeto nuevo.

Una letra, una función local. Una regla práctica muy útil es:

si el objeto cambia, considere cambiar también el símbolo.

No es una ley absoluta.

Hay contextos donde una variable puede reasignarse de manera inequívoca —por ejemplo en un algoritmo o en una sucesión de sustituciones cuidadosamente declaradas—, pero en una demostración fundacional la reutilización innecesaria suele aumentar el costo de lectura.

Pensemos en el siguiente texto:

Sea A un conjunto y sea $x \in A$. Definamos $A = f(x)$. Entonces...

Aunque la gramática sea correcta, la notación es peligrosa.

El símbolo A empezó nombrando un conjunto y después intenta nombrar un valor de una función.

El lector debe borrar mentalmente la primera identidad de A y reemplazarla por otra.

Eso vuelve inestable la historia del argumento.

Es mejor escribir, por ejemplo,

Sea A un conjunto y sea $x \in A$. Definamos $y = f(x)$. Entonces...

Ahora cada símbolo conserva un papel reconocible.

Objeto y valor no son lo mismo. Uno de los conflictos más frecuentes aparece con funciones.

Si escribimos

$$f : X \to Y,$$

el símbolo f representa una **función**.

Para un elemento $x \in X$, el símbolo

$$f(x)$$

representa el **valor de esa función en** x.

Estos objetos están relacionados, pero no son intercambiables.

Por ejemplo, una frase como

«Sea $f(x)$ una función inyectiva»

puede resultar comprensible por costumbre, pero mezcla el nombre del objeto con una de sus evaluaciones.

En C10 preferiremos escribir:

Sea $f : X \to Y$ una función inyectiva.

Y después, cuando aparezca un elemento $x \in X$, utilizar

$$f(x)$$

para su imagen.

Esta distinción importa especialmente cuando la prueba contiene varias funciones, composiciones o igualdades.

Compare:

Como $f = g$, tenemos $f(x) = g(x)$ para todo $x \in X$.

con

Como $f(x) = g(x)$, tenemos $f = g$.

La segunda frase sólo puede ser correcta si la igualdad se ha establecido **para todo** elemento del dominio y ambas funciones tienen dominios y codominios compatibles.

Una notación estable ayuda a no confundir

igualdad de objetos

con

igualdad de valores en un punto.

Convención no significa obligación absoluta. La matemática dispone de muchas convenciones útiles.

Por ejemplo, es frecuente usar:

- n, m, k para enteros o naturales;
- x, y, z para elementos o números variables;
- A, B, C para conjuntos;
- f, g, h para funciones;
- ε, δ para cantidades positivas pequeñas en análisis.

Estas convenciones reducen fricción porque ofrecen expectativas al lector.

Pero no son leyes lógicas.

Podemos llamar u a un conjunto o A a un número si lo declaramos con claridad.

El problema aparece cuando una elección no convencional no aporta ninguna ventaja y obliga al lector a reaprender el código local sin necesidad.

La regla no será entonces

> «use siempre las letras tradicionales».

Será:

> **aproveche las convenciones cuando ayuden y defina con claridad cualquier desviación que pueda producir ambigüedad.**

La notación local debe pagar su costo. A veces una expresión larga aparece muchas veces.

Entonces puede ser útil introducir una abreviatura.

Supongamos que necesitamos repetir con frecuencia

$$(A \cap B) \cup (A \cap C).$$

Podríamos escribir:

> Denotemos por S el conjunto $(A \cap B) \cup (A \cap C)$.

A partir de ahí, usar S puede reducir ruido.

Pero si la expresión sólo aparece una vez, crear el símbolo S puede producir el efecto contrario.

El lector debe ahora recordar una traducción adicional:

$$S \quad \leftrightarrow \quad (A \cap B) \cup (A \cap C).$$

Toda notación nueva tiene un costo cognitivo.

Por eso conviene preguntar:

> **¿este símbolo elimina más trabajo del que introduce?**

Si la respuesta es no, quizá sea mejor conservar la expresión original.

Nombrar puede revelar estructura. Una buena abreviatura no sólo acorta.

También puede hacer visible una idea.

Supongamos que en una prueba aparecen repetidamente dos conjuntos:

$$\{x \in X : P(x)\} \qquad \text{y} \qquad \{x \in X : Q(x)\}.$$

Si la demostración estudia sistemáticamente la relación entre ambos, puede ser útil definir

$$A = \{x \in X : P(x)\}, \qquad B = \{x \in X : Q(x)\}.$$

Entonces una discusión larga sobre predicados puede convertirse en una pregunta estructural sobre

$$A \subseteq B, \qquad A = B, \qquad A \cap B,$$

según el problema.

Aquí la notación no es meramente abreviatura tipográfica.

Organiza el razonamiento.

Pero esa ventaja sólo existe si los símbolos A y B conservan después su significado.

La reutilización puede ser legítima cuando el alcance terminó. No toda repetición de una letra es un error.

Supongamos que una demostración se divide en dos casos completamente separados.

En el primer caso escribimos:

Existe $k \in \mathbb{Z}$ tal que $n = 2k$.

Terminamos ese caso.

En el segundo caso podríamos volver a escribir:

Existe $k \in \mathbb{Z}$ tal que $n = 2k + 1$.

¿Es obligatorio cambiar a otra letra?

No necesariamente.

Si el primer k dejó de estar activo al cerrar el caso y el segundo se introduce de nuevo con claridad, no existe conflicto lógico.

Podemos representarlo así:

$$\boxed{\text{caso 1: } k \text{ local}} \qquad \boxed{\text{caso 2: } k \text{ local}}.$$

El problema no es repetir una letra en todo el universo del texto.

El problema es tener dos significados activos que el lector pueda confundir.

Por eso la estabilidad debe analizarse junto con el **alcance**.

Índices y subíndices también cuentan una historia. Los subíndices son útiles cuando varios objetos cumplen papeles paralelos.

Por ejemplo,

$$x_1, x_2, \ldots, x_n$$

sugiere una familia o lista de objetos relacionados.

En cambio, escribir

$$x,\ y,\ z,\ w,\ u,\ v$$

para seis objetos de la misma clase puede ser menos informativo si la elección de letras no refleja ninguna diferencia conceptual.

Pero tampoco conviene introducir índices sin necesidad.

Si sólo hay dos candidatos, escribir

$$a_1, a_2$$

puede ser útil.

Escribir

$$a_{1,0}, a_{1,1}, a_{2,0}, a_{2,1}$$

cuando no existe una estructura doble real añade una organización que el problema no tiene.

Una notación buena **expresa estructura existente**; no inventa estructura gráfica para parecer más formal.

La notación debe ser compatible con las operaciones que hacemos. Considere una prueba donde A es un conjunto y más adelante aparece

$$A + 1.$$

Tal expresión podría tener significado en algún contexto especializado, pero si nunca se definió una suma entre conjuntos, el lector tiene una alerta legítima.

Del mismo modo, si f es una función y escribimos

$$f + 1,$$

necesitamos un contexto donde esa operación esté definida.

La notación no sólo debe ser estable; también debe ser **tipológicamente compatible** con lo que hacemos con el objeto.

En lenguaje informal podemos preguntarnos:

> «¿este símbolo sigue comportándose como el tipo de objeto que dije que era?»

Esta pregunta descubre muchos errores de escritura e incluso algunos errores matemáticos.

El mismo símbolo en dos niveles puede inducir confusión. Otro caso frecuente ocurre cuando una letra nombra una colección y también uno de sus elementos.

Por ejemplo:

> Sea $A = \{A_1, A_2, A_3\}$. Tomemos $A \in A$.

La frase intenta usar A al mismo tiempo para el conjunto total y para un elemento de ese conjunto.

Aunque pudiera reconstruirse la intención, la escritura es innecesariamente hostil.

Podemos escribir:

> Sea $\mathcal{A} = \{A_1, A_2, A_3\}$. Tomemos $B \in \mathcal{A}$.

Ahora la diferencia de niveles queda visible:

- $\mathcal{A}$: colección;
- B: uno de sus elementos.

No es obligatorio usar letra caligráfica.

Lo esencial es impedir que dos niveles distintos compitan por el mismo nombre dentro del mismo alcance.

Coherencia no significa monotonía visual. También puede ocurrir el defecto contrario: intentar que todo se parezca demasiado.

Si tenemos una función

$$f : X \to Y$$

y un elemento

$$x \in X,$$

no ayuda llamar a la imagen

$$x'$$

si después necesitamos recordar que, en realidad,

$$x' = f(x).$$

A veces es mejor conservar directamente $f(x)$ porque esa notación muestra la dependencia funcional.

En otros contextos, definir

$$y = f(x)$$

sí puede ser conveniente si y tendrá una vida larga en la prueba.

La elección depende de cuánto trabajo ahorre y cuánta estructura conserve.

Cambiar una letra puede cambiar la lectura lógica. Compare las dos redacciones siguientes.

Versión A.

Sea $x \in X$. Como f es sobreyectiva, existe $x \in X$ tal que $f(x) = y$.

Versión B.

Sea $y \in Y$. Como $f : X \to Y$ es sobreyectiva, existe $x \in X$ tal que $f(x) = y$.

En A, la letra x aparece primero como objeto ya fijado y después se reutiliza en una existencia nueva.

Además, y aparece sin introducción.

La idea matemática quizá sea recuperable, pero la dependencia queda dañada.

En B, el orden de nombres refleja la estructura lógica:

$$y \in Y \quad \longrightarrow \quad \exists x \in X \quad \text{tal que} \quad f(x) = y.$$

La notación correcta no es aquí cosmética.

Hace visible qué objeto depende de cuál.

Un nombre estable facilita referencias posteriores. Si un objeto tendrá un papel importante varias líneas después, conviene darle un nombre que pueda recuperarse de forma inequívoca.

Supongamos que construimos

$$z = \frac{x + y}{2}.$$

Si después debemos demostrar tres propiedades distintas de ese mismo objeto, nombrarlo como z puede ser mucho mejor que repetir cada vez

$$\frac{x + y}{2}.$$

El nombre permite escribir:

- primero, z pertenece al dominio pertinente;
- segundo, $x < z$;
- tercero, $z < y$.

La notación actúa como una **referencia persistente**.

Pero, otra vez, esa ventaja desaparece si después reutilizamos z para otro objeto.

Un protocolo de estabilidad notacional. Podemos auditar una prueba con cinco preguntas.

1. ¿Cada símbolo fue introducido antes de utilizarse?

Ésta era la responsabilidad principal de §10.2.

2. ¿Conserva el mismo significado mientras permanece activo?

Si cambia, ¿el cambio fue declarado y terminó el alcance anterior?

3. ¿El tipo de objeto sigue siendo reconocible?

¿Conjunto, elemento, función, valor, número, proposición?

4. ¿La notación nueva reduce realmente complejidad?

¿O obliga al lector a memorizar traducciones innecesarias?

5. ¿Las convenciones elegidas ayudan a leer la estructura?

¿Los nombres paralelos, índices o letras reflejan relaciones reales entre los objetos?

Podemos condensar el protocolo así:

$$\boxed{\text{introducir} \longrightarrow \text{mantener} \longrightarrow \text{usar} \longrightarrow \text{cerrar alcance}}.$$

El criterio de sustitución mental. Existe una prueba práctica de claridad notacional. Elija un símbolo importante y pregunte:

> **¿puedo reemplazar mentalmente este símbolo por su significado original en cada aparición sin obtener contradicciones ni cambios de objeto?**

Si A significa «el conjunto de enteros pares» y en cada línea podemos leer esa frase en lugar de A, la notación es estable.

Si a mitad de la prueba una aparición de A sólo tiene sentido interpretándola como «un entero par particular», hemos detectado un cambio de nivel o de significado.

Este test no sustituye el análisis matemático, pero es muy eficaz para revisar borradores.

Un símbolo debe tener memoria

Una notación auditable permite reconstruir la historia de cada símbolo:

- dónde fue introducido;
- qué objeto representa;
- qué propiedades se le asignaron;
- de qué otros objetos depende;
- durante qué tramo permanece activo.

Reutilizar una letra no es un error por sí mismo. El problema aparece cuando dos significados incompatibles permanecen activos o cuando el cambio ocurre sin señal suficiente para el lector.

Comprobación. Examina las tres versiones siguientes.

A. Función y valor.

Sea $f : X \to Y$ una función. Sea $x \in X$. Como $f(x)$ es inyectiva, de $f(x) = f(y)$ concluimos $x = y$.

B. Reutilización de nombre.

Sea $A \subseteq X$ y sea $x \in A$. Definimos $A = f(x)$. Como $A \in Y$, continuamos...

C. Símbolo local útil.

Sean $x < y$ números reales. Definamos

$$m = \frac{x + y}{2}.$$

Demostraremos que $x < m < y$.

Para cada caso:

1. identifica qué representa cada símbolo;
2. decide si el significado permanece estable;
3. localiza cualquier conflicto entre objeto y valor o entre dos niveles de objetos;
4. reescribe sólo lo necesario;
5. decide si la notación nueva reduce o aumenta el trabajo del lector.

Respuesta razonada.

En **A**, el conflicto aparece en la frase

«$f(x)$ es inyectiva».

La inyectividad es una propiedad de la función f, no del valor particular $f(x)$.

Además, aparece y sin que se haya declarado que pertenece a X.

Una versión adecuada es:

> Sea $f : X \to Y$ una función inyectiva. Sean $x, y \in X$ y supongamos que $f(x) = f(y)$. Por inyectividad de f, concluimos que $x = y$.

La reparación conserva tres niveles:

- f: función;
- x, y: elementos del dominio;
- $f(x), f(y)$: valores de la función.

En **B**, el símbolo A cambia de identidad.

Primero representa un subconjunto de X; después se redefine como el valor $f(x)$.

Ese cambio destruye la posibilidad de recuperar de manera estable qué significaba A en las líneas anteriores.

Podemos escribir:

> Sea $A \subseteq X$ y sea $x \in A$. Definamos $y = f(x)$. Como $y \in Y$, continuamos...

Ahora A conserva su papel de conjunto y y nombra el nuevo objeto.

En **C**, la notación local es útil.

La expresión

$$\frac{x+y}{2}$$

será utilizada como objeto central de varias comparaciones.

Nombrarla m reduce repetición y hace más visible la estructura:

$$x < m < y.$$

El símbolo m tiene además una interpretación natural como punto medio, aunque la prueba no depende de esa palabra.

La diferencia entre los tres casos puede resumirse así:

$$\boxed{\text{notación estable: comprime estructura}}$$

mientras que $\boxed{\text{notación inestable: obliga a reconstruir identidad}}$.

De los nombres a la arquitectura visible. Una vez que los objetos están bien introducidos y sus nombres permanecen estables, aparece una pregunta de escala mayor.

Una demostración no es una sola oración larga.

Puede contener:

- dos direcciones de un bicondicional;
- varios casos;
- una construcción auxiliar;
- una contradicción;

- una cadena de cálculos seguida por una conclusión conceptual.

El lector necesita ver cuándo cambia la función lógica del texto.

Esa estructura no depende sólo de las letras que elegimos, sino también de cómo **agrupamos las inferencias en párrafos y transiciones**.

Ése será el contenido de §10.4.

10.4. Párrafos de prueba y transiciones lógicas

Considere este argumento, cuya matemática ya conocemos:

> Sean A, B, C conjuntos. Si $A \subseteq B$, entonces $A \cap C \subseteq B \cap C$. Si $x \in A \cap C$, está en A y en C, por tanto en B y en C, luego en $B \cap C$. Si $A \cap C \subseteq B \cap C$ para todo conjunto C, tomamos $C = A$, luego $A \subseteq B$.

El texto contiene las dos direcciones de una equivalencia, pero no anuncia esa equivalencia ni señala el cambio de dirección. La primera oración parece fijar el único objetivo; la última empieza otra prueba y obliga a reinterpretar lo anterior. Un lector puede reconstruirlo, aunque una organización más clara le ahorraría ese trabajo.

Separar obligaciones. La siguiente versión conserva el argumento y organiza sus dependencias.

Proposición. Para cualesquiera conjuntos A y B, se tiene $A \subseteq B$ si y sólo si, para todo conjunto C, se cumple $A \cap C \subseteq B \cap C$.

Demostración. Supongamos primero que $A \subseteq B$. Sea C un conjunto arbitrario y sea $x \in A \cap C$. Entonces $x \in A$ y $x \in C$. La inclusión $A \subseteq B$ da $x \in B$; junto con $x \in C$, obtenemos $x \in B \cap C$. Queda probada la inclusión para todo C.

Recíprocamente, supongamos que $A \cap C \subseteq B \cap C$ para todo conjunto C. En particular, podemos elegir $C = A$, con lo que $A = A \cap A \subseteq B \cap A \subseteq B$. Por tanto, $A \subseteq B$.

Cada párrafo responde a una obligación distinta. «Supongamos primero» abre la primera dirección; «recíprocamente» abre la segunda y sustituye la hipótesis disponible. La primera dirección termina antes de que comience la otra. No estamos enseñando de nuevo cómo probar una equivalencia: estamos haciendo visible una arquitectura ya disponible en C06.

La separación también protege las hipótesis. En el segundo párrafo no se permite utilizar $A \subseteq B$ porque era la hipótesis de la dirección anterior y ahora es el objetivo. La transición informa al lector de ese cambio.

Qué hace un párrafo. Un párrafo puede introducir una construcción y verificar sus propiedades, resolver un caso o aplicar un lema para alcanzar una conclusión parcial. La unidad la determina esa función. Si una construcción breve y su verificación caben en cuatro frases, separarlas artificialmente puede volver más difícil seguir el objeto construido. Si una prueba contiene dos direcciones independientes, juntarlas en un bloque largo oculta su frontera.

En una prueba por casos conviene que el inicio de cada párrafo nombre la condición que lo gobierna. Para demostrar $|t| \geq t$ con $t \in \mathbb{R}$, por ejemplo:

> Si $t \geq 0$, la definición de valor absoluto da $|t| = t$.
>
> Si $t < 0$, tenemos $|t| = -t$. Como $-t > 0 > t$, se cumple $|t| > t$.
>
> Los dos casos cubren todos los números reales; en ambos se obtiene $|t| \geq t$.

Los dos primeros párrafos contienen argumentos locales. El tercero cierra la afirmación global al registrar que la división fue exhaustiva. No todas las pruebas por casos necesitan una oración final tan explícita, pero aquí permite distinguir la conclusión local del cierre completo.

Los conectores declaran relaciones. «Como» introduce una razón; «por tanto» anuncia una consecuencia; «para comprobar» declara un objetivo todavía pendiente; «en particular» restringe una afirmación disponible; «recíprocamente» cambia de dirección. Elegir un conector exige saber qué relación existe entre las frases.

Compare:

> Para probar que $u \in B$, usamos $A \subseteq B$ y $u \in A$.

con

> Como $u \in B$, usamos $A \subseteq B$ y $u \in A$.

La primera frase distingue el objetivo de los datos. La segunda presenta ese objetivo como una razón ya disponible. Aunque el lector conozca la intención, el texto le ofrece un orden de dependencia diferente. Un conector no justifica por sí mismo una inferencia; la razón nombrada debe autorizarla.

La palabra «entonces» admite varios usos. En una prueba conviene comprobar que significa consecuencia y que no está sustituyendo una relación más precisa, como «en el segundo caso» o «al aplicar la definición». Tampoco hace falta comenzar cada frase con un conector: dos oraciones pueden pertenecer a un mismo movimiento y resultar inequívocas sin una señal adicional.

Comprobación. Reorganice el texto siguiente sin cambiar su contenido: «Sean $f : X \to Y$ y $g : Y \to Z$ inyectivas. Sean $u, v \in X$ con $g(f(u)) = g(f(v))$. Son iguales. Por inyectividad son iguales. La composición es inyectiva». Identifique qué deben nombrar las dos frases «son iguales» y qué función cumple cada tramo.

Respuesta razonada. Una versión verificable es:

> Sean $u, v \in X$ y supongamos que $(g \circ f)(u) = (g \circ f)(v)$. Esto significa $g(f(u)) = g(f(v))$. Como $f(u), f(v) \in Y$ y g es inyectiva, obtenemos $f(u) = f(v)$.
>
> La inyectividad de f da ahora $u = v$. Puesto que u y v eran elementos arbitrarios de X, la composición $g \circ f : X \to Z$ es inyectiva.

El primer tramo reduce la igualdad de valores de la composición a una igualdad en Y; el segundo usa esa igualdad para cerrar la obligación en X. También sería aceptable conservar toda esta prueba corta en un párrafo. La división elegida deja visibles las dos aplicaciones, pero la auditabilidad depende de nombrarlas y de identificar sus entradas, no del número de saltos de línea.

Una lectura de los comienzos de párrafo debería permitir reconocer qué obligación se está resolviendo. Después, la lectura de sus frases debe proporcionar las razones que sostienen ese movimiento.

10.5. Ecuaciones como parte de la oración

Una fórmula en una línea separada sigue perteneciendo a la oración. Considere la frase «Como $a = b$, obtenemos» seguida del bloque

$$a + c = b + c.$$

El punto cierra la frase que empezó antes del bloque. La posición de la expresión no modifica su función gramatical. Una lectura en voz alta permite comprobar si las palabras y los símbolos forman una afirmación completa.

Una cadena tiene una relación determinada. Sean $r, s \in \mathbb{R}$. La cadena

$$(r + s)^2 = r^2 + 2rs + s^2$$

conecta dos números iguales. Si además conocemos $r, s \geq 0$, podemos deducir $(r + s)^2 \geq r^2 + s^2$, pero no añadir esa desigualdad mediante otro signo igual. Una versión coherente es:

Como $r, s \geq 0$, tenemos $2rs \geq 0$. Al desarrollar el cuadrado, obtenemos

$$(r+s)^2 = r^2 + s^2 + 2rs \geq r^2 + s^2.$$

La mezcla de igualdad y desigualdad es legítima: cada relación afirma lo que corresponde. Si el propósito es destacar la consecuencia lógica, podemos escribirla en prosa después de la identidad.

Los signos de implicación y equivalencia relacionan proposiciones. Para $t \in \mathbb{R}$, la expresión

$$t = 5 \implies t^2 = 25$$

es una implicación válida. La expresión $t = 5 \iff t^2 = 25$ es falsa, porque $t = -5$ satisface la segunda proposición y no la primera. Bajo la hipótesis adicional $t \geq 0$, sí se obtiene esa equivalencia. La exposición debe hacer visible la restricción que autoriza la reversibilidad.

En cambio, «$t \Rightarrow t^2 \Rightarrow 25$» usa números o expresiones numéricas como si fueran afirmaciones. Puede servir como apunte personal de una operación, pero el lector de una prueba necesita saber qué se afirma sobre esos objetos.

Introducir y cerrar los bloques. Compare «Aplicamos» seguido de $a(b+c) = ab+ac$ con «Por distributividad,» seguido de la misma identidad. La segunda apertura permite identificar la razón; la primera deja indeterminado qué se aplica. A menudo conviene escribir toda la oración: «Por distributividad, $a(b+c) = ab+ac$». El bloque se reserva para una fórmula que requiere espacio o merece atención particular.

También se puede continuar una oración después del bloque:

> Si $u \in A \setminus B$, entonces
>
> $$u \in A \quad \text{y} \quad u \notin B,$$
>
> por definición de diferencia de conjuntos.

La coma señala que la justificación posterior pertenece a la misma frase. El lector debe poder reconocer dónde termina la afirmación y cómo se enlaza con su razón.

Cadenas largas y razones locales. Si $f : X \to Y$, $g : Y \to Z$ y $h : Z \to W$, para cada $x \in X$ tenemos, por la definición de composición,

$$\begin{aligned}((h \circ g) \circ f)(x) &= (h \circ g)(f(x)) \\ &= h(g(f(x))) \\ &= h((g \circ f)(x)) \\ &= (h \circ (g \circ f))(x).\end{aligned}$$

Todas las igualdades proceden de la misma definición. Basta nombrar esa razón antes de la cadena; repetirla al final de cada fila añadiría poco. El texto debe completar después la obligación sobre funciones: ambas composiciones van de X a W y coinciden en cada elemento de X, por lo que son iguales. La cadena compara valores; la frase final establece la igualdad de las funciones.

Si una fila utilizara un resultado distinto, la explicación debería quedar junto a ella o en una oración cercana. La razón debe poder localizarse sin recorrer todo el capítulo buscando cuál podría encajar.

Referencias que se reutilizan. Una identidad puede recibir un nombre o una etiqueta cuando será usada otra vez. Supongamos que hemos obtenido

$$p = q + r. \tag{10.5.1}$$

Más adelante, «al sustituir p mediante (10.5.1)» identifica tanto el dato como la operación. Para una identidad usada sólo una vez suele bastar una referencia verbal cercana. Numerar todas las fórmulas puede crear una segunda tarea de memoria sin mejorar la verificación.

Comprobación. Reescriba este apunte como argumento completo: «$x \geq 0$, $y \geq 0$, $x + y = 0 = x = 0 = y = 0$». Determine cuáles son hipótesis, qué debe concluirse y cuáles relaciones pueden escribirse con $=$.

Respuesta razonada. Las hipótesis son $x, y \in \mathbb{R}$, $x, y \geq 0$ y $x + y = 0$. Se quiere demostrar que $x = y = 0$. Una prueba breve es:

> Como $y \geq 0$, tenemos $x \leq x + y = 0$. Junto con $x \geq 0$, esto da $x = 0$. Al sustituir en $x + y = 0$, obtenemos $y = 0$.

La cadena del apunte junta conclusiones sin mostrar por qué se obtienen. El texto corregido conserva la identidad $x + y = 0$ y añade la desigualdad que permite establecer la primera conclusión. Una vez demostradas ambas igualdades, $x = y = 0$ sí es una cadena legítima.

La fórmula final debe expresar una relación matemática precisa y ocupar un lugar reconocible en la oración y en el argumento.

10.6. Referencias a definiciones y resultados previos

La frase «por lo anterior» puede ser suficiente cuando una única igualdad acaba de aparecer. En una prueba que contiene una definición, dos inclusiones y una hipótesis de inyectividad, esa frase puede dejar al lector sin saber qué autoriza el paso. La precisión necesaria depende de cuántos antecedentes son posibles y del papel que cumple el resultado citado.

Identificar la herramienta y sus entradas. Recordemos el resultado de C04: si $f : X \to Y$ es inyectiva, entonces $f(u) = f(v)$ con $u, v \in X$ implica $u = v$. Una aplicación auditable puede escribir:

> Como $u, v \in X$ y $f(u) = f(v)$, la inyectividad de f da $u = v$.

El texto identifica la propiedad utilizada y muestra sus entradas. La mera frase «por C04» obligaría a buscar cuál de sus muchos resultados interviene. Una referencia a capítulo es útil para recuperar una noción, pero una razón local suele necesitar el nombre de esa noción o el enunciado de la herramienta.

Una referencia numerada tampoco reemplaza la comprobación de hipótesis. Si un lema requiere $a \neq 0$, invocarlo con a sin determinar si se anula deja pendiente una condición. El lema puede ser correcto y estar perfectamente localizado; su aplicación necesita todavía ese dato.

Una referencia puede ser verbal. No toda dependencia requiere un número. «Por la definición de preimagen» localiza una herramienta única en el contexto. «Por la igualdad de conjuntos que acabamos de probar» puede bastar si sólo se demostró una. En cambio, «por un teorema conocido» no permite identificar una herramienta sin adivinarla.

Una buena referencia permite recuperar el enunciado empleado y confrontarlo con los datos actuales. Puede hacerlo mediante nombre, número o una formulación breve. La elección debe servir a esa recuperación y evitar códigos que el lector tenga que aprender para seguir una prueba elemental.

Cuando conviene escribir un lema local. Si una misma observación sostiene varios pasos, puede merecer una formulación separada. Por ejemplo:

Lema local. Para $f : X \to Y$ y $S, T \subseteq X$, si $S \subseteq T$, entonces $f(S) \subseteq f(T)$.

Demostración. Si $y \in f(S)$, existe $x \in S$ con $f(x) = y$. Como $S \subseteq T$, ese mismo elemento pertenece a T, por lo que $y \in f(T)$.

Después, «aplicando el lema local a $S = A \cap B$ y $T = A$» permite verificar los dos subconjuntos y su inclusión. La identificación de entradas también impide usar por descuido un lema de imágenes como si fuera de preimágenes.

En una prueba muy corta podría incorporarse la observación directamente. Separarla es útil cuando volverá a emplearse, cuando organiza una dependencia importante o cuando su prueba interrumpiría el argumento principal.

Citar sin esconder ni saturar. Las referencias deben tener una densidad proporcional a la dificultad. Para el lector de C10, escribir $3(k+1) = 3k+3$ no exige remitir a la sección de distributividad. Si una prueba utiliza la cancelación en un cuerpo, en cambio, puede ser relevante indicar que el factor cancelado es no nulo y que la propiedad fue demostrada desde los axiomas en C05.

La precisión tampoco obliga a copiar cada resultado previo. Un recordatorio breve puede exponer la hipótesis decisiva: «Como $a \neq 0$, la cancelación multiplicativa en el cuerpo permite…». Esa frase informa por qué la herramienta está disponible en ese punto.

Comprobación. Audite esta aplicación: «Sean $f : X \to Y$ y $g : Y \to Z$. Si $g \circ f$ es inyectiva, entonces f es inyectiva, por el resultado sobre composiciones». Reescríbala sin usar una referencia que obligue a adivinar y distinga la hipótesis disponible de una que no hace falta.

Respuesta razonada. El resultado «la composición de dos inyecciones es inyectiva» no demuestra esta afirmación mediante una aplicación directa. Una prueba auditable es:

> Sean $u, v \in X$ y supongamos que $f(u) = f(v)$. Al aplicar g a ambos lados, obtenemos $(g \circ f)(u) = (g \circ f)(v)$. La inyectividad de $g \circ f$ da $u = v$. Por tanto, f es inyectiva.

La herramienta invocada es la definición de inyectividad de la composición, cuya hipótesis fue declarada. No es necesario suponer que g es inyectiva. La referencia imprecisa de la primera versión ocultaba esa diferencia entre resultados próximos.

Al revisar una referencia, conviene reconstruir la aplicación completa: enunciado utilizado, objetos que lo instancian, hipótesis verificadas y conclusión que efectivamente proporciona.

10.7. Pronombres, alcance y ambigüedad

«Éste pertenece al dominio» puede ser una afirmación clara si sólo se acaba de introducir un elemento. Si la frase aparece después de mencionar una función, un conjunto y dos elementos, el lector puede tener que decidir cuál es su antecedente. La dificultad afecta la matemática: una misma propiedad puede ser verdadera para un objeto y falsa para otro.

Recuperar el referente. Considere: «Sean $A \subseteq X$, $f : X \to Y$ y $y \in f(A)$. Existe $x \in A$ con $f(x) = y$. Como éste está en el dominio, pertenece a la imagen». La última oración podría referirse a x o a y; además, estar en el dominio no significa pertenecer a la imagen, porque los dos conjuntos pueden tener elementos distintos.

Una versión que conserva la intención pertinente es: «Como $x \in A \subseteq X$, el valor $f(x)$ está definido. La igualdad $f(x) = y$ muestra que $y \in f(X)$». Nombrar x y y permite distinguir el argumento de entrada del valor de salida. Una revisión de pronombres puede revelar que, además de una ambigüedad, había una frase matemática que necesitaba reparación.

No hace falta eliminar todos los pronombres. En «$u = v$; esta igualdad permite sustituir u por v», el referente es único y la referencia resulta económica. El criterio consiste en verificar que la sustitución del pronombre por su antecedente produce la afirmación que el texto pretende sostener.

Referirse a afirmaciones, no sólo a objetos. «Esto prueba el resultado» también necesita un antecedente. En una existencia única, haber construido un candidato sólo cierra la existencia. Si después de esa construcción aparece la frase, el lector puede creer que ya se probó todo. Conviene escribir «la construcción prueba la existencia» y reservar el cierre completo para cuando también se haya establecido la unicidad.

Igualmente, «el caso anterior» debe permitir identificar qué caso terminó y qué propiedad se reutiliza. Si una propiedad dependía de una condición local, no puede trasladarse al nuevo caso sin comprobar esa condición.

Cerrar el alcance de una elección. Los nombres introducidos en una subprueba tienen un alcance. Supongamos que, para cada $y \in Y$, la sobreyectividad de $f : X \to Y$ permite tomar un $x \in X$ con $f(x) = y$. La frase «este x sirve para todos los y» cambia la dependencia: la elección anterior se hizo después de fijar y y puede variar con él.

Una redacción suficiente es: «Fijado $y \in Y$, elegimos $x \in X$ con $f(x) = y$». Si se necesita repetir el argumento para otro valor, se abre otra elección. Aquí no estamos reconstruyendo la teoría de cuantificadores de C03; estamos conservando en la prosa una dependencia que ya fue establecida.

Cuando dos subpruebas tienen nombres locales iguales, el cambio de párrafo y la introducción explícita pueden separar sus alcances. Si ambos candidatos deben compararse después, es preferible nombrarlos x_1, x_2 y mantenerlos activos. El contexto decide qué objetos necesita recordar el lector.

Comparar versiones. Sean $A, B \subseteq X$ y $f : X \to Y$. Para justificar $f(A \cap B) \subseteq f(A) \cap f(B)$, alguien escribe:

> Sea $y \in f(A \cap B)$. Existe $x \in A \cap B$ con $f(x) = y$. Como pertenece a ambos, también pertenece a ambas imágenes.

El primer «pertenece» debe referirse a x; el segundo, a y. El texto cambia de referente sin nombrarlo. Una edición mínima dice:

> Como $x \in A$ y $x \in B$, el mismo testigo x muestra que $y = f(x) \in f(A)$ y $y \in f(B)$. Por tanto, $y \in f(A) \cap f(B)$.

La edición también hace visible que usamos el mismo testigo. Esa información será importante cuando se compare esta inclusión con su recíproca, que no vale para una función arbitraria.

Comprobación. Reescriba: «Sean $a, b \in \mathbb{R}$ con $a < b$. Su promedio está entre ellos porque su diferencia es positiva». Indique a qué diferencia se refiere y compruebe ambas desigualdades.

Respuesta razonada. Definamos $m = (a + b)/2$. Las diferencias pertinentes son

$$m - a = \frac{b - a}{2} > 0, \qquad b - m = \frac{b - a}{2} > 0.$$

Ambas son positivas por $a < b$. La primera demuestra $a < m$ y la segunda $m < b$. Una versión completa puede decir: «Las igualdades anteriores y $b - a > 0$ dan $a < m < b$». El término «su diferencia» ocultaba que había dos obligaciones, cada una asociada a una diferencia distinta.

Una prueba debe conservar el referente de sus palabras con el mismo cuidado que conserva el significado de sus símbolos.

10.8. Cuánto detalle escribir y el uso de «claramente»

Una prueba puede omitir información necesaria o acumular tantos detalles que el lector pierda su dirección. El lector de referencia de este capítulo conoce las herramientas anteriores de FND y puede hacer cálculos elementales; todavía necesita ver qué herramienta organiza el paso decisivo.

Reconstrucción segura. Considere la afirmación $f^{-1}(U \cap V) = f^{-1}(U) \cap f^{-1}(V)$ para $f : X \to Y$ y $U, V \subseteq Y$. Escribir «se sigue claramente» puede resultar suficiente para quien ya usa rutinariamente preimágenes, pero no deja visible el mecanismo. Una versión breve y verificable es:

> Para cada $x \in X$, pertenecer a $f^{-1}(U \cap V)$ significa que $f(x)$ pertenece a U y a V. Esto equivale a que $x \in f^{-1}(U)$ y $x \in f^{-1}(V)$, es decir, a que $x \in f^{-1}(U) \cap f^{-1}(V)$. Ambos conjuntos tienen, por tanto, los mismos elementos.

La expansión proporciona la razón que sostiene la igualdad. No necesita reproducir la definición de función ni desarrollar de nuevo la doble inclusión de C06. El nivel de detalle se concentra en la correspondencia entre pertenencias.

Detectar qué se está comprimiendo. Una omisión es razonable si el lector puede recuperarla mediante una operación conocida y única, sin añadir hipótesis ni encontrar una nueva idea. Por ejemplo, después de $n = 7k$ con $k \in \mathbb{Z}$, escribir $2n = 7(2k)$ permite reconstruir el cálculo de inmediato. En cambio, de $y \in f(A) \cap f(B)$ no puede deducirse $y \in f(A \cap B)$ por un cálculo rutinario: los testigos de ambas imágenes podrían ser distintos.

Esa diferencia proporciona una prueba práctica para la edición. Expanda mentalmente la frase omitida. Si exige elegir entre varias herramientas, demostrar un resultado nuevo o controlar una hipótesis que todavía no aparece, la exposición necesita más información. Si sólo exige distribuir un factor o sustituir una igualdad ya escrita, suele bastar una compresión.

Quitar un adverbio. Al revisar «claramente», «obviamente» o «trivialmente», borre primero la palabra y lea el paso. Si la razón sigue siendo recuperable, el adverbio puede ser innecesario. Si desaparece la única apariencia de justificación, escriba la razón.

Compare «Claramente, $x \neq 0$» con «Como $x > 0$, se tiene $x \neq 0$». La segunda frase registra una dependencia de la que quizá dependa después una división. En otra prueba donde se acaba de fijar $x > 0$ y no existe ambigüedad, podría bastar escribir directamente «dividimos por x». La decisión depende de la distancia y del riesgo de perder la hipótesis.

El mismo problema aparece sin adverbios, en frases como «se obtiene», «es inmediato» o «el cálculo da». El control debe recaer sobre la inferencia omitida, no sobre una lista de palabras prohibidas.

Retirar explicación rutinaria. Para demostrar que $n \mapsto 3n + 2$ es inyectiva en $\mathbb{Z}$, una vez fijados $u, v \in \mathbb{Z}$ con $3u + 2 = 3v + 2$, basta explicar que restar 2 y cancelar el factor no nulo 3 da $u = v$. Enumerar por separado que 2 es entero, que su opuesto existe y que podemos sumar iguales a ambos miembros ocultaría la operación relevante.

Una prueba más extensa puede necesitar un mapa inicial: «primero construiremos el candidato y después verificaremos su unicidad». Esa frase reduce el trabajo de orientación y permite comprimir detalles posteriores. El detalle útil puede consistir en una señal de arquitectura, aunque no añada una fórmula.

Comprobación. Sean $f : X \to Y$ inyectiva y $A, B \subseteq X$. Evalúe el nivel de detalle de «Claramente, $f(A) \cap f(B) \subseteq f(A \cap B)$». Redacte el paso que el lector necesitaría reconstruir y explique por qué es decisivo.

Respuesta razonada. Si $y \in f(A) \cap f(B)$, existen $a \in A$ y $b \in B$ con $f(a) = y = f(b)$. La inyectividad de f da $a = b$. Por tanto, $a \in A \cap B$ y $y = f(a) \in f(A \cap B)$.

La información decisiva es que dos testigos inicialmente distintos coinciden por inyectividad. Ocultarla impide ver dónde se usa la hipótesis que hace verdadera la inclusión. La edición puede comprimirse, pero debe conservar ese encuentro de testigos y su razón.

El detalle suficiente permite al lector verificar el paso sin descubrir por su cuenta la idea que falta en el texto.

10.9. Auditoría por terceros y edición final

El autor conoce la solución y puede completar sin advertirlo una frase incompleta. Otro lector dispone sólo del argumento escrito. Pedirle que reconstruya la prueba permite contrastar la exposición con esa diferencia de información.

Qué pedir al lector. «¿Se entiende?» produce una respuesta demasiado general para editar. Es más útil pedir que el lector identifique los objetos, las hipótesis y el objetivo; explique la razón de los pasos decisivos; y marque el primer lugar donde tuvo que añadir información que el texto no proporcionaba.

Si el lector no puede justificar un paso, todavía debemos determinar qué ocurrió. Puede faltar un prerrequisito que el texto había declarado, puede existir una dependencia expositiva omitida o puede haber una inferencia incorrecta. Estas situaciones requieren acciones diferentes. La crítica es evidencia para revisar; su mera existencia no demuestra que la prueba sea falsa.

Revisar con preguntas separadas. Una lectura final resulta más eficaz si concentra su atención en una clase de problemas cada vez. El orden siguiente evita pulir una frase que después habrá que retirar por una razón matemática.

1. **Matemática y lógica.** Comprobar que las afirmaciones y las inferencias son correctas con las hipótesis disponibles. Verificar que se cumplieron todas las obligaciones del enunciado.
2. **Dependencias.** Localizar definiciones, hipótesis y resultados que autorizan los pasos decisivos. Comprobar las entradas de cada aplicación.
3. **Notación y alcance.** Seguir la historia de cada símbolo y cerrar elecciones e hipótesis locales. Distinguir objetos de sus valores o representaciones.
4. **Prosa.** Leer las oraciones, incluidas sus fórmulas; comprobar referentes, conectores y puntuación.
5. **Arquitectura.** Reconocer direcciones, casos, construcciones y cierres mediante los párrafos y sus transiciones.
6. **Economía.** Retirar repeticiones que no añaden verificabilidad y completar las omisiones que todavía exigen adivinar.

Estas preguntas pueden requerir volver a un pase anterior. Si una corrección de pronombre revela que se usó el objeto equivocado, corresponde reabrir el control matemático. Separar focos de revisión ayuda a prestar atención, pero no convierte la edición en una secuencia irreversible.

Una auditoría completa de una prueba corta. Se quiere demostrar que, para $A \subseteq X$ y $f : X \to Y$, se cumple $A \subseteq f^{-1}(f(A))$. El borrador dice:

> Tome x en el conjunto. Por la definición pertenece a la imagen y entonces a la preimagen. Luego la inclusión.

La idea correcta consiste en utilizar x como testigo de que $f(x) \in f(A)$. La primera frase no identifica el conjunto; la segunda cambia tácitamente de x a $f(x)$ y después vuelve a x. El lector debe reconstruir una parte sustancial del contenido.

Una versión revisada es:

> Sea $x \in A$. Como $A \subseteq X$, el valor $f(x)$ está definido. El propio x es testigo de que $f(x) \in f(A)$. Por definición de preimagen, esto da $x \in f^{-1}(f(A))$. Puesto que x era arbitrario en A, queda probada la inclusión.

El pase matemático comprueba las dos definiciones y la inclusión. El de dependencias identifica el testigo y la condición $A \subseteq X$. El de notación distingue x de $f(x)$. El de prosa verifica que «esto» se refiere a una única pertenencia inmediata. La arquitectura cabe en un párrafo porque sólo hay una obligación de inclusión. La revisión de economía puede comprimir «el valor está definido», pero aquí esa frase sirve para hacer visible la entrada de la función.

Recibir una observación y verificar la edición. Si un lector pide que se explique de dónde salió un testigo, la revisión debe localizar la definición o hipótesis que garantiza su existencia. Añadir una explicación extensa del método de descubrimiento podría no resolver la observación. La edición debe atender al punto donde el lector perdió la dependencia.

Después de corregir, hay que releer el tramo modificado con su contexto. Un cambio de letra puede romper una referencia posterior; separar párrafos puede dejar activa una hipótesis que debía cerrarse; eliminar una frase puede borrar la única justificación de una elección. La versión final es el objeto que se verifica, no una suma de correcciones aisladas.

Comprobación. Un lector propone reemplazar el borrador anterior por «Por definición de imagen y preimagen, $A \subseteq f^{-1}(f(A))$». ¿Es una edición correcta? ¿Es suficiente para el lector de referencia de C10? Produzca una alternativa breve que conserve el mecanismo.

Respuesta razonada. La afirmación es correcta, pero esa versión vuelve a comprimir el testigo y el cambio entre objeto y valor que la actividad pretende auditar. Una alternativa breve es: «Para $x \in A$, el propio x muestra que $f(x) \in f(A)$; por definición de preimagen, $x \in f^{-1}(f(A))$. Esto demuestra la inclusión». Se presuponen aquí $f : X \to Y$ y $A \subseteq X$, declarados en el enunciado. La explicación conserva las dos dependencias y el testigo sin repetir el desarrollo completo.

Una prueba está lista para otro lector cuando su versión escrita permite reconstruir y verificar las dependencias de las que procede la conclusión.

10.10. Ejercicios y soluciones desarrolladas

La tarea principal de estos ejercicios es editar o auditar matemática elemental ya disponible. Una edición debe preservar el enunciado, hacer visibles las dependencias y permitir comprobar la conclusión. Cuando el borrador contiene un error matemático, hay que identificarlo antes de producir la versión corregida.

Ejercicios.

Nivel I — Edición guiada. **Ejercicio 1.** Reescriba la apertura «Como $n = 20k$, $5 \mid n$» para demostrar que todo entero divisible por 20 es divisible por 5. Declare el origen y dominio de k; cierre la obligación con un testigo entero.

Ejercicio 2. Se afirma que $q : \mathbb{Z} \to \mathbb{Z}$, $q(t) = 5t - 4$, es inyectiva. Edite: «Sean x, y, como q es igual, $5x - 4 = 5y - 4$, $x = y$». No suponga la inyectividad que se quiere demostrar.

Ejercicio 3. Sean A, B, C conjuntos. Organice en dos párrafos la prueba de $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$, una dirección por párrafo. Nombre los objetos arbitrarios y el cierre de cada inclusión.

Ejercicio 4. Reescriba y puntúe: «a, b reales positivos, $a/b = 2 \Rightarrow a = 2b = a - b = b$». Determine qué igualdad del apunte es incorrecta y conserve la conclusión válida $a - b = b$.

Ejercicio 5. Sean $f : X \to Y$, $A \subseteq X$ y $u \in A$. Edite: «Su imagen está en ella y éste pertenece a la preimagen». Escriba las pertenencias pertinentes con referentes inequívocos.

Nivel II — Auditoría y comparación. **Ejercicio 6.** Para $a, b \in \mathbb{R}$, compare «Si $a < b$, su punto medio está dentro» con una prueba de $a < (a + b)/2 < b$. Produzca una versión breve que haga visibles las dos razones. Identifique qué detalle rutinario puede omitirse.

Ejercicio 7. Un borrador dice: «Sean a, b, c elementos de un cuerpo con $ac = bc$. Cancelamos c, de modo que $a = b$». Determine la condición faltante y redacte una aplicación auditable de cancelación. Explique qué ocurre si $c = 0$.

Ejercicio 8. Sean $f : X \to Y$ y $U, V \subseteq Y$. Escriba una prueba de $f^{-1}(U \setminus V) = f^{-1}(U) \setminus f^{-1}(V)$ en un párrafo con una cadena de equivalencias de pertenencia. Indique qué definiciones autorizan la cadena.

Ejercicio 9. Para mostrar que todo entero tiene un único opuesto aditivo, un autor construye $-n$ y termina con «Esto prueba el resultado». Separe existencia y unicidad, cierre ambas y explique qué ocultaba el demostrativo.

Ejercicio 10. Sean $f : X \to Y$ y $g : Y \to Z$. Edite: «Si la composición es sobreyectiva, dado z, existe x, luego la otra también». Identifique qué función resulta sobreyectiva y compruebe el dominio del testigo que necesita.

Nivel III — Reescritura e integración. **Ejercicio 11.** Se dispone de la siguiente matemática: para $A \subseteq X$ y $f : X \to Y$ inyectiva, $x \in f^{-1}(f(A))$ implica que existe $a \in A$ con $f(x) = f(a)$; la inyectividad da $x = a$. Redacte la prueba completa de $f^{-1}(f(A)) = A$ en dos direcciones. Explique por qué omitir la inyectividad cambia el resultado.

Ejercicio 12. Se quiere demostrar por inducción que $s_n = 4n+1$ si $s_0 = 1$ y $s_{n+1} = s_n + 4$ para $n \in \mathbb{N}$. Edite «Vale al principio; si vale, sumamos 4; listo». Mantenga visibles base, hipótesis, paso y conclusión sin desarrollar otra vez la teoría de inducción.

Ejercicio 13. Se quiere probar que existe un único par real (x, y) con $x + 3y = 17$ y $2x - y = -1$. La matemática disponible proporciona el candidato $(2, 5)$ y permite probar unicidad al restar las ecuaciones de dos soluciones. Redacte dos párrafos, uno por obligación. Un borrador comienza la unicidad con «Como el único par es $(2, 5)$». Explique el problema de dependencia y evítelo.

Ejercicio 14. Audite dos versiones de la inyectividad de $p : \mathbb{R} \to \mathbb{R}$, $p(t) = t^3$: A, «Claramente es creciente y por tanto inyectiva»; B, «Si $u^3 = v^3$, entonces $(u-v)(u^2+uv+v^2) = 0$; de aquí $u = v$». Produzca una versión que justifique la conclusión sin presuponer un resultado de monotonía no demostrado.

Ejercicio 15. Redacte y audite una prueba completa de que, para $f : X \to Y$ y $A, B \subseteq X$, siempre se cumple $f(A \cup B) = f(A) \cup f(B)$. La prueba debe permitir localizar objetos, testigos, casos y cierres. Proponga después una compresión que conserve esas dependencias.

Soluciones desarrolladas. Solución 1. Sea $n \in \mathbb{Z}$ y supongamos que $20 \mid n$. Por definición de divisibilidad, existe $k \in \mathbb{Z}$ con $n = 20k$. Entonces $n = 5(4k)$ y $4k \in \mathbb{Z}$, por lo que $5 \mid n$. La edición registra la hipótesis que garantiza k y el dominio que convierte $4k$ en un testigo admisible. La frase inicial del borrador podía entenderse, pero no permitía recuperar esos datos sin reconstruirlos.

Solución 2. Sean $x, y \in \mathbb{Z}$ y supongamos que $q(x) = q(y)$. Por la definición de q, tenemos $5x - 4 = 5y - 4$. Al sumar 4 y cancelar el factor no nulo 5, obtenemos $x = y$. Por tanto, q es inyectiva. La igualdad inicial compara los valores $q(x), q(y)$; «q es igual» no identifica dos objetos comparados. La inyectividad aparece como conclusión después de comprobar su condición para elementos arbitrarios.

Solución 3. Sea $x \in A \setminus (B \cup C)$. Entonces $x \in A$ y $x \notin B \cup C$, de modo que $x \notin B$ y $x \notin C$. Por tanto, $x \in A \setminus B$ y $x \in A \setminus C$, lo que prueba la primera inclusión.

Recíprocamente, sea $x \in (A \setminus B) \cap (A \setminus C)$. Tenemos $x \in A$, $x \notin B$ y $x \notin C$; luego $x \notin B \cup C$. Así, $x \in A \setminus (B \cup C)$ y queda probada la segunda inclusión. Las definiciones de diferencia, unión e intersección justifican las pertenencias. Los dos párrafos separan las obligaciones que componen la igualdad.

Solución 4. Sean $a, b > 0$ números reales y supongamos que $a/b = 2$. Como $b \neq 0$, podemos multiplicar la igualdad por b, obteniendo $a = 2b$. Al restar b, resulta $a - b = 2b - b = b$. El apunte escribía $2b = a - b$, igualdad incompatible con $a = 2b$ y $b > 0$. La primera transición es una consecuencia entre proposiciones; la cadena final contiene números iguales. La puntuación separa los dos movimientos sin borrar sus razones.

Solución 5. Como $u \in A$, el elemento u es testigo de que $f(u) \in f(A)$. Por definición de preimagen, esta pertenencia da $u \in f^{-1}(f(A))$. «Su imagen» debía nombrar $f(u)$; «ella», el conjunto $f(A)$; «éste», el elemento u. La versión explícita conserva esos tres referentes y evita afirmar que u pertenece al conjunto de imágenes.

Solución 6. Definamos $m = (a + b)/2$. Como $a < b$, tenemos $b - a > 0$. Las identidades $m - a = (b - a)/2$ y $b - m = (b - a)/2$ muestran que $a < m < b$. Basta escribir las identidades sin desplegar cada suma y resta, pues esas operaciones son rutinarias. En cambio, la positividad de ambas diferencias hace visible por qué se obtienen las dos desigualdades. «Dentro» no especificaba el intervalo ni las obligaciones que debían comprobarse.

Solución 7. Hay que añadir $c \neq 0$. Una aplicación auditable dice: «Como $c \neq 0$, su inverso multiplicativo existe. Multiplicando $ac = bc$ por c^{-1} y usando asociatividad e identidad, obtenemos $a = b$». También puede invocarse directamente la cancelación no nula demostrada en C05, nombrando esa condición. Si $c = 0$, la hipótesis se reduce a $0 = 0$ para cualesquiera a, b y no implica que coincidan. Por ejemplo, $a = 0, b = 1$ satisfacen la hipótesis y contradicen la conclusión.

Solución 8. Para cada $x \in X$, por las definiciones de preimagen y diferencia,

$$\begin{aligned} x \in f^{-1}(U \setminus V) &\iff f(x) \in U \setminus V \\ &\iff f(x) \in U \text{ y } f(x) \notin V \\ &\iff x \in f^{-1}(U) \text{ y } x \notin f^{-1}(V) \\ &\iff x \in f^{-1}(U) \setminus f^{-1}(V). \end{aligned}$$

Los dos conjuntos son subconjuntos de X y tienen exactamente los mismos elementos, luego son iguales. La cadena es reversible porque cada fila despliega una definición; no presupone que f sea inyectiva ni sobreyectiva.

Solución 9. Fijemos $n \in \mathbb{Z}$. El entero $-n$ satisface $n + (-n) = 0$, por lo que existe un opuesto. Si $u, v \in \mathbb{Z}$ satisfacen $n + u = 0 = n + v$, entonces $n + u = n + v$ y la cancelación aditiva da $u = v$. Esto prueba la unicidad. Por tanto, n tiene un único opuesto aditivo. El demostrativo del borrador se refería a la construcción, que sólo cerraba la existencia, pero anunciaba un cierre que incluía una obligación todavía pendiente.

Solución 10. La función que resulta sobreyectiva es g. Sea $z \in Z$. Por sobreyectividad de $g \circ f$, existe $x \in X$ tal que $g(f(x)) = z$. Como $f(x) \in Y$, el elemento $y = f(x)$ es un testigo de que z tiene preimagen por g. Por tanto, $g : Y \to Z$ es sobreyectiva. No se obtiene necesariamente la sobreyectividad de f. Por ejemplo, para $X = \{0\}$, $Y = \{0, 1\}$ y $Z = \{0\}$, la inclusión $f(0) = 0$ no es sobreyectiva, mientras que g constante y $g \circ f$ sí lo son.

Solución 11. Si $x \in A$, entonces $f(x) \in f(A)$ y, por definición de preimagen, $x \in f^{-1}(f(A))$. Queda probada $A \subseteq f^{-1}(f(A))$.

Para la otra inclusión, sea $x \in f^{-1}(f(A))$. Entonces $x \in X$ y existe $a \in A$ con $f(x) = f(a)$. La inyectividad de f, aplicada a $x, a \in X$, da $x = a$, por lo que $x \in A$. Ambas inclusiones prueban la igualdad. Sin inyectividad, la segunda dirección puede fallar: si $X = \{0, 1\}$, $Y = \{0\}$, f es constante y $A = \{0\}$, tenemos $f^{-1}(f(A)) = X \neq A$.

Solución 12. Para $n = 0$, la definición da $s_0 = 1 = 4 \cdot 0 + 1$. Supongamos que $s_k = 4k + 1$ para un $k \in \mathbb{N}$ arbitrario. Por la regla recursiva y la hipótesis inductiva,

$$s_{k+1} = s_k + 4 = (4k + 1) + 4 = 4(k + 1) + 1.$$

La base y el paso inductivo permiten concluir que $s_n = 4n + 1$ para todo $n \in \mathbb{N}$. La regla recursiva justifica la primera igualdad, la hipótesis la segunda y el cálculo las restantes. La hipótesis sólo estuvo disponible en el paso; la conclusión universal procede del principio de inducción de C08.

Solución 13. El par real $(2, 5)$ satisface $2 + 3 \cdot 5 = 17$ y $2 \cdot 2 - 5 = -1$. Por tanto, existe al menos una solución.

Para probar unicidad, sean (u, v) y (r, s) dos soluciones. Al restar sus primeras ecuaciones, obtenemos $(u - r) + 3(v - s) = 0$; al restar las segundas, $2(u - r) - (v - s) = 0$. La segunda igualdad da $v - s = 2(u - r)$. Sustituyendo en la primera, resulta $7(u - r) = 0$, luego $u = r$

y después $v = s$. Ambas obligaciones prueban la existencia de un único par. El borrador presuponía la unicidad que debía demostrar; la introducción de dos soluciones arbitrarias permite verificarla sin asumir la conclusión.

Solución 14. Sean $u, v \in \mathbb{R}$ con $u^3 = v^3$. La factorización de la diferencia de cubos da $(u - v)(u^2 + uv + v^2) = 0$. Si $u^2 + uv + v^2 \neq 0$, la cancelación da $u = v$. Si ese factor es cero, usamos

$$u^2 + uv + v^2 = \left(u + \frac{v}{2}\right)^2 + \frac{3v^2}{4}.$$

Los dos sumandos son no negativos; su suma cero obliga a que ambos sean cero. Así, $v = 0$ y luego $u = 0$, de modo que también $u = v$. Por tanto, p es inyectiva. A omitía una justificación de monotonía; B omitía la razón que controla el segundo factor y, en particular, su posible anulación. La versión completa resuelve esa condición sin añadir un resultado posterior.

Solución 15. Si $y \in f(A \cup B)$, existe $x \in A \cup B$ con $f(x) = y$. Cuando $x \in A$, ese testigo da $y \in f(A)$; cuando $x \in B$, da $y \in f(B)$. Los casos cubren $A \cup B$, luego $y \in f(A) \cup f(B)$.

Recíprocamente, si $y \in f(A) \cup f(B)$, pertenece al menos a una de esas imágenes. En el primer caso existe $a \in A$ con $f(a) = y$; como $a \in A \cup B$, se sigue $y \in f(A \cup B)$. En el segundo existe $b \in B$ con $f(b) = y$ y la misma conclusión se obtiene con $b \in A \cup B$. Ambas inclusiones prueban la igualdad.

Una compresión posible dice: «Un testigo de $y \in f(A \cup B)$ pertenece a A o a B, por lo que muestra $y \in f(A) \cup f(B)$. Recíprocamente, un testigo de pertenencia a cualquiera de esas imágenes pertenece a $A \cup B$ y muestra $y \in f(A \cup B)$». El enunciado proporciona los dominios; la compresión conserva los testigos, las alternativas y las dos direcciones.

La revisión de una prueba termina al verificar que la conclusión responde al enunciado y que el lector puede recuperar las dependencias utilizadas para alcanzarla.

Capítulo 11 — Familias, productos, cocientes y tamaño

En el capítulo 4 aprendimos a expresar afirmaciones mediante conjuntos, relaciones y funciones. Ahora necesitamos organizar conjuntos que varían con un parámetro, reunir condiciones, conservar varias coordenadas o identificar objetos según una relación. También necesitamos decir con precisión cuándo dos conjuntos tienen igual tamaño, aun cuando no podamos terminar de contar sus elementos.

Estas operaciones aparecen en álgebra, cálculo y matemática discreta. Aquí construiremos su lenguaje común. Las pruebas usan pertenencia, funciones y las obligaciones ya estudiadas; no requieren teoría axiomática avanzada de conjuntos ni categorías.

La pregunta que guiará el capítulo es: **¿qué datos conserva una construcción y qué hay que comprobar para utilizarla?** Una familia conserva etiquetas; una intersección conserva condiciones simultáneas; un producto conserva coordenadas; un cociente conserva clases; una biyección conserva la correspondencia exacta entre elementos. Confundir estas funciones puede hacer falsa una afirmación aparentemente familiar.

Mantendremos $\mathbb{N} = \{0, 1, 2, \ldots\}$. Cuando trabajemos con complementos o con una intersección de índice vacío, declararemos un conjunto ambiente. No usaremos «el conjunto de todos los conjuntos».

11.1. Familias indexadas: etiquetas y valores

Una **familia indexada de conjuntos** $(A_i)_{i \in I}$ asigna a cada elemento i de un conjunto I un conjunto A_i. El conjunto I es el conjunto de índices; i es una etiqueta y A_i es el valor correspondiente. Se puede expresar formalmente como una función cuyo dominio es I y cuyos valores son los conjuntos considerados. Para una familia ya dada, basta como codominio su conjunto de valores: no necesitamos reunir todos los conjuntos posibles.

Por ejemplo, con $I = \{\text{norte}, \text{sur}, \text{centro}\}$, fijemos

$$A_{\text{norte}} = \{1, 4\}, \qquad A_{\text{sur}} = \{4\}, \qquad A_{\text{centro}} = \{1, 4\}.$$

Los tres índices son distintos, aunque dos valores coinciden. La familia no es el conjunto $\{\{1, 4\}, \{4\}\}$: ese conjunto no indica qué valor corresponde a cada etiqueta. La diferencia importará especialmente al formar productos, donde cada etiqueta determina una coordenada.

Una sucesión $(A_n)_{n \in \mathbb{N}}$ es una familia con índices naturales. No toda familia es una sucesión: un conjunto de índices puede ser finito, estar formado por pares o ser un intervalo real. Por ejemplo, $(\{x \in \mathbb{R} : x < t\})_{t \in \mathbb{R}}$ tiene un conjunto de índices distinto de $\mathbb{N}$.

Reindexar exige una función. Sea $r : J \to I$. La familia $(A_{r(j)})_{j \in J}$ está **reindexada** por r. Si r no es sobreyectiva, algunos índices originales pueden no aparecer. Si r no es inyectiva, un mismo índice puede aparecer varias veces. Una biyección cambia las etiquetas sin perder ni repetir posiciones.

En el ejemplo anterior, si $J = \{0, 1\}$ y $r(0) = \text{norte}$, $r(1) = \text{centro}$, la familia reindexada tiene ambos valores iguales a $\{1, 4\}$ y no incluye la posición sur. No podemos afirmar que recupera toda la familia original.

Una igualdad de familias con el mismo conjunto de índices significa $A_i = B_i$ para cada $i \in I$. Igualdad de sus conjuntos de valores es una afirmación más débil: no identifica qué

conjunto aparece en cada posición. Si los dominios índice son distintos, antes de comparar hay que declarar cómo se relacionan sus etiquetas.

Comprobación. Sean $U_0 = \{a\}$, $U_1 = \{b\}$, $V_0 = \{b\}$ y $V_1 = \{a\}$, con $a \neq b$. ¿Son iguales las familias?

Respuesta. No: sus valores en el índice 0 difieren. Sus conjuntos de valores sí coinciden. La biyección que intercambia 0 y 1 permite obtener una familia a partir de la otra por reindexación; no establece igualdad con las etiquetas originales.

11.2. Uniones e intersecciones: alguna condición o todas

Sea $(A_i)_{i\in I}$ una familia de subconjuntos de un conjunto X. La **unión** reúne los elementos que pertenecen a alguno de los conjuntos:

$$\bigcup_{i\in I} A_i = \{x \in X : \exists i \in I \ (x \in A_i)\}.$$

La **intersección**, relativa al ambiente declarado X, reúne los que pertenecen a todos:

$$\bigcap_{i\in I}^{X} A_i = \{x \in X : \forall i \in I \ (x \in A_i)\}.$$

Usaremos el superíndice X cuando haga falta recordar el ambiente, especialmente con índice vacío. Con índice no vacío podemos omitirlo: elegir $i_0 \in I$ obliga a que cualquier elemento de la intersección pertenezca a A_{i_0}, y ampliar el ambiente no añade elementos.

En la familia norte/sur/centro de §11.1 la unión es $\{1, 4\}$ y la intersección es $\{4\}$. Repetir el valor $\{1, 4\}$ no cambia ninguna de las dos operaciones. Para demostrarlo no contamos índices: comprobamos qué condiciones satisface cada elemento.

El índice vacío no significa lo mismo en ambas operaciones. Si $I = \varnothing$, no existe un índice que pueda servir de testigo. Por tanto,

$$\bigcup_{i\in\varnothing} A_i = \varnothing.$$

En cambio, no hay ninguna condición indexada que incumplir. Todo $x \in X$ satisface la condición universal vacía; así,

$$\bigcap_{i\in\varnothing}^{X} A_i = X.$$

Esto no define un conjunto de todos los conjuntos ni de todos los objetos posibles. Si no se ha declarado X, escribir una intersección vacía deja pendiente una convención necesaria. El capítulo 3 explica la semántica de los cuantificadores vacíos; aquí aplicamos esa semántica a una construcción de conjuntos.

Inclusiones y testigos. Para cada $k \in I$ se cumple

$$\bigcap_{i\in I}^{X} A_i \subseteq A_k \subseteq \bigcup_{i\in I} A_i.$$

Si x está en la intersección, la condición universal se puede aplicar al índice k. Si $x \in A_k$, ese mismo k es testigo para la unión. Ambas justificaciones requieren disponer de $k \in I$; no podemos extraer uno de un conjunto índice vacío.

Comprobación. En ambiente $X = \{a, b\}$, ¿puede ocurrir que la intersección vacía tenga más elementos que la unión vacía?

Respuesta. Sí. La intersección vacía es X y la unión vacía es $\varnothing$. No contradice las inclusiones anteriores: éstas están formuladas para cada $k \in I$, y aquí no existe tal índice.

11.3. Operaciones finitas e infinitas: controlar las condiciones

Las reglas de las operaciones indexadas se comprueban traduciendo pertenencia. La diferencia entre finito e infinito puede cambiar el resultado, aunque no cambie la forma de esa traducción.

Si $A_i \subseteq B_i \subseteq X$ para cada $i \in I$, entonces

$$\bigcup_{i\in I} A_i \subseteq \bigcup_{i\in I} B_i, \qquad \bigcap_{i\in I}^{X} A_i \subseteq \bigcap_{i\in I}^{X} B_i.$$

En la primera inclusión, un testigo k con $x \in A_k$ también da $x \in B_k$. En la segunda, para cada índice la pertenencia a A_i implica pertenencia a B_i. El ambiente común importa con índice vacío: las dos intersecciones son entonces el mismo X.

De Morgan conserva el ambiente. Bajo lógica clásica, para cualquier conjunto índice, incluido el vacío,

$$X \setminus \bigcup_{i\in I} A_i = \bigcap_{i\in I}^{X} (X \setminus A_i),$$
$$X \setminus \bigcap_{i\in I}^{X} A_i = \bigcup_{i\in I} (X \setminus A_i).$$

Para la primera igualdad, un elemento de X no está en la unión exactamente cuando no pertenece a ningún A_i; eso equivale a pertenecer a cada complemento relativo. Para la segunda, no satisfacer todas las pertenencias equivale a que alguna falle. En ambos argumentos fijamos primero $x \in X$. Si I es vacío, las igualdades se reducen respectivamente a $X \setminus \varnothing = X$ y $X \setminus X = \varnothing$.

También, si $B \subseteq X$,

$$B \cap \bigcup_{i\in I} A_i = \bigcup_{i\in I} (B \cap A_i).$$

En efecto, estar a la izquierda significa estar en B y en algún A_i. Ese mismo índice testimonia la pertenencia al lado derecho; en sentido inverso, el testigo proporciona ambas pertenencias. No hace falta que I sea finito.

Un límite de la intuición finita. Consideremos en $\mathbb{N}$ los conjuntos $T_n = \{m \in \mathbb{N} : m \geq n\}$. Cada T_n es no vacío, y cualquier cantidad finita no vacía de ellos tiene intersección no vacía: si M es el mayor de sus índices, la intersección es T_M.

Pero la intersección de todos los T_n es vacía. Si un natural m perteneciera a ella, tendría que pertenecer a T_{m+1}; eso exigiría $m \geq m + 1$, lo cual es falso. El índice que detecta el fallo depende del candidato m. No se está buscando un índice único que excluya a todos los candidatos a la vez.

Por tanto, «todas las intersecciones finitas son no vacías» no autoriza aquí «la intersección total es no vacía». Otros cursos estudiarán hipótesis que sí permiten conclusiones de esa forma; no las suponemos en este libro.

Reindexación sobreyectiva. Si $r : J \to I$ es sobreyectiva, la unión y la intersección de $(A_{r(j)})_{j\in J}$ coinciden con las de $(A_i)_{i\in I}$, con el mismo ambiente. Todo índice original tiene una preimagen, y cada índice nuevo apunta a uno original. En una unión esto permite transportar testigos; en una intersección permite recuperar cada condición. La repetición de índices no añade condiciones distintas.

Comprobación. ¿Basta que r sea inyectiva para conservar una unión?

Respuesta. No. Tomemos $I = \{0, 1\}$, $J = \{0\}$, $r(0) = 0$, $A_0 = \{a\}$ y $A_1 = \{b\}$, con $a \neq b$. La unión original es $\{a, b\}$ y la reindexada es $\{a\}$. La inyectividad impide repeticiones, pero no garantiza cubrir todos los índices.

11.4. Productos finitos: conservar las coordenadas

El producto $A \times B$ del capítulo 4 reúne pares ordenados (a, b) con $a \in A$ y $b \in B$. Para $n \geq 1$, el producto de $A_1, \ldots, A_n$ reúne tuplas $(a_1, \ldots, a_n)$ con $a_i \in A_i$ en cada posición. Dos tuplas son iguales exactamente cuando coinciden todas sus coordenadas.

También podemos describir una tupla como una función $f : \{1, \ldots, n\} \to \bigcup_{i=1}^{n} A_i$ tal que $f(i) \in A_i$. La correspondencia envía $(a_1, \ldots, a_n)$ a la función $f(i) = a_i$; su inversa lee los valores $f(1), \ldots, f(n)$. No confunde un objeto a_i con la posición i.

Para un conjunto índice finito de etiquetas cualesquiera usamos directamente esta descripción funcional. Si $I = \{\text{entrada}, \text{salida}\}$, un elemento de $\prod_{i\in I} A_i$ especifica un valor admisible de entrada y uno de salida. Aunque ambos factores sean el mismo conjunto, las posiciones siguen siendo distintas.

Vacío en el índice y vacío en un factor. El **producto de una familia vacía** contiene exactamente una función, la función vacía $e_\varnothing : \varnothing \to \varnothing$. Su grafo no tiene pares y no hay restricciones de coordenadas que verificar. Es un conjunto con un elemento:

$$\prod_{i\in\varnothing} A_i = \{e_\varnothing\}.$$

Si identificáramos una función con su grafo, podríamos escribir también $\{\varnothing\}$. Sin embargo, en este libro (§4.5 y §4.7) declaramos dominio y codominio como parte de los datos de cada función. Bajo **esa convención**, el elemento $e_\varnothing$ no es literalmente el conjunto vacío: es la función de $\varnothing$ en $\varnothing$ cuyo grafo es vacío. El producto, en cualquier caso, tiene un elemento, no cero. Esta precisión será útil en §11.5, cuando comparemos productos vacíos y potencias $A^\varnothing$ con codominios declarados distintos.

En cambio, si $k \in I$ y $A_k = \varnothing$, el producto es vacío. Un elemento del producto tendría que asignar a k un valor en A_k, lo cual es imposible. La conclusión requiere un factor vacío con índice existente; no se aplica a la familia sin índices.

Si hay un número finito de posiciones y todos los factores son no vacíos, el producto es no vacío. Para justificarlo, empezamos con la tupla vacía. Al añadir un factor no vacío elegimos un elemento de ese factor y lo añadimos a una tupla ya construida. Inducción sobre el número de factores da una tupla completa. Cada etapa hace una elección dentro de una construcción finita; no estamos afirmando existencia para una familia arbitraria infinita.

Asociar y cambiar orden. La función $((a, b), c) \mapsto (a, (b, c))$ es una biyección de $(A \times B) \times C$ a $A \times (B \times C)$: su inversa devuelve $((a, b), c)$. Esta correspondencia permite identificar ambas presentaciones cuando la convención esté declarada; no afirma igualdad literal de pares anidados.

Análogamente, $(a, b) \mapsto (b, a)$ da una biyección entre $A \times B$ y $B \times A$. Las coordenadas intercambian posición y dominio. No se puede conservar la misma fórmula de una función que utiliza la primera coordenada sin ajustar esa fórmula al intercambio.

Comprobación. Sean $A = \{u, v\}$ y $B = \{w\}$, con tres objetos distintos. ¿Son iguales $A \times B$ y $B \times A$?

Respuesta. No: (u, w) pertenece al primero y no al segundo. Tienen igual número de elementos y la función de intercambio es una biyección. Igualdad de conjuntos y existencia de una biyección son afirmaciones diferentes.

11.5. Productos indexados: un elemento es una selección completa

Sea $(A_i)_{i\in I}$ una familia de conjuntos y sea $U = \bigcup_{i\in I} A_i$. Definimos su **producto indexado** mediante

$$\prod_{i\in I} A_i = \{f : I \to U : \forall i \in I\ (f(i) \in A_i)\}.$$

La definición vale para índices finitos o infinitos. Un elemento del producto es una sola función que asigna simultáneamente un valor permitido a cada posición. Con índice vacío $U = \varnothing$ y existe una única función de $\varnothing$ en $\varnothing$, como establecimos en §11.4.

Si todos los factores son un mismo conjunto A, usamos A^I para el conjunto de funciones $I \to A$; tiene la misma descripción de coordenadas. Con índice vacío, el producto usa la función vacía con codominio $\varnothing$, mientras $A^{\varnothing}$ usa codominio A: sus únicos grafos coinciden, y la identificación es por grafo, no una igualdad de funciones con codominios declarados distintos. Para $I = \mathbb{N}$, un elemento de $A^{\mathbb{N}}$ es una sucesión de elementos de A. Si I es vacío, la función vacía existe cualquiera que sea A.

Construir una selección. Definamos $A_n = \{n, n + 2\}$ para $n \in \mathbb{N}$. La fórmula $f(n) = n + 2$ define un elemento del producto. Hay que comprobar dos cosas: que proporciona una función en todo $\mathbb{N}$ y que su valor pertenece al factor correcto para cada n. La segunda es inmediata por la definición de A_n; la primera se apoya en una regla única definida en cada natural.

No basta decir «cada factor tiene algún elemento». Esa afirmación tiene forma $\forall i \in I\ \exists a \in A_i$. Un elemento del producto exige una función f que reúna todas las selecciones. Para familias arbitrarias, el principio que garantiza esa existencia es el **axioma de elección**. No lo desarrollaremos ni usaremos aquí como una hipótesis tácita. En los ejemplos donde necesitamos un elemento, lo construiremos por una fórmula o una regla ya justificada. Con factores finitos en número, §11.4 dio una construcción finita.

Qué permite recuperar una coordenada. Para cada $k \in I$ existe una función de proyección

$$\pi_k : \prod_{i\in I} A_i \to A_k, \qquad \pi_k(f) = f(k).$$

Es función porque cada f tiene un único valor en k, y ese valor pertenece a A_k. No afirmamos sin hipótesis que la proyección sea sobreyectiva. Si ya tenemos un elemento g del producto, podemos probarlo: para $a \in A_k$, definimos $f(k) = a$ y $f(i) = g(i)$ en los demás índices. Esa modificación produce un elemento del producto cuya proyección es a. Si el producto está vacío pero A_k no lo está, la proyección no es sobreyectiva.

La construcción distingue entre contar con una selección completa y limitarse a saber que cada factor tiene elementos. Esa diferencia no es un detalle de notación.

Comprobación. Para $I = \mathbb{N}$ y $A_n = \{n + 1\}$, determine el producto.

Respuesta. Contiene exactamente la función $f(n) = n + 1$. La fórmula produce una selección admisible. Cualquier otra función del producto debe tener en cada posición el único valor del factor, por lo que coincide con f. No se necesita un principio de elección para esta familia: la regla ya está dada.

11.6. Equivalencias y particiones: agrupar sin solapamientos

Una relación $\sim$ sobre un conjunto X es una **relación de equivalencia** si es reflexiva, simétrica y transitiva: para $x, y, z \in X$, se cumplen $x \sim x$, $x \sim y \Rightarrow y \sim x$ y ($x \sim y$ y $y \sim z$) $\Rightarrow x \sim z$.

La **clase de equivalencia** de $x \in X$ es

$$[x] = \{y \in X : y \sim x\}.$$

La reflexividad garantiza $x \in [x]$, de modo que ninguna clase de un elemento existente es vacía. Además,

$$x \sim y \quad \Longleftrightarrow \quad [x] = [y].$$

Para probar la dirección directa, supongamos $x \sim y$. Si $z \in [x]$, tenemos $z \sim x$ y por transitividad $z \sim y$, así que $z \in [y]$. La simetría da $y \sim x$ y el mismo argumento prueba la inclusión opuesta. Para la dirección inversa, si $[x] = [y]$, la pertenencia $x \in [x]$ implica $x \in [y]$, esto es, $x \sim y$.

Dos clases que se encuentran son la misma. Si $z \in [x] \cap [y]$, entonces $z \sim x$ y $z \sim y$. Por simetría $x \sim z$ y por transitividad $x \sim y$. La equivalencia anterior da $[x] = [y]$. Por tanto, dos clases distintas son disjuntas. Toda persona que manipule clases necesita esta propiedad: no basta que los conjuntos se parezcan o tengan un elemento en común sin haber probado equivalencia.

Una **partición** de X es un conjunto $\mathcal{P}$ de subconjuntos no vacíos de X, disjuntos dos a dos, cuya unión es X. Cada elemento pertenece a exactamente un bloque. Las clases de una relación de equivalencia forman una partición: son no vacías, cubren X por reflexividad y son iguales o disjuntas por el argumento anterior. Al reunirlas como conjunto, las repeticiones $[x] = [y]$ aparecen una sola vez.

A la inversa, dada una partición, definimos $x \sim y$ cuando ambos pertenecen al mismo bloque. Es reflexiva porque cada elemento está en un bloque; simétrica porque compartir bloque es simétrico; transitiva porque los dos bloques que contienen a y deben ser el mismo. Sus clases son exactamente los bloques de la partición.

Cuando $X = \varnothing$, la partición vacía cumple las condiciones: su unión es vacía y no tiene bloques vacíos. La relación vacía es una equivalencia sobre X, puesto que no existen elementos que violen ninguna de las tres condiciones. No aparece una clase vacía: no hay elementos de los que formar clases.

Ejemplo por una característica compartida. En $\mathbb{Z} \times \mathbb{Z}$, declaremos $(a, b) \sim (c, d)$ si $a + b = c + d$. Igualdad de las sumas prueba las tres propiedades. La clase de (a, b) reúne todos los pares cuya suma es $a + b$. Un par distinto puede pertenecer a la misma clase; por ejemplo, $(2, 5)$ y $(4, 3)$. El criterio identifica sumas, no coordenadas individuales. El capítulo siguiente en la ruta algebraica podrá añadir estructuras a cocientes; aquí sólo construimos sus conjuntos.

Comprobación. En la equivalencia por suma del ejemplo, un editor propone reemplazar las clases de $(2, 5)$ y $(4, 3)$ por su unión, como si estuviera reuniendo dos bloques distintos. ¿Qué debe corregir?

Respuesta. Ambos pares tienen suma 7, por lo que sus clases ya son el mismo bloque. La unión no reúne dos bloques distintos: devuelve ese único bloque. La partición recoge clases como conjuntos, sin duplicar un bloque por cada representante que lo nombra.

11.7. Cocientes: comprobar independencia del representante

El **conjunto cociente** de X por una equivalencia $\sim$ es

$$X/\sim = \{[x] : x \in X\}.$$

Sus elementos son clases, no los elementos originales de X. La proyección $q : X \to X/\sim$, definida por $q(x) = [x]$, es sobreyectiva: cada clase del cociente tiene por definición al menos un representante. Para justificar esa sobreyectividad tomamos una clase particular y un representante de ella; no elegimos simultáneamente un representante de todas las clases.

Qué significa que una regla esté bien definida. Antes de preguntar si una fórmula define una función, debemos precisar **qué son sus entradas y sus salidas**. Hay al menos dos controles diferentes.

El primero es el **control de dominio y unicidad del valor**. Una regla propuesta como $f : A \to B$ debe asignar a cada $x \in A$ un valor que exista, pertenezca a B y sea único. Cuando la regla se presenta mediante operaciones, también debemos comprobar que éstas están definidas para las entradas permitidas. Esto **no exige un algoritmo para calcular cada valor**: la existencia y la unicidad son requisitos matemáticos, distintos de la computabilidad. Por ejemplo, la fórmula

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

define una función $\mathbb{R} \setminus \{1\} \to \mathbb{R}$, pero **no** define una función $\mathbb{R} \to \mathbb{R}$ mediante la misma regla, porque en $x = 1$ el denominador es cero. No se corrige ese defecto demostrando que la fórmula funciona en muchos valores: la definición de función exige una salida para **cada** entrada del dominio declarado.

El segundo control aparece cuando una entrada puede escribirse con **representantes distintos**, como ocurre en un conjunto cociente. Aun si la fórmula puede evaluarse para cada representante, habrá que demostrar que todos los representantes de una misma clase producen **el mismo valor**. De otro modo no se ha definido una función *sobre las clases*. La independencia del representante no es una condición adicional de inyectividad o sobreyectividad; es anterior a ambas, porque determina si la aplicación misma existe.

De este modo, «estar bien definida» no significa simplemente «tener una fórmula conocida»: significa que **la regla determina inequívocamente un objeto del tipo anunciado para cada entrada legítima**. En el caso de los cocientes, veremos enseguida cómo se traduce esta exigencia a una implicación precisa.

Una fórmula con clases puede ser ambigua. Si proponemos $\overline{F}([x]) = F(x)$, dos representantes x e y de una misma clase deben dar el mismo resultado. La condición precisa es

$$x \sim y \quad \Longrightarrow \quad F(x) = F(y).$$

Supongamos que $F : X \to Y$ satisface esa condición. Para cada clase C, existe un valor $F(x)$ con $x \in C$, y ese valor es único: cualquier otro $y \in C$ verifica $y \sim x$, así que $F(y) = F(x)$. Existencia y unicidad por clase definen una función $\overline{F} : X/\sim \to Y$. Se dice que F **desciende al cociente**.

La identidad $\overline{F}(q(x)) = F(x)$ vale para cada x, es decir, $\overline{F} \circ q = F$. Además, la función descendida es única con esa propiedad. Si $G \circ q = F$ y $C = [x]$, entonces $G(C) = G(q(x)) = F(x) = \overline{F}(C)$. Cada clase tiene representante, de modo que las funciones coinciden en todo su dominio. La prueba usa representantes locales para verificar una propiedad, no una función global de elección de representantes.

Una fórmula que sí desciende y otra que no. Para la equivalencia por suma de §11.6, $F(a,b) = a+b$ es constante en clases y desciende a una función del cociente en $\mathbb{Z}$. En cambio, $H(a,b) = a$ no desciende: $(2,5)$ y $(4,3)$ son equivalentes, pero dan valores 2 y 4. El escrito «$\overline{H}([(a,b)]) = a$» no define una función del cociente. No es un problema de falta de inyectividad: falta un valor único para una misma entrada.

Si después queremos definir una operación entre clases, habrá que comprobar independencia de cada representante que interviene. No basta con que la fórmula se pueda calcular en representantes particulares. Esa obligación será decisiva en álgebra; aquí la reconocemos sin desarrollar cocientes de grupos o anillos.

Dos clases como entradas: la independencia se comprueba dos veces. Sea $Q = (\mathbb{Z}\times\mathbb{Z})/\!\sim$ el cociente por igualdad de sumas de §11.6. Queremos proponer una función $\Theta : Q \times Q \to \mathbb{Z}$. Si escribimos dos de sus entradas como $C = [(a,b)]$ y $D = [(c,d)]$, una fórmula posible es

$$\Theta(C,D) = (a+b) + (c+d).$$

¿Depende de cómo escribamos cada clase? Supongamos que elegimos otros representantes (a',b') y (c',d') de C y D, respectivamente. Por la definición de equivalencia,

$$a+b = a'+b', \qquad c+d = c'+d'.$$

Sumar esas dos igualdades da

$$(a+b)+(c+d) = (a'+b')+(c'+d').$$

Por tanto, cambiar el representante de **cualquiera de las dos entradas**, o de ambas a la vez, conserva el resultado. La regla determina una función $Q\times Q \to \mathbb{Z}$ bien definida. Su salida es un entero, no una clase: no estamos introduciendo aquí una operación algebraica sobre Q.

En cambio, la propuesta $\Lambda(C,D) = a+c$ no está bien definida. Para la primera clase podemos escribir tanto $[(2,5)]$ como $[(4,3)]$, que son iguales porque ambas sumas valen 7. Fijando como segunda entrada $D = [(0,0)]$, la fórmula daría respectivamente 2 y 4. Una misma pareja de clases recibiría dos valores distintos, aun cuando en cada cálculo individual no haya ninguna operación ilegítima.

Comprobación. ¿Define $\Omega(C,D) = (a+b)-(c+d)$ una función de $Q\times Q$ en $\mathbb{Z}$?

Respuesta razonada. Sí. Si cambiamos las representaciones de C y D, las dos sumas permanecen iguales a sus valores iniciales, y por tanto también permanece igual su diferencia. Todas las operaciones son legítimas en $\mathbb{Z}$ y para cada pareja de clases se determina un único entero. No necesitamos elegir de antemano un representante privilegiado de cada clase.

Comprobación. En el cociente por suma, ¿desciende $K(a,b) = 2(a+b)+1$?

Respuesta. Sí. Si $(a,b)\sim(c,d)$, sus sumas son iguales y por ello $2(a+b)+1 = 2(c+d)+1$. La función descendida tiene dominio el cociente y codominio $\mathbb{Z}$; no es sobreyectiva en ese codominio porque sólo alcanza enteros impares. Buena definición y sobreyectividad son comprobaciones distintas.

11.8. Igual tamaño, finitud e infinitud

Decimos que A y B son **equipotentes**, o que tienen igual tamaño, si existe una biyección $A \to B$. Escribiremos $A \approx B$. La afirmación no exige que los elementos sean los mismos ni que los conjuntos sean iguales. En conjuntos finitos corresponde al resultado de contar; en conjuntos infinitos proporciona una definición que no necesita terminar una lista.

La equipotencia es reflexiva, simétrica y transitiva. La identidad prueba reflexividad; la inversa de una biyección prueba simetría; componer dos biyecciones prueba transitividad.

Se aplican las propiedades de funciones de §4.7. No formaremos un cociente de «todos los conjuntos» por esta relación: sólo comparamos los conjuntos de los que estamos hablando.

Contar finitamente y demostrar que el número es único. Para $n \in \mathbb{N}$ fijamos $I_n = \{0, \ldots, n-1\}$, con $I_0 = \varnothing$. Un conjunto A es **finito** si hay alguna biyección $I_n \to A$. En ese caso escribimos $|A| = n$. Para que esa notación sea legítima tenemos que demostrar que el número no depende de cómo se haya contado.

La herramienta es el siguiente resultado, que merece una demostración completa porque la unicidad del tamaño finito dependerá de él.

Lema (no caben más entradas distintas que posiciones disponibles). Si existe una inyección $u : I_m \to I_n$, entonces $m \leq n$.

Demostración. Procedemos por inducción sobre n; en cada etapa la afirmación comprende **todos** los posibles tamaños m del dominio.

Base $n = 0$. El codominio es $I_0 = \varnothing$. Si existe una función $I_m \to \varnothing$, su dominio necesariamente es vacío. Por tanto $m = 0$ y se cumple $m \leq 0$.

Paso de n a $n+1$. Supongamos que el lema vale para cualquier inyección con codominio I_n. Sea $u : I_m \to I_{n+1}$ una inyección. Si $m = 0$, ya tenemos $m \leq n+1$. Supongamos $m \geq 1$ y fijemos la última posición del dominio, $m-1$. Su imagen es algún índice $j = u(m-1) \in I_{n+1}$.

Podemos intercambiar las etiquetas j y n del codominio, dejando fijas las demás. Esa permutación $\sigma : I_{n+1} \to I_{n+1}$ es una biyección (si $j = n$, es la identidad). La composición $v = \sigma \circ u$ sigue siendo inyectiva y satisface $v(m-1) = n$.

Ahora retiremos la última posición del dominio. Para todo $k \in I_{m-1}$ tenemos $k \neq m-1$, de modo que la inyectividad de v implica $v(k) \neq n$. Por tanto, la restricción de v a I_{m-1} toma valores **sólo** en I_n y sigue siendo inyectiva:

$$v|_{I_{m-1}} : I_{m-1} \longrightarrow I_n.$$

Por la hipótesis inductiva, $m - 1 \leq n$; sumando 1 obtenemos $m \leq n+1$. Concluye la inducción. □

Lectura de la prueba. No hemos supuesto que el codominio tenga exactamente tantas posiciones como el dominio. Hemos elegido una entrada existente, trasladado su imagen a una posición conveniente y retirado ambas sin destruir la inyectividad. Lo que queda es un problema del mismo tipo, con un codominio de tamaño menor: por eso puede aplicarse la hipótesis inductiva. La permutación sólo cambia nombres de posiciones; no crea ni elimina imágenes.

Una biyección $I_m \to I_n$ y su inversa dan $m \leq n$ y $n \leq m$; por tanto $m = n$. Si hay biyecciones $I_m \to A$ e $I_n \to A$, componemos la primera con la inversa de la segunda y aplicamos el resultado. Así el tamaño finito está bien definido.

Un subconjunto de un conjunto finito también es finito. Veamos por qué sin presuponer que podemos «contarlo» antes de haber construido la biyección. Basta probarlo para subconjuntos de cada I_n y transportar después el resultado mediante una biyección.

Procedemos por inducción sobre n. Para $n = 0$, el único subconjunto de I_0 es $\varnothing$, equipotente a I_0. Supongamos que todo subconjunto de I_n es finito y tomemos $S \subseteq I_{n+1}$. Si $n \notin S$, entonces $S \subseteq I_n$ y aplicamos la hipótesis inductiva. Si $n \in S$, el conjunto $T = S \setminus \{n\}$ está contenido en I_n y, por hipótesis inductiva, existe alguna biyección $h : I_k \to T$. La ampliamos a I_{k+1} mediante la misma regla en las primeras k posiciones y asignando a la nueva posición k el valor n. La ampliación es biyectiva sobre S: las imágenes anteriores eran exactamente T y la nueva imagen n no pertenece a T. Así también S es finito. Si A es finito y $S \subseteq A$, transportamos S por la inversa de una biyección $I_n \to A$ y aplicamos lo demostrado.

Comprobación. ¿Por qué no basta afirmar que la restricción $v|_{I_{m-1}}$ sigue siendo inyectiva, sin explicar además cuál es su codominio?

Respuesta razonada. La hipótesis inductiva se refiere a inyecciones **en** I_n, no a inyecciones en I_{n+1}. Debemos comprobar que ninguna de las entradas restantes llega a n, para poder restringir de verdad el codominio a I_n. Esa comprobación es el paso que conecta el argumento con la hipótesis inductiva.

Ser infinito no significa tener un último número muy grande. Un conjunto es **infinito** si no es finito. Los naturales son infinitos. Si hubiera una biyección $\mathbb{N} \to I_n$, su restricción a I_{n+1} sería una inyección en I_n, contraria al principio finito anterior. La afirmación excluye cada posible tamaño finito mediante un argumento, no mediante una lista muy larga.

La función $f : \mathbb{N} \to \mathbb{N} \setminus \{0\}$, $f(n) = n + 1$, es una biyección: su inversa envía $m \geq 1$ a $m - 1$. Por tanto un conjunto infinito puede ser equipotente a un subconjunto propio. Esto no cambia la inclusión: $\mathbb{N} \setminus \{0\} \neq \mathbb{N}$.

Para un conjunto finito ocurre lo contrario, y podemos justificarlo con el lema anterior. Sea A finito, con $|A| = n$, y tomemos un subconjunto propio $S \subsetneq A$. Si $n = 0$, $A = \varnothing$ y no existe tal subconjunto propio, de modo que el caso sólo puede presentarse con $n \geq 1$. Fijemos una biyección $e : I_n \to A$ y elijamos $a \in A \setminus S$. El índice $j = e^{-1}(a)$ no pertenece a $e^{-1}(S)$. La permutación de I_n que intercambia j con $n-1$ lleva **todos los índices de** $e^{-1}(S)$ a I_{n-1}, pues la única posición que podría transformarse en $n - 1$ es j, que hemos excluido.

Ya sabemos que S es finito; sea $|S| = k$ y fijemos una biyección $h : I_k \to S$. Al componer h con e^{-1} y con la permutación anterior, obtenemos una inyección $I_k \to I_{n-1}$. Nuestro lema obliga a que $k \leq n - 1 < n$. Por tanto, ningún conjunto finito puede ser equipotente a un subconjunto propio. La intuición «quitar elementos reduce necesariamente el tamaño» requiere la hipótesis de finitud, y ahora hemos probado exactamente dónde interviene.

Comprobación. Una función inyectiva $I_5 \to A$ y otra inyectiva $A \to I_5$ están dadas. ¿Qué tamaño tiene A?

Respuesta. La segunda identifica A con un subconjunto de I_5, así que A es finito, digamos de tamaño k. Transportando ambas inyecciones a conjuntos de índices, el principio finito da $5 \leq k \leq 5$. Por ello $|A| = 5$. La justificación usa finitud; no invoca un teorema general sobre dos inyecciones entre conjuntos infinitos.

11.9. Numerabilidad y convenciones de tamaño

Un conjunto es **numerablemente infinito** si es equipotente a $\mathbb{N}$. Diremos que es **a lo sumo numerable** si admite una inyección en $\mathbb{N}$. Usaremos «numerable» como abreviatura de «a lo sumo numerable», no como sinónimo de infinito. Así, $\varnothing$ y todos los conjuntos finitos son numerables. Cuando necesitemos infinitud, lo diremos.

Estas convenciones importan al comparar fuentes: algunos autores reservan «numerable» para conjuntos infinitos. Una diferencia terminológica no es una diferencia matemática si las condiciones están explícitas.

Enumerar un subconjunto infinito de los naturales. Todo subconjunto infinito $S \subseteq \mathbb{N}$ es numerablemente infinito. Por buen orden y recursión, tomamos $s_0 = \min S$ y s_{n+1} como el menor elemento de S mayor que s_n. Ese conjunto de candidatos es no vacío: si ninguno fuera mayor, S estaría contenido en el conjunto finito $\{0, \ldots, s_n\}$ y sería finito por §11.8.

La sucesión es estrictamente creciente, luego inyectiva. Además $s_n \geq n$, por inducción. Para comprobar que no omite $m \in S$, supongamos que nunca lo alcanza. El primer término es menor que m; mientras un término sea menor que m, la regla del menor candidato hace que el siguiente sea a lo sumo m, y bajo la suposición de omisión será menor. Se seguiría $s_n < m$ para todo n, contradiciendo $s_m \geq m$. Así la función $n \mapsto s_n$ es sobreyectiva en S.

Si un conjunto infinito A admite una inyección $j : A \to \mathbb{N}$, su imagen $S = j(A)$ es infinita: una cuenta finita de S y la biyección $j : A \to S$ harían finito a A. La enumeración anterior, seguida de la inversa de j con codominio restringido a S, da una biyección $\mathbb{N} \to A$. Hemos

probado que «a lo sumo numerable» equivale a «finito o numerablemente infinito» sin usar elección de representantes arbitrarios.

Una enumeración explícita. Los enteros son numerablemente infinitos. Una biyección $e : \mathbb{N} \to \mathbb{Z}$ se define por

$$e(0) = 0, \qquad e(2k+1) = k+1, \qquad e(2k+2) = -(k+1) \quad (k \in \mathbb{N}).$$

Cada natural distinto de cero cae en exactamente una de las dos formas; los valores recorren respectivamente los enteros positivos y negativos, sin repetirlos ni incluir de nuevo el cero. La inversa envía $z > 0$ a $2z - 1$, $z < 0$ a $-2z$ y 0 a 0. Escribir el dominio, la regla y su verificación aporta más que decir «se pueden poner en una lista».

Un tamaño que no cabe en una lista natural. No todo conjunto infinito es numerable. Sea $D = \{0,1\}^{\mathbb{N}}$, el conjunto de sucesiones binarias. Si alguien propone una lista $d_0, d_1, \ldots$ que pretenda cubrir D, definimos una nueva sucesión por

$$b(n) = 1 - d_n(n) \qquad (n \in \mathbb{N}).$$

Cada valor es 0 o 1, por lo que $b \in D$. Para cada n, la sucesión b difiere de d_n en la coordenada n; no aparece en la lista. No existe, por tanto, sobreyección $\mathbb{N} \to D$.

Esto excluye tanto una cuenta infinita como una finita: D es no vacío, y una lista finita que cubriera D podría prolongarse repitiendo valores para producir tal sobreyección. Por la equivalencia demostrada arriba, D no es a lo sumo numerable. No necesitamos comparar todos los tamaños infinitos ni construir una teoría de cardinales para obtener esta conclusión.

Conjuntos, clases y ambientes. Los índices, dominios, familias y cocientes de este capítulo son conjuntos especificados. Un ambiente X fija el complemento y la intersección vacía; no es un universo que contenga automáticamente todo objeto matemático. «Clase de equivalencia» significa aquí un subconjunto concreto $[x] \subseteq X$, no una clase propia de teoría axiomática de conjuntos. Cuando una teoría posterior necesite otras convenciones, deberá declararlas allí.

Comprobación. ¿Basta que una lista de sucesiones binarias permita calcular cada entrada para que incluya todas las sucesiones?

Respuesta. No. Que la lista esté bien definida permite calcular $d_n(n)$ y construir $b(n) = 1 - d_n(n)$, pero esa construcción exhibe una sucesión fuera de la lista. La buena definición de una enumeración propuesta no demuestra su sobreyectividad.

11.10. Ejercicios y soluciones desarrolladas

Los primeros cinco ejercicios controlan etiquetas y pertenencia; los siguientes cinco construyen productos, particiones y cocientes; los últimos cinco combinan buena definición y tamaño. En cada solución conviene separar los datos conservados por la construcción de las propiedades que todavía hay que probar.

Ejercicios. Ejercicio 1. Tres etiquetas, dos valores. Sea $I = \{r, s, t\}$, con $A_r = \{0, 2\}$, $A_s = \{2\}$ y $A_t = \{0, 2\}$. Determine unión, intersección y conjunto de valores de la familia. Un editor cambia A_r por $\{2\}$ y A_s por $\{0, 2\}$, dejando A_t igual. ¿Qué construcciones permanecen iguales y cuál cambia?

Ejercicio 2. Un ambiente omitido. Una familia tiene índice vacío. Dos lectores calculan su intersección usando ambientes $X = \{a, b\}$ e $Y = \{a, b, c\}$, con objetos distintos. Compare sus resultados y sus uniones. ¿Puede resolver la diferencia diciendo que uno olvidó un factor?

Ejercicio 3. Una condición que se aleja. Para $n \in \mathbb{N}$, sea $A_n = \{m \in \mathbb{N} : m \geq n + 2\}$. Determine unión e intersección total, y la intersección de A_0, A_1, A_2, A_3. Explique por qué las dos intersecciones no tienen que coincidir.

Ejercicio 4. Una condición que se amplía. Sea $B_n = \{m \in \mathbb{N} : m \le 2n+1\}$. Determine unión e intersección de toda la familia. En una prueba de la unión, alguien escribe «para cada m elijo $n = m$». Compruebe si ese testigo sirve, incluidos los valores iniciales.

Ejercicio 5. Repetición y pérdida de posiciones. Use la familia del ejercicio 1. Sea $J = \{0, 1, 2, 3\}$ y $r : J \to I$ dada por $r(0) = r(1) = r$, $r(2) = s$, $r(3) = t$; en esta fórmula la letra r a la derecha de las igualdades representa la etiqueta de I. Compare unión e intersección antes y después de reindexar. Si eliminamos del dominio las posiciones 2 y 3, ¿qué cambia? Proponga un nombre para la función que evite el doble uso de r.

Ejercicio 6. Un registro con tres campos. Para $I = \{p, q, r\}$, los factores son $X_p = \{1, 4\}$, $X_q = \{0, 3\}$ y $X_r = \{2\}$. Liste los elementos del producto dando sus valores en cada etiqueta. Determine las imágenes de las tres proyecciones y explique por qué el conjunto unión de los factores no describe el mismo objeto.

Ejercicio 7. Dos vacíos distintos. Compare el producto sin posiciones con el producto de la familia $Y_p = \{7\}$, $Y_q = \varnothing$. Escriba sus elementos si existen. Un informe les asigna a ambos tamaño 0; corrija el informe y justifique cada corrección.

Ejercicio 8. Selección y modificación. Sea $X_n = \{2n, 2n+3\}$ para $n \in \mathbb{N}$. Construya un elemento f del producto usando siempre el segundo valor. Modifíquelo sólo en la coordenada 4 para que tome allí el primer valor, obteniendo g. Compruebe que ambos pertenecen al producto y determine si $f = g$.

Ejercicio 9. Bloques y nombres repetidos. Sea $X = \{a, b, c, d, e, f\}$, con todos los objetos distintos, y $\mathcal{P} = \{\{a, d\}, \{b, e, f\}, \{c\}\}$. Defina la equivalencia por pertenecer a un mismo bloque y liste el cociente. Una lista de sus elementos es $[a], [b], [c], [d], [e], [f]$. Explique por qué no es una cuenta de seis elementos distintos.

Ejercicio 10. Un cociente de registros enteros. En $\mathbb{Z}^2$, defina $(a, b) \sim (c, d)$ si $2a + 5b = 2c + 5d$. Compruebe que es una equivalencia y que $h(a, b) = 2a + 5b$ induce una biyección $\overline{h} : \mathbb{Z}^2/\!\sim \to \mathbb{Z}$. Para la sobreyectividad, construya un representante a partir de cada entero t.

Ejercicio 11. Qué información permite el cociente. En el cociente del ejercicio 10, examine las propuestas $F([(a, b)]) = a + b$ y $G([(a, b)]) = 3(2a + 5b) - 4$. Decida cuáles definen funciones a $\mathbb{Z}$ y determine la imagen de las que sí lo hacen. Distinga buena definición de sobreyectividad.

Ejercicio 12. Dos cotas finitas. Se dan una inyección $I_5 \to A$ y una inyección $A \to I_7$. Demuestre que A es finito y determine todos los tamaños posibles. Para cada tamaño posible, muestre un conjunto y funciones que cumplan los datos.

Ejercicio 13. Un registro que comienza en 3. Sea $S = \{4n + 3 : n \in \mathbb{N}\}$. Demuestre que S es numerablemente infinito mediante una biyección explícita. Un lector propone $u(n) = 4n - 1$ desde $\mathbb{N}$ a S. Localice el defecto y dé dos reparaciones: una de la fórmula y otra del dominio.

Ejercicio 14. Dos canales de un mismo cociente. Sea Q el cociente del ejercicio 10, y sean p, q etiquetas distintas. Use la biyección $\overline{h} : Q \to \mathbb{Z}$ y la enumeración $e : \mathbb{N} \to \mathbb{Z}$ de §11.9 para construir una biyección $Q \times \{p, q\} \to \mathbb{N}$. Declare todas las funciones auxiliares y verifique que cada natural tiene exactamente una preimagen.

Ejercicio 15. Selecciones completas y listas propuestas. En el producto $P = \prod_{k \in \mathbb{N}} \{2k, 2k + 3\}$, alguien afirma que basta enumerar todas sus funciones como $f_0, f_1, \ldots$ porque cada coordenada ofrece sólo dos opciones. Construya, a partir de cualquier lista propuesta, una función $g \in P$ que no aparezca en ella. Explique por qué la no vaciedad de P y la finitud de los factores no demuestran numerabilidad.

Soluciones desarrolladas. Solución 1. La unión es $\{0, 2\}$: el índice r ya aporta ambos elementos y ningún factor añade otro. La intersección es $\{2\}$, pues 2 está en los tres factores y 0 falta en A_s. El conjunto de valores es $\{\{0, 2\}, \{2\}\}$. Tras el cambio permanecen iguales las tres construcciones: sigue habiendo valores de ambos tipos y el único elemento común es 2. Cambia la familia, porque cambia su valor en las etiquetas r y s. Las operaciones no recuperan todas las asignaciones originales.

Solución 2. En ambiente X la intersección vacía es $\{a, b\}$ y en ambiente Y es $\{a, b, c\}$. Ambas uniones son $\varnothing$, pues no existe índice testigo. No hay un factor olvidado: las familias no tienen posiciones. La diferencia corresponde al ambiente en la definición de intersección vacía. Para compararlas hay que acordar ese ambiente o mostrarlo como superíndice; no declarar erróneo un cálculo correcto bajo otra convención explícita.

Solución 3. La unión es $\{m \in \mathbb{N} : m \geq 2\}$. Todo elemento de algún A_n satisface $m \geq n + 2 \geq 2$; a la inversa, si $m \geq 2$, pertenece a A_0. La intersección total es vacía: para un candidato m el índice $n = m$ exigiría $m \geq m + 2$. La intersección de los cuatro factores es $A_3 = \{m : m \geq 5\}$, porque la condición más restrictiva es la de índice 3. Las cuatro condiciones no cubren las exigencias arbitrariamente grandes de la familia completa. Cada candidato que satisface las primeras puede fallar en una condición posterior.

Solución 4. La unión es $\mathbb{N}$. Para $m \in \mathbb{N}$ el testigo $n = m$ sirve, porque $m \leq 2m+1$; también para $m = 0$. No hay elementos fuera de $\mathbb{N}$ por definición de cada factor. La intersección es $\{0, 1\}$. Pertenecer a todos los B_n exige pertenecer a $B_0 = \{0, 1\}$; esos dos valores satisfacen $m \leq 2n + 1$ para todo $n \geq 0$. El testigo de la unión puede variar con m; no estamos fijando una cota uniforme para todos los naturales.

Solución 5. La función cubre las tres etiquetas de I, así que la unión sigue siendo $\{0, 2\}$ y la intersección $\{2\}$. Repetir el índice r no modifica las condiciones de pertenencia. Si dejamos sólo las posiciones $0, 1$, ambos valores son $\{0, 2\}$: la unión no cambia, pero la intersección se amplía a $\{0, 2\}$ porque desaparece la condición restrictiva de etiqueta s. La pérdida de sobreyectividad permite un cambio, aunque no obliga a que cambien todas las operaciones. Un nombre como $\rho : J \to I$ evita usar r simultáneamente para función y etiqueta.

Solución 6. Describiendo una función por el triple de valores $(f(p), f(q), f(r))$, los elementos son $(1, 0, 2)$, $(1, 3, 2)$, $(4, 0, 2)$ y $(4, 3, 2)$. Cada elección de las dos primeras coordenadas, combinada con el único valor de la tercera, da exactamente una función. Las imágenes de las proyecciones son respectivamente $\{1, 4\}$, $\{0, 3\}$ y $\{2\}$. La unión de factores es $\{0, 1, 2, 3, 4\}$; sus elementos son valores individuales y no registros con tres coordenadas. Ni su tipo de elemento ni su tamaño coincide con el del producto.

Solución 7. El producto sin posiciones contiene la función vacía y tiene tamaño 1. La condición de pertenencia no exige ningún valor y hay un único grafo con dominio vacío. El segundo producto es vacío y tiene tamaño 0: una función admisible tendría que asignar a q un valor en $\varnothing$, imposible. El informe confundió ausencia de posiciones con presencia de una posición sin valores. La convención del producto vacío coincide además con la cuenta finita de elecciones: hay una única asignación sin campos.

Solución 8. Definimos $f(n) = 2n + 3$ para todo $n \in \mathbb{N}$, y $g(4) = 8$, $g(n) = 2n + 3$ si $n \neq 4$. Ambas reglas son funciones en todo el dominio. En cada posición, f toma el segundo valor del factor; g toma el primero en posición 4 y el segundo en las restantes. Así ambos pertenecen al producto. No son iguales, pues $f(4) = 11$ y $g(4) = 8$. Cambiar una coordenada no exige volver a elegir todas las demás: se usa la selección f ya construida.

Solución 9. Los tres bloques son no vacíos, disjuntos dos a dos y cubren X, por lo que forman una partición. Compartir bloque define una equivalencia por §11.6. El cociente es $\{\{a, d\}, \{b, e, f\}, \{c\}\}$ y tiene tamaño 3. Se cumple $[a] = [d]$, $[b] = [e] = [f]$ y $[c] = \{c\}$. La lista de seis nombres repite clases; contar nombres de representantes no cuenta elementos distintos del cociente. La proyección envía cada elemento a su bloque y no es inyectiva, por ejemplo porque $a \neq d$ pero $[a] = [d]$.

Solución 10. Igualdad de los valores $2a + 5b$ es reflexiva, simétrica y transitiva, así que la relación es una equivalencia. La función h es constante exactamente en esas clases; por §11.7 desciende a $\overline{h}([(a, b)]) = 2a + 5b$. Si dos clases tienen la misma imagen, sus representantes satisfacen la relación y las clases son iguales; esto prueba inyectividad. Para $t \in \mathbb{Z}$, el par $(-2t, t)$ tiene valor $2(-2t) + 5t = t$. Su clase es una preimagen de t, lo que prueba sobreyectividad. La inversa es $t \mapsto [(-2t, t)]$: cualquier representante con valor t pertenece a esa clase, de modo que la fórmula devuelve la entrada correspondiente.

Solución 11. La primera propuesta falla: $(0, 0)$ y $(5, -2)$ tienen ambos valor ponderado 0, pero sus sumas son respectivamente 0 y 3. Son representantes de la misma clase que producirían respuestas distintas; F no define función. Para G, equivalencia de representantes da igualdad de sus valores ponderados y por ello de $3(2a + 5b) - 4$. Sí define una función. Su imagen es $\{3t - 4 : t \in \mathbb{Z}\}$, porque $\overline{h}$ alcanza todo entero. No es todo $\mathbb{Z}$: por ejemplo 0 exigiría $t = 4/3$, que no es entero. La comprobación de buena definición no prueba sobreyectividad; son obligaciones distintas.

Solución 12. La inyección $A \to I_7$ identifica A con su imagen, un subconjunto de un conjunto finito. Por §11.8, A es finito; sea $|A| = k$. Transportando las inyecciones mediante una biyección $I_k \to A$, obtenemos una inyección $I_5 \to I_k$ y otra $I_k \to I_7$. El principio finito da $5 \leq k \leq 7$. Los tamaños posibles son $5, 6, 7$. Para cada uno, tome $A = I_k$: la inclusión de I_5 en I_k y la inclusión de I_k en I_7 cumplen los datos. Esto muestra que ninguna de las cotas se puede mejorar sin información adicional.

Solución 13. La función $v : \mathbb{N} \to S$, $v(n) = 4n+3$, está bien definida y es sobreyectiva por la descripción de S. Si $v(n) = v(m)$, cancelar 4 en los enteros da $n = m$, luego es inyectiva. Su inversa es $s \mapsto (s - 3)/4$, que es natural precisamente por pertenecer s a S. Así S es numerablemente infinito. La regla propuesta u da $u(0) = -1 \notin S$ y ni siquiera define una función con el codominio declarado. Reparación de fórmula: usar v. Reparación de dominio: restringir u a $\mathbb{N} \setminus \{0\}$; si $n \geq 1$, $4n - 1 = 4(n - 1) + 3$, y cada elemento de S se obtiene una única vez con $n = k + 1$. Ese dominio también es numerablemente infinito por §11.8.

Solución 14. Sea $j = e^{-1} : \mathbb{Z} \to \mathbb{N}$, cuya fórmula se verificó en §11.9. Definimos $c : Q \to \mathbb{N}$ por $c = j \circ \overline{h}$; es una biyección por composición. Ahora sea $H : Q \times \{p, q\} \to \mathbb{N}$ dada por $H(C, p) = 2c(C)$ y $H(C, q) = 2c(C) + 1$. Cada natural es de forma única $2n$ o $2n + 1$, con $n \in \mathbb{N}$. En el primer caso su preimagen es $(c^{-1}(n), p)$; en el segundo, $(c^{-1}(n), q)$. La unicidad de la paridad y de $c^{-1}(n)$ prueba simultáneamente inyectividad y sobreyectividad. La función auxiliar c usa clases, no representantes elegidos libremente, por lo que la fórmula queda bien definida.

Solución 15. Para la lista propuesta definimos $g(k) = 2k + 3$ si $f_k(k) = 2k$, y $g(k) = 2k$ si $f_k(k) = 2k + 3$. Las dos alternativas son exhaustivas y excluyentes porque f_k pertenece al producto y los valores del factor son distintos. La regla define una función en todo $\mathbb{N}$ y siempre toma un valor del factor de índice k, así que $g \in P$. En esa coordenada difiere de f_k; por tanto $g \neq f_k$ para cada k y la lista no cubre P. El producto es no vacío por la fórmula del ejercicio 8, pero esa existencia no proporciona una enumeración. Si P fuera finito no vacío o numerablemente infinito, habría una lista natural sobreyectiva, que el argumento excluye. Por §11.9, P no es a lo sumo numerable. La finitud de cada factor no limita el conjunto de selecciones completas al tamaño de un único factor.

Cierre del lenguaje estructural. Las construcciones de este capítulo permiten organizar objetos sin perder de vista sus condiciones de uso. Antes de operar con una familia hay que identificar sus índices; antes de tomar un complemento, el ambiente; antes de usar un producto, las coordenadas admisibles; antes de calcular en clases, la independencia del representante; antes de comparar tamaños, la función y las propiedades que justifican la comparación.

El libro ha construido un lenguaje para afirmar, definir, demostrar, explorar, refutar y presentar razonamientos. Los cursos de álgebra, cálculo, análisis y matemática discreta podrán

usarlo en contextos propios. La obligación permanece: fijar el dominio, hacer visibles las hipótesis y justificar exactamente la conclusión que se afirma.

Capítulo 12 — ¿Cómo puede una colección convertirse en un número?

12.1. Cuando todavía no hemos dado nombre a una cantidad

Sobre una mesa hay cuatro piedras; en otra, cuatro conchas. Los objetos son diferentes y tampoco existe una manera privilegiada de ordenar unos junto a otros. Sin embargo, podemos poner una piedra frente a cada concha hasta agotar ambas colecciones. Esa posibilidad de formar parejas, sin dejar objetos sobrantes en ninguna mesa, permite afirmar que tienen la misma cantidad incluso antes de haber elegido la palabra «cuatro».

Probemos con una colección de tres lápices y otra de cinco botones. Podemos emparejar cada lápiz con un botón, pero quedarán dos botones sin pareja. La dificultad no depende del orden en que hagamos los emparejamientos. Tampoco cambiará si trasladamos los objetos, los reemplazamos por otros o modificamos sus colores. Todo ello sugiere que la cantidad es una propiedad que queremos conservar al olvidar casi todo lo demás acerca de una colección.

Una correspondencia que empareja cada elemento de A con uno y sólo uno de B, y recíprocamente, es una **biyección** $f : A \to B$. Escribiremos

$$A \approx B \quad \Longleftrightarrow \quad \text{existe una biyección } f : A \longrightarrow B.$$

La existencia de una biyección se puede comprobar sin asignar números a los objetos: basta producir el emparejamiento. En una colección de piedras y conchas de igual cantidad, el emparejamiento puede hacerse de muchas maneras; cualquiera de ellas sirve para establecer la relación $A \approx B$.

Esta relación tiene tres propiedades que hacen posible construir a partir de ella. Cada colección se corresponde consigo misma por la aplicación identidad, de modo que $A \approx A$. Si una biyección lleva A sobre B, su inversa lleva B sobre A; por eso $A \approx B$ implica $B \approx A$. Finalmente, si $f : A \to B$ y $g : B \to C$ son biyectivas, su composición $g \circ f : A \to C$ también lo es: de $A \approx B$ y $B \approx C$ se sigue $A \approx C$. Son, respectivamente, la **reflexividad**, la **simetría** y la **transitividad**.

Para comprender por qué las tres propiedades importan, pensemos en un conjunto X de objetos, provisto de una relación $\sim$ que las cumpla. Dado $x \in X$, reunimos todos los objetos relacionados con él en un solo conjunto:

$$[x] = \{y \in X : y \sim x\}.$$

Ésta es la **clase de equivalencia** de x. El propio x pertenece a su clase por reflexividad, pero otros objetos también pueden pertenecer a ella. Una clase admite diversos nombres, según el representante escogido. Si $x \sim y$, la transitividad muestra que cada elemento de $[x]$ pertenece a $[y]$; la simetría permite demostrar la inclusión inversa. De aquí obtenemos $[x] = [y]$. Recíprocamente, si ambas clases son iguales, entonces $x \in [x] = [y]$ y, por definición, $x \sim y$. Hemos establecido una regla que utilizaremos constantemente:

$$\boxed{x \sim y \quad \Longleftrightarrow \quad [x] = [y].}$$

Podemos comprobar la necesidad de la transitividad con una comparación tentadora: declarar que dos colecciones están relacionadas cuando comparten alguna marca. Las colecciones $\{p\}$ y $\{p, q\}$ comparten p, y $\{p, q\}$ y $\{q\}$ comparten q, pero $\{p\}$ y $\{q\}$ no comparten ninguna. Esta comparación no es transitiva y, por eso, no podría organizar las colecciones en clases que se comporten como bloques de una partición.

Si dos clases tienen un elemento en común, la simetría y la transitividad obligan a que sean la misma clase. Así, las clases forman una **partición** del conjunto de partida: todos sus elementos quedan agrupados, cada clase es no vacía y dos clases distintas nunca se solapan. El **conjunto cociente** reúne esos grupos como elementos de un conjunto nuevo:

$$X/\!\sim \;= \{[x] : x \in X\}.$$

El capítulo 11 mostró cómo las clases de equivalencia forman particiones y conjuntos cociente. Con las colecciones de este ejemplo podemos dar un paso más: si x es una colección y $[x]$ reúne todas las que pueden emparejarse exactamente con ella, quizá la cantidad que buscamos pueda identificarse con esa clase. Para que esa idea se convierta en una construcción matemática necesitamos fijar primero de qué colecciones podemos hablar y evitar que la palabra «todas» nos lleve fuera de un conjunto bien definido.

12.2. Primer momento de construcción: los naturales como clases de colecciones finitas

Podemos tener dos colecciones diferentes que admitan el mismo emparejamiento: tres lápices y tres conchas, o tres marcas escritas sobre una hoja y otras tres sobre una piedra. Si queremos llamar *tres* a algo que no dependa de las marcas particulares, la relación $A \approx B$ nos ofrece una posibilidad exacta: tomar **como número** la clase de todas las colecciones que son biyectivas con una colección elegida. La palabra «todas» exige, sin embargo, una precaución. No podemos formar indiscriminadamente un conjunto con todas las colecciones de cualquier procedencia; trabajaremos dentro de un conjunto previamente fijado de marcas.

Elijamos un conjunto infinito U cuyos elementos llamaremos *marcas*. Pensamos en él como una reserva que siempre permite escoger marcas nuevas después de haber empleado una cantidad finita de ellas. Necesitamos, entre sus subconjuntos, aquellos que puedan construirse comenzando con el vacío y agregando marcas una tras otra. **Ésta será nuestra noción de colección finita.** Su formulación conjuntista puede darse sin introducir los naturales que todavía queremos construir.

En efecto, consideremos las familias $\mathcal{G} \subseteq \mathcal{P}(U)$ que cumplen estas dos condiciones: $\varnothing \in \mathcal{G}$, y si $A \in \mathcal{G}$ y $u \in U$, entonces $A \cup \{u\} \in \mathcal{G}$. La intersección de todas esas familias vuelve a satisfacer ambas condiciones; la llamaremos $\mathcal{F}(U)$. De este modo,

$$\mathcal{F}(U) = \bigcap \{\mathcal{G} \subseteq \mathcal{P}(U) : \varnothing \in \mathcal{G}, \quad \forall A \in \mathcal{G}\ \forall u \in U,\ A \cup \{u\} \in \mathcal{G}\}\,.$$

La fórmula es una manera rigurosa de indicar que no incorporamos colecciones porque nos parezcan pequeñas ni porque ya sepamos contar sus elementos. Incorporamos el vacío y todo lo que sea obligatorio al agregar una marca a una colección admitida. Esta descripción, que utiliza una teoría de conjuntos de fondo y un conjunto infinito U, no pretende deducir la existencia de U desde cero. Su ventaja es impedir un círculo lógico: *no hemos definido «finito» como «tener n elementos», para luego definir n mediante las colecciones finitas.* La propiedad de ser la menor familia cerrada bajo adjunción proporciona además el principio inductivo que usaremos en el capítulo.

Sobre $\mathcal{F}(U)$ conservamos la relación $A \approx B$ cuando existe una biyección $A \to B$. Cada clase $[A]$ agrupa las colecciones que pueden emparejarse exactamente con A. Si $A = \{p, q, r\}$, pertenecen a $[A]$ las ternas de marcas cualesquiera de U, aunque sus elementos no coincidan con p, q, r. El vacío tiene una clase especial: una biyección $\varnothing \to B$ sólo es posible cuando

$B = \varnothing$; por eso $[\varnothing] = \{\varnothing\}$. En cambio, la clase de una colección unitaria contiene muchos representantes diferentes.

Estamos en condiciones de realizar el **primer momento de construcción numérica**:

$$\boxed{\mathbb{N}_U = \mathcal{F}(U)/\approx.}$$

Cada elemento de $\mathbb{N}_U$ es una clase de colecciones. Designamos $0 = [\varnothing]$, $1 = [\{p\}]$, $2 = [\{p, q\}]$ para $p \neq q$, y así sucesivamente. Los nombres tradicionales se aplican a clases, aunque para exhibirlas o calcular con ellas elijamos representantes. Con dos colecciones $A = \{p, q\}$ y $B = \{r, s\}$ de marcas distintas, la afirmación $A = B$ es falsa, mientras que $[A] = [B]$ es verdadera: basta una biyección que lleve p a r y q a s. El número que llamamos dos es esta clase completa; ninguna de las dos colecciones se convierte por sí sola en el número.

Conviene mantener separadas tres afirmaciones. $A \in \mathcal{F}(U)$ dice que A es una colección finita de marcas; $[A] \in \mathbb{N}_U$ dice que su clase es un número natural de esta realización; $B \in [A]$ significa que B y A admiten una biyección. Confundir A con $[A]$ borra justamente la construcción que intentábamos comprender. Una vez fijada esta realización podremos escribir simplemente $\mathbb{N}$ cuando no haya ambigüedad, sin olvidar que otras construcciones conjuntistas de los naturales producen objetos distintos aunque aritméticamente equivalentes.

> [!note] Dos caminos hacia la idea de número En 1895, Georg Cantor formuló una pregunta que se parece mucho a la que nos ha acompañado: ¿qué queda de una colección cuando prescindimos de la naturaleza de sus elementos y del orden en que aparecen? Llamó *número cardinal* a la noción resultante y relacionó la igualdad de cardinales con la posibilidad de establecer una correspondencia uno a uno entre conjuntos. En el primer apartado de sus *Beiträge zur Begründung der transfiniten Mengenlehre* (§1) aparece expresamente esa relación.
>
> Richard Dedekind había elegido otro punto de partida en *Was sind und was sollen die Zahlen?* (1888). En vez de caracterizar la cantidad mediante emparejamientos, investigó los sistemas organizados por una operación sucesor y la menor cadena generada a partir de un elemento inicial. Los dos planteamientos iluminan aspectos complementarios: Cantor ayuda a comprender por qué colecciones diferentes pueden compartir una cantidad; Dedekind ayuda a reconocer qué estructura sostiene la sucesión de los números y la inducción. Nuestra construcción de $\mathbb{N}_U$ mediante clases de subconjuntos de un universo fijo es una **realización moderna elegida con fines pedagógicos**: no debe confundirse literalmente con las definiciones históricas de ninguno de los dos.
>
> Fuentes: Cantor, *Beiträge*, *Mathematische Annalen* 46 (1895), pp. 481–512, §1; *Stanford Encyclopedia of Philosophy*, «Dedekind's Contributions to the Foundations of Mathematics».

Comprobación 1. Si $A = \{\alpha, \beta, \gamma\}$ y $B = \{u, v, w\}$ tienen tres elementos distintos cada uno, ¿es verdadera la igualdad $A = B$? ¿Qué ocurre con $[A] = [B]$? ¿Qué objeto hemos llamado 3?

Solución razonada. No se puede afirmar $A = B$; de hecho, si sus marcas son diferentes, los conjuntos también lo son. Sí existe una biyección entre A y B, por ejemplo la que asigna sucesivamente $\alpha \mapsto u$, $\beta \mapsto v$ y $\gamma \mapsto w$. Por eso $[A] = [B]$. El símbolo 3 nombra esta clase, cuyos elementos son colecciones finitas de tres marcas, y no una colección particular.

12.3. Hacer aritmética con los naturales que acabamos de construir

Reconocer que dos colecciones tienen la misma cantidad es apenas el comienzo. Queremos reunir colecciones, formar filas de objetos, hablar de la cantidad siguiente y comparar

cantidades. Para que esas operaciones pertenezcan realmente a $\mathbb{N}_U$, su resultado debe depender de la **clase**, sin cambiar cuando elijamos otros representantes.

Para sumar las clases de dos colecciones A y B, elegimos representantes A' y B' cuyas marcas sean distintas entre sí, y reunimos ambas colecciones. Siempre podemos hacerlo porque U contiene marcas nuevas fuera de cualquier colección finita. Proponemos

$$\boxed{[A] + [B] = [A' \cup B'], \qquad A' \approx A, \quad B' \approx B, \quad A' \cap B' = \varnothing.}$$

¿Por qué la exigencia de que sean disjuntas? Si $A = \{p, q\}$ y $B = \{q, r\}$, su unión ordinaria tiene tres elementos, aunque las cantidades que queríamos reunir son dos y dos. Tomar marcas nuevas evita contar una misma marca dos veces. Podemos, por ejemplo, reemplazar la segunda colección por $B' = \{s, t\}$, disjunta de A, y obtener una colección de cuatro marcas.

Falta el argumento que permite abandonar esta elección concreta. Supongamos que también utilizamos C', D' como representantes disjuntos de las mismas clases. Una biyección $f : A' \to C'$ y otra $g : B' \to D'$ se pueden reunir en una función sobre $A' \cup B'$: utilizamos f en la primera parte y g en la segunda. Como los dominios y las imágenes son disjuntos, la función resultante es una biyección de $A' \cup B'$ sobre $C' \cup D'$. Por eso ambas uniones pertenecen a la misma clase. **La suma está bien definida.** El lector puede comprobar que aquí hemos demostrado la independencia de dos elecciones a la vez: qué colección representa a cada sumando y qué marcas utilizamos para que los conjuntos reunidos no se solapen.

El producto corresponde a organizar una colección de filas A y columnas B. Cada casilla lleva el par ordenado (a, b), con $a \in A$ y $b \in B$, de modo que la cuadrícula queda representada por $A \times B$. Podemos justificar que esta colección de casillas es finita sin recurrir todavía a la multiplicación numérica: si A está vacía, el producto está vacío; y cada vez que agregamos una marca a a A, añadimos al producto una copia de B, formada por los pares (a, b) con $b \in B$. Como B es finita y sólo agregamos un número finito de marcas a A, la construcción inductiva de finitud garantiza que $A \times B$ también es finito. Podemos, por tanto, copiar sus casillas mediante marcas distintas de U. Si $C \subseteq U$ es una de esas copias, definimos

$$\boxed{[A] \cdot [B] = [C] \quad \text{cuando } C \approx A \times B.}$$

Puede parecer más complicado que multiplicar cantidades conocidas, pero tiene una ventaja: si cambiamos A por una colección biyectiva A' y B por otra B', las biyecciones se combinan y producen una biyección $A \times B \to A' \times B'$, dada por $(a, b) \mapsto (f(a), g(b))$. Las cuadrículas tienen por tanto la misma clase, cualquiera que sea la copia de marcas utilizada. También el producto está bien definido. Una biyección entre los dos conjuntos de filas y otra entre los dos conjuntos de columnas producen el emparejamiento de casillas que necesitamos; no hemos dado por supuesto que multiplicar números naturales ya estuviera disponible.

Podemos expresar la idea de «un número más» sin recurrir a una suma ya conocida. Dada una colección A, elegimos $u \in U \setminus A$ y definimos el **sucesor** de su clase por

$$\boxed{S([A]) = [A \cup \{u\}].}$$

La marca nueva no debe ser alguna que ya pertenezca a A, porque entonces la unión no añadiría nada. Si elegimos otra marca nueva, o reemplazamos A por una colección equivalente, el emparejamiento se prolonga enviando la marca añadida de una colección a la de la otra; obtenemos clases iguales. En particular, $S(0) = 1$, $S(1) = 2$, y así sucesivamente. Para ver que S es inyectiva, consideremos una biyección entre dos colecciones obtenidas añadiendo una marca nueva. Si hace falta, podemos intercambiar dos imágenes para que la primera marca nueva corresponda a la segunda; al quitar esas dos marcas, queda una biyección entre las colecciones originales. El sucesor nunca es cero, porque una colección con al menos una marca no puede estar en biyección con la colección vacía.

Hay además un principio de inducción encerrado en la propia construcción. Si una propiedad de las clases vale para 0 y se conserva cada vez que añadimos una marca nueva, vale para todas las clases de $\mathbb{N}_U$: la familia de colecciones sobre las que la propiedad es verdadera contiene la colección vacía y es estable bajo la adjunción de una marca, de modo que contiene toda la familia $\mathcal{F}(U)$. Éste es el motivo por el cual la inducción puede describirse como una ley del sistema recién construido, y también explica por qué su uso no depende de una lista escrita de números.

Las leyes familiares de la suma y el producto —asociatividad, conmutatividad y distributividad— pueden obtenerse mediante biyecciones entre uniones disjuntas y productos de colecciones; algunas requieren más detalle que el empleado aquí. Tampoco hemos desarrollado aún toda la teoría del orden. La construcción ya nos proporciona objetos y operaciones bien definidos; el tratamiento completo de sus leyes exige demostrar los resultados correspondientes, no darlos por ciertos por la semejanza con la aritmética habitual.

Comprobación 2. ¿Por qué no podemos definir $[A]+[B] = [A\cup B]$ sin ninguna condición? ¿Qué cambio sencillo repara el problema?

Solución razonada. Si $A = B = \{p\}$, la fórmula propuesta daría $[A\cup B] = [\{p\}] = 1$ para la suma de las clases de dos colecciones unitarias. Si sustituimos el segundo representante por $\{q\}$, con $q \neq p$, obtendríamos $[\{p, q\}] = 2$. La misma pareja de clases recibiría dos resultados distintos. Basta exigir que los representantes elegidos sean disjuntos, y después demostrar que todas esas elecciones disjuntas conducen a la misma clase.

12.4. Definir una operación no demuestra sus leyes

La suma que acabamos de definir obliga a escoger representantes disjuntos; la elección puede hacerse siempre, pero su razón matemática merece una imagen que podamos conservar en las pruebas. Una vez entendida esa reunión de colecciones, será posible estudiar las leyes de la suma sin confundir dos cuestiones: que una operación exista sobre las clases y que satisfaga las identidades que esperamos de ella.

La unión disjunta: reunir sin confundir. La dificultad se ve si renunciamos por un momento a escoger representantes disjuntos. Tomemos dos colecciones de marcas,

$$A = \{p, q\}, \qquad B = \{q, r\}.$$

Cada colección tiene dos elementos. Sin embargo, $A \cup B = \{p, q, r\}$ tiene solamente tres, porque una misma marca, q, pertenece a ambas colecciones y en un conjunto no se escribe dos veces. La unión ordinaria responde correctamente a la pregunta por las marcas que aparecen en alguna de las dos colecciones; al sumar queremos conservar las dos aportaciones, incluso si los representantes elegidos comparten una marca.

Podemos lograrlo sin exigir que A y B sean disjuntos desde el comienzo. A cada elemento de A le adjuntamos la etiqueta 0, y a cada elemento de B la etiqueta 1. Las etiquetas permiten distinguir dos apariciones de una misma marca según su procedencia. La colección que resulta recibe el nombre de **unión disjunta**, y se representa mediante el símbolo $\sqcup$:

$$\boxed{A \sqcup B = (A \times \{0\}) \cup (B \times \{1\}).}$$

En nuestro ejemplo obtenemos

$$A \sqcup B = \{(p, 0), (q, 0), (q, 1), (r, 1)\}.$$

Los pares $(q, 0)$ y $(q, 1)$ son distintos, aunque llevan la misma marca en su primera coordenada. Contamos así cuatro elementos, como corresponde a reunir dos colecciones de dos elementos cada una. La unión se llama *disjunta* porque las dos copias etiquetadas, $A\times\{0\}$ y $B\times\{1\}$, no tienen ningún elemento en común. Si A y B ya eran disjuntos, podemos

olvidarnos de las etiquetas y obtener una biyección natural entre $A \sqcup B$ y $A \cup B$. Son conjuntos naturalmente correspondientes, aunque sus elementos no tengan literalmente la misma forma.

En nuestra realización $\mathbb{N}_U = \mathcal{F}(U)/\approx$, los pares etiquetados de $A \sqcup B$ y las casillas de $A \times B$ no tienen por qué pertenecer a U. Cuando pongamos corchetes alrededor de esos conjuntos entenderemos la clase de **cualquier copia finita de ellos formada con marcas de** U. Dos copias así son biyectivas, y el número que representan es el mismo. Con esta convención, la suma de la sección anterior admite la escritura más cómoda

$$\boxed{[A] + [B] = [A \sqcup B].}$$

Ya comprobamos que la clase resultante no cambia si sustituimos A y B por representantes equivalentes. Eso demuestra que la suma está bien definida. Aún queda por averiguar si la nueva operación es conmutativa y asociativa: para lo primero necesitaremos comparar $A \sqcup B$ con $B \sqcup A$; para lo segundo, las dos formas de reunir tres colecciones. Demostrar esas leyes supone un trabajo adicional al de definir correctamente la operación.

12.5. Las leyes de los naturales nacen de las correspondencias

Reunir colecciones de distintas maneras. Si reunimos las colecciones A y B como dos grupos etiquetados, podemos intercambiar sus etiquetas: a cada $(a, 0)$ le hacemos corresponder $(a, 1)$, y a cada $(b, 1)$ le hacemos corresponder $(b, 0)$ en la reunión de orden inverso. Esta regla es biyectiva, porque puede deshacerse mediante el mismo intercambio. De ahí se deduce

$$[A] + [B] = [B] + [A].$$

En una suma de tres cantidades parece haber dos decisiones de agrupación, $(A \sqcup B) \sqcup C$ y $A \sqcup (B \sqcup C)$. Las etiquetas que aparecen en los conjuntos anidados son diferentes, pero cada elemento procede de exactamente una de las colecciones originales. La regla que conserva el elemento y su procedencia —primera, segunda o tercera colección— establece una biyección entre ambos conjuntos. Por ello

$$([A] + [B]) + [C] = [A] + ([B] + [C]).$$

Las clases de $A \sqcup \varnothing$ y de A coinciden: basta olvidar la etiqueta de los elementos de A. Como $0 = [\varnothing]$, obtenemos $[A] + 0 = [A]$. Estas pruebas muestran de dónde proceden la conmutatividad, la asociatividad y el cero aditivo; no necesitamos presuponer las mismas leyes para las cantidades que estamos definiendo.

Para la multiplicación observemos una cuadrícula con filas indexadas por A y columnas indexadas por B. Intercambiar filas y columnas define la biyección $(a, b) \mapsto (b, a)$ entre $A \times B$ y $B \times A$, de modo que $[A][B] = [B][A]$. Si las casillas tienen, además, una tercera coordenada, las correspondencias

$$((a, b), c) \longleftrightarrow (a, (b, c))$$

entre $(A \times B) \times C$ y $A \times (B \times C)$ prueban la asociatividad del producto. Si E es una colección de una sola marca, las parejas (a, e) están en correspondencia biyectiva con los elementos $a \in A$; así, $[A] \cdot 1 = [A]$. En cambio, $A \times \varnothing$ es vacío, y por eso $[A] \cdot 0 = 0$.

La distributividad puede reconocerse en una cuadrícula cuyas columnas provienen de dos grupos diferentes. Cada elemento de $A \times (B \sqcup C)$ pertenece a una de dos regiones, según la etiqueta de su segunda coordenada. En la primera región corresponde a una casilla de $A \times B$ y en la segunda a una de $A \times C$. Ésa es una biyección explícita

$$A \times (B \sqcup C) \longrightarrow (A \times B) \sqcup (A \times C),$$

que envía $(a, (b, 0))$ a $((a, b), 0)$ y $(a, (c, 1))$ a $((a, c), 1)$. Al tomar clases se obtiene

$$[A]([B]+[C]) = [A][B] + [A][C].$$

Las leyes que acabamos de demostrar justifican que $\mathbb{N}_U$, con estas operaciones, sea un **semianillo conmutativo con unidad**. La terminología reúne en un solo nombre las propiedades ya probadas; no hace falta conocerla para comprender las biyecciones. Una vez probadas estas leyes, podremos preguntarnos por la capacidad que todavía nos falta: efectuar una resta cuando la cantidad sustraída es mayor que la cantidad inicial.

Por qué se pueden cancelar sumandos. Consideremos una igualdad de clases $[A] + [C] = [B] + [C]$. Es natural querer concluir $[A] = [B]$, como hacemos cuando cancelamos una cantidad igual en ambos miembros de una ecuación. Sin embargo, la cancelación no se desprende de la mera asociatividad y conmutatividad: necesita una propiedad de las **colecciones finitas**. Vamos a demostrarla sin introducir todavía números negativos.

La igualdad de las sumas afirma que existe una biyección

$$f : A \sqcup C \longrightarrow B \sqcup C.$$

Si C es vacío, la correspondencia requerida entre A y B ya está ante nosotros. Si C contiene una marca c, atendamos a la copia etiquetada de c en el dominio y a la correspondiente copia de c en el codominio. Puede ocurrir que f lleve una a la otra; en ese caso, al retirar ambas marcas queda una biyección entre $A \sqcup (C \setminus \{c\})$ y $B \sqcup (C \setminus \{c\})$. Si f envía la marca elegida a otro elemento del codominio, podemos **intercambiar allí las dos imágenes**: componemos f con la permutación que intercambia su imagen actual y la copia de c en el codominio. La nueva función sigue siendo una biyección y ya empareja las dos copias de c, por lo que podemos retirarlas.

El procedimiento reduce en una marca la colección común. La definición inductiva de finitud asegura que podemos repetir esta operación hasta eliminar todas las marcas de C; formalmente, es una inducción sobre su construcción por adjunción de marcas. Queda una biyección $A \to B$, y concluimos

$$\boxed{[A] + [C] = [B] + [C] \quad \Longrightarrow \quad [A] = [B].}$$

Esta **ley de cancelación aditiva** es el puente que necesitaremos enseguida. La relación sobre pares de naturales con la que formamos los enteros emplea igualdades de sumas cruzadas, y su transitividad requiere cancelar una cantidad común. Esa cancelación pertenece legítimamente a la aritmética natural recién construida.

Comparar cantidades finitas. El orden también puede describirse sin contar previamente los elementos de una colección. Diremos que $[A] \leq [B]$ cuando existe una aplicación inyectiva de A en B. La definición sólo depende de las clases, porque una inyección sigue siendo inyección si antes y después componemos con biyecciones que cambian los representantes. Además, una inyección $j : A \to B$ identifica A con una parte de B; el complemento de esa parte es una colección finita C. Esto permite escribir $[B] = [A] + [C]$. Recíprocamente, si $[B] = [A] + [C]$, la inclusión de A en $A \sqcup C$, seguida de una biyección hacia un representante de $[B]$, proporciona una inyección. Hemos demostrado la caracterización

$$\boxed{n \leq m \quad \Longleftrightarrow \quad \exists k \in \mathbb{N}\ (m = n + k).}$$

La comparación es total: dadas dos colecciones finitas, una admite una inyección en la otra. Podemos probarlo mediante inducción sobre la primera colección. Si está vacía, la función vacía es una inyección de ella en cualquier otra. Para el paso inductivo consideremos una colección no vacía A y otra B. Si B es vacía, tenemos una inyección de B en A. Si ambas contienen marcas, elegimos $a \in A$ y $b \in B$, y aplicamos la hipótesis inductiva a $A \setminus \{a\}$ y $B \setminus \{b\}$. Si la primera colección reducida se inyecta en la segunda, extendemos la aplicación

enviando a a b, y obtenemos una inyección $A \to B$; si la dirección es la inversa, la extendemos enviando b a a y conseguimos una inyección $B \to A$. En cualquier caso existe una comparación entre las clases originales.

La reflexividad de $\leq$ procede de la identidad, y su transitividad de la composición de inyecciones. Para la antisimetría supongamos $n \leq m$ y $m \leq n$. Existen $k, \ell \in \mathbb{N}$ tales que $m = n + k$ y $n = m + \ell$. Sustituyendo, obtenemos $n = n + (k + \ell)$; la cancelación aditiva da $0 = k + \ell$. Pero una unión disjunta sólo es vacía cuando sus dos partes lo son, así que $k = \ell = 0$ y $m = n$. Hemos probado un orden total sin recurrir a la aritmética de enteros que todavía no existe.

Comprobación 3. Alguien intenta demostrar la cancelación natural diciendo: «Si $n + k = m + k$, restamos k y queda $n = m$». Señala qué paso tiene que justificarse en nuestra construcción y explica cómo se consigue sin restar.

Solución razonada. El paso «restamos k» presupone justamente la cancelación que se pretende demostrar. Representamos la igualdad mediante una biyección $A \sqcup C \to B \sqcup C$, emparejamos una por una las marcas del grupo común —reajustando la biyección mediante un intercambio cuando sea necesario— y las retiramos. La finitud permite terminar el procedimiento y obtener una biyección entre A y B, de donde $n = m$.

12.6. Una equivalencia *entre* naturales: los restos no construyen los naturales

Construidos los naturales, podemos utilizarlos para formar un cociente diferente. Pensemos en el reparto de ocho fichas en montones de tres: sobran dos. Si comenzamos con once fichas, sobran también dos. Lo mismo sucede con dos, cinco, catorce y todos los números que aumentan de tres en tres. Si nuestro problema sólo registra el resto, ocho y once resultan indistinguibles **para ese problema**, aunque representan cantidades distintas en $\mathbb{N}$.

Definimos $a \sim_3 b$ cuando a y b dejan el mismo resto al dividir por tres. El resto se apoya en un resultado que podemos justificar a partir de la inducción y el orden que ya construimos. Fijado un natural $m > 0$, todo $n \in \mathbb{N}$ se escribe de una única manera como $n = mq + r$, donde $q \in \mathbb{N}$ y $0 \leq r < m$. Para la existencia comenzamos con $n = 0$, $q = r = 0$, y avanzamos por inducción: al pasar de n a $n + 1$, o bien aumentamos el resto r en una unidad si $r + 1 < m$, o bien, cuando $r + 1 = m$, lo devolvemos a cero y aumentamos q en una unidad. Para la unicidad, dos cocientes diferentes estarían ordenados; si $q < q'$, entonces $mq + r < m(q + 1) \leq mq' \leq mq' + r'$, de modo que dos representaciones del mismo n no pueden tener cocientes distintos. Una vez fijado el cociente, la cancelación aditiva obliga a que sus restos coincidan. Ésta es la **división euclídea** en los naturales, y para $m = 3$ permite separarlos en tres conjuntos:

$$C_0 = \{0, 3, 6, 9, \ldots\},$$
$$C_1 = \{1, 4, 7, 10, \ldots\},$$
$$C_2 = \{2, 5, 8, 11, \ldots\}.$$

La relación es reflexiva, simétrica y transitiva porque estas tres propiedades las posee la igualdad de los restos. Al formar las clases obtenemos

$$\mathbb{N}/{\sim_3} = \{C_0, C_1, C_2\}.$$

El signo de cociente es el mismo que utilizamos en la construcción de $\mathbb{N}_U$, pero los dos procedimientos tienen resultados y propósitos distintos. Antes tomamos colecciones finitas y, al identificarlas cuando tenían igual cantidad, **obtuvimos los propios números naturales**. Aquí tomamos números naturales ya disponibles y los identificamos cuando comparten un resto; hemos obtenido un conjunto de sólo tres clases. Es un ejemplo de equivalencia *en* $\mathbb{N}$, distinto de la construcción de $\mathbb{N}$ *mediante* equivalencia.

Podemos probar, además, que la suma descendente de los representantes define una operación sobre estas nuevas clases:

$$\boxed{[a]+[b]=[a+b] \quad \text{en } \mathbb{N}/\!\sim_3.}$$

La expresión merece una prueba. Si $a \sim_3 a'$ y $b \sim_3 b'$, existen $p, p', q, q' \in \mathbb{N}$ y restos $r, s \in \{0,1,2\}$ tales que

$$a = 3p + r, \quad a' = 3p' + r,$$
$$b = 3q + s, \quad b' = 3q' + s.$$

Entonces $a + b = 3(p + q) + (r + s)$ y $a' + b' = 3(p' + q') + (r + s)$. Si $r + s < 3$, las sumas tienen ese resto; si $r + s \geq 3$, ambas tienen el resto $r + s - 3$. Los representantes elegidos para las dos clases pueden variar, pero el resto de la suma no cambia. Esto prueba $[a+b] = [a'+b']$ y, por tanto, que la suma está bien definida.

La aparente regla $[a] \mapsto a$, en cambio, no define una función sobre el cociente: la entrada $[1] = [4]$ produciría dos salidas diferentes, 1 y 4. Una fórmula sobre clases sólo tiene sentido si todo representante de una misma clase conduce al mismo resultado.

Comprobación 4. ¿Cuántos elementos tiene $\mathbb{N}/\!\sim_5$? ¿Son iguales 2 y 12 como naturales? ¿Son iguales sus clases de restos? ¿Por qué ese cociente no es otra construcción de todo $\mathbb{N}$?

Solución razonada. El cociente tiene cinco elementos, uno por cada resto $0, 1, 2, 3, 4$. Los números 2 y 12 son diferentes, pero sus clases son iguales porque dejan el mismo resto al dividir por cinco. El cociente no es un nuevo modelo de todos los naturales: identifica números que eran distintos y sólo conserva cinco posibilidades.

La construcción ha dejado preparadas las herramientas que necesitaremos al estudiar los enteros. Sabemos que las clases pueden representar cantidades, que las operaciones sobre ellas necesitan independencia de los representantes y que el orden y la cancelación aditiva se justifican desde la aritmética natural. Si más adelante representamos un saldo mediante dos naturales, podremos utilizar esas propiedades sin suponer de antemano que la resta siempre está permitida.

12.7. Ejercicios

Las pruebas anteriores han usado emparejamientos, elecciones de representantes y comparaciones que cambian de significado si se altera una hipótesis. Los ejercicios piden recuperar esos argumentos y ensayarlos en situaciones nuevas, sin depender de la explicación inmediata. En todos ellos, salvo indicación contraria, U es el conjunto de marcas fijado en §12.2 y $\mathbb{N} = \mathcal{F}(U)/\!\approx$.

Comprender los representantes y las clases. Ejercicio 12.1. Dos conjuntos distintos, un mismo número. Sean $A = \{p, q\}$ y $B = \{r, s\}$, con cuatro marcas diferentes. Determina cuáles de las expresiones $A = B$, $A \approx B$, $[A] = [B]$, $A \in [B]$ y $[A] \in \mathbb{N}$ son verdaderas. Indica en cada caso qué tipo de objeto aparece a cada lado.

Ejercicio 12.2. Una relación que parece razonable. En $X = \mathcal{P}(\{p, q\}) \setminus \{\varnothing\}$ definamos $A \,R\, B$ si $A \cap B \neq \varnothing$. Decide si es reflexiva, simétrica y transitiva; construye un ejemplo concreto que permita justificar cada respuesta. Explica por qué no se puede formar un cociente por clases de equivalencia usando R.

Ejercicio 12.3. Particiones y cocientes. En $X = \{a, b, c, d, e\}$ consideremos la partición $\mathcal{P} = \{\{a, c\}, \{b, d, e\}\}$. Define $x \sim y$ cuando pertenecen a un mismo bloque. Encuentra $[a]$, $[c]$, $[d]$, y $X/\!\sim$. Demuestra que $[a] = [c]$ y que $[a] \cap [d] = \varnothing$, justificando el paso desde la relación hasta las clases.

Ejercicio 12.4. La proyección que olvida el representante. Para una equivalencia $\sim$ en X, definimos $q : X \to X/\sim$ mediante $q(x) = [x]$. Demuestra que q es sobreyectiva. Prueba además que q es inyectiva si y sólo si cada clase tiene un único elemento. Aplica el resultado a las colecciones finitas de marcas de U.

Ejercicio 12.5. Finitud sin números previamente construidos. Sea P una propiedad de subconjuntos de U que se cumple para $\varnothing$ y que satisface: si vale para A y $u \in U$, vale también para $A \cup \{u\}$. Demuestra, usando exclusivamente la definición de $\mathcal{F}(U)$ como menor familia cerrada, que P se cumple para todas las colecciones finitas. Explica por qué esta prueba no depende de suponer que ya existe la sucesión $0, 1, 2, \ldots$.

De las colecciones a las operaciones. Ejercicio 12.6. La misma suma, dos resultados aparentes. Calcula $A \cup B$ y $A \sqcup B$ cuando $A = B = \{p, q\}$, indicando los elementos del segundo conjunto mediante pares ordenados. Explica por qué la primera operación no puede definir $[A] + [B]$ sin restricciones sobre los representantes y por qué la segunda sí permite hacerlo.

Ejercicio 12.7. La independencia del representante en un producto. Supongamos que $f : A \to A'$ y $g : B \to B'$ son biyecciones entre colecciones finitas. Construye explícitamente una función $h : A \times B \to A' \times B'$, escribe su inversa y demuestra que es una biyección. Concluye que $[A][B]$ es independiente de los representantes escogidos.

Ejercicio 12.8. Una regla que no desciende al cociente. Fijemos dos marcas distintas $p, q \in U$. Alguien propone definir $T([A]) = 1$ si $p \in A$ y $T([A]) = 0$ en caso contrario. ¿Define esto una función $T : \mathbb{N} \to \{0, 1\}$? Justifica la respuesta con dos representantes de una misma clase. Diseña una condición sobre A que sí dependa exclusivamente de $[A]$.

Ejercicio 12.9. El sucesor no pierde información. Sean A, B colecciones finitas y $u \notin A$, $v \notin B$. Supongamos que $A \cup \{u\} \approx B \cup \{v\}$. Demuestra que $A \approx B$, explicando qué debe hacerse cuando la biyección inicial no envía u a v. Concluye que S es inyectiva.

Ejercicio 12.10. Una distributividad que puede verse. Escribe la biyección natural

$$A \times (B \sqcup C) \longrightarrow (A \times B) \sqcup (A \times C)$$

y su inversa. Comprueba por separado las entradas provenientes de B y de C, y deduce la ley distributiva $n(m + k) = nm + nk$ para las clases correspondientes.

Demostrar y reparar. Ejercicio 12.11. Cancelar sin restar. Se sabe que $[A] + [C] = [B] + [C]$, con C finita. Demuestra que $[A] = [B]$ sin utilizar restas de números naturales. Escribe la demostración por inducción sobre la adjunción de marcas a C, aclarando cómo se corrige una biyección que no empareje la marca añadida consigo misma.

Ejercicio 12.12. Comparar colecciones por inyecciones. Sea $j : A \to B$ una inyección entre colecciones finitas. Construye una colección C tal que $[B] = [A] + [C]$. Demuestra también el recíproco: si $[B] = [A] + [C]$ para alguna C, existe una inyección de A en B. ¿Por qué esta equivalencia permite definir un orden sobre las clases?

Ejercicio 12.13. Dividir sin negativos. Demuestra por inducción que, para $m \in \mathbb{N}$ con $m > 0$ y para cada $n \in \mathbb{N}$, existen $q, r \in \mathbb{N}$ con $n = mq + r$ y $0 \leq r < m$. Justifica también la unicidad del par (q, r) sin utilizar enteros negativos.

Ejercicio 12.14. Aritmética de restos. Para $m = 4$, escribe las cuatro clases de congruencia de $\mathbb{N}/\sim_4$ y calcula $[3] + [3]$ de dos maneras, usando respectivamente $(3, 3)$ y $(7, 11)$ como representantes de las entradas. Demuestra en general que la regla $[a] + [b] = [a + b]$ está bien definida en $\mathbb{N}/\sim_4$.

Ejercicio 12.15. Un argumento defectuoso. Una persona sostiene: «Si $A \approx B$, entonces $A \cup C \approx B \cup C$, porque sólo hemos añadido el mismo conjunto a ambos». Refuta la afirmación con conjuntos finitos explícitos e identifica el dato que se ha perdido. Repara el argumento usando $A \sqcup C$ y $B \sqcup C$, y explica por qué este contraste importa al construir $\mathbb{N}$ como cociente.

12.8. Soluciones desarrolladas

Soluciones 12.1–12.5. Solución 12.1. Los conjuntos A y B tienen elementos diferentes, de modo que $A = B$ es falso; ambos son conjuntos de marcas y no deben identificarse por el hecho de tener el mismo tamaño. Existe la biyección $p \mapsto r$, $q \mapsto s$, así que $A \approx B$ es verdadero. Por la definición de las clases y la propiedad de las equivalencias, $[A] = [B]$ también es verdadero. La expresión $A \in [B]$ dice que el conjunto A pertenece a la clase cuyo representante es B; equivale precisamente a $A \approx B$, por lo que es verdadera. Finalmente, $[A] \in \mathbb{N}$ es verdadero porque $[A]$ es una de las clases que forman el cociente $\mathcal{F}(U)/\!\approx$. La distinción que conviene retener es ésta: A y B son colecciones, $[A]$ es un conjunto de colecciones, y $\mathbb{N}$ es el conjunto de tales clases.

Solución 12.2. La relación es reflexiva sobre X, porque cada $A \in X$ es no vacío y satisface $A \cap A = A \neq \varnothing$. Es simétrica, pues $A \cap B = B \cap A$. No es transitiva: con $A = \{p\}$, $B = \{p, q\}$ y $C = \{q\}$ tenemos $A\,R\,B$ y $B\,R\,C$, pero $A \cap C = \varnothing$ y, por tanto, A no está relacionado con C. Si intentáramos formar $[A]_R = \{D \in X : D\,R\,A\}$, el conjunto B pertenecería tanto a $[A]_R$ como a $[C]_R$, pero ambas clases no serían iguales. No obtendríamos una partición, requisito del cociente por equivalencia que hemos definido. La falla está localizada en la transitividad, no en una elección desafortunada de símbolos.

Solución 12.3. Por definición, las clases son los bloques de la partición: $[a] = [c] = \{a, c\}$ y $[d] = \{b, d, e\}$. Así,

$$X/\!\sim \;= \{\{a, c\}, \{b, d, e\}\}.$$

Puesto que a y c pertenecen al mismo bloque, se cumple $a \sim c$, y para una relación de equivalencia eso implica $[a] = [c]$. En cambio, a y d pertenecen a bloques distintos. Estos bloques son disjuntos por hipótesis de partición, por lo que $[a] \cap [d] = \varnothing$. El cociente tiene dos elementos, y cada uno de esos elementos es a su vez un subconjunto de X.

Solución 12.4. Todo elemento K de $X/\!\sim$ es, por definición, una clase $[x]$ de algún $x \in X$. En consecuencia $q(x) = K$, lo que prueba la sobreyectividad. Supongamos que q es inyectiva y que $y \in [x]$. Entonces $x \sim y$, de donde $q(x) = q(y)$. La inyectividad obliga a $x = y$. Como x pertenece a $[x]$ por reflexividad, resulta $[x] = \{x\}$. Recíprocamente, si cada clase es unitaria y $q(x) = q(y)$, ambas entradas pertenecen a la misma clase $\{x\} = \{y\}$, luego $x = y$ y q es inyectiva. En $\mathcal{F}(U)$ no ocurre esto: las colecciones $\{p\}$ y $\{q\}$, con $p \neq q$, son representantes distintos de la misma clase; por tanto, la proyección que envía una colección finita a su número no es inyectiva.

Solución 12.5. Consideremos $\mathcal{G} = \{A \subseteq U : P(A)\}$. La hipótesis inicial asegura $\varnothing \in \mathcal{G}$. La segunda hipótesis dice que $\mathcal{G}$ permanece cerrada al agregar cualquier marca $u \in U$ a cualquiera de sus elementos. Por lo tanto, $\mathcal{G}$ es una de las familias cuya intersección define $\mathcal{F}(U)$. Como esa intersección está contenida en cada familia participante, $\mathcal{F}(U) \subseteq \mathcal{G}$, y toda colección finita cumple P. No fue necesario atribuir un número a cada colección ni invocar propiedades de suma o de sucesor numérico; bastó la caracterización de la menor familia cerrada bajo adjunción. Este argumento explica de dónde procede una forma de inducción sobre colecciones finitas.

Soluciones 12.6–12.10. Solución 12.6. La unión ordinaria es $A \cup B = \{p, q\}$. Con etiquetas,

$$A \sqcup B = \{(p, 0), (q, 0), (p, 1), (q, 1)\}.$$

La primera colección tiene dos elementos y la segunda cuatro. La suma buscada debe reunir dos cantidades de dos elementos y conservar las cuatro aportaciones. Si usáramos $[A \cup B]$ sin exigir $A \cap B = \varnothing$, el resultado variaría al cambiar de representantes: usando $B' = \{r, s\}$ con r, s ajenos a A, la unión ordinaria de A y B' tendría cuatro elementos. La unión disjunta no presenta ese defecto: una biyección entre representantes se prolonga, conservando la etiqueta,

a una biyección de las uniones disjuntas. Por eso $[A \sqcup B]$ depende únicamente de las dos clases sumadas.

Solución 12.7. Definimos $h(a,b) = (f(a), g(b))$. Cada valor pertenece a $A' \times B'$ porque $f(a) \in A'$ y $g(b) \in B'$. Las inversas f^{-1} y g^{-1} existen por la biyectividad; por ello

$$h^{-1}(a', b') = (f^{-1}(a'), g^{-1}(b')).$$

La composición de las dos reglas en cualquier orden devuelve cada par original, de modo que h es una biyección. Si $A \approx A'$ y $B \approx B'$, los productos cartesianos también son equivalentes. Al pasar a las clases, $[A \times B] = [A' \times B']$, donde cada expresión designa la clase de una copia finita dentro de U. Esto prueba la buena definición de la multiplicación.

Solución 12.8. No define una función sobre las clases. Tomemos $A = \{p\}$ y $B = \{q\}$, con $p \neq q$. Existe una biyección entre ambas colecciones, así que $[A] = [B]$. La regla daría $T([A]) = 1$ porque $p \in A$, pero $T([B]) = 0$ porque $p \notin B$, es decir, dos resultados distintos para una misma entrada. Una condición que sí depende únicamente de la clase es «A es vacía»: ninguna colección no vacía es biyectiva con $\varnothing$. La asignación $F([A]) = 0$ si $A = \varnothing$ y $F([A]) = 1$ en caso contrario está bien definida, pues la vaciedad se conserva bajo biyecciones.

Solución 12.9. Sea $f : A \cup \{u\} \to B \cup \{v\}$ una biyección. Si $f(u) = v$, la restricción de f a A es una biyección de A en B, porque el valor v tiene como preimagen únicamente a u. Si $f(u) = w \in B$, existe un único $t \in A$ con $f(t) = v$. Podemos intercambiar en el codominio w y v mediante una permutación τ, y la biyección $\tau \circ f$ envía u a v y t a w; sus demás valores permanecen bien definidos. En ambos casos, al retirar u y v queda una biyección $A \to B$, así que $[A] = [B]$. Puesto que $S([A]) = S([B])$ implica la hipótesis, concluimos que S es inyectiva. La posibilidad de intercambiar dos valores es el paso que la expresión «retirar una marca» no debe esconder.

Solución 12.10. Los elementos de $B \sqcup C$ son pares $(b, 0)$ o $(c, 1)$. Definimos

$$\Phi(a, (b, 0)) = ((a, b), 0),$$
$$\Phi(a, (c, 1)) = ((a, c), 1).$$

No hay ambigüedad porque las etiquetas 0 y 1 distinguen los dos casos. La inversa se obtiene recuperando la segunda etiqueta: $\Phi^{-1}((a,b),0) = (a,(b,0))$ y $\Phi^{-1}((a,c),1) = (a,(c,1))$. Cada composición es la identidad en su dominio, luego Φ es biyectiva. Al pasar a clases, el lado izquierdo representa $n(m+k)$ y el derecho $nm + nk$, lo que demuestra $n(m+k) = nm + nk$ sin presuponer la distributividad numérica.

Soluciones 12.11–12.15. Solución 12.11. Sea $f : A \sqcup C \to B \sqcup C$ la biyección que expresa la igualdad de clases. Inducimos sobre la construcción finita de C. Para $C = \varnothing$, la función f se reduce a una biyección entre las copias etiquetadas de A y B, con lo que $A \approx B$. En el paso, escribamos $C = C_0 \cup \{c\}$ con $c \notin C_0$, y distinguamos el elemento etiquetado $(c, 1)$ del dominio y el del codominio. Si $f(c, 1) = (c, 1)$, al suprimirlo la restricción es una biyección $A \sqcup C_0 \to B \sqcup C_0$. Si no, $f(c, 1) = y$ y existe un único elemento x con $f(x) = (c, 1)$. Componiendo con la permutación del codominio que intercambia y y $(c, 1)$ obtenemos una biyección que ya empareja ambos elementos etiquetados. Al retirarlos queda otra biyección $A \sqcup C_0 \to B \sqcup C_0$, a la que se aplica la hipótesis inductiva. Por consiguiente, $A \approx B$ y $[A] = [B]$. En ningún punto se utilizó una resta de naturales; la cancelación se obtuvo de una construcción finita de biyecciones.

Solución 12.12. La imagen $j(A)$ es un subconjunto de B en biyección con A. Definamos $C = B \setminus j(A)$, que es finito por ser subconjunto de una colección finita. Las partes $j(A)$ y C son disjuntas y su unión es B; de aquí $[B] = [j(A)] + [C] = [A] + [C]$. Recíprocamente, si $[B] = [A] + [C]$, existe una biyección $f : A \sqcup C \to B$. La aplicación $a \mapsto f(a, 0)$ es una inyección de A en B, porque es la restricción de una biyección a la copia etiquetada de A. La relación depende únicamente de las clases: cualquier cambio de representantes se transporta mediante

biyecciones antes y después de la inyección. La reflexividad se obtiene de la identidad; la transitividad, de la composición de inyecciones; la antisimetría y totalidad necesitan además la finitud y se demuestran en §12.5. Esa precisión impide tomar la definición por una prueba automática de todas las propiedades de orden.

Solución 12.13. Para $n = 0$, tomamos $q = r = 0$; la desigualdad $0 < m$ se cumple por hipótesis. Supongamos que $n = mq + r$ con $0 \le r < m$. Al sumar uno tenemos $n + 1 = mq + (r + 1)$. Como $r < m$, se cumple $r + 1 \le m$. Si $r + 1 < m$, conservamos q y usamos el resto $r + 1$. Si $r + 1 = m$, escribimos $n + 1 = m(q + 1) + 0$. En ambos casos existe la representación requerida. Para la unicidad, supongamos $mq + r = mq' + r'$ con $0 \le r, r' < m$. Si $q < q'$, entonces $q + 1 \le q'$ y, como $m > 0$, se deduce

$$mq + r < m(q + 1) \le mq' \le mq' + r',$$

contradicción con la igualdad. El caso $q' < q$ conduce a la contradicción simétrica. Así $q = q'$ y la cancelación aditiva en $mq + r = mq + r'$ da $r = r'$. Hemos usado orden, inducción y cancelación de naturales, ya disponibles, pero ninguna resta negativa.

Solución 12.14. Las cuatro clases son $C_0 = \{0, 4, 8, \ldots\}$, $C_1 = \{1, 5, 9, \ldots\}$, $C_2 = \{2, 6, 10, \ldots\}$ y $C_3 = \{3, 7, 11, \ldots\}$. Los representantes 3 y 7 corresponden a la primera clase de entrada, y 3 y 11 a la segunda, pues dejan todos resto 3. Con los primeros obtenemos $[3] + [3] = [6] = C_2$; con los otros, $[7] + [11] = [18] = C_2$. Para justificar la independencia en general, escribamos $a = 4p + r$, $a' = 4p' + r$ y $b = 4q + s$, $b' = 4q' + s$ con los mismos restos respectivos $r, s \in \{0, 1, 2, 3\}$. Entonces

$$a + b = 4(p + q) + (r + s), \qquad a' + b' = 4(p' + q') + (r + s).$$

El resto de ambas sumas coincide, ya sea $r + s$ si es menor que 4 o $r + s - 4$ si está entre 4 y 6. Por ello $a + b \sim_4 a' + b'$ y la regla $[a] + [b] = [a + b]$ está bien definida.

Solución 12.15. Tomemos tres marcas distintas p, q, r y elijamos $A = \{p\}$, $B = \{q\}$ y $C = \{p\}$. Tenemos $A \approx B$ mediante $p \mapsto q$. Sin embargo, $A \cup C = \{p\}$, mientras que $B \cup C = \{p, q\}$; una colección tiene una marca y la otra dos, de modo que no admiten una biyección. La unión ordinaria ha olvidado que una aparición de p procedía de A y otra de C. Para repararlo, definamos $h : A \sqcup C \to B \sqcup C$ enviando $(a, 0)$ a $(f(a), 0)$, donde $f : A \to B$ es la biyección dada, y dejando fijas las entradas $(c, 1)$. La regla es una biyección, con inversa que utiliza f^{-1} en la copia etiquetada 0 y la identidad en la copia 1. Por tanto $A \sqcup C \approx B \sqcup C$. Ésta es la propiedad que permite sumar clases sin que una coincidencia de marcas cambie el resultado.

Capítulo 13 — ¿De dónde salen los enteros?

13.1. Segundo momento de construcción: los enteros a partir de los naturales

El capítulo anterior nos permitió construir cantidades naturales y justificar sus operaciones a partir de colecciones. También mostró una limitación que el uso corriente de los números puede hacernos olvidar: $7 - 4$ se interpreta dentro de $\mathbb{N}$ como el natural que, sumado a 4, produce 7, mientras que no existe ningún natural que, sumado a 7, produzca 4. Estamos acostumbrados a escribir $4 - 7 = -3$; falta construir el objeto que pueda recibir ese nombre, utilizando sólo las propiedades de $\mathbb{N}$ que ya demostramos.

Un camino consiste en registrar por separado dos cantidades naturales. Pensemos en una cuenta donde anotamos unidades a favor y unidades en contra. El par $(4, 7)$ expresa cuatro unidades a favor y siete en contra; el par $(2, 5)$ describe una cuenta diferente, pero el saldo que queremos conservar es el mismo. Lo mismo sucede con $(0, 3)$ y $(10, 13)$. Podemos agregar la misma cantidad a ambas columnas o retirarla de ambas, siempre que sea posible hacerlo, y la comparación que nos interesa no cambiará. En cambio, si cambiamos una sola columna, alteramos el saldo. Las distintas cuentas sugieren un saldo común, aunque todavía no contamos con un número de nuestro sistema que lo exprese.

Cada par es una representación posible; ninguno tiene por qué convertirse en la representación oficial. Elegir $(0, 3)$ como la forma más breve puede ayudar a calcular, pero la construcción no debe depender de esa preferencia. Buscamos una relación que diga cuándo **dos pares describen el mismo saldo**, utilizando exclusivamente la suma natural que construimos en el capítulo anterior.

Sería tentador declarar equivalentes (a, b) y (c, d) cuando $a - b = c - d$. El problema es que, si $b > a$ o $d > c$, las restas todavía no están definidas dentro de los naturales. Podemos expresar la misma comparación sin efectuar ninguna de ellas: trasladamos las cantidades a ambos lados de una igualdad y exigimos

$$\boxed{(a, b) \sim (c, d) \quad \Longleftrightarrow \quad a + d = b + c.}$$

La condición sólo utiliza sumas de naturales. Con ella, $(4, 7) \sim (2, 5)$ porque $4 + 5 = 7 + 2$, y $(4, 7) \sim (0, 3)$ porque $4 + 3 = 7 + 0$. En cambio, $(4, 7) \not\sim (4, 6)$: al comparar las sumas cruzadas aparecen $4 + 6$ y $7 + 4$, que son diferentes. Las cuentas pueden parecer parecidas sin representar el mismo saldo.

También resulta esclarecedor escribir la relación como igualdad de dos sumas completas: $(a, b) \sim (c, d)$ significa que la cantidad a favor de la primera cuenta, junto con la cantidad en contra de la segunda, equilibra exactamente la cantidad en contra de la primera junto con la cantidad a favor de la segunda. El símbolo $\sim$ compara pares; todavía no hemos definido una resta para expresarlos mediante un solo número.

Antes de formar clases, la nueva relación debe superar las tres pruebas que ya conocemos. Es **reflexiva**, pues $(a, b) \sim (a, b)$ equivale a $a + b = b + a$. Es **simétrica** porque la igualdad $a + d = b + c$ sigue siendo válida si intercambiamos los miembros. Para la **transitividad** hace falta enlazar dos comparaciones. Supongamos que $(a, b) \sim (c, d)$ y $(c, d) \sim (e, f)$; entonces

$$a + d = b + c, \qquad c + f = d + e.$$

Si sumamos estas igualdades, reordenamos sus términos y agrupamos la cantidad que aparece en ambos lados, obtenemos

$$a + f + (c + d) = b + e + (c + d).$$

La suma de los naturales, cuya construcción por uniones disjuntas estudiamos anteriormente, tiene la propiedad de cancelación: de $u + t = v + t$ se deduce $u = v$. Aplicándola con $t = c + d$ llegamos a $a + f = b + e$, que es justamente $(a, b) \sim (e, f)$. Hemos demostrado la transitividad valiéndonos únicamente de propiedades que los naturales ya poseen. Esto importa: la construcción de un nuevo conjunto numérico no debe apoyarse en operaciones que todavía intenta justificar.

Podemos formar, por consiguiente, las clases de equivalencia de los pares. Éste es el **segundo momento de construcción numérica**:

$$\boxed{\mathbb{Z} = (\mathbb{N} \times \mathbb{N})/\!\sim.}$$

En esta construcción, un **entero** es una clase de pares de naturales. A diferencia de las clases por restos, que agrupaban naturales ya existentes, estas clases proporcionan objetos nuevos con los que podremos efectuar restas antes imposibles. Llamamos -3 a la clase que contiene $(0, 3)$; también contiene $(4, 7)$, $(2, 5)$ y los demás pares equivalentes. La clase de $(3, 0)$ recibe el nombre de 3, y contiene, entre otros, $(4, 1)$ y $(5, 2)$. El cero es la clase de $(0, 0)$, que incluye todos los pares (n, n). Los nombres familiares adquieren así un significado preciso dentro del conjunto recién definido.

Los naturales reaparecen en el nuevo sistema mediante la función

$$\iota : \mathbb{N} \longrightarrow \mathbb{Z}, \qquad \iota(n) = [(n, 0)].$$

Esta función es inyectiva. Si $[(n, 0)] = [(m, 0)]$, los representantes son equivalentes, y la igualdad que los relaciona es $n + 0 = 0 + m$, de donde $n = m$. Podemos entonces identificar el natural n con el entero $[(n, 0)]$ sin confundir los dos objetos en la construcción estricta. Todo entero, además, puede nombrarse mediante un par de la forma $(n, 0)$ o $(0, n)$. En efecto, dados $a, b \in \mathbb{N}$, o bien $a = b + n$ para algún natural n, en cuyo caso $(a, b) \sim (n, 0)$, o bien $b = a + n$ para algún natural n, y entonces $(a, b) \sim (0, n)$. Si ambas cantidades son iguales, las dos formas representan el cero. La costumbre de escribir $\mathbb{N} \subseteq \mathbb{Z}$ se justifica por esta identificación, aun cuando los naturales originales y las clases sean formalmente objetos de distinta naturaleza.

La suma de enteros. La imagen de las cuentas nos ofrece una regla para sumar dos saldos: reunimos las cantidades a favor y, por separado, las cantidades en contra. Si las clases están representadas por (a, b) y (c, d), la operación propuesta es

$$\boxed{[(a, b)] + [(c, d)] = [(a + c, b + d)].}$$

Con las clases por restos aprendimos que una fórmula puede depender indebidamente del representante; con los pares que definen enteros debemos verificar de nuevo esa independencia. Supongamos que $(a, b) \sim (a', b')$ y $(c, d) \sim (c', d')$. Tenemos, pues, $a+b' = b+a'$ y $c+d' = d+c'$. Sumando las dos igualdades y ordenando sus términos obtenemos

$$(a + c) + (b' + d') = (b + d) + (a' + c').$$

Ésta es la condición de equivalencia entre $(a + c, b + d)$ y $(a' + c', b' + d')$. Hemos probado que la clase de la suma sólo depende de las clases sumadas. Podemos, por ejemplo, calcular

$$[(4, 7)] + [(6, 2)] = [(10, 9)] = [(1, 0)],$$

que, con los nombres habituales, se lee $(-3)+4=1$. La clase $[(0,0)]$ actúa como cero de la suma. Y si intercambiamos las coordenadas de (a,b) obtenemos (b,a), cuya clase es el opuesto aditivo, porque

$$[(a,b)]+[(b,a)]=[(a+b,a+b)]=[(0,0)].$$

Tenemos así una razón interna para la existencia de los negativos: cada entero posee otro que, sumado al primero, da cero. Una resta que carecía de respuesta en $\mathbb{N}$ puede realizarse en $\mathbb{Z}$ sumando el opuesto correspondiente.

La multiplicación y el orden. La multiplicación requiere una regla algo menos evidente. Ya hemos construido la suma y los opuestos; además, queremos que los naturales insertados mediante $n\mapsto[(n,0)]$ continúen multiplicándose como antes y que la distributividad no se pierda. Como orientación podemos efectuar un cálculo **provisional**, usando la familiar escritura de diferencias: $(a-b)(c-d)=ac+bd-ad-bc$. Esta igualdad sirve aquí para *descubrir* qué par debería representar el producto; su validez dentro de nuestro nuevo sistema todavía tendrá que demostrarse. Nos sugiere registrar $ac+bd$ como primera coordenada y $ad+bc$ como segunda. La fórmula que queremos definir emplea solamente operaciones ya disponibles en $\mathbb{N}$:

$$\boxed{[(a,b)]\cdot[(c,d)]=[(ac+bd,ad+bc)].}$$

La interpretación ha servido para encontrar la fórmula, pero no basta para probar que ésta tiene sentido sobre clases. Fijemos, primero, el par (c,d) y supongamos $(a,b)\sim(a',b')$. Por definición $a+b'=b+a'$. Multiplicando esta igualdad por c y por d obtenemos

$$ac+b'c=bc+a'c,\qquad ad+b'd=bd+a'd.$$

Sumemos la primera igualdad con la segunda leída en el sentido opuesto. Después de reordenar los sumandos aparece

$$(ac+bd)+(a'd+b'c)=(ad+bc)+(a'c+b'd).$$

Es exactamente la condición que afirma que $(ac+bd,ad+bc)$ y $(a'c+b'd,a'd+b'c)$ pertenecen a la misma clase. Por tanto, cambiar el primer representante no altera el producto. La expresión $ac+bd$ es simétrica respecto de los dos pares y lo mismo sucede con el par de términos $ad+bc$ tras intercambiarlos; por esa conmutatividad de la fórmula, el argumento vale también cuando cambiamos únicamente el segundo representante. Si cambian los dos, podemos sustituirlos uno después del otro. El producto está bien definido.

También queremos comparar enteros. La clase $[(a,b)]$ debe quedar por debajo de $[(c,d)]$ cuando el primer saldo sea menor o igual que el segundo. La comparación se puede formular de nuevo sin restar:

$$\boxed{[(a,b)]\le[(c,d)]\quad\Longleftrightarrow\quad a+d\le b+c.}$$

Podemos probar primero un caso: $[(4,7)]\le[(6,2)]$ porque $4+2\le7+6$. Si reemplazamos el primer representante por $(0,3)$, la desigualdad correspondiente sigue siendo verdadera, $0+2\le3+6$. No basta ese ejemplo: debe mantenerse para cualquier cambio legítimo de representantes.

Comprobemos que los representantes no afectan a esta desigualdad. Si $(a,b)\sim(a',b')$ y $(c,d)\sim(c',d')$, las igualdades $a+b'=b+a'$ y $c+d'=d+c'$ permiten deducir

$$(a+d)+(b'+c')=(a'+d')+(b+c).$$

Supongamos $a+d \leq b+c$. Al sumar $b'+c'$ obtenemos $(a+d)+(b'+c') \leq (b+c)+(b'+c')$; sustituimos el primer miembro por la expresión igual que acabamos de obtener y cancelamos el sumando común $b+c$. Resulta $a'+d' \leq b'+c'$. Intercambiar los pares primados con los no primados demuestra la implicación contraria. Esta comparación, igual que las dos operaciones, es independiente de los representantes.

Las fórmulas construidas permiten trabajar con las clases como enteros. El examen de la asociatividad, la distributividad, la compatibilidad del orden y las demás leyes llegará inmediatamente después: probar que las expresiones están bien definidas era una obligación previa, y por sí sola no demuestra todas esas propiedades.

Comprobación 5. Decide si $(8,11)$ y $(5,8)$ representan el mismo entero. Explica por qué el entero no es ninguno de esos dos pares tomado aisladamente y encuentra el opuesto aditivo de su clase.

Solución razonada. Las sumas cruzadas coinciden: $8+8=11+5$. Los pares son distintos, pero pertenecen a una única clase, que llamamos -3. Su opuesto es la clase de $(11,8)$, obtenida al intercambiar las coordenadas, y coincide con $[(3,0)]$, que llamamos 3. El entero es la clase completa, no el par elegido para representarla.

13.2. Los enteros heredan unas leyes y adquieren otras

Al construir $\mathbb{Z}$, una clase $[(a,b)]$ representa un saldo; la relación $(a,b) \sim (c,d)$ equivale a $a+d=b+c$. Acabamos de demostrar que la suma y el producto de estas clases están bien definidos. Es legítimo utilizar ese resultado, pero no podemos declarar concluidas las leyes de las operaciones sin estudiar qué sucede al sumar o multiplicar **tres** clases, o al combinar ambas operaciones.

El grupo aditivo y la multiplicación. Para la suma usamos

$$[(a,b)] + [(c,d)] = [(a+c, b+d)].$$

Cambiar el orden de los pares no altera el resultado, porque la suma de naturales es conmutativa; reagrupar tres pares tampoco lo altera, porque esa suma es asociativa. La clase $[(0,0)]$ es neutra y cada clase $[(a,b)]$ tiene por opuesta la clase $[(b,a)]$:

$$[(a,b)] + [(b,a)] = [(a+b, a+b)] = [(0,0)].$$

Esto sí representa una novedad frente a $\mathbb{N}$: toda clase posee una opuesta aditiva. De aquí se deduce que $\mathbb{Z}$, con la suma, es un **grupo abeliano**. La palabra «grupo» expresa conjuntamente los resultados que acabamos de establecer y, sobre todo, la posibilidad de resolver cualquier ecuación $x+z=w$ mediante $x=w+(-z)$.

Para multiplicar recordemos la definición

$$[(a,b)]\,[(c,d)] = [(ac+bd, ad+bc)].$$

La conmutatividad se obtiene intercambiando los dos pares y utilizando la conmutatividad del producto natural. La clase $[(1,0)]$ es unidad, pues el producto con ella produce $[(a,b)]$. Para la asociatividad tomemos, además, (e,f). Si multiplicamos primero los dos primeros pares, el representante final tiene por coordenadas

$$ace+bde+adf+bcf,$$
$$acf+bdf+ade+bce.$$

Si multiplicamos primero el segundo y el tercero, obtenemos las mismas expresiones, quizá con los sumandos en otro orden. La igualdad es una consecuencia de la asociatividad, conmutatividad y distributividad de las operaciones naturales demostradas antes. Así, el producto de enteros es asociativo.

Queda la distributividad. Al multiplicar (a,b) por la suma $(c,d)+(e,f)=(c+e,d+f)$, las dos coordenadas son

$$a(c+e)+b(d+f)=ac+ae+bd+bf,$$
$$a(d+f)+b(c+e)=ad+af+bc+be.$$

Es exactamente el mismo par que obtenemos al sumar $(ac+bd,ad+bc)$ y $(ae+bf,af+be)$, representantes de los productos parciales. La distributividad por el otro lado se sigue de la conmutatividad de la multiplicación. Con estas pruebas, $\mathbb{Z}$ es un **anillo conmutativo con unidad**. Hemos construido el sistema en el que las restas ya no tropiezan con la falta de un opuesto.

Ceros, signos y ausencia de divisores de cero. Un entero puede representarse como $[(n,0)]$ o como $[(0,n)]$, con $n\in\mathbb{N}$: en un par (a,b), la totalidad del orden natural nos permite escribir $a=b+n$ o $b=a+n$, y cancelar los términos comunes en la relación de equivalencia. El primer tipo de representante corresponde a los enteros no negativos y el segundo a los no positivos. Cuando $n=0$, ambas formas designan el mismo cero; si $n>0$, designan dos enteros distintos y opuestos.

Esta forma de escoger representantes revela la regla de los signos sin imponerla desde fuera. Para $m,n>0$,

$$[(m,0)]\,[(n,0)]=[(mn,0)],$$
$$[(0,m)]\,[(n,0)]=[(0,mn)],$$
$$[(0,m)]\,[(0,n)]=[(mn,0)].$$

El producto mn es distinto de cero: si M y N son colecciones no vacías que representan m y n, el producto cartesiano $M\times N$ contiene al menos un par. En consecuencia, un producto de dos enteros no nulos tampoco puede ser cero. La fórmula de los signos y el hecho de que $\mathbb{Z}$ **no tiene divisores de cero** se desprenden de la construcción, aunque los enteros no nulos no tengan siempre inversos multiplicativos dentro de $\mathbb{Z}$.

Una aritmética de fortunas y deudas en la historia. La comparación de saldos que utilizamos para motivar los enteros tiene una historia matemática sugerente. En el *Brāhmasphuṭasiddhānta* (628), Brahmagupta enunció reglas para cantidades positivas, negativas y cero mediante el lenguaje de **fortunas y deudas**. Entre sus reglas figura la de que el producto de dos deudas es una fortuna; también distingue qué ocurre al sumar o sustraer cantidades de signos diferentes. No es difícil imaginar la utilidad de esas expresiones al hacer cuentas: una deuda de cinco unidades y una fortuna de tres dejan un saldo que debe poder tratarse coherentemente en operaciones posteriores.

Esta observación histórica ilumina el *problema* que hemos estudiado, aunque hay que distinguirla de nuestra solución. La construcción de $\mathbb{Z}$ mediante clases de pares de naturales es una **fundamentación moderna** presentada aquí con el lenguaje de la teoría de conjuntos; no afirmamos que Brahmagupta empleara relaciones de equivalencia ni que su obra contuviera esta construcción. Sus reglas muestran, en cambio, que trabajar rigurosamente con cantidades negativas fue una necesidad matemática concreta muchos siglos antes de que se adoptaran las formalizaciones actuales.

Fuentes de contraste histórico: MacTutor, «Brahmagupta» y NRICH, «The History of Negative Numbers». Las reglas se mencionan mediante paráfrasis; no se reproducen citas literales de una traducción del sánscrito.

Comprobación 6. Justifica, sin dar por supuesta una tabla de signos, que el producto del entero representado por $(0,4)$ y el representado por $(0,3)$ es el entero representado por $(12,0)$. Explica después por qué la prueba no depende de haber elegido esos pares particulares.

Solución razonada. La fórmula del producto da $(0,4)(0,3) = (12,0)$, porque $0{\cdot}0+4{\cdot}3 = 12$ y $0\cdot 3 + 4\cdot 0 = 0$. El resultado se lee $(-4)(-3) = 12$. La independencia de los representantes se demostró al definir el producto en la sección anterior: si sustituimos $(0,4)$ y $(0,3)$ por pares equivalentes, el producto pertenece a la misma clase que $[(12,0)]$.

13.3. El orden de los enteros y el sentido de la ampliación

Que los enteros puedan restarse no impide compararlos. Para clases $x = [(a,b)]$ e $y = [(c,d)]$ propusimos la condición

$$[(a,b)] \le [(c,d)] \quad \Longleftrightarrow \quad a + d \le b + c.$$

Ya probamos que el resultado no depende de los representantes; debemos examinar si posee las propiedades de un orden y si respeta las operaciones. La reflexividad es la desigualdad $a + b \le b + a$, que en realidad es una igualdad. Si $x \le y$ e $y \le x$, obtenemos $a + d \le b + c$ y $b + c \le a + d$. Por antisimetría del orden natural, $a + d = b + c$ y, por definición de nuestras clases, $x = y$.

Para la transitividad tomemos también $z = [(e,f)]$ y supongamos $x \le y$ e $y \le z$, es decir,

$$a + d \le b + c, \qquad c + f \le d + e.$$

Al sumar las desigualdades resulta $a + f + (c + d) \le b + e + (c + d)$. La cancelación del término común $c + d$ en el **orden natural** nos da $a + f \le b + e$, de donde $x \le z$. Esa cancelación de desigualdades se deduce de la caracterización $p \le q \iff q = p + t$ y de la cancelación aditiva ya demostrada. Finalmente, el orden es total: dados $a + d$ y $b + c$, el orden natural decide cuál de las dos sumas es menor o igual, y con ello decide $x \le y$ o $y \le x$.

La suma conserva las comparaciones. Si añadimos $w = [(e,f)]$ a ambos miembros, la condición para $x + w \le y + w$ es

$$(a + e) + (d + f) \le (b + f) + (c + e),$$

que equivale a $a + d + e + f \le b + c + e + f$ y, tras cancelar, a la condición original $x \le y$. También se conserva la desigualdad al multiplicar por un entero no negativo: si $x \le y$, la traslación por $-x$ da $0 \le y - x$. Todo entero no negativo se representa como $[(n,0)]$, de modo que tanto $y - x$ como un multiplicador $w \ge 0$ admiten representantes de ese tipo. Su producto es entonces de la forma $[(mn,0)]$ y es no negativo. La distributividad proporciona $yw - xw = (y - x)w \ge 0$, es decir, $xw \le yw$.

Hemos obtenido así un **anillo conmutativo totalmente ordenado**, sin divisores de cero, en el que el orden es compatible con la suma y con la multiplicación por elementos no negativos. La inclusión usual de los naturales se entiende mediante la función $\iota(n) = [(n,0)]$, que conserva suma, producto y orden y es inyectiva. Gracias a ella utilizamos el mismo nombre n cuando consideramos un natural como entero. No afirmamos que las clases de colecciones finitas y las clases de pares sean literalmente el mismo conjunto: hemos demostrado por qué podemos tratarlas como representaciones compatibles de la misma aritmética natural dentro de una estructura más amplia.

Hay una propiedad adicional que ilumina el sentido de esta ampliación. Supongamos que G es un grupo abeliano y que $h : \mathbb{N} \to G$ conserva la suma y envía 0 al elemento neutro de G. Toda función aditiva $\widetilde{h} : \mathbb{Z} \to G$ que prolongue a h **tiene que** satisfacer

$$\widetilde{h}([(a,b)]) = h(a) - h(b),$$

porque $[(a,b)] = [(a,0)] - [(b,0)]$. Esta fórmula no sólo es obligatoria: también define efectivamente una función sobre las clases. Si $(a,b) \sim (c,d)$, entonces $a + d = b + c$ en $\mathbb{N}$; al aplicar h obtenemos $h(a) + h(d) = h(b) + h(c)$ y, utilizando los opuestos disponibles en G,

concluimos $h(a) - h(b) = h(c) - h(d)$. La fórmula respeta la suma por la conmutatividad de G y prolonga a h, puesto que $\widetilde{h}([(n,0)]) = h(n)$. Por tanto, esa prolongación existe y es única.

Esta **propiedad universal** da un alcance nuevo al procedimiento de los pares: cualquier interpretación aditiva de los naturales en un grupo abeliano se prolonga de una única manera a los enteros. La fórmula $\widetilde{h}([(a,b)]) = h(a) - h(b)$ expresa esa compatibilidad; su independencia de los representantes muestra por qué la construcción funciona en cualquier grupo abeliano de llegada, y no sólo en nuestras cuentas de saldos.

Comprobación 7. Sea $h : \mathbb{N} \to \mathbb{Z}$ la aplicación $h(n) = 2n$. Utiliza la propiedad anterior para determinar el valor que debe tener su prolongación aditiva en la clase $[(3,8)]$. Explica por qué el cálculo no depende del representante.

Solución razonada. La fórmula obliga a asignar $h(3) - h(8) = 6 - 16 = -10$. La clase $[(3,8)]$ es la misma que $[(0,5)]$, y con este representante la fórmula da $h(0) - h(5) = 0 - 10 = -10$. La coincidencia de este ejemplo está garantizada en general porque la equivalencia de pares $a + d = b + c$ se transforma, al aplicar h, en una igualdad que permite cancelar en el grupo abeliano de llegada.

La construcción conserva la suma, el producto y el orden natural, y permite resolver toda ecuación aditiva $x + z = w$. Aún no garantiza resolver una ecuación como $2x = 1$: ésa será la dificultad que nos lleve a los racionales. Mientras tanto, los enteros recién construidos ofrecen otro uso de la equivalencia, que ya encontramos entre los naturales: agrupar números por el resto de una división, conservando sólo la información pertinente a ese cálculo.

13.4. Las congruencias en los enteros: construir un cociente no es ampliar el sistema

Una vez construidos los enteros, podemos plantear la misma pregunta que hicimos con los naturales: ¿cuándo dejan dos números un mismo resto al dividirlos por tres, cuatro o cualquier natural positivo? La división euclídea sigue funcionando. Ya la probamos para naturales y podemos extenderla a cualquier entero: si $z = -n$ con $n > 0$ y $n = mq + r$, $0 \le r < m$, entonces $z = m(-q)$ cuando $r = 0$, mientras que si $r > 0$ podemos escribir $z = m(-q-1) + (m - r)$ y el nuevo resto vuelve a estar entre 0 y $m - 1$. La unicidad se prueba como antes: dos cocientes enteros diferentes, ordenados uno antes que otro, darían valores separados al menos por m, imposibles de compensar con restos menores que m. Por ejemplo, al dividir -2 por 3, escribimos

$$-2 = 3(-1) + 1.$$

Su resto es 1, exactamente el mismo que el de 1 y el de 4. El cociente de la división puede ser negativo, pero elegimos el resto entre 0 y $m - 1$ cuando el divisor m es positivo. Para todo $z \in \mathbb{Z}$ existen un único cociente $q \in \mathbb{Z}$ y un único resto $r \in \{0, 1, \dots, m-1\}$ tales que $z = mq + r$.

Con los enteros podemos expresar la comparación de restos mediante una divisibilidad. Diremos que x e y son **congruentes módulo** m, y escribiremos

$$\boxed{x \equiv y \pmod{m} \quad \Longleftrightarrow \quad m \mid (x - y),}$$

siempre con $m \in \mathbb{N}$ y $m > 0$. La condición de la derecha significa que $x - y = mk$ para algún entero k. Para ver que coincide con la definición por restos, escribamos $x = mq + r$ e $y = mp + s$, con $0 \le r, s < m$. Si $m \mid (x-y)$, también $m \mid (r-s)$, pues $r - s = (x-y) - m(q-p)$. Como $r - s$ está estrictamente entre $-m$ y m, el único múltiplo de m que puede ser es cero: $r = s$. A la inversa, si $r = s$, entonces $x - y = m(q - p)$, y m divide la diferencia. Las dos formulaciones expresan, pues, la misma relación.

También aquí la relación es de equivalencia. Todo x es congruente consigo mismo porque $m \mid 0$; de $m \mid (x-y)$ se sigue $m \mid (y-x)$ por cambio de signo; y de $m \mid (x-y)$ junto con $m \mid (y-z)$ se deduce $m \mid (x-z)$ sumando las dos diferencias. Por ejemplo,

$$[-2] = [1] = [4] \qquad \text{en } \mathbb{Z}/\equiv_3.$$

El cociente $\mathbb{Z}/\equiv_3$ tiene otra vez tres clases, pero sus elementos son conjuntos de enteros y, por ello, incluyen representantes negativos. La clase de 1 en $\mathbb{N}/\sim_3$ y la de 1 en $\mathbb{Z}/\equiv_3$ no son literalmente el mismo conjunto. Existe, sin embargo, una correspondencia natural entre ellas: a cada clase de naturales le asignamos la clase de enteros con ese mismo resto. La correspondencia es biyectiva, pues cada clase entera posee uno y sólo uno de los representantes 0, 1 o 2. De este modo podemos comparar ambos cocientes sin confundir sus elementos como conjuntos.

Esta última comparación permite advertir algo que la notación puede esconder. Para construir $\mathbb{Z}$ tomamos **pares** de naturales y los reunimos en clases para obtener números de un sistema nuevo; en $\mathbb{Z}/\equiv_m$ tomamos **enteros ya construidos** y los reunimos por una propiedad aritmética, el resto. La definición abstracta de equivalencia es idéntica en ambos casos, mientras que los representantes y el propósito matemático de la identificación cambian. En las clases de saldos buscamos hacer posible cualquier resta; en las de congruencia aprendemos a calcular conservando sólo una información de cada número. Una clase es un objeto matemático genuino en ambos casos, y una regla sobre ella exige la misma prueba de independencia del representante.

Comprobación 8. Determina el resto de -7 y el de 5 al dividir por cuatro en $\mathbb{Z}$, y decide si sus clases son iguales.

Solución razonada. Las divisiones euclídeas se escriben $-7 = 4(-2)+1$ y $5 = 4(1)+1$. Ambos enteros dejan resto 1, por lo que $-7 \equiv 5 \pmod 4$ y $[-7] = [5]$ en $\mathbb{Z}/\equiv_4$. Esta igualdad entre clases no dice que los enteros -7 y 5 sean iguales.

13.5. Ejercicios

En estos ejercicios una misma clase deberá reconocerse a partir de pares diferentes, y una regla sobre representantes tendrá que superar las pruebas de buena definición, compatibilidad y orden que acabamos de utilizar. Suponemos $\mathbb{N}$ con cero, $(a,b) \sim (c,d) \iff a+d = b+c$ en $\mathbb{N}\times\mathbb{N}$, y escribimos $[(a,b)]$ para la clase correspondiente en $\mathbb{Z}$. Cuando los corchetes designen una clase de congruencia, indicaremos expresamente el módulo.

Comprender las clases y construir el nuevo sistema. Ejercicio 13.1. Cinco representaciones de una misma deuda. Determina si los pares $(0,4)$, $(2,6)$, $(5,9)$, $(3,7)$ y $(2,5)$ pertenecen a una misma clase. Justifica la respuesta utilizando sólo sumas naturales y explica qué entero representa la clase mayoritaria.

Ejercicio 13.2. Un error de clasificación. Una persona define $(a,b)\,R\,(c,d)$ si $a = c$, porque considera que «la primera cantidad» debe determinar el saldo. ¿Es R una relación de equivalencia sobre $\mathbb{N}^2$? ¿Por qué, pese a ello, su conjunto cociente no construye el sistema de saldos buscado? Da pares concretos que R identifique indebidamente y pares que R separe aunque tengan el mismo saldo.

Ejercicio 13.3. Lo que exige la transitividad. Sean $(a,b),(c,d),(e,f) \in \mathbb{N}^2$ tales que $(a,b)\sim(c,d)$ y $(c,d)\sim(e,f)$. Reconstruye la prueba completa de $(a,b)\sim(e,f)$ e indica el único paso que utiliza cancelación en $\mathbb{N}$.

Ejercicio 13.4. Naturales que se conservan. Prueba que $\iota(n) = [(n,0)]$ es inyectiva y verifica directamente que $\iota(n+m) = \iota(n)+\iota(m)$ y $\iota(nm) = \iota(n)\iota(m)$. Explica por qué estas identidades justifican tratar $\mathbb{N}$ como una parte de $\mathbb{Z}$ sin afirmar una igualdad literal entre las realizaciones conjuntistas.

Ejercicio 13.5. Representantes canónicos y cero. Demuestra que toda clase $[(a,b)]$ contiene algún par $(n,0)$ o algún par $(0,n)$ y que ambas presentaciones sólo pueden ocurrir simultáneamente para la clase cero. ¿Cuál es el criterio para decidir si $[(a,b)]$ es cero utilizando únicamente naturales?

Operaciones y orden: probar antes de utilizar. Ejercicio 13.6. Sumar de dos maneras. Calcula $[(4,9)]+[(7,2)]$ y después vuelve a calcular la suma utilizando los representantes $(1,6)$ y $(10,5)$. ¿Qué igualdad garantiza que ambos cálculos representan el mismo entero? Explica por qué dos ejemplos no sustituyen la prueba general de buena definición.

Ejercicio 13.7. Encontrar el opuesto. Define $-[(a,b)]=[(b,a)]$. Comprueba que la regla está bien definida: si $(a,b)\sim(c,d)$, demuestra que $(b,a)\sim(d,c)$. Luego prueba que $[(a,b)]+(-[(a,b)])=0$.

Ejercicio 13.8. El producto sin restas ocultas. Dados $x=[(a,b)]$ e $y=[(c,d)]$, demuestra que el producto $[(ac+bd,ad+bc)]$ es independiente de ambos representantes. Puedes probar primero la independencia en el primer factor, luego en el segundo y finalmente combinar los dos cambios. Indica por qué no basta escribir $(a-b)(c-d)$.

Ejercicio 13.9. Asociar tres productos. Sean $x=[(a,b)]$, $y=[(c,d)]$ y $z=[(e,f)]$. Calcula un representante de $(xy)z$ y otro de $x(yz)$, organiza por separado ambas coordenadas y comprueba que coinciden mediante las leyes ya demostradas en $\mathbb{N}$.

Ejercicio 13.10. La regla de los signos no es un axioma nuevo. Para naturales positivos m,n, utiliza la fórmula del producto de pares para demostrar $(-m)(-n)=mn$ y $(-m)n=-(mn)$. Aclara por qué la positividad de mn permite concluir que un producto de dos enteros no nulos no puede ser cero.

Ejercicio 13.11. Comparar representaciones. Decide si $[(4,9)]<[(2,5)]$, si $[(4,9)]=[(2,5)]$ y si $[(4,9)]>[(2,5)]$. Repite la comparación reemplazando $(4,9)$ por $(1,6)$ y explica por qué el resultado no puede cambiar.

Ejercicio 13.12. Trasladar una desigualdad. Partiendo exclusivamente de la definición $[(a,b)]\le[(c,d)]\iff a+d\le b+c$, prueba que para todo $w\in\mathbb{Z}$ se cumple $x\le y\iff x+w\le y+w$. Identifica la propiedad del orden natural que permite eliminar los sumandos comunes.

Generalizar, diagnosticar y regresar a los restos. Ejercicio 13.13. Una prolongación obligatoria. Sea G un grupo abeliano, y sea $h:\mathbb{N}\to G$ aditiva, con $h(0)=0_G$. Demuestra que $H([(a,b)])=h(a)-h(b)$ está bien definida y es aditiva. Demuestra asimismo que es la única aplicación aditiva $H:\mathbb{Z}\to G$ cuya restricción a los naturales incorporados es h.

Ejercicio 13.14. Congruencias y representantes. En $\mathbb{Z}/\equiv_5$ demuestra que $[-8]=[2]$ y calcula $[-8]+[7]$ mediante dos pares de representantes distintos. Prueba en general que $[x]+[y]=[x+y]$ está bien definida para la congruencia módulo cinco.

Ejercicio 13.15. Identificar demasiado. Un lector afirma: «Como -1 y 4 son congruentes módulo cinco, son el mismo entero, y por eso podemos reemplazar -1 por 4 en cualquier ecuación sobre $\mathbb{Z}$». Explica el error con un contraejemplo, distingue la igualdad en $\mathbb{Z}$ de la igualdad de clases en $\mathbb{Z}/\equiv_5$ y da un contexto en el que la sustitución sí sea legítima.

13.6. Soluciones desarrolladas

Soluciones 13.1–13.5. Solución 13.1. Para comparar los cuatro primeros pares con $(0,4)$ debemos comprobar, respectivamente, las igualdades $0+6=4+2$, $0+9=4+5$ y $0+7=4+3$. Todas son verdaderas. En cambio, $(2,5)$ no es equivalente a $(0,4)$, porque $2+4=6$ y $5+0=5$ son distintos. Los primeros cuatro pares pertenecen a la clase que llamamos -4; el quinto representa -3. El cálculo no ha necesitado efectuar ninguna resta

fuera de los naturales para decidir la equivalencia, y muestra por qué la expresión «pares parecidos» no determina una clase.

Solución 13.2. R es reflexiva, simétrica y transitiva porque corresponde a la igualdad de las primeras coordenadas. Sin embargo, identifica $(2,0)$ con $(2,9)$, aunque las cuentas tienen saldos distintos: para la relación correcta, $2+9 \neq 0+2$. También separa $(2,5)$ de $(0,3)$, pese a que $2+3=5+0$. El hecho de que una relación cumpla los axiomas de equivalencia garantiza la existencia de un cociente, pero **no garantiza que el cociente resuelva el problema de representación que nos interesa**. Es necesario escoger la equivalencia adecuada a la información que queremos conservar.

Solución 13.3. Las dos equivalencias nos proporcionan $a+d=b+c$ y $c+f=d+e$. Sumando, $a+d+c+f=b+c+d+e$, y por conmutatividad y asociatividad podemos reordenar los términos como $a+f+(c+d)=b+e+(c+d)$. Aplicamos la cancelación natural, demostrada en el capítulo 12, y obtenemos $a+f=b+e$. Esa igualdad es exactamente la definición de $(a,b)\sim(e,f)$. El paso de cancelación es el único en que eliminamos información común, y no supone que existan opuestos de naturales.

Solución 13.4. La igualdad $\iota(n)=\iota(m)$ significa $(n,0)\sim(m,0)$, esto es, $n+0=0+m$; por tanto $n=m$, y ι es inyectiva. Por definición de suma, $[(n,0)]+[(m,0)]=[(n+m,0)]=\iota(n+m)$. Por definición de producto, $[(n,0)][(m,0)]=[(nm+0,0+0)]=[(nm,0)]=\iota(nm)$. La inyección preserva las operaciones; esa estructura compartida justifica la identificación habitual. Formalmente, un natural del capítulo 12 era una clase de colecciones finitas y un entero es una clase de pares de naturales, de modo que no son literalmente el mismo objeto conjuntista.

Solución 13.5. Como el orden natural es total, o bien $b\leq a$, o bien $a\leq b$. En el primer caso existe n con $a=b+n$, y $a+0=b+n$ prueba $(a,b)\sim(n,0)$. En el segundo, existe n con $b=a+n$, y $a+n=b+0$ prueba $(a,b)\sim(0,n)$. Si una clase tuviera ambas presentaciones, $(n,0)\sim(0,m)$ obligaría a $n+m=0$; la suma de naturales sólo es cero cuando ambos son cero, así que representa $[(0,0)]$. Finalmente, $[(a,b)]=0$ si y sólo si $(a,b)\sim(0,0)$, condición que se reduce a $a=b$.

Soluciones 13.6–13.10. Solución 13.6. Con los primeros representantes obtenemos $(4+7,9+2)=(11,11)$, por lo que el resultado es cero. Con los otros dos obtenemos $(1+10,6+5)=(11,11)$, el mismo representante. Los pares usados eran equivalentes: $(4,9)\sim(1,6)$ porque $4+6=9+1$ y $(7,2)\sim(10,5)$ porque $7+5=2+10$. La igualdad que garantiza el resultado es $[(a+c,b+d)]=[(a'+c',b'+d')]$ siempre que cada par sea sustituido por otro equivalente. Dos cálculos ilustran la definición; sólo el argumento general con las sumas cruzadas establece que la operación funciona para **todos** los representantes.

Solución 13.7. Si $(a,b)\sim(c,d)$, entonces $a+d=b+c$. Al conmutar los términos y los miembros tenemos $b+c=a+d$, que es la condición $(b,a)\sim(d,c)$: cambiar de representante no modifica el opuesto. La suma de una clase con su opuesta vale $[(a,b)]+[(b,a)]=[(a+b,a+b)]$. Como $(a+b,a+b)\sim(0,0)$ por reflexividad de la igualdad, esa clase es cero. Hemos construido el opuesto antes de utilizarlo para definir la resta $x-y=x+(-y)$.

Solución 13.8. Fijemos (c,d) y cambiemos (a,b) por (a',b'), con $a+b'=b+a'$. Tras multiplicar por c y d y sumar las igualdades oportunamente orientadas resulta $(ac+bd)+(a'd+b'c)=(ad+bc)+(a'c+b'd)$, que prueba $(ac+bd,ad+bc)\sim(a'c+b'd,a'd+b'c)$. La fórmula del producto es conmutativa al intercambiar ambos pares, gracias a las leyes naturales; por tanto, el mismo argumento permite cambiar sólo (c,d). Si cambiamos ambos representantes, primero sustituimos el de la izquierda y luego el de la derecha: por transitividad, el producto inicial y el final están en la misma clase. Escribir $(a-b)(c-d)$ como si ya poseyéramos la aritmética entera supondría lo que estamos tratando de construir; sirve para adivinar una fórmula, pero la prueba debe hacerse mediante igualdades de naturales.

Solución 13.9. El producto de (a,b) y (c,d) se representa por $(ac+bd, ad+bc)$. Multiplicándolo por (e,f) obtenemos

$$(ace+bde+adf+bcf,\ acf+bdf+ade+bce).$$

Si multiplicamos antes (c,d) y (e,f), el par es $(ce+df, cf+de)$, y al multiplicarlo por (a,b) obtenemos

$$(ace+adf+bcf+bde,\ acf+ade+bce+bdf).$$

Ambos pares son iguales después de reordenar los sumandos mediante conmutatividad y asociatividad de las operaciones naturales. La demostración de la asociatividad para las clases se obtiene, por tanto, sin presuponerla para los enteros.

Solución 13.10. El entero m está representado por $(m,0)$ y el negativo $-m$ por $(0,m)$. La fórmula del producto proporciona $(0,m)(0,n)=(mn,0)$ y $(0,m)(n,0)=(0,mn)$. La primera clase es el entero positivo mn y la segunda su opuesto, lo que demuestra ambas reglas. Si $m,n>0$, el producto cartesiano de dos colecciones finitas no vacías es no vacío, así que $mn>0$. Todo entero no nulo es positivo o negativo de uno positivo, y las fórmulas muestran que su producto con otro no nulo también lo es en módulo; por consiguiente no es cero. La prueba requiere que m,n sean estrictamente positivos: si uno es cero, el producto sí es cero.

Soluciones 13.11–13.15. Solución 13.11. Por la definición del orden, $[(4,9)] \le [(2,5)]$ equivale a $4+5 \le 9+2$, es decir, $9 \le 11$, que es verdadera. La igualdad requeriría $9=11$, falsa; en consecuencia, $[(4,9)] < [(2,5)]$. Para el representante alternativo $(1,6)$ obtenemos $1+5 \le 6+2$, esto es $6 \le 8$, también verdadera. Y $(4,9) \sim (1,6)$ porque $4+6=9+1$. La comparación depende de las clases, como exige la demostración de buena definición del orden, y por ello no puede variar al escoger otro representante.

Solución 13.12. Sean $x=[(a,b)]$, $y=[(c,d)]$ y $w=[(e,f)]$. La condición $x+w \le y+w$ es $(a+e)+(d+f) \le (b+f)+(c+e)$. Asociando y conmutando términos se convierte en $a+d+(e+f) \le b+c+(e+f)$. En el orden natural, sumar una misma cantidad a ambos miembros preserva y refleja la desigualdad: si $q=p+t$, entonces $q+k=(p+k)+t$, y el recíproco se obtiene cancelando k en la igualdad que expresa esa caracterización. Por tanto, la última desigualdad equivale a $a+d \le b+c$, que define $x \le y$. El paso delicado es cancelar *desigualdades* usando la caracterización natural y la cancelación aditiva, sin suponer una resta todavía no justificada.

Solución 13.13. Si $(a,b) \sim (c,d)$, sabemos que $a+d=b+c$. Aplicamos h y usamos su aditividad: $h(a)+h(d)=h(b)+h(c)$. En el grupo abeliano G podemos sumar los opuestos de $h(b)$ y $h(d)$ a ambos miembros y obtener $h(a)-h(b)=h(c)-h(d)$. Por tanto $H([(a,b)])$ no depende del representante. Además,

$$\begin{aligned} H([(a,b)]+[(c,d)]) &= H([(a+c,b+d)]) \\ &= h(a+c)-h(b+d) \\ &= (h(a)-h(b))+(h(c)-h(d)), \end{aligned}$$

pues G es abeliano, y la función es aditiva. Para $n \in \mathbb{N}$, $H([(n,0)]) = h(n)-h(0)=h(n)$, de modo que prolonga a h. Si K fuese otra prolongación aditiva, la identidad $[(a,b)] = [(a,0)]-[(b,0)]$ obligaría a $K([(a,b)]) = h(a)-h(b) = H([(a,b)])$; la prolongación es única.

Solución 13.14. Como $-8-2=-10$ es múltiplo de cinco, tenemos $[-8]=[2]$ en el cociente módulo cinco. Usando -8 y 7 obtenemos $[-8]+[7]=[-1]=[4]$. Con los representantes equivalentes 2 y 12 obtenemos $[2]+[12]=[14]=[4]$. Para cualquier módulo cinco, si $x \equiv x'$ y $y \equiv y'$, existen enteros k,ℓ tales que $x-x'=5k$ e $y-y'=5\ell$; por tanto $(x+y)-(x'+y')=5(k+\ell)$ y las sumas son congruentes. Así $[x+y]$ depende únicamente de las clases de x y y.

Solución 13.15. La igualdad de clases $[-1] = [4]$ en el cociente módulo cinco significa que $5 \mid (-1 - 4)$; no es la igualdad $-1 = 4$ en $\mathbb{Z}$. Por ejemplo, en la ecuación $x + 1 = 0$, el entero $x = -1$ es solución y $x = 4$ no lo es, pues $4 + 1 = 5 \neq 0$. En cambio, si el objetivo es resolver una congruencia $x + 1 \equiv 0 \pmod 5$, ambos representantes son igualmente válidos, porque $-1 + 1 = 0$ y $4 + 1 = 5$ pertenecen a la misma clase módulo cinco. La sustitución es legítima en operaciones que respetan la congruencia, después de justificar que están bien definidas sobre el cociente; no lo es en cualquier ecuación entera.

Capítulo 14 — ¿Por qué las fracciones forman un cuerpo ordenado?

La construcción de los enteros permitió efectuar cualquier resta, pero dejó pendientes divisiones tan sencillas como la que pide resolver $2x = 1$. En la aritmética que aprendimos antes de este libro escribimos $x = 1/2$ y aceptamos sin reparos que $1/2$, $2/4$ y $-3/-6$ designan un solo número. Hemos llegado al momento de justificar esa costumbre: los enteros servirán para fabricar representaciones y una relación de equivalencia decidirá cuáles deben pertenecer a una misma clase. La prueba decisiva volverá a ser que sumar, multiplicar, dividir y comparar esas clases no depende de la representación escogida.

14.1. Tercer momento de construcción: los racionales a partir de los enteros

Consideremos las dos escrituras que han motivado el problema, $1/2$ y $2/4$. Si las sustituimos por los pares $(1, 2)$ y $(2, 4)$, obtenemos objetos distintos de $\mathbb{Z} \times \mathbb{Z}$, aunque queremos que designen un solo número. Como todavía no hemos construido ese número, el criterio para identificar los pares tendrá que expresarse mediante las operaciones enteras disponibles; de ahí surgirá la relación que necesitamos.

Partimos de los pares (a, b) con $a, b \in \mathbb{Z}$ y $b \neq 0$. El segundo componente se declara no nulo porque no queremos introducir una división por cero. Llamemos $D = \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$. Podemos comparar dos pares por una igualdad que sólo utiliza el producto de enteros:

$$\boxed{(a, b) \sim (c, d) \quad \Longleftrightarrow \quad ad = bc.}$$

En el ejemplo anterior, $(1, 2) \sim (2, 4)$ porque $1 \cdot 4 = 2 \cdot 2$, y $(2, 4) \sim (-3, -6)$ porque $2(-6) = 4(-3)$. Estamos identificando escrituras que queremos utilizar para expresar la misma razón, pero antes de llamar «números» a sus clases tenemos que demostrar que realmente hemos definido una equivalencia.

La reflexividad se sigue de $ab = ba$ y la simetría de poder leer una igualdad de productos en cualquiera de los dos sentidos. La transitividad requiere un cálculo que merece atención. Supongamos que $(a, b) \sim (c, d)$ y $(c, d) \sim (e, f)$. Las hipótesis nos dan

$$ad = bc, \qquad cf = de.$$

Multiplicamos la primera igualdad por f y utilizamos la segunda para sustituir cf:

$$adf = bcf = bde.$$

Como $d \neq 0$ y en los enteros es lícito cancelar un factor no nulo, concluimos que $af = be$, que es exactamente la condición $(a, b) \sim (e, f)$. La exigencia de denominadores no nulos interviene aquí de verdad. Si admitiéramos $d = 0$, esa cancelación dejaría de estar justificada; es fácil incluso construir ejemplos en que la transitividad falla al permitir denominadores nulos. Para comprobarlo, admitamos provisionalmente todos los pares de enteros y tomemos $(1, 1)$, $(0, 0)$ y $(1, 2)$. Según la regla de productos cruzados, el primero es equivalente al segundo porque $1 \cdot 0 = 1 \cdot 0$, y el segundo es equivalente al tercero porque $0 \cdot 2 = 0 \cdot 1$; pero $(1, 1)$ no es equivalente a $(1, 2)$, pues $1 \cdot 2 \neq 1 \cdot 1$. Un único par con denominador cero ha permitido

conectar representaciones que la relación no puede enlazar transitivamente. Excluirlo es una decisión matemática necesaria para construir el cociente.

Al reunir los pares equivalentes obtenemos el **tercer momento de construcción numérica**:

$$\boxed{\mathbb{Q} = D/\!\sim = (\mathbb{Z} \times (\mathbb{Z} \setminus \{0\}))/\!\sim.}$$

Una fracción escrita con dos enteros es, en esta realización, un **representante**; el racional es su **clase completa**. La clase $[(1,2)]$ incluye $(2,4)$, $(-3,-6)$ y cuantos pares tengan productos cruzados iguales. La escritura habitual a/b será un nombre más breve para $[(a,b)]$, siempre con $b \neq 0$.

Sumar y multiplicar clases racionales. Con las representaciones ordinarias de fracciones conocemos dos fórmulas. Podemos emplearlas como candidatas para operar sobre las clases:

$$\boxed{\begin{aligned}[(a,b)] + [(c,d)] &= [(ad+bc, bd)],\\ [(a,b)] \cdot [(c,d)] &= [(ac, bd)].\end{aligned}}$$

Antes de considerar los representantes, debe estar garantizado que las salidas pertenecen a D: como $b \neq 0$ y $d \neq 0$, también $bd \neq 0$. Ésta es una propiedad del producto de enteros que se obtiene de su construcción y de las leyes aritméticas demostradas para él; no bastaría conocer que b y d son distintos de cero si estuviéramos trabajando en una estructura con divisores de cero.

Probemos la independencia de los representantes. Supongamos que $(a,b) \sim (a',b')$ y $(c,d) \sim (c',d')$. Es decir,

$$ab' = a'b, \qquad cd' = c'd.$$

Para la suma debemos comparar los pares $(ad+bc, bd)$ y $(a'd' + b'c', b'd')$. La condición que buscamos es

$$(ad+bc)b'd' = (a'd' + b'c')bd.$$

Podemos obtenerla sin recurrir a ninguna «igualdad de fracciones» que todavía no esté definida. Desarrollamos el primer miembro y reemplazamos ab' por $a'b$ y cd' por $c'd$:

$$\begin{aligned}(ad+bc)b'd' &= ab'dd' + cd'bb'\\ &= a'bdd' + c'dbb'\\ &= (a'd' + b'c')bd.\end{aligned}$$

La suma produce, por tanto, la misma clase para cualquier elección de los dos representantes. Para el producto, la comprobación es aún más breve:

$$(ac)(b'd') = (ab')(cd') = (a'b)(c'd) = (a'c')(bd).$$

Esa igualdad es exactamente la equivalencia entre los pares (ac, bd) y $(a'c', b'd')$. Quedan bien definidas ambas operaciones.

Los enteros se incorporan al nuevo sistema mediante

$$\iota_{\mathbb{Z}} : \mathbb{Z} \longrightarrow \mathbb{Q}, \qquad \iota_{\mathbb{Z}}(z) = [(z,1)].$$

Si $\iota_{\mathbb{Z}}(z) = \iota_{\mathbb{Z}}(w)$, entonces $(z,1) \sim (w,1)$ y la igualdad cruzada se reduce a $z = w$; luego la aplicación es inyectiva. Es correcto identificar el entero z con el racional $z/1$ cuando trabajamos

dentro de $\mathbb{Q}$, tal como antes identificamos cada natural con su entero correspondiente. También puede comprobarse directamente con las fórmulas que $\iota_{\mathbb{Z}}$ conserva sumas y productos.

Simplificar una fracción significa cambiar de representante. En la aritmética escolar se aprende a simplificar $6/8$ y obtener $3/4$. Esa operación se comprende mejor después de construir las clases. Los pares $(6, 8)$ y $(3, 4)$ son equivalentes porque $6 \cdot 4 = 8 \cdot 3$; por ello dan dos nombres del mismo elemento de $\mathbb{Q}$. Simplificar no transforma un número racional en otro: sustituye un representante por otro más cómodo. Del mismo modo, $(-3, -4)$ representa la clase de $3/4$, y $(3, -4)$ representa la de $-3/4$. Los cambios de signo deben realizarse en ambos componentes si deseamos conservar la clase.

Elegir representantes con numerador y denominador sin factores comunes simplifica los cálculos, pero la clase ya está definida antes de escoger esa escritura reducida. En una demostración podemos cambiar de representante siempre que comprobemos que la nueva pareja conserva la relación de productos cruzados; lo que se simplifica es el trabajo sobre el número, no el número mismo.

Comprobación 9. Comprueba que $(1, 2)$ y $(3, 6)$ representan un mismo racional. Calcula la suma de sus clases con la de $(1, 3)$ utilizando primero el representante $(1, 2)$ y luego $(3, 6)$. Indica qué argumento general impide que los resultados dependan de esa elección.

Solución razonada. Son equivalentes porque $1 \cdot 6 = 2 \cdot 3$. Con el primer representante obtenemos $[(1, 2)] + [(1, 3)] = [(5, 6)]$; con el segundo, $[(3, 6)] + [(1, 3)] = [(15, 18)]$. Ambos resultados pertenecen a la misma clase, pues $5 \cdot 18 = 6 \cdot 15$. El cálculo comprueba un caso; la prueba general de buena definición de la suma, realizada mediante productos cruzados, garantiza la coincidencia para todos los representantes permitidos.

Una vez comprobadas estas dos operaciones, la pregunta aritmética que nos hizo ampliar los enteros se vuelve precisa: ¿existe una clase que, multiplicada por la de 2, produzca la clase de 1? Resolveremos esa cuestión mediante inversos de clases, sin recurrir a fracciones ya conocidas.

Una perspectiva histórica: la igualdad de razones. La pregunta por cuándo dos razones deben considerarse iguales es mucho más antigua que nuestra construcción conjuntista. En el libro V de los *Elementos*, Euclides presenta una teoría de proporciones entre magnitudes. Su definición 5 caracteriza la igualdad de razones mediante la comparación de múltiplos correspondientes de las magnitudes, y permite trabajar incluso con magnitudes que no poseen una medida común. Allí se trata de razones de magnitudes; no se define $\mathbb{Q}$ como un conjunto de clases de pares de enteros. El vínculo con nuestro recorrido es intelectual: tanto en aquella teoría como aquí debemos precisar qué quiere decir que distintas presentaciones expresen una misma relación y demostrar que esa igualdad se comporta como esperamos. Esta comparación histórica deja a salvo las diferencias de objetos y métodos.

Referencia de la nota: Euclides, *Elementos*, libro V, definiciones 3 y 5; traducción y texto consultados en Euclid's Elements — Book V y Los Elementos — Libro V. La nota no atribuye a Euclides las clases de equivalencia modernas ni la formalización conjuntista de los racionales.

14.2. Dividir dentro de los racionales

Las fórmulas obtenidas permiten sumar y multiplicar clases de pares enteros, pero el motivo inicial de pasar de $\mathbb{Z}$ a $\mathbb{Q}$ era resolver ecuaciones como $2x = 1$. Necesitamos producir, para cada clase $x \neq 0$, otra clase y con $xy = 1$. La regla escolar de intercambiar numerador y denominador sugiere cómo hacerlo; la dificultad está en que una misma clase admite infinitas escrituras, y el resultado deberá seguir siendo el mismo aunque escojamos otra.

Conservaremos la notación $D = \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ y la equivalencia $(a, b) \sim (c, d)$ cuando $ad = bc$. Las operaciones con clases ya quedaron bien definidas en la sección precedente, de modo que podemos utilizarlas sin volver a justificar en cada línea los cambios de representante que allí demostramos.

Cero, uno y el significado de una clase no nula. La clase $[(0,1)]$ desempeña el papel del cero. En efecto, para cualquier $(a,b) \in D$,

$$[(a,b)] + [(0,1)] = [(a,b)],$$

pues la fórmula de la suma produce precisamente el representante (a,b). De igual manera, la clase $[(1,1)]$ es la unidad multiplicativa, porque

$$[(a,b)]\,[(1,1)] = [(a,b)].$$

Estas dos clases son diferentes: si fueran iguales, la equivalencia de los pares exigiría $0 \cdot 1 = 1 \cdot 1$, es decir, $0 = 1$ en $\mathbb{Z}$. Utilizaremos desde aquí los nombres $0 = [(0,1)]$ y $1 = [(1,1)]$, recordando que las igualdades ocurren dentro de $\mathbb{Q}$.

Hay una comprobación especialmente útil antes de hablar de inversos. Una clase $[(a,b)]$ es cero exactamente cuando $a = 0$. Para justificarlo, escribamos la condición de equivalencia con $(0,1)$:

$$[(a,b)] = [(0,1)] \quad\Longleftrightarrow\quad (a,b) \sim (0,1) \quad\Longleftrightarrow\quad a \cdot 1 = b \cdot 0 \quad\Longleftrightarrow\quad a = 0.$$

El denominador ya sabemos que no puede ser cero. Así, una clase racional distinta de cero tiene un representante con **ambas coordenadas no nulas**, y cualquier otro representante suyo también tiene numerador no nulo. Esta observación autoriza una operación que, hasta aquí, sólo podíamos anticipar por la costumbre de invertir fracciones.

Invertir una clase. Sea $x = [(a,b)] \neq 0$. Como $a \neq 0$, el par (b,a) pertenece a D y podemos proponer

$$\boxed{x^{-1} = [(b,a)].}$$

Antes de multiplicar, hemos de considerar una dificultad propia del cociente: ¿qué sucede si en lugar de (a,b) habíamos elegido otro representante (c,d) de x? La equivalencia inicial nos dice $ad = bc$. Precisamente esa igualdad, leída como $bc = ad$, afirma que $(b,a) \sim (d,c)$. Los pares obtenidos al intercambiar coordenadas representan, pues, una misma clase. La inversión **está bien definida** en $\mathbb{Q} \setminus \{0\}$.

Al multiplicar obtenemos

$$[(a,b)]\,[(b,a)] = [(ab,ba)] = [(1,1)],$$

porque $(ab,ba) \sim (1,1)$: sus productos cruzados son iguales por conmutatividad, y $ab \neq 0$ asegura que el par empleado pertenece a D. Por consiguiente, todo racional no nulo tiene inverso multiplicativo. La ecuación $2x = 1$, que carecía de solución entera, ya se resuelve dentro del nuevo sistema: el entero 2 se representa por $[(2,1)]$ y su inverso es $[(1,2)]$.

El símbolo $\frac{a}{b}$ puede usarse desde este momento como abreviatura de la clase $[(a,b)]$. La notación no elimina la clase: si escribimos $\frac{1}{2} = \frac{2}{4}$, afirmamos la igualdad entre las clases de $(1,2)$ y $(2,4)$, aunque los pares sigan siendo distintos.

Comprobación 10. Considera los representantes $(2,3)$ y $(-4,-6)$ de una misma clase racional. Calcula la inversa empleando cada uno. Explica por qué las respuestas tienen que representar el mismo racional e indica qué falla si intentamos invertir la clase cero.

Solución razonada. Los pares son equivalentes, pues $2(-6) = 3(-4)$. Con el primer representante obtenemos $[(3,2)]$; con el segundo, $[(-6,-4)]$. Estas clases coinciden porque $3(-4) = 2(-6)$. En general la equivalencia $(a,b) \sim (c,d)$ se conserva al intercambiar ambas coordenadas, siempre que a y c sean no nulos. Para el cero esa condición falla: un representante $(0,b)$ se transformaría en $(b,0)$, que queda fuera de D por tener denominador cero. No existe inversa multiplicativa de 0, pues cualquier producto por 0 vuelve a ser 0 y $0 \neq 1$.

14.3. Las leyes del cuerpo: lo que las fórmulas deben demostrar

Hemos logrado definir operaciones con clases y producir la inversa de cada racional no nulo. La palabra **cuerpo**, sin embargo, reúne varias afirmaciones que requieren justificación. Entre ellas se encuentran la asociatividad de la suma y del producto, sus elementos neutros, los opuestos aditivos, la conmutatividad, la distributividad y la existencia de inversos para todos los elementos distintos de cero. La experiencia de los enteros muestra cómo estas leyes pueden estudiarse a partir de las operaciones que ya conocemos.

Para abreviar los cálculos pondremos $x = [(a, b)]$, $y = [(c, d)]$ y $z = [(e, f)]$, con $b, d, f \neq 0$. Las igualdades entre clases se obtendrán utilizando representantes concretos; no suponemos que los distintos pares elegidos sean literalmente iguales.

Sumar tres fracciones. Si calculamos primero $x+y$ y después añadimos z, la definición da

$$\begin{aligned}(x + y) + z &= [(ad + bc, bd)] + [(e, f)] \\ &= [((ad + bc)f + ebd, bdf)] \\ &= [(adf + bcf + bde, bdf)].\end{aligned}$$

Si, por el contrario, sumamos primero $y + z$, resulta

$$\begin{aligned}x + (y + z) &= [(a, b)] + [(cf + de, df)] \\ &= [(adf + b(cf + de), bdf)] \\ &= [(adf + bcf + bde, bdf)].\end{aligned}$$

Ambos recorridos llegan a un mismo representante, gracias a las leyes distributiva, asociativa y conmutativa de los enteros. La suma racional es, por tanto, **asociativa**. Al intercambiar x e y obtenemos $(cb + da, db)$ en lugar de $(ad + bc, bd)$; las operaciones enteras muestran que ambos pares son iguales. La suma es también **conmutativa**.

Ya comprobamos que $0 = [(0, 1)]$ es neutro. El opuesto aditivo de $[(a, b)]$ es $[(-a, b)]$, porque

$$[(a, b)] + [(-a, b)] = [(ab - ba, b^2)] = [(0, b^2)] = 0.$$

La última igualdad se deduce de la relación de equivalencia: $(0, b^2) \sim (0, 1)$, y $b^2 \neq 0$. La clase opuesta también puede escribirse como $[(a, -b)]$; la igualdad de estas clases se demuestra mediante productos cruzados. Así, $(\mathbb{Q}, +)$ es un grupo abeliano: podemos sumar, reagrupar y resolver cualquier ecuación $x + y = z$ sumando el opuesto de y.

Multiplicar y distribuir. La multiplicación ofrece una cuenta más corta. Las dos formas de agrupar tres clases conducen al representante (ace, bdf):

$$\begin{aligned}(xy)z &= [(ac, bd)]\,[(e, f)] = [(ace, bdf)], \\ x(yz) &= [(a, b)]\,[(ce, df)] = [(ace, bdf)].\end{aligned}$$

El intercambio de los dos factores cambia (ac, bd) por (ca, db), que es el mismo par. Tenemos asociatividad y conmutatividad; el elemento neutro es $1 = [(1, 1)]$. En la sección anterior construimos, además, el inverso multiplicativo de cualquier clase no nula.

Para comprobar la distributividad comparemos $x(y + z)$ con $xy + xz$. El primer miembro tiene como representante

$$\begin{aligned}x(y + z) &= [(a, b)]\,[(cf + de, df)] \\ &= [(acf + ade, bdf)].\end{aligned}$$

El segundo puede expresarse con el **mismo denominador**. En efecto, las clases $[(ac, bd)]$ y $[(acf, bdf)]$ coinciden, y lo mismo ocurre con $[(ae, bf)]$ y $[(aed, bdf)]$. Si sumamos esas dos representaciones obtenemos

$$\begin{aligned} xy + xz &= [(acf, bdf)] + [(aed, bdf)] \\ &= [((acf + aed)bdf, (bdf)^2)] \\ &= [(acf + aed, bdf)]. \end{aligned}$$

La simplificación de la última clase está justificada por la equivalencia de pares, pues $bdf \neq 0$. Los dos resultados coinciden, y queda probada la distributividad. La correspondiente identidad $(y + z)x = yx + zx$ se sigue de la conmutatividad del producto.

Los cálculos anteriores prueban que $\mathbb{Q}$, con las operaciones definidas mediante pares y con $0 \neq 1$, **es un cuerpo conmutativo**. El nombre resume las propiedades; las razones matemáticas están en las pruebas. Los enteros se incorporan mediante la función $z \mapsto [(z, 1)]$, que es inyectiva y conserva suma, producto, cero y unidad, como demostramos al construir las operaciones racionales. Al trabajar dentro de $\mathbb{Q}$ identificaremos un entero con su imagen, sin afirmar que las clases de pares de naturales y las clases de pares de enteros son literalmente el mismo tipo de conjunto.

Comprobación 11. Una persona observa que la regla de suma de fracciones está bien definida y concluye que necesariamente es asociativa. ¿Qué falta en ese razonamiento? Comprueba después, con representantes, que $\frac{1}{2} + (\frac{1}{3} + \frac{1}{6}) = (\frac{1}{2} + \frac{1}{3}) + \frac{1}{6}$.

Solución razonada. La buena definición asegura que cambiar las escrituras de las entradas no cambia la clase resultante; la asociatividad compara dos órdenes diferentes de aplicación de la operación. Son obligaciones independientes. En el ejemplo, $\frac{1}{3} + \frac{1}{6} = \frac{1}{2}$, de modo que la primera agrupación produce 1. Para la otra, $\frac{1}{2} + \frac{1}{3} = \frac{5}{6}$ y $\frac{5}{6} + \frac{1}{6} = 1$. La igualdad de estos cálculos ilustra la propiedad; la prueba algebraica anterior, que usa tres fracciones arbitrarias, es la que la establece en general.

14.4. Ordenar las clases sin dejarse engañar por los signos

En la recta que conocemos por la aritmética elemental, $\frac{1}{2}$ queda a la derecha de 0 y $-\frac{1}{2}$ a su izquierda. Si queremos recuperar ese orden desde nuestra construcción, necesitamos decidir qué significa que **una clase de pares sea positiva**. El signo de un representante puede inducir a error: $[(1, -2)]$ y $[(-1, 2)]$ son la misma clase, aunque el primer numerador sea positivo y el segundo negativo. La positividad del numerador por sí sola no sirve para clasificar los racionales mientras admitamos denominadores de ambos signos.

Todo racional tiene, sin embargo, algún representante con denominador positivo. Si $b > 0$, utilizamos (a, b); si $b < 0$, lo sustituimos por $(-a, -b)$, que es equivalente porque $a(-b) = b(-a)$ y tiene denominador positivo. Esto permite fijar una regla sin imponer a cada clase una fracción reducida ni elegir un representante privilegiado.

Definición. Diremos que $x \in \mathbb{Q}$ es **positivo**, y escribiremos $x > 0$, cuando exista un representante (a, b) de x con $b > 0$ y $a > 0$. Diremos que $x < 0$ cuando $-x > 0$; las clases restantes serán examinadas mediante la igualdad con cero.

Hay que verificar que la palabra «positivo» tiene sentido para la clase. Supongamos que $[(a, b)] = [(c, d)]$, con $b, d > 0$. La equivalencia significa $ad = bc$. Si $a > 0$, entonces $ad > 0$, y la igualdad obliga a $bc > 0$; como $b > 0$, se sigue $c > 0$. El mismo argumento, intercambiando los pares, prueba la implicación contraria. **Todos los representantes con denominador positivo tienen numeradores del mismo signo.** La definición no depende del par escogido.

Podemos clasificar por completo los racionales utilizando uno de esos representantes: según $a > 0$, $a = 0$ o $a < 0$, la clase es positiva, es cero o es negativa. Los tres casos son mutuamente excluyentes porque así ocurre en el orden de $\mathbb{Z}$. Esta tricotomía permitirá definir la comparación entre dos racionales sin volver a examinar cada posible representación.

Comparar dos clases. Definimos

$$\boxed{x < y \quad \Longleftrightarrow \quad y - x > 0.}$$

La diferencia está bien definida porque ya hemos construido la suma y los opuestos aditivos. Si elegimos representantes $x = [(a, b)]$, $y = [(c, d)]$ con $b, d > 0$, las fórmulas nos dicen

$$y - x = [(bc - ad, bd)].$$

Como $bd > 0$, la condición de positividad se reduce a una comparación de enteros:

$$\boxed{[(a, b)] < [(c, d)] \quad \Longleftrightarrow \quad ad < bc \qquad (b > 0,\ d > 0).}$$

Ésta es la conocida comparación mediante productos cruzados, ahora con sus hipótesis a la vista. No podemos aplicar la misma desigualdad sin atender a los signos de los denominadores. Por ejemplo, la clase escrita $\frac{1}{-2}$ es igual a $\frac{-1}{2}$ y resulta menor que $\frac{1}{3}$; al utilizar los representantes de denominador positivo, la comparación correcta es $(-1)3 < 1 \cdot 2$. Multiplicar una desigualdad por un denominador negativo sin cambiar su sentido conduciría a un resultado falso.

La relación $<$ es transitiva. Si $x < y$ e $y < z$, tanto $y - x$ como $z - y$ son positivos. La suma de dos racionales positivos es positiva: con denominadores positivos, la fórmula de la suma produce denominador positivo y numerador $ad + bc > 0$. Por tanto,

$$z - x = (z - y) + (y - x) > 0,$$

y concluimos $x < z$. La tricotomía aplicada a $y - x$ asegura que, para dos racionales, ocurre exactamente una de las tres posibilidades $x < y$, $x = y$ o $y < x$. De ellas resulta el orden total no estricto $x \leq y$ si $x < y$ o $x = y$.

También comprobamos la compatibilidad con las operaciones. Si sumamos w a ambos miembros de una comparación, la diferencia permanece igual:

$$(y + w) - (x + w) = y - x.$$

Por eso $x < y$ implica $x + w < y + w$. Si $w > 0$, el producto $(y - x)w$ es positivo: en denominadores positivos, el producto de numeradores positivos y el de denominadores positivos vuelven a ser positivos. Como $(y - x)w = yw - xw$ por distributividad, se sigue $xw < yw$. Para $w < 0$ sucede lo contrario, puesto que $-w > 0$ y multiplicar por -1 invierte la desigualdad. Estas propiedades son exactamente las que permiten hablar de un **cuerpo ordenado**.

La inclusión $\mathbb{Z} \hookrightarrow \mathbb{Q}$ respeta el orden. Para dos enteros m, n tomamos $[(m, 1)]$ y $[(n, 1)]$; la comparación de sus clases da $m < n$ si y sólo si $m \cdot 1 < n \cdot 1$, que es la comparación original. Así, el nuevo orden prolonga el orden entero sin alterar sus resultados.

Comprobación 12. Compara las clases $[(1, -2)]$ y $[(1, 3)]$ sin utilizar una regla de productos cruzados que ignore el signo del denominador. Decide, además, si $[(-2, -4)]$ es positiva o negativa y justifica que tu respuesta no depende de esa representación.

Solución razonada. Sustituimos $(1, -2)$ por el par equivalente $(-1, 2)$. Sus denominadores son positivos, y $(-1) \cdot 3 = -3 < 2 = 1 \cdot 2$; por tanto $[(1, -2)] < [(1, 3)]$. La clase $[(-2, -4)]$ coincide con $[(2, 4)]$, cuyo numerador y denominador son positivos, así que la clase es positiva. Cualquier otro representante con denominador positivo tendrá numerador positivo por la igualdad de productos cruzados; la conclusión pertenece a la clase, no a una escritura particular.

14.5. Una recta densamente poblada que todavía tiene una ausencia

Una de las primeras propiedades que descubrimos al trabajar con fracciones es la posibilidad de encontrar nuevas fracciones entre otras dos. Si $x, y \in \mathbb{Q}$ y $x < y$, la media aritmética

$$m = \frac{x+y}{2}$$

es racional: podemos sumar x e y y multiplicar su suma por el inverso del racional no nulo 2. Además,

$$m - x = \frac{y-x}{2} > 0, \qquad y - m = \frac{y-x}{2} > 0.$$

Por consiguiente, $x < m < y$. Podemos repetir el procedimiento entre x y m, entre m e y, y tantas veces como queramos. Decimos que **los racionales son densos en su propio orden**: entre dos racionales distintos siempre hay otro racional. Es una propiedad notable, pero deja abierta una pregunta diferente: que podamos intercalar números indefinidamente ¿garantiza que exista un racional para toda posición que nuestras operaciones y desigualdades parecen describir?

Consideremos la ecuación

$$x^2 = 2.$$

En $\mathbb{Q}$ podemos sumar, multiplicar, ordenar y dividir por todo número no nulo. Ninguna de esas posibilidades asegura, por sí sola, que la ecuación tenga solución. De hecho, **no existe un racional cuyo cuadrado sea** 2. Para demostrarlo aprovechemos precisamente la construcción mediante clases.

Si hubiera una solución racional, podríamos escoger una representación $[(a, b)]$ con $a, b \in \mathbb{Z}$, $b > 0$ y $a > 0$. Entre todos los denominadores positivos de representaciones de una solución, el principio de buen orden de los naturales nos permitiría elegir uno mínimo. De $[(a, b)]^2 = 2 = [(2, 1)]$ se obtiene, por la igualdad de clases,

$$a^2 = 2b^2.$$

El cuadrado de un entero impar es impar: si escribimos $a = 2k + 1$, su cuadrado es $2(2k^2 + 2k) + 1$. Por tanto, de que a^2 sea par deducimos que a es par, digamos $a = 2c$. Al sustituirlo en la igualdad y cancelar 2 resulta

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

por lo que b también es par, digamos $b = 2d$ con $d > 0$. Ahora (c, d) representa la misma clase que $(a, b) = (2c, 2d)$ y posee un denominador positivo menor: $d < b$. Hemos contradicho la elección de b como denominador mínimo. La imposibilidad de resolver $x^2 = 2$ dentro de $\mathbb{Q}$ no proviene de una limitación de nuestros símbolos para escribir fracciones; se deduce de la propia aritmética de enteros con la que formamos sus clases.

Una colección que tiene cotas superiores, pero ninguna mínima. Podemos hacer visible la insuficiencia de $\mathbb{Q}$ sin dar por supuesto ningún número irracional. Consideremos el conjunto

$$\boxed{S = \{q \in \mathbb{Q} : q > 0 \text{ y } q^2 < 2\}.}$$

Contiene al 1, porque $1^2 < 2$, y está acotado superiormente por 2: ningún $q \geq 2$ pertenece a S, ya que entonces $q^2 \geq 4 > 2$. Un lector que conoce la recta real imaginará que el borde superior de S se sitúa donde el cuadrado vale exactamente dos. Todavía no hemos construido

los reales, de modo que esa imagen sólo sirve para orientar la búsqueda. Preguntaremos directamente si **algún racional** puede ser la menor de las cotas superiores de S.

Supongamos que existe una, y llamémosla r. Como $1 \in S$, necesariamente $r \geq 1$, de modo que $r > 0$. Ya hemos demostrado que $r^2 \neq 2$ para todo racional. La elección se reduce entonces a dos casos: $r^2 < 2$ o $r^2 > 2$. Nos será útil construir un nuevo racional a partir de r mediante la fórmula

$$\boxed{t = \frac{2(r+1)}{r+2}.}$$

Está bien definida porque $r + 2 > 0$. Dos cálculos elementales explican por qué escogimos esa expresión:

$$\boxed{t - r = \frac{2 - r^2}{r+2}, \qquad 2 - t^2 = \frac{2(2-r^2)}{(r+2)^2}.}$$

En ambos miembros aparece el mismo factor $2 - r^2$, que determina simultáneamente de qué lado de r cae t y de qué lado de 2 cae su cuadrado. Si $r^2 < 2$, entonces $t - r > 0$ y $2 - t^2 > 0$. Como $t > r > 0$, resulta $t \in S$ y $t > r$: hemos encontrado un elemento de S mayor que una supuesta cota superior, lo cual es imposible.

Si $r^2 > 2$, los signos se invierten: $t < r$ y $t^2 > 2$, mientras que $t > 0$ porque numerador y denominador son positivos. Pero cualquier racional positivo cuyo cuadrado sea mayor que dos es una cota superior de S. En efecto, si $s \in S$ y $s \geq t > 0$, la monotonía del cuadrado sobre los racionales no negativos daría $s^2 \geq t^2 > 2$, en contradicción con $s^2 < 2$. Así que t sería una cota superior estrictamente menor que r, contraria a la supuesta minimalidad de r.

Ningún caso es posible. Hemos probado que S es un conjunto no vacío de racionales, acotado superiormente en $\mathbb{Q}$, que **no posee supremo en** $\mathbb{Q}$. La palabra *supremo* designa la menor de las cotas superiores, si existe. La propiedad que falta a los racionales es, por tanto, mucho más precisa que una vaga sensación de que hay «huecos» entre fracciones: los racionales son densos y, sin embargo, no garantizan supremos para todos sus subconjuntos no vacíos y acotados superiormente.

Comprobación 13. En el argumento anterior, ¿por qué no basta afirmar que $x^2 = 2$ carece de solución racional para concluir sin más que S no tiene supremo racional? Explica el papel de $t = 2(r+1)/(r+2)$ en los dos casos de la prueba.

Solución razonada. Saber que ningún racional satisface $r^2 = 2$ excluye un candidato para el borde, pero todavía hace falta demostrar que cualquier racional con cuadrado menor que dos queda **demasiado abajo** para ser una cota superior, y que cualquiera con cuadrado mayor que dos queda **demasiado arriba** para ser la menor. La fórmula de t resuelve ambas obligaciones: cuando $r^2 < 2$ produce un elemento mayor que r que sigue perteneciendo a S; cuando $r^2 > 2$ produce una cota superior positiva menor que r. Estas dos mejoras incompatibles con ser supremo cubren todas las posibilidades racionales.

La construcción de los reales responderá a una dificultad de este género. Una ruta posible utiliza sucesiones racionales de Cauchy y vuelve a formar clases de equivalencia; otra organiza cortes en el orden racional. Cada método tendrá que explicar por qué los objetos nuevos completan lo que $\mathbb{Q}$ deja sin resolver. La densidad de $\mathbb{Q}$ y la ausencia de ciertos supremos ofrecen la razón matemática para continuar, aunque antes merece la pena reconocer con precisión qué hemos conseguido mediante las tres ampliaciones.

14.6. Lo que significa haber construido los números

El procedimiento que permitió pasar de una colección a un número natural volvió a aparecer en las dos ampliaciones siguientes, aunque con representantes de naturaleza distinta:

$$\boxed{\begin{aligned}\mathbb{N}_U &= \mathcal{F}(U)/\approx,\\ \mathbb{Z} &= (\mathbb{N}_U \times \mathbb{N}_U)/\sim,\\ \mathbb{Q} &= (\mathbb{Z} \times (\mathbb{Z} \setminus \{0\}))/\sim.\end{aligned}}$$

La primera igualdad identifica colecciones finitas que admiten una biyección y las trata como cantidades naturales; la segunda identifica pares de naturales que expresan un mismo saldo; la tercera identifica pares de enteros con denominador no nulo que expresan una misma razón. En cada caso, antes de operar comprobamos que las reglas elegidas no dependen de la representación. Las inclusiones que escribimos habitualmente como $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q}$ se realizan mediante aplicaciones inyectivas que respetan las operaciones, y esa precisión evita identificar literalmente conjuntos construidos con clases de objetos de naturaleza diferente.

Los cocientes de restos $\mathbb{N}/\sim_m$ y $\mathbb{Z}/\equiv_m$ responden a otra pregunta: una vez construidos esos números, podemos **identificarlos por una propiedad** y obtener nuevos objetos, generalmente en cantidad finita. Por eso tiene importancia distinguir «construir los naturales mediante clases» de «estudiar clases de equivalencia en los naturales». Las dos ideas utilizan el mismo mecanismo, pero sirven a problemas distintos.

Con los enteros obtuvimos opuestos aditivos; con los racionales, inversos multiplicativos para cada elemento distinto de cero. La prueba de que las operaciones y el orden pertenecen a las clases —sin depender del par elegido— acompaña ambas ampliaciones. Y el conjunto acotado que acabamos de estudiar muestra exactamente dónde termina la tercera: el cuerpo ordenado $\mathbb{Q}$ permite aproximarse indefinidamente a un borde que no tiene por qué existir como racional. Esa dificultad motiva la construcción posterior de los reales.

La pregunta que acompañará los ejercicios es siempre la misma, aunque cambien los objetos: ¿qué parte del cálculo depende de una representación concreta y qué parte pertenece a la clase? Para responder habrá que utilizar las leyes del cuerpo y el orden que ya demostramos, e identificar las hipótesis que permiten cada paso.

14.7. Ejercicios

Para todos los problemas siguientes usamos $D = \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$, con $(a, b) \sim (c, d)$ cuando $ad = bc$; las operaciones sobre las clases son las que hemos definido en las secciones anteriores. Cuando se pida justificar que una regla funciona, habrá que distinguir la validez de una cuenta particular de la independencia respecto de cualquier representante.

Representaciones, equivalencias y operaciones. Ejercicio 14.1. El denominador cero y la transitividad. Sobre $\mathbb{Z} \times \mathbb{Z}$ se propone $(a, b)\, R\, (c, d)$ cuando $ad = bc$, permitiendo denominadores nulos. Examina los pares $(1, 1)$, $(0, 0)$ y $(1, 2)$. Demuestra qué propiedad de equivalencia falla y explica por qué excluir denominadores nulos repara precisamente el paso de la demostración que ha fallado.

Ejercicio 14.2. Una clase con signos diferentes. Demuestra que $(3, -8)$, $(-3, 8)$ y $(9, -24)$ representan el mismo racional. Comprueba si $(-3, -8)$ pertenece a esa clase. Explica por qué basta verificar productos cruzados y por qué los pares siguen siendo distintos.

Ejercicio 14.3. Cero y unidad son clases. Demuestra directamente que $[(a, b)] = 0$ si y sólo si $a = 0$, y que $[(a, b)] = 1$ si y sólo si $a = b$. Usa estos resultados para explicar por qué $[(0, 7)]$ y $[(7, 7)]$ no son la misma clase.

Ejercicio 14.4. Buena definición de la suma. Dados $(a, b) \sim (a', b')$ y $(c, d) \sim (c', d')$, con los cuatro denominadores no nulos, demuestra por productos cruzados que

$$[(ad + bc, bd)] = [(a'd' + b'c', b'd')].$$

Detalla las dos sustituciones que autorizan las hipótesis y comprueba que los pares resultantes pertenecen a D.

Ejercicio 14.5. Buena definición del producto. Demuestra que $(a,b)\sim(a',b')$ y $(c,d)\sim(c',d')$ implican $(ac,bd)\sim(a'c',b'd')$. Después señala el error de quien propone multiplicar clases mediante $[(a,b)]\star[(c,d)]=[(a+c,bd)]$ sin comprobar nada: encuentra dos representantes de una misma primera clase con los que la regla da resultados diferentes, usando como segunda entrada $[(1,1)]$.

Estructura de cuerpo y orden. Ejercicio 14.6. Invertir sin confundir las operaciones. Para $x=[(-3,5)]$, determina $-x$ y x^{-1} mediante representantes admisibles. Comprueba por separado $x+(-x)=0$ y $xx^{-1}=1$. Explica por qué 0 no tiene inverso multiplicativo y por qué 1 sí lo tiene.

Ejercicio 14.7. Dos rutas hacia una misma suma. Calcula $[(1,2)]+[(-3,4)]$ y repite el cálculo utilizando para la primera entrada $(2,4)$ y para la segunda $(6,-8)$. Demuestra la igualdad entre los resultados sin hablar de números decimales.

Ejercicio 14.8. Distributividad comprobada en un caso y en general. Verifica $\frac{1}{2}(\frac{1}{3}+\frac{1}{4})=\frac{1}{2}\cdot\frac{1}{3}+\frac{1}{2}\cdot\frac{1}{4}$ con pares enteros. Después explica qué identidad entre representantes prueba $x(y+z)=xy+xz$ para tres racionales cualesquiera, indicando dónde se utiliza la distributividad de $\mathbb{Z}$.

Ejercicio 14.9. Una regla de orden falsa. Alguien afirma que $[(a,b)]<[(c,d)]$ equivale siempre a $ad<bc$, aunque b o d sean negativos. Refuta esa afirmación con los racionales $[(1,-2)]$ y $[(1,3)]$. Formula la versión correcta e indica por qué puede elegirse un denominador positivo sin cambiar la clase.

Ejercicio 14.10. Multiplicar numerador y denominador por un entero. Sean $a,b,k\in\mathbb{Z}$ con $b\neq0$ y $k\neq0$. Prueba que $[(a,b)]=[(ka,kb)]$. Explica por qué no se puede extender esta regla a $k=0$ y qué efecto tiene un $k<0$ sobre la comparación por productos cruzados si se olvidan las hipótesis de signo.

Ejercicio 14.11. Orden y operaciones. Supón $x<y$ en $\mathbb{Q}$. Demuestra, utilizando diferencias y positividad, que $x+w<y+w$ para todo w y que $xw<yw$ si $w>0$. Prueba también que la desigualdad se invierte cuando $w<0$. Especifica qué resultado algebraico autoriza pasar de $(y-x)w$ a $yw-xw$.

Densidad, límites y diagnóstico. Ejercicio 14.12. El medianto entre dos racionales. Sean $a,b,c,d\in\mathbb{Z}$ con $b,d>0$ y $a/b<c/d$. Demuestra que

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

Comprueba el resultado entre $3/7$ y $4/9$. Explica por qué este procedimiento, aunque pueda repetirse indefinidamente, no demuestra que todos los subconjuntos acotados de $\mathbb{Q}$ tengan supremo.

Ejercicio 14.13. No existe un mayor racional. Demuestra que, dado $x\in\mathbb{Q}$, existe un natural n con $n>x$. Parte de un representante (a,b) con $b>0$ y evita recurrir a la recta real. Deduce que $\mathbb{Q}$ carece de elemento máximo. ¿En qué sentido esta propiedad es distinta de la densidad?

Ejercicio 14.14. La imposibilidad de cuadrar hasta dos. Supón que $x\in\mathbb{Q}$ satisface $x^2=2$. Expresa x mediante un par de enteros con denominador positivo mínimo entre los que representan la clase y utiliza la paridad para construir un representante con denominador menor. Explica por qué la elección mínima está justificada y dónde se obtiene la contradicción.

Ejercicio 14.15. Candidatos a supremo. Sea $S=\{q\in\mathbb{Q}:q>0,\ q^2<2\}$. Para $r>0$ define $t=2(r+1)/(r+2)$. Verifica las identidades

$$t-r=\frac{2-r^2}{r+2},\qquad 2-t^2=\frac{2(2-r^2)}{(r+2)^2}.$$

Calcula t cuando $r = 1$ y cuando $r = 3/2$. Utiliza ambos casos generales $r^2 < 2$ y $r^2 > 2$ para probar que S no tiene supremo racional; explica por qué saber solamente que $x^2 = 2$ no tiene solución racional no bastaría para esa conclusión.

14.8. Soluciones desarrolladas

Soluciones 14.1–14.5. Solución 14.1. Con los pares propuestos se tiene $(1,1)R(0,0)$ porque $1 \cdot 0 = 1 \cdot 0$, y $(0,0)R(1,2)$ porque $0 \cdot 2 = 0 \cdot 1$. Sin embargo, $(1,1)$ no está relacionado con $(1,2)$, pues $1 \cdot 2 \neq 1 \cdot 1$. Por tanto, falla la transitividad; no basta que la regla tenga simetría y reflexividad para producir una partición. En D no puede aparecer $(0,0)$. Más generalmente, en la prueba de transitividad de D, las hipótesis $ad = bc$ y $cf = de$ implican $adf = bde$. Para obtener $af = be$ cancelamos d, operación válida únicamente porque $d \neq 0$ en un anillo sin divisores de cero. La restricción de dominio no es ornamental: cierra el argumento que el par $(0,0)$ destruía.

Solución 14.2. Entre $(3,-8)$ y $(-3,8)$ los productos cruzados son $3 \cdot 8 = 24$ y $(-8)(-3) = 24$. Entre $(3,-8)$ y $(9,-24)$ son $3(-24) = -72$ y $(-8)9 = -72$. Por transitividad, los tres pertenecen a la misma clase. En cambio, $(3,-8)$ no es equivalente a $(-3,-8)$, puesto que $3(-8) = -24$ y $(-8)(-3) = 24$. Al tomar clases identificamos las representaciones equivalentes; los pares, que son objetos de D, siguen siendo diferentes cuando cambian sus coordenadas.

Solución 14.3. La clase de (a,b) es cero si $(a,b) \sim (0,1)$. Esto equivale a $a \cdot 1 = b \cdot 0$, luego a $a = 0$. Del mismo modo, $[(a,b)] = [(1,1)]$ equivale a $a \cdot 1 = b \cdot 1$, esto es, $a = b$. El par $(0,7)$ representa cero porque su numerador es cero, mientras que $(7,7)$ representa uno porque ambas coordenadas son iguales; no pueden ser clases iguales ya que $0 \neq 1$ en los enteros. Ninguno de los dos argumentos necesita interpretar las clases como cocientes numéricos antes de que éstos se hayan construido.

Solución 14.4. Las hipótesis se escriben $ab' = a'b$ y $cd' = c'd$. Como $b, d, b', d' \neq 0$ y $\mathbb{Z}$ no tiene divisores de cero, también $bd \neq 0$ y $b'd' \neq 0$: ambos pares propuestos pertenecen a D. Calculamos

$$\begin{aligned}(ad + bc)b'd' &= ab'dd' + cd'bb' \\ &= a'bdd' + c'dbb' \\ &= (a'd' + b'c')bd.\end{aligned}$$

La primera sustitución usa $ab' = a'b$ y la segunda $cd' = c'd$. La igualdad final es la condición de equivalencia entre las dos parejas que representan las sumas. No se ha supuesto que las fracciones fueran iguales por las reglas escolares; se ha demostrado que la operación sobre el cociente respeta sus representantes.

Solución 14.5. De $ab' = a'b$ y $cd' = c'd$ resulta

$$(ac)(b'd') = (ab')(cd') = (a'b)(c'd) = (a'c')(bd).$$

Con denominadores productos no nulos, la igualdad prueba la equivalencia pedida. En cambio, la operación $\star$ no es bien definida. Tomemos dos representantes de la primera clase, $(1,2)$ y $(2,4)$, y la segunda entrada $[(1,1)]$. La regla propuesta produce $[(2,2)]$ usando el primero y $[(3,4)]$ usando el segundo. Estas clases son diferentes: $2 \cdot 4 = 8$, mientras que $2 \cdot 3 = 6$. Por tanto, la entrada del cociente era la misma pero la salida cambió; ésa es la prueba decisiva del defecto.

Soluciones 14.6–14.10. Solución 14.6. Para $x = [(-3,5)]$, el opuesto aditivo es $-x = [(3,5)]$; su inverso multiplicativo es $x^{-1} = [(5,-3)]$, o equivalentemente $[(-5,3)]$. Al sumar $x + (-x)$ obtenemos $[(-3 \cdot 5 + 5 \cdot 3, 25)] = [(0,25)] = 0$. El producto xx^{-1} se representa por $[(-15,-15)]$, que es la clase de uno. No podemos invertir el cero porque un representante suyo tiene numerador cero y al intercambiar coordenadas aparecería un denominador nulo;

además, la identidad $0z = 0$ en el cuerpo impide que $0z = 1$. Para $1 = [(1, 1)]$, la inversa es la propia unidad, pues $1 \cdot 1 = 1$.

Solución 14.7. Con los primeros representantes,

$$[(1, 2)] + [(-3, 4)] = [(1 \cdot 4 + 2(-3), 2 \cdot 4)] = [(-2, 8)].$$

Usando $(2, 4)$ y $(6, -8)$, que representan respectivamente $1/2$ y $-3/4$ (en el segundo caso ambos productos cruzados valen 24), obtenemos

$$[(2, 4)] + [(6, -8)] = [(2(-8) + 4 \cdot 6, 4(-8))] = [(8, -32)].$$

Las clases de $(-2, 8)$ y $(8, -32)$ son iguales, pues $(-2)(-32) = 8 \cdot 8 = 64$. Un error frecuente sería modificar sólo el signo del numerador al sustituir representantes: para conservar la clase deben seguir coincidiendo los productos cruzados.

Solución 14.8. Primero $[(1, 3)] + [(1, 4)] = [(7, 12)]$; al multiplicar por $[(1, 2)]$ resulta $[(7, 24)]$. Por separado los productos son $[(1, 6)]$ y $[(1, 8)]$, cuya suma es $[(14, 48)] = [(7, 24)]$. Para la demostración general, si $x = [(a, b)]$, $y = [(c, d)]$ y $z = [(e, f)]$, la expresión $x(y + z)$ tiene representante $(acf + ade, bdf)$. Los productos xy y xz admiten respectivamente los representantes (acf, bdf) y (ade, bdf); su suma es equivalente a $(acf + ade, bdf)$. La descomposición $a(cf + de) = acf + ade$ utiliza la distributividad de los enteros, ya establecida en el capítulo 13. La coincidencia numérica es un ejemplo; la identidad de clases obtenida con letras prueba la ley.

Solución 14.9. La clase $[(1, -2)]$ equivale a $[(-1, 2)]$, de modo que es negativa y por ello menor que $[(1, 3)]$, que es positiva. Pero si se usa sin cuidado la regla $ad < bc$ con $(1, -2)$ y $(1, 3)$, se obtiene $3 < -2$, que es falso. La formulación correcta para representantes con $b, d > 0$ es

$$[(a, b)] < [(c, d)] \iff ad < bc.$$

Toda clase admite un representante de denominador positivo: cuando el denominador es negativo, multiplicamos ambas coordenadas por -1. La equivalencia se conserva y el orden puede compararse de forma consistente.

Solución 14.10. Los productos cruzados de (a, b) y (ka, kb) son $a(kb) = b(ka)$ por conmutatividad y asociatividad; $kb \neq 0$ porque $k, b \neq 0$ y $\mathbb{Z}$ no tiene divisores de cero. Si $k = 0$, el nuevo par sería $(0, 0)$ y quedaría fuera de D. Si $k < 0$, el denominador cambia de signo, por lo que una fórmula de comparación de productos cruzados derivada para denominadores positivos no se puede aplicar directamente. El valor racional permanece idéntico, pero para usar el criterio $ad < bc$ entre clases debemos escoger representantes de denominador positivo (o justificar por separado el signo del producto de denominadores).

Soluciones 14.11–14.15. Solución 14.11. Como $x < y$, la diferencia $y - x$ es positiva. Para cualquier w,

$$(y + w) - (x + w) = y - x > 0,$$

por lo que $x + w < y + w$. Si $w > 0$, el producto $(y - x)w$ es positivo, pues el producto de dos racionales positivos posee representantes de numerador y denominador positivos. Por distributividad y existencia de opuestos,

$$(y - x)w = yw - xw > 0,$$

lo cual implica $xw < yw$. Si $w < 0$, entonces $-w > 0$ y $x(-w) < y(-w)$; al tomar opuestos se obtiene $yw < xw$, es decir, $xw > yw$. La distributividad racional —que demostramos a partir de las leyes enteras— es el paso que enlaza la diferencia con los productos.

Solución 14.12. La hipótesis $a/b < c/d$, con $b, d > 0$, equivale a $ad < bc$. Como $b(b+d)$ es positivo,

$$\frac{a}{b} < \frac{a+c}{b+d} \iff a(b+d) < b(a+c) \iff ad < bc.$$

Del mismo modo,

$$\frac{a+c}{b+d} < \frac{c}{d} \iff d(a+c) < c(b+d) \iff ad < bc.$$

Para $3/7 < 4/9$ tenemos $3 \cdot 9 = 27 < 28 = 4 \cdot 7$ y el mediante $7/16$, con $3/7 < 7/16 < 4/9$ porque $3 \cdot 16 = 48 < 49 = 7 \cdot 7$ y $7 \cdot 9 = 63 < 64 = 4 \cdot 16$. Repetir la construcción proporciona infinitos racionales intermedios, pero la existencia de puntos intermedios no garantiza que un conjunto acotado tenga una menor cota superior. El conjunto S del §14.5 demuestra precisamente esa diferencia.

Solución 14.13. Sea $x = [(a, b)]$ con $b > 0$; como b es un entero positivo, $b \geq 1$. Tomemos $n = |a| + 1$, que es un natural positivo. Si $a \leq 0$, se tiene $x \leq 0 < n$. Si $a > 0$, entonces $a < |a| + 1 = n$ y, como $b \geq 1$, tenemos $ab^{-1} \leq a < n$ en el orden racional. Por consiguiente existe un natural mayor que x y, con más razón, un racional mayor. No hay racional máximo. La propiedad probada compara cualquier racional con naturales suficientemente grandes; la densidad compara dos racionales ya dados y produce otro entre ambos. Son dos afirmaciones de naturaleza distinta.

Solución 14.14. Si $x^2 = 2$, la clase x no es cero y puede representarse como $[(a, b)]$ con $b > 0$; la ecuación equivale a $a^2 = 2b^2$. Los denominadores positivos de representantes de esta clase constituyen un subconjunto no vacío de $\mathbb{N}_{>0}$, por lo que el buen orden permite elegir uno mínimo, digamos b. La igualdad muestra que a^2 es par. Si a fuera impar, $a = 2k + 1$ y $a^2 = 2(2k^2 + 2k) + 1$ sería impar; por ello $a = 2c$. Sustituyendo se obtiene $4c^2 = 2b^2$, o $b^2 = 2c^2$, de donde $b = 2d$ para algún entero $d > 0$. Entonces (c, d) pertenece a D, representa la misma clase y tiene un denominador positivo $d < b$. La contradicción no depende de que sepamos previamente qué número debería ocupar el lugar de la raíz cuadrada: sale de la imposibilidad de tener un representante mínimo y, al mismo tiempo, otro con denominador menor.

Solución 14.15. Los cálculos por denominador común dan

$$t - r = \frac{2(r+1) - r(r+2)}{r+2} = \frac{2 - r^2}{r+2},$$
$$2 - t^2 = \frac{2(r+2)^2 - 4(r+1)^2}{(r+2)^2} = \frac{2(2 - r^2)}{(r+2)^2}.$$

Para $r = 1$, resulta $t = 4/3 > 1$ y $t^2 = 16/9 < 2$. Para $r = 3/2$, resulta $t = 10/7 < 3/2$ y $t^2 = 100/49 > 2$. Supongamos que r es el supremo racional de S. Como $1 \in S$, es $r \geq 1$; por la imposibilidad de $r^2 = 2$, quedan los casos $r^2 < 2$ y $r^2 > 2$. En el primero, $t > r$ y $t^2 < 2$, así que $t \in S$, contradiciendo que r sea cota superior. En el segundo, $t < r$, $t > 0$ y $t^2 > 2$; si hubiera $s \in S$ con $s \geq t$, el orden del cuadrado en los racionales no negativos daría $s^2 \geq t^2 > 2$, contradicción. Así t es una cota superior estrictamente menor que r. Como ambos casos contradicen la definición de supremo, S carece de supremo en $\mathbb{Q}$. La mera inexistencia de una solución racional de $r^2 = 2$ no prueba estas dos mejoras de los candidatos; por eso se necesitan los cálculos con t.

Epílogo — Cuando la aproximación pide nuevos números

Busquemos un número positivo cuyo cuadrado sea 2. Sabemos ya que 1 resulta demasiado pequeño y 2 demasiado grande. El punto que buscamos, si existe, debe estar entre ambos. Podemos estrechar ese intervalo sin recurrir a ninguna operación que no conozcamos: tomamos su punto medio, $3/2$, observamos que su cuadrado supera a 2 y conservamos la mitad izquierda. El nuevo intervalo es $[1, 3/2]$. Su punto medio, $5/4$, tiene cuadrado menor que 2; conservamos entonces $[5/4, 3/2]$. Podemos continuar del mismo modo, comparando siempre el cuadrado de una fracción con 2.

Cada decisión se toma dentro de $\mathbb{Q}$. Si llamamos a_n y b_n a los extremos después de n divisiones, obtenemos intervalos encajados cuyos extremos son racionales y cumplen

$$a_n^2 < 2 < b_n^2, \qquad b_n - a_n = 2^{-n}.$$

El intervalo inicial corresponde a $n = 0$. Ningún punto medio detiene el procedimiento, pues en el capítulo anterior demostramos que un racional no puede tener cuadrado igual a 2. Al cabo de muchas divisiones, los extremos están extraordinariamente próximos. La diferencia entre ellos puede hacerse menor que cualquier racional positivo que fijemos de antemano.

Sigamos sólo los extremos izquierdos. Forman una sucesión de racionales, y para $m \geq n$ ambos términos a_n y a_m pertenecen al intervalo $[a_n, b_n]$. Por tanto,

$$|a_m - a_n| \leq 2^{-n}.$$

Esta desigualdad dice algo preciso: a partir de una etapa suficientemente avanzada, todos los términos posteriores están tan cerca unos de otros como deseemos. Una sucesión con esta propiedad recibe el nombre de **sucesión de Cauchy**. El nombre nos será útil más adelante; por el momento importa reconocer qué hemos logrado. Hemos encontrado un procedimiento racional que controla su propio error, aunque los racionales no contengan el resultado al que parece dirigirse.

En efecto, si los extremos izquierdos convergieran a un racional r, los derechos tendrían el mismo límite, porque su diferencia tiende a cero. Las desigualdades $a_n^2 < 2 < b_n^2$, junto con las reglas elementales de los límites para sumas y productos, obligarían a que $r^2 = 2$. Ya sabemos que eso es imposible dentro de $\mathbb{Q}$. Nuestra sucesión puede estabilizar sus aproximaciones con toda la precisión imaginable y, sin embargo, carecer de límite racional.

En el capítulo 14 encontramos la misma dificultad al buscar el supremo del conjunto de racionales positivos cuyos cuadrados son menores que 2. Podemos aproximar un borde desde dentro y demostrar que ninguna fracción ocupa ese lugar. La **completitud** será la propiedad que permita comprender qué le falta a $\mathbb{Q}$ y por qué la ampliación siguiente necesita algo más que nuevas reglas de cálculo. En el sistema de los números reales, ese conjunto tendrá un supremo, al que llamaremos $\sqrt{2}$; las aproximaciones anteriores podrán converger a él.

También los reales necesitan una construcción que justifique su existencia. Una posibilidad parte precisamente de sucesiones racionales de Cauchy y reúne en clases las que representan una misma aproximación; otra utiliza cortes del orden racional. Hemos aprendido en este libro qué exige un procedimiento semejante: habrá que definir las relaciones pertinentes, demostrar sus propiedades, comprobar que las operaciones no dependen de los representantes y establecer las propiedades que hacen del resultado un sistema numérico adecuado. En

Análisis para matemáticos, esa exigencia conducirá a una construcción concreta mediante **cortaduras de Dedekind**: los reales surgirán allí de conjuntos de racionales cuidadosamente definidos, y su propiedad de completitud deberá ser demostrada.

En *Cálculo para matemáticos* comenzaremos precisamente con los números reales, considerados a partir de los axiomas de cuerpo, del orden y de la completitud. Desde esa base podremos estudiar qué sucede con las pendientes de secantes cuando sus puntos se acercan, cómo se determina una velocidad instantánea y de qué manera las sumas de áreas conducen a la integral. Las construcciones de este libro permiten comprender de dónde provienen los números con los que trabajaremos; el cálculo investigará el cambio y la acumulación en ese nuevo ámbito.

En *Análisis para matemáticos* podremos volver a la construcción de la recta completa y demostrar por qué las cortaduras racionales permiten recuperar el supremo que faltaba en $\mathbb{Q}$. Con los reales establecidos, las preguntas sobre sucesiones, distancias, continuidad y compacidad adquirirán un alcance más amplio: servirán para estudiar qué significa aproximarse y en qué circunstancias es legítimo afirmar que existe un límite. La sucesión de intervalos de este epílogo será entonces uno de los primeros ejemplos de una teoría mucho más general.

La sucesión de intervalos con la que empezamos permanece como una pequeña promesa. Cada extremo está construido con los números que ya conocemos; cada paso se justifica mediante una comparación que sabemos realizar. Lo que falta es dar existencia matemática al destino de esas aproximaciones. Esa exigencia abre el camino de los reales y, con ellos, el del cálculo y el análisis.

Apéndice A — Glosario lógico y demostrativo

Este glosario sirve para recuperar una distinción al leer una prueba. La sección indicada contiene la exposición y sus ejemplos.

Término	Sentido en este libro	Recuperación
Proposición	Afirmación susceptible de verdad o falsedad una vez fijada su interpretación.	§1.2
Fórmula abierta	Expresión cuyo valor de verdad depende de una asignación a sus variables libres.	§3.1
Dominio	Conjunto en el que se permiten los objetos o las asignaciones considerados.	§3.2
Hipótesis	Afirmación disponible como dato de una demostración, con su alcance declarado.	§5.3
Validez	Imposibilidad de que las premisas sean verdaderas y la conclusión falsa en una interpretación admitida.	§2.2
Implicación	Relación expresada por un condicional; se distingue del acto de inferir una conclusión.	§2.3
Equivalencia	Dos afirmaciones que se implican en ambas direcciones bajo el contexto fijado.	§6.3
Variable libre o ligada	Libre: recibe valor del contexto. Ligada: está bajo un cuantificador que determina su uso.	§3.4
Testigo	Objeto del dominio que satisface la propiedad de una afirmación existencial.	§3.3
Unicidad	A lo sumo un objeto satisface la propiedad; para existencia única hay que demostrar además existencia.	§3.7
Contraejemplo	Instancia admisible que hace falsa una afirmación universal.	§9.1

Término	Sentido en este libro	Recuperación
Definición	Fija el significado de una expresión; no garantiza por sí sola la existencia de objetos que la satisfagan.	§5.2
Buena definición	Una expresión designa un único objeto del tipo anunciado para cada elección admisible de sus parámetros; cuando hay distintas representaciones, el resultado es independiente de la elegida.	§1.4; §11.7
Inducción	Principio para demostrar una propiedad de los naturales mediante base y paso.	§8.2
Recursión	Definición de objetos por un valor inicial y una regla sucesiva autorizada.	§8.8
Auditabilidad	Posibilidad de reconstruir y verificar las dependencias de un escrito con los antecedentes declarados.	§10.1

Los términos estructurales «familia», «producto», «cociente» y «numerable» se consultan en el capítulo 11. Convención global: $\mathbb{N} = \{0, 1, 2, \dots\}$. Un índice inicial diferente se declara expresamente.

Apéndice B — Formas de objetivo y estrategias iniciales

La forma del objetivo permite reconocer qué obligaciones hay que cerrar. La elección de una idea concreta requiere además leer las hipótesis. Esta tabla remite al desarrollo de los capítulos 6 y 7.

Objetivo	Obligación	Primer paso posible	Recuperación
$P \Rightarrow Q$	Obtener Q bajo P.	Fijar los datos de P y desplegar la definición pertinente.	§6.2
$P \Leftrightarrow Q$	Ambas implicaciones.	Separar direcciones y sus hipótesis disponibles.	§6.3
$A = B$	Ambas inclusiones, si son conjuntos.	Tomar un elemento arbitrario de uno y traducir pertenencia.	§6.3
$\forall x \in D\ P(x)$	Ningún elemento admisible queda fuera.	Fijar $x \in D$ arbitrario, sin añadir propiedades.	§3.2
$\exists x \in D\ P(x)$	Un testigo admisible y verificado.	Examinar qué restricciones permitirían construirlo.	§6.7
$\exists! x \in D\ P(x)$	Existencia y a lo sumo uno.	Separar construcción y comparación de dos candidatos.	§6.8
Propiedad para todo natural	Base y paso o una arquitectura equivalente adecuada.	Declarar inicio y qué índices permite usar la hipótesis.	§8.4
Refutar un universal	Un caso con hipótesis verdadera y conclusión falsa.	Probar bordes del dominio y supuestos frágiles.	§9.2

Si no aparece un camino, §7.3 permite buscar una condición suficiente; §7.7 ayuda a localizar el bloqueo. Un intento exploratorio sólo se incorpora a la prueba final cuando sus dependencias quedan justificadas.

Apéndice C — Errores frecuentes y pruebas defectuosas

Las pruebas defectuosas y sus reparaciones están desarrolladas en el capítulo 9. Este índice permite elegir el lugar que conviene releer; no añade un segundo catálogo de ejemplos.

Síntoma en el escrito	Qué comprobar	Recuperación
Se usa un condicional en sentido inverso.	Distinguir recíproco y contraposición.	§9.3
Se fija un testigo antes de los parámetros de que depende.	Orden y alcance de los cuantificadores.	§9.4
Se cancela, divide o aplica una función sin verificar condiciones.	Dominio, no nulidad y restricciones de la operación.	§9.5
Una transformación produce candidatos adicionales.	Distinguir implicación y equivalencia; verificar en el enunciado inicial.	§9.6
La razón invocada contiene la conclusión.	Dependencias y circularidad.	§9.7
Muchos casos calculados se presentan como prueba universal.	Qué casos cubre realmente el argumento.	§9.8
Una letra cambia de función o un pronombre pierde referente.	Nombres estables y alcance.	§10.7
«Claramente» oculta el paso decisivo.	Escribir la razón que autoriza la inferencia.	§10.8

Una corrección debe volver a comprobarse desde el enunciado: cambiar una frase puede reparar la exposición, pero cambiar una hipótesis cambia el problema. El pase completo de revisión se encuentra en §10.9.

Apéndice D — Directorio de soluciones desarrolladas

Las soluciones forman parte de cada capítulo. Este directorio permite acceder a ellas sin reproducir un segundo corpus. Conviene intentar el ejercicio antes de consultar su solución.

Capítulo	Pares finales	Acceso
1	12	Soluciones del capítulo 1
2	12	Soluciones del capítulo 2
3	12	Soluciones del capítulo 3
4	15	Soluciones del capítulo 4
5	15	Soluciones del capítulo 5
6	15	Soluciones del capítulo 6
7	15	Soluciones del capítulo 7
8	15	Soluciones del capítulo 8
9	15	Soluciones del capítulo 9
10	15	Soluciones del capítulo 10
11	15	Soluciones del capítulo 11
12	15	Soluciones del capítulo 12
13	15	Soluciones del capítulo 13
14	15	Soluciones del capítulo 14

Las comprobaciones intermedias conservan sus respuestas en el lugar donde se plantean.

Esta versión amplía el directorio editorial a **201 ejercicios finales y 201 soluciones**, distribuidos en 14 capítulos.

GNU Free Documentation License

GNU Free Documentation License
Version 1.3, 3 November 2008

Copyright (C) 2000, 2001, 2002, 2007, 2008 Free Software Foundation, Inc. https://fsf.org/
Everyone is permitted to copy and distribute verbatim copies of this license document, but changing it is not allowed.

0. PREAMBLE

The purpose of this License is to make a manual, textbook, or other functional and useful document "free" in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondarily, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

1. APPLICABILITY AND DEFINITIONS

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The "Document", below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as "you". You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A "Modified Version" of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A "Secondary Section" is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document's overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The "Invariant Sections" are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed

to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The "Cover Texts" are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A "Transparent" copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not "Transparent" is called "Opaque".

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The "Title Page" means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.

The "publisher" means any person or entity that distributes copies of the Document to the public.

A section "Entitled XYZ" means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as "Acknowledgements", "Dedications", "Endorsements", or "History".) To "Preserve the Title" of such a section when you modify the Document means that it remains a section "Entitled XYZ" according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

2. VERBATIM COPYING

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

3. COPYING IN QUANTITY

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

4. MODIFICATIONS

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission. B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement. C. State on the Title page the name of the publisher of the Modified Version, as the publisher. D. Preserve all the copyright notices of the Document. E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices. F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below. G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice. H. Include an unaltered copy of this License. I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence. J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise

the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission. K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein. L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles. M. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version. N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section. O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties–for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

5. COMBINING DOCUMENTS

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled "History" in the various original documents, forming one section Entitled "History"; likewise combine any sections Entitled "Acknowledgements", and any sections Entitled "Dedications". You must delete all sections Entitled "Endorsements".

6. COLLECTIONS OF DOCUMENTS

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

7. AGGREGATION WITH INDEPENDENT WORKS

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an "aggregate" if the copyright resulting from the compilation is not used to limit the legal rights of the compilation's users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document's Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

8. TRANSLATION

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled "Acknowledgements", "Dedications", or "History", the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

9. TERMINATION

You may not copy, modify, sublicense, or distribute the Document except as expressly provided under this License. Any attempt otherwise to copy, modify, sublicense, or distribute it is void, and will automatically terminate your rights under this License.

However, if you cease all violation of this License, then your license from a particular copyright holder is reinstated (a) provisionally, unless and until the copyright holder explicitly and finally terminates your license, and (b) permanently, if the copyright holder fails to notify you of the violation by some reasonable means prior to 60 days after the cessation.

Moreover, your license from a particular copyright holder is reinstated permanently if the copyright holder notifies you of the violation by some reasonable means, this is the first time you have received notice of violation of this License (for any work) from that copyright holder, and you cure the violation prior to 30 days after your receipt of the notice.

Termination of your rights under this section does not terminate the licenses of parties who have received copies or rights from you under this License. If your rights have been terminated and not permanently reinstated, receipt of a copy of some or all of the same material does not give you any rights to use it.

10. FUTURE REVISIONS OF THIS LICENSE

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See https://www.gnu.org/licenses/.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License "or any later version" applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation. If the Document specifies that a proxy can decide which future versions of this License can be used, that proxy's public statement of acceptance of a version permanently authorizes you to choose that version for the Document.

11. RELICENSING

"Massive Multiauthor Collaboration Site" (or "MMC Site") means any World Wide Web server that publishes copyrightable works and also provides prominent facilities for anybody to edit those works. A public wiki that anybody can edit is an example of such a server. A "Massive Multiauthor Collaboration" (or "MMC") contained in the site means any set of copyrightable works thus published on the MMC site.

"CC-BY-SA" means the Creative Commons Attribution-Share Alike 3.0 license published by Creative Commons Corporation, a not-for-profit corporation with a principal place of business in San Francisco, California, as well as future copyleft versions of that license published by that same organization.

"Incorporate" means to publish or republish a Document, in whole or in part, as part of another Document.

An MMC is "eligible for relicensing" if it is licensed under this License, and if all works that were first published under this License somewhere other than this MMC, and subsequently incorporated in whole or in part into the MMC, (1) had no cover texts or invariant sections, and (2) were thus incorporated prior to November 1, 2008.

The operator of an MMC Site may republish an MMC contained in the site under CC-BY-SA on the same site at any time before August 1, 2009, provided the MMC is eligible for relicensing.

ADDENDUM: How to use this License for your documents

To use this License in a document you have written, include a copy of the License in the document and put the following copyright and license notices just after the title page:

```
Copyright (c)  YEAR  YOUR NAME.
Permission is granted to copy, distribute and/or modify this document
under the terms of the GNU Free Documentation License, Version 1.3
or any later version published by the Free Software Foundation;
with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts.
A copy of the license is included in the section entitled "GNU
Free Documentation License".
```

If you have Invariant Sections, Front-Cover Texts and Back-Cover Texts, replace the "with... Texts." line with this:

```
with the Invariant Sections being LIST THEIR TITLES, with the
Front-Cover Texts being LIST, and with the Back-Cover Texts being LIST.
```

If you have Invariant Sections without Cover Texts, or some other combination of the three, merge those two alternatives to suit the situation.

If your document contains nontrivial examples of program code, we recommend releasing these examples in parallel under your choice of free software license, such as the GNU General Public License, to permit their use in free software.