Capítulo 1

Teoría axiomática de conjuntos

En este capítulo se presenta la teoría axiomática de conjuntos, siguiendo el sistema de axiomas de Zermelo y Fraenkel (ZF).

Introducción

Históricamente, la teoría de conjuntos surgió en los años 1870 a partir de los trabajos de Georg Cantor (1845–1918) en Análisis. Es estudiando las propiedades de las series trigonométricas que Cantor fue conducido a introducir números enteros más allá de los enteros naturales: los *números ordinales*. En este estudio, Cantor también descubrió que los conjuntos infinitos de números reales no tienen todos el mismo "tamaño"; es así que introdujo la jerarquía (infinita) de los *cardinales infinitos* y que empezó a estudiar la noción de conjunto en sí misma, con el fin de clasificar los conjuntos infinitos. En particular, es en 1878 que enunció su conjetura más notable, la *hipótesis del continuo*, que fue resuelta sólo en 1963 por Paul Cohen (1934–2007, medalla Fields 1966), utilizando trabajos anteriores de Kurt Gödel (1906–1978).

Rápidamente, Cantor y sus sucesores entendieron que la noción de conjunto era tan general que permitía reconstruir todos los objetos matemáticos conocidos como conjuntos puros. En 1872, Richard Dedekind (1831–1916) ya había realizado un paso notable, mostrando cómo reconstruir los números reales como conjuntos particulares de números racionales: las *cortaduras de Dedekind*. La teoría de conjuntos así permitía realizar un sueño antiguo de los matemáticos: unificar todas las ramas de la matemática en una teoría única[1].

Sin embargo, la teoría de conjuntos naciente todavía contenía paradojas (en particular: la paradoja de Buralli-Forti, 1897), que se trataba de eliminar mediante una axiomatización adecuada. La primera axiomatización de la teoría de conjuntos fue propuesta en 1903 por Gottlob Frege (1848–1925) —el fundador de la lógica moderna, a quién se debe el *cálculo de predicados*. Desgraciadamente, esta primera axiomatización era inconsistente, como lo mostró Bertrand Russell (1872–1970). Una axiomatización corregida fue propuesta en 1908 por Ernst Zermelo (1871–1953) —a quién se debe notablemente el *axioma de elección*— y completada en 1922 por Abraham Fraenkel (1891–1965) y Thoralf Skolem (1887–1963) —que introdujeron independientemente el *esquema de remplazo*. Así nació la teoría de conjuntos moderna, dicha de Zermelo-Fraenkel (notación: ZF).

[1]De hecho, la teoría de conjuntos no es el único marco unificador posible. También se puede basar la matemática sobre la teoría de tipos (Russell 1910, Martin-Löf 1984), cuyos objetos fundamentales son las funciones.

Hoy no se sabe si la teoría de conjuntos de Zermelo-Fraenkel es consistente o no —y no se puede demostrar que es consistente en razón de los límites fundamentales puestos por el segundo teorema de incompletitud de Gödel (1931)—, pero en casi cien años de existencia, nunca se encontró ninguna contradicción en dicha teoría.

1.1. Conceptos primitivos

1.1.1. El universo $\mathscr{U}$ de los conjuntos

En lo siguiente, se trabaja adentro de un universo $\mathscr{U}$ no vacío (en el sentido intuitivo), cuyos objetos son llamados *conjuntos*. Evitaremos decir que el universo $\mathscr{U}$ es un conjunto, pues los conjuntos son por definición los objetos adentro de $\mathscr{U}$, que no hay que confundir con el propio universo $\mathscr{U}$. En lugar de eso, diremos a veces que $\mathscr{U}$ es la *colección* o la *clase* de todos los conjuntos, utilizando estas expresiones con su sentido intuitivo.

La estructura del universo $\mathscr{U}$ es descrita mediante dos relaciones primitivas:

- La relación de igualdad $x = y$ («x es igual a y»), que es regida por las reglas usuales de la igualdad en matemática (véase Sección 1.1.3).
- La relación de pertenencia $x \in y$ («x pertenece a y»), que es regida por los axiomas de Zermelo y Fraenkel dados en las Secciones 1.2–1.7 más abajo.

Así, la teoría de conjuntos de Zermelo-Fraenkel es una teoría de *conjuntos puros* en la cual sólo se manipulan conjuntos; y cuando se escribe $x = y$ o $x \in y$, se entiende implícitamente que los dos objetos x, y son conjuntos. En particular, la relación de pertenencia $x \in y$ es una relación entre dos conjuntos, lo que significa que por diseño, los elementos de un conjunto sólo pueden ser (otros) conjuntos, y esto recursivamente. En la práctica, esta restricción no hace ningún problema, pues veremos que todos los objetos matemáticos usuales (los enteros, los reales, las n-uplas, las funciones, etc.) pueden ser reconstruidos como conjuntos puros adentro de $\mathscr{U}$. De tal modo que el universo $\mathscr{U}$ pueda ser considerado más generalmente como el universo de todos los objetos matemáticos, y la palabra *conjunto* como sinónimo de *objeto matemático*.

1.1.2. El lenguaje de la teoría de conjuntos

En la teoría de conjuntos, sólo se consideran aserciones matemáticas —o *fórmulas*— construidas a partir de las dos relaciones primitivas $x = y$ y $x \in y$ mediante las conectivas lógicas $\neg$ (negación), $\land$ (conjunción), $\lor$ (disyunción), $\Rightarrow$ (implicación) y $\Leftrightarrow$ (equivalencia lógica), así como de los cuantificadores $\forall$ (universal) y $\exists$ (existencial); el dominio de cada cuantificación siendo implícitamente el universo $\mathscr{U}$.

Todas las otras notaciones de la teoría de conjuntos (y más generalmente de la matemática) son definidas a partir de este lenguaje de base mediante abreviaturas. Por ejemplo, las versiones negadas de la igualdad y de la pertenencia son definidas por:

$$x \neq y \;\equiv\; \neg(x = y) \qquad \text{y} \qquad x \notin y \;\equiv\; \neg(x \in y)$$

Y las dos cuantificaciones restringidas a un conjunto a son definidas por

$$(\forall x \in a)\, \phi(x) \;\equiv\; \forall x\, (x \in a \Rightarrow \phi(x)) \qquad \text{(«para todo } x \text{ en } a,\ \phi(x)\text{»)}$$
$$(\exists x \in a)\, \phi(x) \;\equiv\; \exists x\, (x \in a \land \phi(x)) \qquad \text{(«existe } x \text{ en } a \text{ tal que } \phi(x)\text{»)}$$

donde $\phi(x)$ es cualquier fórmula (del lenguaje de la teoría de conjuntos) que depende de x. Así, cada aserción matemática, como por ejemplo la aserción

existe un cuerpo totalmente ordenado, arquimediano y completo,

se puede desplegar en una fórmula del lenguaje de base, en general muy larga.

1.1.3. Reglas de la igualdad

En teoría de conjuntos como en cualquier teoría matemática, la relación de igualdad $x = y$ es regida por las dos reglas introducidas por Gottfried Leibniz (1646–1716):

- El *principio de identidad*, que expresa que todo objeto es igual a sí mismo:

  Identidad $\qquad \forall x\,(x = x)$

- El principio de *sustitución de los iguales*, que expresa que dos objetos iguales cumplen las mismas propiedades, es decir:

  Sustitución $\qquad \forall x\,\forall y\,(x = y \wedge \phi(x) \Rightarrow \phi(y))$

  donde $\phi(x)$ es cualquier fórmula del lenguaje que depende de la variable x.

Precisemos que en el esquema anterior, la fórmula $\phi(x)$ puede depender de otras variables, que tienen un papel de parámetros, y son implícitamente cuantificadas universalmente.

Las reglas de la igualdad implican en particular que:

Proposición 1.1 (Relación de equivalencia). — *La igualdad es una relación de equivalencia:*

(Reflexividad)	$\forall x\,(x = x)$
(Simetría)	$\forall x\,\forall y\,(x = y \Rightarrow y = x)$
(Transitividad)	$\forall x\,\forall y\,\forall z\,(x = y \wedge y = z \Rightarrow x = z)$

Demostración. (Reflexividad) Es dada por la regla de identidad.

(Simetría) Sea la fórmula $\phi(x) \equiv x = z$, parametrizada por un conjunto z cualquiera. Por la regla de sustitución, tenemos que $x = y \wedge x = z \Rightarrow y = z$. Como esta fórmula vale para todo z, se puede sustituir z por x, lo que nos da: $x = y \wedge x = x \Rightarrow y = x$. Pero como $x = x$ siempre se cumple (principio de identidad), la fórmula anterior se simplifica en $x = y \Rightarrow y = x$.

(Transitividad) Ya vimos que $x = y \wedge x = z \Rightarrow y = z$ para todos x, y, z (por el principio de sustitución), es decir $y = x \wedge y = z \Rightarrow x = z$ (intercambiando las variables x e y). Utilizando la simetría de la igualdad entre x e y, se deduce que $x = y \wedge y = z \Rightarrow x = z$. □

1.2. Axioma de extensionalidad

1.2.1. Motivación

La regla de sustitución de los iguales expresa que dos conjuntos iguales cumplen las mismas propiedades, lo que implica en particular que tienen los mismos elementos:

$$\forall a\,\forall b\,[a = b \;\Rightarrow\; \forall x\,(x \in a \Leftrightarrow x \in b)]\,.$$

En cambio, la recíproca de este teorema no es consecuencia de las reglas de la igualdad, y es la razón por la cual constituye el primer axioma de la teoría de conjuntos:

Axioma 1 (Extensionalidad). — *Dos conjuntos que tienen los mismos elementos son iguales:*

$$\forall a\, \forall b\, [\forall x\, (x \in a \Leftrightarrow x \in b) \;\Rightarrow\; a = b]\,.$$

Intuitivamente, el axioma de extensionalidad expresa que un conjunto sólo depende de su «contenido», y no de la manera en que es construido (el «continente»)[2].

1.2.2. Relación de inclusión

En teoría de conjuntos, la relación de inclusión $a \subseteq b$ es definida mediante la abreviatura

$$a \subseteq b \;\equiv\; \forall x\, (x \in a \Rightarrow x \in b)\,.$$

Se demuestra que:

Proposición 1.2 (Orden de inclusión). — *La inclusión es una relación de orden:*

(Reflexividad)	$\forall a\, (a \subseteq a)$
(Transitividad)	$\forall a\, \forall b\, \forall c\, (a \subseteq b \wedge b \subseteq c \Rightarrow a \subseteq c)$
(Antisimetría)	$\forall a\, \forall b\, \forall c\, (a \subseteq b \wedge b \subseteq a \Rightarrow a = b)$

Demostración. Las propiedades de reflexividad y de transitividad son obvias (las fórmulas correspondientes son tautologías del cálculo de predicados), mientras la propiedad de antisimetría sigue inmediatamente del axioma de extensionalidad. □

De hecho, la propiedad de antisimetría de la inclusión es *tautológicamente* equivalente al axioma de extensionalidad, que también hubiéramos podido formular así:

Axioma 1 (Extensionalidad). — *La relación de inclusión es antisimétrica:*

$$\forall a\, \forall b\, (a \subseteq b \wedge b \subseteq a \Rightarrow a = b)\,.$$

1.2.3. Predicados colectivizantes

Se recuerda que un *predicado* es una fórmula $\phi(x)$ que depende de una variable x distinguida. (Tal fórmula puede depender de otras variables, que tienen un papel de parámetros.) Se dice que un predicado $\phi(x)$ es *colectivizante* cuando existe un conjunto cuyos elementos son exactamente los objetos x del universo que cumplen la propiedad $\phi(x)$:

$$\phi(x) \text{ colectivizante } \;\equiv\; \exists a\, \forall x\, (x \in a \Leftrightarrow \phi(x))\,.$$

Es claro que:

[2]Es fácil imaginar una teoría de conjuntos que no satisface la propiedad de extensionalidad: por ejemplo, una teoría de conjuntos «coloreados» donde cada conjunto lleva un color además de su contenido, de tal modo que se pueda construir dos conjuntos distintos con el mismo contenido, simplemente tiñéndolos con distintos colores. Los informáticos conocen muy bien este problema, pues la casi totalidad de los métodos de representación de los conjuntos finitos en la máquina (por ejemplo: con listas finitas o árboles finitos) permiten representar el mismo conjunto de diversas maneras, lo que necesita trabajar a través de une relación de equivalencia adecuada con el fin de ignorar las diferencias de representación. El mismo problema aparece en la teoría de modelos, donde muchos modelos de la teoría de conjuntos (por ejemplo: los modelos de P-nombres en *forcing*) naturalmente vienen sin la propiedad de extensionalidad. De nuevo, se necesita cocientar tales modelos con la relación de equivalencia adecuada, con el fin de restaurar la propiedad perdida.

Proposición 1.3 (Unicidad). — *Si un predicado $\phi(x)$ es colectivizante, entonces el conjunto a tal que $\forall x\,(x \in a \Leftrightarrow \phi(x))$ es único; se escribe $\{x : \phi(x)\}$.*

Demostración. Sean dos conjuntos a y a' tales que $\forall x\,(x \in a \Leftrightarrow \phi(x))$ y $\forall x\,(x \in a' \Leftrightarrow \phi(x))$. Esto implica que $\forall x\,(x \in a \Leftrightarrow x \in a')$, luego $a = a'$ por extensionalidad. □

Recordemos que la noción de conjunto fue introducida en matemática con el fin de manipular los predicados como «ciudadanos de primera clase», mediante el mecanismo que asocia a cada predicado $\phi(x)$ el conjunto $\{x : \phi(x)\}$. Gracias a este mecanismo de *cosificación*, el conjunto $\{x : \phi(x)\}$ se puede manipular como cualquier objeto matemático, y se puede estudiar mediante nuevos predicados, que también se pueden cosificar, con el fin de ser estudiados por otros predicados, etc.[3] Desgraciadamente, no se puede adoptar tal mecanismo de cosificación sin restricciones, pues existen predicados no colectivizantes:

Ejemplos 1.4 (Predicados colectivizantes y no colectivizantes).

(1) Dado un conjunto a cualquiera, es claro que el predicado $\phi(x) \equiv x \in a$ (parametrizado por a) es colectivizante, ya que $a = \{x : x \in a\}$.

(2) Sea $\phi(x)$ el predicado definido por $\phi(x) \equiv x \notin x$. Este predicado (debido a Russell) no es colectivizante, porque si existiera un conjunto a tal que

$$\forall x\,(x \in a \Leftrightarrow x \notin x)\,,$$

tendríamos en particular que $a \in a \Leftrightarrow a \notin a$, lo que es absurdo.[4]

Así, la primera tarea de una teoría de conjuntos es definir cuáles predicados son colectivizantes sin poner en peligro la consistencia de la teoría, mediantes *axiomas de existencia* adecuados. Es precisamente el objeto de los axiomas de ZF que siguen.

1.3. Axioma de pares

El primer axioma de existencia de ZF expresa que el predicado $\phi(x) \equiv x = a \lor x = b$ (parametrizado por dos objetos a, b cualesquiera) es colectivizante:

Axioma 2 (Pares). — *Para todos objetos a y b, existe un conjunto cuyos elementos son a y b:*

$$\forall a\,\forall b\,\exists c\,(x \in c \Leftrightarrow x = a \lor x = b)\,.$$

El conjunto c es único (por extensionalidad); le llama *el par no ordenado formado por a y b*, y se escribe $\{a, b\}$ (en lugar de $\{x : x = a \lor x = b\}$).

Se dice que el par $\{a, b\}$ es *no ordenado* pues cumple la igualdad $\{a, b\} = \{b, a\}$, que resulta de que ambos conjuntos tienen los mismos elementos. Más generalmente, se demuestra que:

[3] Desde el punto de vista de la lógica, una teoría de conjuntos sirve esencialmente a transformar objetos de orden superior —es decir: predicados sobre predicados sobre predicados, etc.— en objetos de primer orden.

[4] Históricamente, la primera axiomatización de la teoría de conjuntos por Frege (1903) sólo contenía el axioma de extensionalidad y un esquema de axiomas que permitía construir el conjunto $\{x : \phi(x)\}$ para cualquier predicado $\phi(x)$. Su inconsistencia fue demostrada por Russell utilizando el contra-ejemplo anterior.

Proposición 1.5 (Igualdad entre pares). — *Para todos a, b, c, d:*

$$\{a, b\} = \{c, d\} \;\Leftrightarrow\; (a = c \wedge b = d) \vee (a = d \wedge b = c)\,.$$

Demostración. ($\Leftarrow$) En el caso donde $a = c$ y $b = d$, tenemos que $\{a, b\} = \{c, d\}$ (por sustitución de los iguales). Y en el caso donde $a = d$ y $b = c$, tenemos que $\{a, b\} = \{d, c\}$ (ídem), lo que implica que $\{a, b\} = \{c, d\}$ (pues $\{d, c\} = \{c, d\}$).

($\Rightarrow$) Supongamos que $\{a, b\} = \{c, d\}$. De esta hipótesis, se deduce que:

- $a \in \{c, d\}$, entonces $a = c \vee a = d$;
- $b \in \{c, d\}$, entonces $b = c \vee b = d$;
- $c \in \{a, b\}$, entonces $c = a \vee c = b$;
- $d \in \{a, b\}$, entonces $d = a \vee d = b$.

Así se obtiene una conjunción de cuatro disyunciones, cada una con dos alternativas:

$$(a = c \vee a = d) \wedge (b = c \vee b = d) \wedge (c = a \vee c = b) \vee (d = a \vee d = b)$$

Distribuyendo las disyunciones con las conjunciones, se distinguen los siguientes 16 casos:

$$\begin{array}{llll}
(1) & a = c \wedge b = c \wedge c = a \wedge d = a & (9) & a = d \wedge b = c \wedge c = a \wedge d = a \\
(2) & a = c \wedge b = c \wedge c = a \wedge d = b & (10) & a = d \wedge b = c \wedge c = a \wedge d = b \\
(3) & a = c \wedge b = c \wedge c = b \wedge d = a & (11) & a = d \wedge b = c \wedge c = b \wedge d = a \\
(4) & a = c \wedge b = c \wedge c = b \wedge d = b & (12) & a = d \wedge b = c \wedge c = b \wedge d = b \\
(5) & a = c \wedge b = d \wedge c = a \wedge d = a & (13) & a = d \wedge b = d \wedge c = a \wedge d = a \\
(6) & a = c \wedge b = d \wedge c = a \wedge d = b & (14) & a = d \wedge b = d \wedge c = a \wedge d = b \\
(7) & a = c \wedge b = d \wedge c = b \wedge d = a & (15) & a = d \wedge b = d \wedge c = b \wedge d = a \\
(8) & a = c \wedge b = d \wedge c = b \wedge d = b & (16) & a = d \wedge b = d \wedge c = b \wedge d = b
\end{array}$$

Luego se observa que:

- En el caso (6), se deduce que $a = c \wedge b = d$.
- En el caso (11), se deduce que $a = d \wedge b = c$.
- En todos los otros casos, se deduce que $a = b = c = d$, lo que implica trivialmente que $(a = c \wedge b = d) \vee (a = d \wedge b = c)$ (ambas alternativas siendo verdaderas).

Al final, tenemos que $(a = c \wedge b = d) \vee (a = d \wedge b = c)$ en todos los casos. □

1.3.1. Conjuntos unitarios

Dado un objeto a cualquiera, el axioma de pares implica que el predicado $\phi(x) \equiv x = a$ (equivalente a $x = a \vee x = a$) es colectivizante. Luego, existe un (único) conjunto cuyo único elemento es a; se llama el *conjunto unitario formado por* a y se escribe $\{a\}$. Por construcción, tenemos que $\{a\} = \{a, a\}$, y es claro que $\{a\} = \{b\}$ si y sólo si $a = b$ para todos a y b.

1.3.2. El par ordenado

Dado dos objetos a y b cualesquiera, se define el par ordenado (a, b) a partir del par no ordenado (y del conjunto unitario) por[5]

(Par ordenado) $$(a, b) := \{\{a\}, \{a, b\}\}.$$

El interés de tal definición es que rompe la simetría entre a y b, lo que permite mantener el orden entre las dos componentes a y b:

Proposición 1.6 (Igualdad entre pares ordenados). — *Para todos a, b, c, d:*

$$(a, b) = (c, d) \;\Leftrightarrow\; a = c \wedge b = d\,.$$

Demostración. ($\Leftarrow$) Obvio, por sustitución de los iguales.
($\Rightarrow$) Supongamos que $(a, b) = (c, d)$, es decir: $\{\{a\}, \{a, b\}\} = \{\{c\}, \{c, d\}\}$. Según la Prop. 1.5 aplicada a la igualdad anterior, se distinguen dos casos:

- $\{a\} = \{c\}$ y $\{a, b\} = \{c, d\}$. La primera igualdad implica que $a = c$. Aplicando de nuevo la Prop. 1.5 a la segunda igualdad $\{a, b\} = \{c, d\}$, se distinguen los siguientes casos:
  - $a = c$ y $b = d$. Es lo que queríamos demostrar.
  - $a = d$ y $b = c$. Como $a = c$, tenemos que $a = b = c = d$, entonces $a = c$ y $b = d$.
- $\{a\} = \{c, d\}$ y $\{a, b\} = \{c\}$. Esto implica que $a = b = c = d$, luego $a = c$ y $b = d$.

En todos los casos, tenemos que $a = c$ y $b = d$. □

Más generalmente, se definen las ternas ordenadas, las cuádruplas, etc. por

$$(x, y, z) = (x, (y, z)), \qquad (x, y, z, u) = (x, (y, (z, u))), \qquad \text{etc.}$$

1.4. El esquema de comprensión

El siguiente axioma de ZF no es estrictamente hablando un axioma, pero un *esquema de axiomas*, que define un axioma particular para cada predicado $\phi(x)$:

Axioma 3 (Esquema de comprensíon). — *Para todo conjunto a, existe un conjunto b cuyos elementos son los elementos $x \in a$ que cumplen $\phi(x)$:*

$$\forall a\, \exists b\, \forall x\, (x \in b \;\Leftrightarrow\; x \in a \wedge \phi(x))$$

(donde $\phi(x)$ es cualquier predicado de la teoría de conjuntos).

El conjunto b definido así es único, y se escribe $\{x \in a : \phi(x)\}$.

Intuitivamente, el esquema de comprensión expresa que la intersección (en el sentido intuitivo) entre un conjunto a y una clase (definida por un predicado $\phi(x)$) forma un conjunto. Más generalmente, el esquema de comprensión permite demostrar que cualquier clase incluida (en el sentido intuitivo) en un conjunto forma un conjunto:

[5] Esta definición del par ordenado es debida a Kazimierz Kuratowski (1886–1980).

Proposición 1.7 (Criterio de colectivización). – *Un predicado $\phi(x)$ es colectivizante si y sólo si existe un conjunto a que contiene (al menos) todos los objetos x que cumplen $\phi(x)$:*

$$\phi(x) \text{ colectivizante } \Leftrightarrow \exists a \, \forall x \, (\phi(x) \Rightarrow x \in a) .$$

Demostración. ($\Rightarrow$) Si el predicado $\phi(x)$ es colectivizante, entonces el conjunto $a = \{x : \phi(x)\}$ satisface obviamente la propiedad deseada.
($\Leftarrow$) Supongamos que a es tal que $\forall x \, (\phi(x) \Rightarrow x \in a)$. Por comprensión, se define el conjunto $a' = \{x \in a : \phi(x)\}$, y se verifica inmediatamente que $a' = \{x : \phi(x)\}$. □

Una consecuencia importante del resultado anterior es que:

Corolario 1.8. — *No existe ningún conjunto de todos los conjuntos:*

$$\neg \exists U \, \forall x \, (x \in U) .$$

Demostración. En efecto, si existiera tal conjunto U, cualquier predicado sería colectivizante, incluso el predicado de Russell $\phi(x) \equiv x \notin x$, lo que es absurdo. □

1.4.1. Diferencia conjuntista y conjunto vacío

Dado dos conjuntos a y b, se define la *diferencia conjuntista* $a - b$ por

$$a - b \; := \; \{x \in a : x \notin b\} .$$

En el caso particular donde $a = b$, se observa que el conjunto $a - a$ es vacío, en el sentido que para todo x, tenemos que $x \notin (a - a)$. Además, como $a - a = b - b$ para todos a y b (por extensionalidad), el conjunto vacío $a - a$ no depende de a, y se escribe $\varnothing$. Otra definición posible (y equivalente) del conjunto vacío es $\varnothing := \{x \in a : x \neq x\}$.

El conjunto vacío es el conjunto más pequeño del universo en el sentido de la inclusión, y tenemos que $\varnothing \subseteq a$ para todo conjunto a. En particular, la inclusión $a \subseteq \varnothing$ implica la igualdad $a = \varnothing$ (por antisimetría). Más generalmente, para todos a y b tenemos la equivalencia:

$$a - b = \varnothing \;\Leftrightarrow\; a \subseteq b .$$

1.4.2. Intersección

Dado dos conjuntos a y b, se define la intersección binaria $a \cap b$ por

$$a \cap b \; := \; \{x \in a : x \in b\} .$$

Se verifica fácilmente que la operación $a \cap b$ es idempotente, conmutativa y asociativa:

$$a \cap a = a, \qquad a \cap b = b \cap a, \qquad (a \cap b) \cap c = a \cap (b \cap c) .$$

Además el conjunto vacío es elemento absorbente: $a \cap \varnothing = \varnothing \cap a = \varnothing$.

Más generalmente, si a es un conjunto (de conjuntos) no vacío, se llama la *intersección de todos los elementos de* a y se escribe $\bigcap a$ el conjunto definido por:

$$\bigcap a \; := \; \{x : (\forall y \in a)(x \in y)\} .$$

El conjunto $\bigcap a$ existe por la Prop. 1.7, pues la colección definida por el predicado $\phi(x) \equiv (\forall y \in a)(x \in y)$ está incluida en cualquier elemento del conjunto a, que es no vacío por hipótesis. En otro lado, cuando $a = \varnothing$, el predicado $\phi(x) \equiv (\forall y \in \varnothing)(x \in y)$ no es colectivizante (se cumple para todos los objetos del universo) y la intersección $\bigcap \varnothing$ no existe.

$$
\begin{array}{rcl@{\qquad\qquad}rcl}
a \cup a &=& a & a \cap a &=& a \\
a \cup b &=& b \cup a & a \cap b &=& b \cap a \\
(a \cup b) \cup c &=& a \cup (b \cup c) & (a \cap b) \cap c &=& a \cap (b \cap c) \\[1ex]
(a \cap b) \cup c &=& (a \cup c) \cap (b \cup c) & (a \cup b) \cap c &=& (a \cap c) \cup (b \cap c) \\
(a \cup b) - c &=& (a - c) \cup (b - c) & (a \cap b) - c &=& (a - c) \cap (b - c) \\
a - (b \cup c) &=& (a - b) \cap (a - c) & a - (b \cap c) &=& (a - b) \cup (a - c) \\[1ex]
(a - b) - c &=& a - (b \cup c) & a - (b - c) &=& (a - b) \cup (a \cap c) \\[1ex]
a \cup \varnothing &=& \varnothing \cup a \;=\; a & a \cap \varnothing &=& \varnothing \cap a \;=\; \varnothing \\[1ex]
a - \varnothing &=& a & a \subseteq b &\Leftrightarrow& a \cup b = b \\
\varnothing - a &=& \varnothing & &\Leftrightarrow& a \cap b = a \\
a - a &=& \varnothing & &\Leftrightarrow& a - b = \varnothing
\end{array}
$$

Cuadro 1.1: Propiedades de las operaciones binarias $a \cup b$, $a \cap b$ y $a - b$

1.5. Axioma de unión

Mientras la intersección (binaria y generalizada) se puede construir mediante el esquema de comprensión, la construcción de la unión necesita un axioma específico:

Axioma 4 (Unión)**.** — *Para todo conjunto a, existe un conjunto b cuyos elementos son los elementos de los elementos de a:*

$$\forall a\, \exists b\, \forall x\, (x \in b \;\Leftrightarrow\; (\exists y \in a)\, (x \in y))\,.$$

De nuevo, el conjunto b es único; se llama la *unión (de los elementos)* de a y se escribe $\bigcup a$.

La unión binaria $a \cup b$ es definida para todos conjuntos a y b por $a \cup b = \bigcup\{a, b\}$. Se verifica fácilmente que esta operación es idempotente, conmutativa, asociativa

$$a \cup a = a, \qquad a \cup b = b \cup a, \qquad (a \cup b) \cup c = a \cup (b \cup c)\,,$$

y que el conjunto vació es neutro: $a \cup \varnothing = \varnothing \cup a = a$. (Las propiedades notables de las tres operaciones $a \cup b$, $a \cap b$ y $a - b$ son recordadas en el Cuadro 1.1.)

Combinada con el axioma de pares, la operación de unión binaria también permite definir conjuntos finitos de tamaño arbitrario: $\{a, b, c\} = \{a, b\} \cup \{c\}$, $\{a, b, c, d\} = \{a, b, c\} \cup \{d\}$, etc.

1.6. Axioma del conjunto potencia

Axioma 5 (Conjunto potencia)**.** — *Para todo conjunto a, existe un conjunto b cuyos elementos son los subconjuntos de a:*

$$\forall a\, \exists b\, \forall x\, (x \in b \;\Leftrightarrow\; x \subseteq a)\,.$$

El conjunto b (que es único) se llama el *conjunto potencia de a* y se escribe $\mathfrak{P}(a)$.

El axioma del conjunto potencia es uno de los axiomas más potentes de ZF, que permite construir muchos conjuntos importantes en matemática, tales que el producto cartesiano, el conjunto cociente, así como todos los conjuntos usuales de funciones y de relaciones.

1.6.1. Producto cartesiano, grafos y relaciones

Proposición y definición 1.9 (Producto cartesiano)**.** — *Dado dos conjuntos A y B, existe un (único) conjunto, escrito $A \times B$, cuyos elementos son los pares (x, y) tales que $x \in A$ e $y \in B$:*

$$A \times B \ := \ \{z : (\exists x \in A)\,(\exists y \in B)\, z = (x, y)\}\,.$$

Demostración. Se aplica la Prop. 1.7 al predicado $\phi(z) \equiv (\exists x \in A)\,(\exists y \in B)\, z = (x, y)$, observando que $\phi(z)$ implica $z \in \mathfrak{P}(\mathfrak{P}(A \cup B))$ para todo z. □

Es claro que $A \times \varnothing = \varnothing \times B = \varnothing$ para todos A y B. En otro lado, la operación $A \times B$ no es ni conmutativa ($A \times B \neq B \times A$ en general) ni asociativa ($(A \times B) \times C \neq A \times (B \times C)$ en general), aunque existan biyecciones canónicas entre $A \times B$ y $B \times A$ y entre $(A \times B) \times C$ y $A \times (B \times C)$, anticipándose a las definiciones de las nociones de función y de biyección.

Grafos y relaciones Se llama un *grafo* cualquier conjunto de pares:

$$G \text{ grafo} \ \equiv \ (\forall z \in G)\, \exists x\, \exists y\, z = (x, y)\,.$$

Dado un grafo G, se escriben

$$\begin{array}{rcl} \mathrm{pr}_1(G) & := & \{x : \exists y\, (x, y) \in G\} \\ \mathrm{pr}_2(G) & := & \{y : \exists x\, (x, y) \in G\} \end{array}$$

(La existencia de ambos conjuntos $\mathrm{pr}_1(G)$ y $\mathrm{pr}_1(G)$ sigue de la Prop. 1.7, pues los dos predicados $\phi_1(x) \equiv \exists y\, (x, y) \in G$ y $\phi_2(y) \equiv \exists x\, (x, y) \in G$ utilizados para definirlos cumplen las implicaciones $\phi_1(x) \Rightarrow x \in \bigcup\bigcup G$ y $\phi_2(y) \Rightarrow y \in \bigcup\bigcup G$ para todos x e y.) Para todo grafo G, tenemos por construcción la inclusión $G \subseteq \mathrm{pr}_1(G) \times \mathrm{pr}_2(G)$.

Todo grafo G define una relación binaria, dada por la fórmula $\phi(x, y) \equiv (x, y) \in G$ (parametrizada por G). Recíprocamente, si $\phi(x, y)$ es una relación binaria (es decir: una formula que depende de dos variables distinguidas x e y, y posiblemente de otros parámetros), se dice que G es *el grafo de la relación* $\phi(x, y)$ cuando G es el conjunto de todos los pares (x, y) tales que $\phi(x, y)$. En este caso, el grafo G (que es único) se escribe $G = \{(x, y) : \phi(x, y)\}$.

Del mismo modo que un predicado no siempre define un conjunto, una relación binaria $\phi(x, y)$ no siempre admite un grafo. Por ejemplo, la relación de igualdad $x = y$, la relación de pertenencia $x \in y$ y la relación de inclusión $x \subseteq y$ no tienen grafo. (Si tal grafo G existiera, su primera proyección $\mathrm{pr}_1(G)$ sería el conjunto de todos los conjuntos.)

En cambio, si $\phi(x, y)$ es una relación binaria entre dos conjuntos A y B —es decir: si la relación $\phi(x, y)$ implica que $x \in A$ e $y \in B$ para todos x e y—, entonces esta relación admite un grafo, que es definido como un subconjunto del producto cartesiano $A \times B$. Así, como toda relación $\phi(x, y)$ entre A y B se puede representar por un subconjunto $G = \{(x, y) : \phi(x, y)\} \subseteq A \times B$, y como todo subconjunto $G \subseteq A \times B$ define a su vez una relación $\phi(x, y) \equiv (x, y) \in G$, se tiene la costumbre de identificar las relaciones entre A y B con los elementos de $\mathfrak{P}(A \times B)$.

1.6.2. Funciones y conjuntos de funciones

Se llama una *función* todo *grafo funcional*, es decir, todo conjunto de pares f tal que las dos condiciones $(x, y) \in f$ y $(x, y') \in f$ impliquen $y = y'$ (para todos x, y, y'). Formalmente:

$$\begin{array}{rcll} f \text{ función} & \equiv & (\forall z \in f)\, \exists x\, \exists y\, z = (x, y) \quad \wedge & (f \text{ es un grafo}) \\ & & \forall x\, \forall y\, \forall y'\, ((x, y) \in f \wedge (x, y') \in f \Rightarrow y = y') & (\text{el grafo } f \text{ es funcional}) \end{array}$$

Dado una función f, se llaman el *dominio* y la *imagen* de f los dos conjuntos $\mathrm{dom}(f)$ y $\mathrm{img}(f)$ definidos por $\mathrm{dom}(f) = \mathrm{pr}_1(f)$ y $\mathrm{img}(f) = \mathrm{pr}_2(f)$.

Para toda función f y para todo elemento $x \in \mathrm{dom}(f)$, existe un único elemento $y \in \mathrm{img}(f)$ tal que $(x, y) \in f$; este objeto y se llama la *imagen* de x por f, y se escribe $f(x)$. Un modo sencillo de definir la notación $f(x)$ consiste en escribir

$$f(x) \; := \; \bigcup \{y \in \mathrm{img}(f) : (x, y) \in f\}\,.$$

(En efecto, cuando $x \in \mathrm{dom}(f)$, el conjunto $\{y \in \mathrm{img}(f) : (x, y) \in f\}$ es unitario, y el operador de unión naturalmente extrae su único elemento.)

Funciones de A hasta B Dado dos conjuntos A y B, se llama una *función de A hasta B* toda función f tal que $\mathrm{dom}(f) = A$ y $\mathrm{img}(f) \subseteq B$. Se observa que las condiciones $\mathrm{dom}(f) = A$ y $\mathrm{img}(f) \subseteq B$ son asimétricas: mientras el dominio $A = \mathrm{dom}(f)$ es determinado por f, el codominio B es un conjunto *convencional* que no se puede deducir de f (sólo se requiere que $\mathrm{img}(f) \subseteq B$). En lo siguiente, la aserción «f es una función de A a B» se escribe $f : A \to B$:

$$f : A \to B \;\equiv\; f \text{ función} \land \mathrm{dom}(f) = A \land \mathrm{img}(f) \subseteq B\,.$$

Proposición y definición 1.10 (Conjunto de funciones)**.** — *Para todos conjuntos A y B, existe un conjunto de todas las funciones de A a B, que se escribe B^A.*

Demostración. Se aplica la Prop. 1.7 con el predicado $\phi(f) \equiv (f : A \to B)$, observando que la aserción $\phi(f)$ implica que $f \in \mathfrak{P}(A \times B)$ para todo f. □

Operaciones sobre las funciones Se dice que dos funciones f y g son componibles (en este orden) cuando $\mathrm{img}(f) \subseteq \mathrm{dom}(g)$. Cuando son, se define la *función compuesta $g \circ f$* por

$$(g \circ f)(x) \; := \; g(f(x)) \qquad \text{(para todo } x \in \mathrm{dom}(f))$$

En particular, dos funciones $f : A \to B$ y $g : B \to C$ son siempre componibles (en este orden), y su compuesta tiene el tipo $g \circ f : A \to C$. La operación de composición es obviamente asociativa, ojalá que las funciones involucradas sean componibles:

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

(si $\mathrm{img}(f) \subseteq \mathrm{dom}(g)$ y $\mathrm{img}(g) \subseteq \mathrm{dom}(h)$). Dado un conjunto A, se define la *función identidad* $\mathrm{id}_A : A \to A$ por

$$\mathrm{id}_A(x) \; := \; x \qquad \text{(para todo } x \in A)$$

Para toda función $f : A \to B$, tenemos que

$$f \circ \mathrm{id}_A \; = \; \mathrm{id}_B \circ f \; = \; f\,.$$

Observación 1.11 (Estructura categórica)**.** — Las operaciones anteriores equipan naturalmente el universo $\mathscr{U}$ de los conjuntos con una estructura de *categoría*, que se llama la *categoría de los conjuntos*, y se escribe **Set**. En esta categoría:

- Los *objetos* son los conjuntos.

- Las *flechas* entre dos objetos A y B son las funciones $f : A \to B$.
- La *compuesta de dos flechas* $f : A \to B$ y $g : B \to C$ es la función $g \circ f : A \to C$.
- La *flecha identidad* en un objeto A es la función $\mathrm{id}_A : A \to A$.

Esta estructura cumple obviamente los dos axiomas de las categorías:

(Asociatividad)	$(h \circ g) \circ f = h \circ (g \circ f)$	(si $f : A \to B, g : B \to C, h : C \to D$)
(Identidad)	$f \circ \mathrm{id}_A \ = \ \mathrm{id}_B \circ f \ = \ f$	(si $f : A \to B$)

Funciones notables Dado dos conjuntos A, B y una función $f : A \to B$, se dice que:

- f es *inyectiva* (notación: $f : A \hookrightarrow B$) si: $(\forall x, x' \in A)\,(f(x) = f(x') \Rightarrow x = x')$;
- f es *sobreyectiva* (notación: $f : A \twoheadrightarrow B$) si: $(\forall y \in B)\,(\exists x \in A)\, y = f(x)$;
- f es *biyectiva* (notación: $f : A \xrightarrow{\sim} B$) si f es inyectiva y sobreyectiva.

Toda función $f : A \to B$ inyectiva (resp. sobreyectiva, biyectiva) también se llama una *inyección* (resp. una *sobreyección*, una *biyección*). Toda biyección $f : A \to B$ es invertible (respecto a la operación de composición), y su inversa $f^{-1} : B \to A$ es definida por:

$$f^{-1} \ := \ \{(y, x) : (x, y) \in f\}\,.$$

Se verifica inmediatamente que $f^{-1} \circ f = \mathrm{id}_A$ y $f \circ f^{-1} = \mathrm{id}_B$. Cuando existe una biyección $f : A \xrightarrow{\sim} B$, se dice que los dos conjuntos A y B son *equipotentes*.

Observación 1.12. — Recordemos que en la definición de la fórmula

$$f : A \to B \ \equiv \ f \text{ función} \wedge \mathrm{dom}(f) = A \wedge \mathrm{img}(f) \subseteq B\,,$$

el dominio A es determinado por el grafo f mientras el codominio B es convencional, ojalá que $\mathrm{img}(f) \subseteq B$. En particular, toda función $f : A \to B$ también se puede ver como una función $f : A \to \mathrm{img}(f)$ (sin cambiar el grafo subyacente), que es siempre sobreyectiva, como función de tipo $A \to \mathrm{img}(f)$. En este sentido, toda inyección $f : A \to B$, vista como una biyección $f : A \to \mathrm{img}(f)$, es invertible, y su inversa $f^{-1} : \mathrm{img}(f) \to A$ cumple $f^{-1} \circ f = \mathrm{id}_A$.

Otras operaciones Dado una función $f : A \to B$ y dos subconjuntos $X \subseteq A$, $Y \subseteq B$, se llama:

- la *imagen* de X por f el subconjunto $f(X) \subseteq B$ definido por

$$f(X) \ := \ \{y \in B : (\exists x \in X)\, y = f(x)\};$$

- la *preimagen* de $Y \subseteq B$ por f el subconjunto $f^{-1}(Y) \subseteq A$ definido por

$$f^{-1}(Y) \ := \ \{x \in A : f(x) \in Y\}\,;$$

- la *restricción* de f a X la función $f_{\restriction X} : X \to B$ definida por

$$f_{\restriction X} \ := \ f \cap (X \times B)\,.$$

Por construcción, tenemos que $\mathrm{dom}(f_{\restriction X}) = X$ y $\mathrm{img}(f_{\restriction X}) = f(X)$.

En razón del carácter convencional del codominio, también se puede ver la función $f_{\restriction X}$ como una función de tipo $X \to Y$ para todo subconjunto $Y \subseteq B$ tal que $f(X) \subseteq Y$.

Pegado de funciones En general, la unión $f \cup g$ de dos funciones f y g no es una función. En lo siguiente, se dice que dos funciones f y g son *compatibles* cuando coinciden sobre la intersección de sus dominios:

$$f \text{ y } g \text{ compatibles} \ \equiv\ f\restriction_{\mathrm{dom}(f)\cap\mathrm{dom}(g)} = g\restriction_{\mathrm{dom}(f)\cap\mathrm{dom}(g)}\,.$$

(En particular, cuando $\mathrm{dom}(f) = \mathrm{dom}(g)$: f y g son compatibles si y sólo si $f = g$.)

Se verifica sin dificultad que dos funciones f y g son compatibles si y sólo si se pueden «pegar», en el sentido que el grafo $f \cup g$ también es una función. En este caso, tenemos que $\mathrm{dom}(f \cup g) = \mathrm{dom}(f) \cup \mathrm{dom}(g)$ e $\mathrm{img}(f \cup g) = \mathrm{img}(f) \cup \mathrm{img}(g)$.

En la teoría de conjuntos, se necesita frecuentemente pegar conjuntos (o familias) de funciones con múltiples dominios. Para ello, se utiliza el siguiente lema:

Lema 1.13 (Pegado de funciones)**.** — *Dado un conjunto C de funciones, el grafo*

$$F \ :=\ \bigcup C$$

es una función si y sólo si las funciones en C son compatibles a pares. En este caso, la función «pegada» $F = \bigcup C$ cumple $\mathrm{dom}(F) = \bigcup\{\mathrm{dom}(f) : f \in C\}$ *e* $\mathrm{img}(F) = \bigcup\{\mathrm{img}(f) : f \in C\}$.

(Se deja la demostración en ejercicio al lector.)

1.6.3. Familias de conjuntos

Dado un conjunto I, se llama una *familia de conjuntos indizada por I* toda función F tal que $\mathrm{dom}(F) = I$. (No se supone nada sobre la imagen de F.) La imagen de un elemento $i \in I$ por F se escribe F_i en lugar de $F(i)$, y la familia si misma se escribe $(F_i)_{i\in I}$ en lugar de F.

Dado una familia de conjuntos $F = (F_i)_{i\in I}$, se escriben

$$\begin{aligned}\bigcup_{i\in I} F_i &= \bigcup \mathrm{img}(F) = \{x : (\exists i \in I)\, x \in F_i\}\\ \bigcap_{i\in I} F_i &= \bigcap \mathrm{img}(F) = \{x : (\forall i \in I)\, x \in F_i\}\,,\end{aligned}$$

la intersección $\bigcap_{i\in I} F_i$ siendo definida sólo cuando $I \neq \varnothing$.

Producto cartesiano (generalizado) de una familia de conjuntos Sea $(A_i)_{i\in I}$ una familia de conjuntos indizada por un conjunto I cualquiera. Se llama el *producto cartesiano (generalizado)* de la familia $(A_i)_{i\in I}$ y se escribe $\prod_{i\in I} A_i$ el conjunto formado por todas las familias $(a_i)_{i\in I}$ tales que $a_i \in A_i$ para todo $i \in I$:

$$\prod_{i\in I} A_i \ :=\ \Big\{(a_i)_{i\in I} : (\forall i \in I)\, a_i \in A_i\Big\}.$$

(Este conjunto es un subconjunto del conjunto de todas las funciones de I a $\bigcup_{i\in I} A_i$, lo que justifica su existencia.) Para todo $i \in I$, se define la *proyección* $\pi_i : (\prod_{i\in I} A_i) \to A_i$ por

$$\pi_i((a_i)_{i\in I}) \ :=\ a_i \qquad\qquad (\text{para todo } (a_i)_{i\in I} \in \textstyle\prod_{i\in I} A_i)$$

El producto cartesiano $\prod_{i\in I} A_i$ satisface la siguiente propiedad universal:

Proposición 1.14. — *Para todo conjunto X dado con una familia de funciones $(f_i : X \to A_i)_{i\in I}$, existe una única función $h : X \to (\prod_{i\in I} A_i)$ tal que $\pi_i \circ h = f_i$ para todo $i \in I$:*

$$\begin{tikzcd} X \arrow[d, dashed, "h"'] \arrow[dr, "f_i"] & \\ (\prod_{i\in I} A_i) \arrow[r, "\pi_i"'] & A_i \end{tikzcd}$$

La función $h : X \to (\prod_{i\in I} A_i)$ es caracterizada para todo $x \in X$ por:

$$h(x) \ = \ (f_i(x))_{i\in I}\,.$$

(Se deja la demostración en ejercicio al lector.)

Observaciones 1.15. — (1) El producto cartesiano $\prod_{i\in I} A_i$ es una generalización natural del producto cartesiano finitario $A_1 \times \cdots \times A_n$ a las familias de conjuntos $(A_i)_{i\in I}$ indizadas por un conjunto I cualquiera, finito o infinito —anticipándose a la distinción finito/infinito que introduciremos en la sección ???. Intuitivamente, se pueden ver los elementos del producto cartesiano generalizado $\prod_{i\in I} A_i$ como «tuplas generalizadas»

$$(a_i)_{i\in I} = (a_i, a_j, a_k, \ldots) \ \in \ A_i \times A_j \times A_k \times \cdots\,,$$

que contienen tantas componentes como elementos $i, j, k, \ldots \in I$ (cada componente a_i viviendo en el conjunto A_i correspondiente). En particular, cuando el conjunto I es finito —típicamente, cuando $I = \{1, \ldots, n\}$—, existe una biyección canónica entre el producto cartesiano generalizado $\prod_{i\in I} A_i$ y el producto cartesiano usual $A_1 \times \cdots \times A_n$, la cual es definida por:

$$\begin{array}{rrcl} f & : \ (\prod_{i\in\{1,\ldots,n\}} A_i) & \xrightarrow{\sim} & A_1 \times \cdots \times A_n \\ & (a_i)_{i\in\{1,\ldots,n\}} & \mapsto & (a_1, \ldots, a_n)\,. \end{array}$$

(2) En el caso particular donde $I = \varnothing$ (la familia de conjuntos $(A_i)_{i\in I}$ es vacía), tenemos que $\prod_{i\in I} A_i = \{\varnothing\}$. Así, el producto $\prod_{i\in\varnothing} A_i$ no es vacío, pero es un conjunto unitario cuyo único elemento es la familia vacía $\varnothing$, que también se puede ver como la única «0-upla» ().

(3) Cuando $A_i = A$ para todo $i \in I$ (la familia de conjuntos $(A_i)_{i\in I}$ es constante), el producto cartesiano $\prod_{i\in I} A_i$ se reduce al conjunto de las funciones de I a A: $\prod_{i\in I} A_i \ = \ A^I$.

Suma directa de una familia de conjuntos Sea $(A_i)_{i\in I}$ una familia de conjuntos indizada por un conjunto I cualquiera. Se llama la *suma directa* de la familia $(A_i)_{i\in I}$ y se escribe $\sum_{i\in I} A_i$ el conjunto formado por todos los pares (i, a) donde $i \in I$ y $a \in A_i$:

$$\sum_{i\in I} A_i \ := \ \{(i, a) \ : \ i \in I \ \wedge \ a \in A_i\} \ = \ \biguplus_{i\in I} (\{i\} \times A_i)\,.$$

(Este conjunto es un subconjunto del conjunto $I \times \bigcup_{i\in I} A_i$, lo que justifica su existencia.) Para todo $i \in I$, se define el *constructor* $\sigma_i : A_i \to (\prod_{i\in I} A_i)$ por

$$\sigma_i(a) \ := \ (i, a) \qquad\qquad \text{(para todo } a \in A_i\text{)}$$

La suma directa $\sum_{i\in I} A_i$ satisface la propiedad universal «dual» del producto cartesiano generalizado $\prod_{i\in I} A_i$, que es el siguiente:

Proposición 1.16. — *Para todo conjunto X dado con una familia de funciones $(f_i : A_i \to X)_{i\in I}$, existe una única función $h : (\sum_{i\in I} A_i) \to X$ tal que $h \circ \sigma_i = f_i$ para todo $i \in I$:*

$$\begin{tikzcd}
X & \\
(\sum_{i\in I} A_i) \arrow[u, dotted, "h"] & A_i \arrow[l, "\sigma_i"] \arrow[ul, "f_i"']
\end{tikzcd}$$

La función $h : (\sum_{i\in I} A_i) \to X$ es caracterizada para todo $(i, a) \in \sum_{i\in I} A_i$ por:

$$h((i, a)) \ = \ f_i(a)\,.$$

(Se deja la demostración en ejercicio al lector.)

Observaciones 1.17. — (1) Cuando $I = \{1, \dots, n\}$ para algún entero natural n, se escribe

$$A_1 + \cdots + A_n \ := \sum_{i\in\{1,\dots,n\}} A_i \ = \ (\{1\} \times A_1) \uplus \cdots \uplus (\{n\} \times A_n)\,.$$

(2) En el caso particular donde $I = \varnothing$ (la familia de conjuntos $(A_i)_{i\in I}$ es vacía), la suma directa $\sum_{i\in I} A_i$ también es vacía: $\sum_{i\in\varnothing} A_i = \varnothing$.
(3) Cuando $A_i = A$ para todo $i \in I$ (la familia de conjuntos $(A_i)_{i\in I}$ es constante), la suma directa $\sum_{i\in I} A_i$ se reduce al producto cartesiano de I por A: $\sum_{i\in I} A_i \ = \ I \times A$.

1.7. Axioma del infinito

1.7.1. Conjuntos Dedekind-infinitos

Los axiomas y esquemas de axiomas que presentamos hasta ahora sólo permiten definir conjuntos *hereditariamente finitos*, es decir: conjuntos finitos cuyos elementos también son finitos, y esto recursivamente. (Esta noción será definida formalmente en la sección ???.)

Antes de introducir el axioma que permite construir conjuntos infinitos, recordemos que:

Definición 1.18 (Conjunto Dedekind-infinito)**.** — Un conjunto A es *Dedekind-infinito* si existe una función $f : A \to A$ que es inyectiva y no sobreyectiva.

Proposición 1.19. — *Si un conjunto A es Dedekind-infinito, entonces todo conjunto B para el cual existe una inyección $f : A \hookrightarrow B$ es Dedekind-infinito igualmente.*

Demostración. El conjunto A siendo Dedekind-infinito, se toma una función $g : A \to A$ que es inyectiva y no sobreyectiva, y se considera la función $h : B \to B$ definida por:

$$h(y) \ = \ \begin{cases} f(g(f^{-1}(y))) & \text{si } y \in \operatorname{img}(f) \\ y & \text{si no} \end{cases} \qquad \text{(para todo } y \in B\text{)}$$

Se verifica fácilmente que la función $h : B \to B$ es inyectiva y no sobreyectiva. □

La proposición anterior implica en particular que todo conjunto que contiene un subconjunto Dedekind-infinito también es Dedekind-infinito. Como tales conjuntos no se pueden construir a partir de los axiomas anteriores, se añade el siguiente axioma:

Axioma 6 (Axioma del infinito)**.** — *Existe un conjunto Dedekind-infinito.*

1.7.2. Estructuras aritméticas

Definición 1.20 (Estructura aritmética)**.** — Una *estructura aritmética* es una terna (N, o, s) formada por un conjunto N, un elemento $o \in N$ y una función $s : N \to N$, tales que:

(1) la función $s : N \to N$ es inyectiva;
(2) $o \notin \mathrm{img}(s)$;
(3) para todo $P \subseteq N$: si $o \in P$ y $s(P) \subseteq P$, entonces $P = N$ (principio de inducción).

Es claro que el conjunto de base N de cualquier estructura aritmética (N, o, s) es Dedekind-infinito (según las propiedades de la función s). Más generalmente, se demuestra que todo conjunto Dedekind-infinito contiene un subconjunto con una estructura aritmética:

Proposición 1.21. — *Para todo conjunto Dedekind-infinito A, existe un subconjunto $N \subseteq A$, un elemento $o \in N$ y una función $s : N \to N$ tales que (N, o, s) es una estructura aritmética.*

Demostración. Sea A un conjunto Dedekind-infinito. Se considera una función $f : A \to A$ inyectiva y no sobreyectiva, y se toma un elemento $o \in A$ tal que $o \notin \mathrm{img}(f)$. Sean

$$C = \{P \in \mathfrak{P}(A) : o \in P \land f(P) \subseteq P\} \qquad \text{y} \qquad N = \bigcap C$$

(El conjunto C no es vacío, pues $A \in C$.) Se verifica fácilmente que $o \in N$ y $f(N) \subseteq N$, lo que permite definir la función $s : N \to N$ por $s = f_{\restriction N}$. Es claro que (1) la función $s : N \to N$ es inyectiva, y (2) $o \notin \mathrm{img}(s)$. Para establecer el ítem (3), se considera un subconjunto $P \subseteq N$ tal que $o \in P$ y $s(P) \subseteq P$. Como $s = f_{\restriction N}$ y $P \subseteq N$, tenemos que $f(P) = s(P) \subseteq P$, de tal modo que $P \in C$ (por definición de C). Esto implica que $N\,(= \bigcap C) \subseteq P$, luego $P = N$. □

Combinada con el axioma del infinito, la proposición anterior implica la existencia de estructuras aritméticas. Ahora, se trata de demostrar que todas las estructuras aritméticas son isomorfas entre ellas —y en particular, que todos los conjuntos de base de tales estructuras son equipotentes. Para ello, se demuestra la siguiente propiedad universal, que establece el mecanismo de *definición de función por recursión*:

Proposición 1.22 (Definición de función por recursión)**.** — *Sea (N, o, s) una estructura aritmética. Para todo conjunto X dado con un elemento $x_0 \in X$ y una función $f : X \to X$, existe una única función $h : N \to X$ tal que $h(o) = x_0$ y $h \circ s = f \circ h$:*

$$\begin{tikzcd} o \arrow[d, mapsto, dotted] & \in & N \arrow[r, "s"] \arrow[d, dotted, "h"'] & N \arrow[d, dotted, "h"] \\ x_0 & \in & X \arrow[r, "f"'] & X \end{tikzcd}$$

Dicho de otro modo, la propiedad universal expresa que para todos X, $x_0 \in X$ y $f : X \to X$, existe una única función $h : N \to X$ que cumple las dos condiciones:

$$h(0) = x_0 \qquad \text{y} \qquad h(n+1) = f(h(n)) \qquad\qquad \text{para todo } n \in N$$

(utilizando las notaciones sugestivas $0 := o$ y $n + 1 := s(n)$ en N).

Demostración. (*Unicidad*) Sean dos funciones $h, h' : N \to X$ tales que $h(o) = h'(o) = x_0$, $h \circ s = f \circ h$ y $h' \circ s = f \circ h'$. Definiendo $P = \{n \in N : h(n) = h'(n)\} \subseteq N$, se verifica fácilmente que $o \in P$ y $s(P) \subseteq P$, luego $P = N$, lo que implica que las funciones h y h' son iguales.

(*Existencia*) Se dice que un subconjunto $G \subseteq N \times X$ es *interesante* cuando G cumple las siguientes dos condiciones:

(*i*) $(\forall x \in X)\,((o, x) \in G \Rightarrow x = x_0)$

(*ii*) $(\forall n \in N)\,(\forall x' \in X)\,((s(n), x') \in G \Rightarrow (\exists x \in X)\,((n, x) \in G \wedge x' = f(x)))$

Sea $H = \bigcup\{G \in \mathfrak{P}(N \times X) : G \text{ interesante}\}$ la unión de todos los subconjuntos interesantes de $N \times X$. Se verifica fácilmente que el subconjunto $H \subseteq N \times X$ es interesante; por construcción, es el subconjunto interesante más grande de $N \times X$. Se verifica igualmente que:

(1) $(o, x_0) \in H$. En efecto, el conjunto $\{(o, x_0)\}$ es interesante, entonces $\{(o, x_0)\} \subseteq H$.

(2) Si $(n, x) \in H$, entonces $(s(n), f(x)) \in H$. En efecto, dado $(n, x) \in H$, es obvio que el conjunto $H' := H \cup \{(s(n), f(x))\}$ es interesante, luego $H' = H$ por maximalidad.

Ahora, se trata de demostrar que para todo $n \in N$, existe un único $x \in X$ tal que $(n, x) \in H$. De nuevo, se define $P = \{n \in N : (\exists! x \in X)(n, x) \in H\}$; es obvio que $o \in P$ (por (*i*) y (1)) y $s(P) \subseteq P$ (por (*ii*) y (2)), de tal modo que $P = N$. Esto demuestra que el subconjunto $H \subseteq N \times X$ es el grafo de una función $h\,(= H) : N \to X$, que cumple obviamente las condiciones deseadas. □

Teorema 1.23 (Isomorfismo). — *Si (N, o, s) y (N', o', s') son dos estructuras aritméticas, entonces existe una única biyección $h : N \to N'$ tal que $h(o) = o'$ y $h \circ s = s' \circ h$.*

Demostración. Aplicando la Prop. 1.22 a la estructura aritmética (N, o, s) con $X = N'$, $x_0 = o'$ y $f = s'$, se obtiene una función $h : N \to N'$ tal que $h(o) = o'$ y $h \circ s = s' \circ h$. Aplicando de nuevo la Prop. 1.22 a la estructura aritmética (N', o', s') con $X = N$, $x_0 = o$ y $f = s$, se obtiene otra función $h' : N' \to N$ tal que $h'(o') = o$ y $h' \circ s' = s \circ h'$. Se verifica fácilmente que $h' \circ h = \mathrm{id}_N$ y $h \circ h' = \mathrm{id}_{N'}$, lo que implica que la función $h : N \to N'$ es biyectiva. La unicidad de tal isomorfismo es obvia por la Prop. 1.22. □

Gracias al teorema anterior, se puede fijar una estructura aritmética cualquiera $(\mathbb{N}, 0, s)$, y definir los *enteros naturales* como los elementos del conjunto $\mathbb{N}$.

En el capítulo 2, daremos una formulación alternativa del axioma del infinito, así como una construcción más canónica del conjunto $\mathbb{N}$, como primer ordinal límite.

1.7.3. La teoría de conjuntos de Zermelo

La teoría inducida por los axiomas anteriores se llama la *teoría de conjuntos de Zermelo*, y se escribe Z. Se recuerda que sus axiomas son:

- el *axioma de extensionalidad* (Axioma 1 p. 12),
- el *axioma de pares* (Axioma 2 p. 13),
- los *axiomas de comprensión* (Axioma 3 p. 15),
- el *axioma de unión* (Axioma 4 p. 17),
- el *axioma del conjunto potencia* (Axioma 5 p. 17), y
- el *axioma del infinito* (Axioma 6 p. 23).

En la práctica, la teoría de Zermelo es (más que) suficiente para desarrollar la casi totalidad de las matemáticas usuales. Sin embargo, la teoría de los ordinales y de los cardinales (Capítulo 2) necesita introducir un último esquema: *el esquema de remplazo*.

1.8. Esquema de remplazo

1.8.1. Imagen de una relación funcional y total

Los axiomas de Zermelo permiten construir la imagen de una función f cualquiera, por:

$$\mathrm{img}(f) \;=\; \{y : \exists x\,(x,y) \in f\}$$

(En efecto, el predicado $\psi(y) \equiv \exists x\,(x,y) \in f$ es colectivizante por la Prop. 1.7 p. 16, pues $\psi(y)$ implica $y \in \bigcup\bigcup f$ para todo y.) Desgraciadamente, tal construcción ya no es posible en la teoría de Zermelo si se remplaza la función f por una relación *funcional y total*[6] sobre un conjunto a, es decir, por una fórmula $\phi(x,y)$ tal que:

$$(\forall x \in a)\,\exists! y\,\phi(x,y)\,.$$

Aquí, el problema viene de que los axiomas de Zermelo no permiten demostrar en general que el conjunto a tiene una imagen a través de la relación funcional y total $\phi(x,y)$, aunque la *clase* imagen $\psi(y) \equiv (\exists x \in a)\,\phi(x,y)$ tenga intuitivamente un tamaño no mayor que el del conjunto a (por unicidad del objeto y asociado a cada $x \in a$). El ejemplo típico de relación funcional y total que no tiene imagen en la teoría de Zermelo es el siguiente:

Ejemplo 1.24. — Sea $(\mathbb{N}, 0, s)$ una estructura aritmética fijada. Para todo $n \in \mathbb{N}$, se escriben

$$[0..n] = \{m \in N : m \leq n\} \qquad \text{y} \qquad [0..n) = \{m \in N : m < n\}$$

(anticipándose a la definición de las relaciones de orden $\leq$ y $<$ en $\mathbb{N}$). Dado un conjunto a que parametriza la construcción, se considera la relación $\phi(n,y)$ definida por la fórmula

$$\begin{aligned}\phi(n,y) \;\equiv\; n \in \mathbb{N} \;\land\; \exists f\,[&f \text{ función} \;\land\; \mathrm{dom}(f) = [0..n] \;\land\; f(0) = a \;\land\\ &(\forall i \in [0..n))\, f(s(i)) = \mathfrak{P}(f(i)) \;\land\; f(n) = y \quad]\end{aligned}$$

Se verifica fácilmente que:

(1) $\forall y\,(\phi(0,y) \Leftrightarrow y = a)$ («a es la única imagen de 0 por ϕ»)
(2) $(\forall n \in \mathbb{N})\,\forall y\,[\phi(n,y) \;\Rightarrow\; \forall y'\,(\phi(s(n),y') \Leftrightarrow y' = \mathfrak{P}(y))]$
(«Si y es imagen de n por ϕ, entonces $\mathfrak{P}(y)$ es la única imagen de $s(n)$ por ϕ»)

Por una inducción obvia, sigue de los ítems (1) y (2) que para todo $n \in \mathbb{N}$, existe un único objeto y tal que $\phi(n,y)$. Intuitivamente, el único objeto y asociado al entero $n \in \mathbb{N}$ por la relación $\phi(n,y)$ es el *n-esimo conjunto potencia de a*, es decir:

$$y \;=\; \mathfrak{P}^n(a) \;=\; \underbrace{\mathfrak{P}(\cdots\mathfrak{P}}_{n\text{ veces}}(a)\cdots)\,.$$

(Insistamos en que tal escritura con puntos suspensivos es puramente informal, y que sólo se puede formalizar en la teoría de conjuntos con una fórmula tal que $\phi(n,y)$.)

[6] Aquí, se dice que la relación $\phi(x,y)$ es *total (sobre a)* en el sentido que $(\forall x \in a)\,\exists y\,\phi(x,y)$. Esta noción de totalidad es distinta de la noción de totalidad que definiremos en la Sección 1.9 para las relaciones homogéneas sobre un conjunto A. Para distinguir las dos nociones de totalidad, se utiliza a veces la terminología de *relación total a la izquierda (sobre a)* para indicar una relación $\phi(x,y)$ tal que $(\forall x \in a)\,\exists y\,\phi(x,y)$. En general, el contexto permite determinar de cuál noción de relación total se trata.

Sin embargo, se puede demostrar por métodos metamatemáticos que la fórmula

$$\exists b\, \forall y\, (y \in b \;\Leftrightarrow\; (\exists n \in \mathbb{N})\, \phi(x, y))$$

que expresa la existencia del conjunto imagen $b = \{\mathfrak{P}^n(a) : n \in \mathbb{N}\}$ es una fórmula indecidible en la teoría de Zermelo (es decir: una fórmula que se puede ni demostrar ni refutar), aunque la clase imagen sea intuitivamente numerable. De hecho, este problema no es específico a la operación $y \mapsto \mathfrak{P}(y)$ (conjunto potencia), pues el mismo problema se encuentra cuando se itera n veces las operaciones $y \mapsto \{y\}$ (conjunto unitario) y $y \mapsto \bigcup y$ (unión) para todo $n \in \mathbb{N}$.

La discusión anterior justifica el siguiente esquema:

Axioma 7 (Esquema de remplazo). — *Para todo conjunto a, si la relación $\phi(x, y)$ es funcional y total sobre a, entonces existe un conjunto b que contiene al menos todas las imágenes de los elementos de a por la relación $\phi(x, y)$:*

$$\forall a\, [(\forall x \in a)\, \exists! y\, \phi(x, y) \;\Rightarrow\; \exists b\, (\forall x \in a)\, (\exists y \in b)\, \phi(x, y)]$$

(donde $\phi(x, y)$ es cualquier relación binaria de la teoría de conjuntos).

Proposición 1.25 (Imagen y grafo de una relación funcional y total). — *Cada relación binaria $\phi(x, y)$ que es funcional y total sobre un conjunto A tiene imagen y grafo:*

$$\mathrm{img}(\phi) \;=\; \{y : (\exists x \in A)\, \phi(x, y)\} \qquad y \qquad \mathrm{gr}(\phi) \;=\; \{(x, y) : x \in A \wedge \phi(x, y)\}\,.$$

En particular, el grafo $f = \mathrm{gr}(\phi)$ es una función de dominio A y de imagen $\mathrm{img}(\phi)$.

Demostración. Según el esquema de remplazo aplicado al conjunto A y a la relación $\phi(x, y)$, existe un conjunto B_0 tal que $(\forall x \in A)(\exists y \in B_0)\, \phi(x, y)$. Luego se definen por comprensión:

$$\mathrm{img}(\phi) \;:=\; \{y \in B_0 : (\exists x \in A)\, \phi(x, y)\} \quad \text{y} \quad \mathrm{gr}(\phi) \;:=\; \{(x, y) \in A \times B_0 : \phi(x, y)\}. \qquad \square$$

Volviendo a la relación $\phi(n, y)$ del ejemplo 1.24 (la cual es funcional y total sobre $\mathbb{N}$), la proposición anterior implica la existencia del conjunto

$$\mathrm{img}(\phi) \;=\; \{\mathfrak{P}^n(a) : n \in \mathbb{N}\}$$

así como de la función f de dominio $\mathbb{N}$ definida por

$$f(0) = a \qquad \text{y} \qquad f(s(n)) = \mathfrak{P}(f(n)) \qquad\qquad \text{para todo } n \in \mathbb{N}\,.$$

1.8.2. Aplicación: clausura transitiva de un conjunto

Se dice que un conjunto a es *transitivo* cuando todo elemento de a está incluido en a:

$$\begin{aligned} a \text{ transitivo} \;&\equiv\; (\forall x \in a)\, x \subseteq a \\ &\equiv\; (\forall x \in a)(\forall y \in x)\, y \in a \end{aligned}$$

Por una inducción obvia, es claro que todo conjunto transitivo a contiene los elementos de sus elementos, los elementos de los elementos de sus elementos, y esto recursivamente:

$$x_n \in x_{n-1} \in \cdots \in x_2 \in x_1 \in a \quad\Rightarrow\quad x_1, x_2, \ldots, x_{n-1}, x_n \in a\,.$$

Además, se verifica fácilmente que:

Proposición 1.26 (Propiedades de los conjuntos transitivos)**.**

(1) Un conjunto a es transitivo si y sólo si $(\bigcup a) \subseteq a$;

(2) Si $(a_i)_{i\in I}$ es una familia de conjuntos transitivos, entonces los conjuntos $\bigcup_{i\in I} a_i$ y $\bigcap_{i\in I} a_i$ (éste siendo definido sólo si $I \neq \varnothing$) son transitivos igualmente.

Gracias al esquema de remplazo, se puede demostrar que:

Proposición y definición 1.27 (Clausura transitiva)**.** — *Para todo conjunto a, existe un conjunto transitivo minimal b (en el sentido de la inclusión) tal que $a \subseteq b$. Este conjunto se llama la* clausura transitiva *de a, y se escribe* $\mathrm{Cl}(a)$*.*

Demostración. Siguiendo el ejemplo 1.24 (con el operador de unión en lugar del operador de potencia), se considera la relación binaria definida por

$$\begin{aligned}\phi(n,y) \;\equiv\; n \in \mathbb{N} \;\wedge\; \exists f\,[&f \text{ función} \;\wedge\; \mathrm{dom}(f) = [0..n] \;\wedge\; f(0) = a \;\wedge\\ &(\forall i \in [0..n))\, f(s(i)) = \textstyle\bigcup f(i) \;\wedge\; f(n) = y \quad]\end{aligned}$$

Es claro que la relación $\phi(n,y)$ es funcional y total sobre $\mathbb{N}$; por la Prop. 1.25, ella define una función f de dominio $\mathbb{N}$ que satisface por construcción las dos condiciones:

$$f(0) = a \qquad \text{y} \qquad f(s(n)) = \bigcup f(n) \qquad\qquad \text{para todo } n \in \mathbb{N}\,.$$

Ahora se define $\mathrm{Cl}(a) = \bigcup_{n\in\mathbb{N}} f(n)$, y se verifica que:

- $a \subseteq \mathrm{Cl}(a)$. Obvio, pues $a = f(0) \subseteq \mathrm{Cl}(a)$.
- *a es transitivo.* En efecto, si $x \in \mathrm{Cl}(a) = \bigcup_{n\in\mathbb{N}} f(n)$, tenemos que $x \in f(n)$ para algún $n \in \mathbb{N}$, de tal modo que $x \subseteq \bigcup f(n) = f(s(n)) \subseteq \mathrm{Cl}(a)$.
- $\mathrm{Cl}(a)$ *es el conjunto transitivo más pequeño tal que* $a \subseteq \mathrm{Cl}(a)$. En efecto, si b es un conjunto transitivo tal que $a \subseteq b$, se demuestra por una inducción obvia que $f(n) \subseteq b$ para todo $n \in \mathbb{N}$, de tal modo que $\mathrm{Cl}(a) \subseteq b$. □

Intuitivamente, tenemos que

$$\mathrm{Cl}(a) \;=\; \bigcup_{n\in\mathbb{N}}\Big(\bigcup\nolimits^{n} a\Big) \;=\; \bigcup_{n\in\mathbb{N}}\Big(\underbrace{\bigcup\cdots\bigcup}_{n \text{ veces}} a\Big).$$

1.8.3. La teoría de conjuntos de Zermelo-Fraenkel

Por definición, la *teoría de conjuntos de Zermelo-Fraenkel* (notación: ZF) es la teoría inducida por todos los axiomas anteriores, es decir: por los axiomas de Zermelo (Sección 1.7.3), más los axiomas de remplazo (Sección 1.8).

Algunos autores consideran (particularmente en la teoría de modelos) que el sistema ZF también incluye el *axioma de fundación* (o *axioma de regularidad*) que presentaremos en el capítulo siguiente (Sección 2.4.4). En la práctica, este axioma no cambia mucho la expresividad de la teoría, a diferencia del axioma de elección (Sección 2.2).

1.9. Relaciones de equivalencia y relaciones de orden

Para concluir este capítulo introductorio a la teoría axiomática de conjuntos (según la axiomática de Zermelo-Fraenkel), se recuerda aquí la terminología y los principales resultados acerca de las relaciones de equivalencia y de las relaciones de orden.

1.9.1. Propiedades notables de las relaciones

Sea A un conjunto. Se llama una *relación binaria (homogénea[7]) sobre A* cualquier relación binaria $R \subseteq A \times A$. Se dice que tal relación es:

- *reflexiva* si $(\forall x \in A)\, x R x$;
- *irreflexiva* si $(\forall x \in A)\, \neg(x R x)$;
- *simétrica* si $(\forall x, y \in A)\, (x R y \Rightarrow y R x)$;
- *antisimétrica* si $(\forall x, y \in A)\, (x R y \land y R x \Rightarrow x = y)$;
- *transitiva* si $(\forall x, y, z \in A)\, (x R y \land y R z \Rightarrow x R z)$;
- *total*[8] si $(\forall x, y \in A)\, (x R y \lor y R x)$.

Además, se dice que una relación $R \subseteq A \times A$ es:

- un *preorden* si R es reflexiva y transitiva;
- un *orden* (en el sentido amplio) si R es reflexiva, transitiva y antisimétrica;
- un *orden estricto* si R es irreflexiva y transitiva;
- una *relación de equivalencia* si R es reflexiva, simétrica y transitiva.

Dado una relación binaria $R \subseteq A \times A$ y un subconjunto $X \subseteq A$, se llama la *relación inducida por R sobre X* la relación binaria $R_{\restriction X} \subseteq X \times X$ definida por $R_{\restriction X} := R \cap (X \times X)$.

Se observa que si la relación R es reflexiva, irreflexiva, simétrica, antisimétrica, transitiva, total, un preorden, un orden, un orden estricto, o una relación de equivalencia sobre A, entonces la relación inducida $R_{\restriction X}$ sobre el subconjunto $X \subseteq A$ todavía cumple dicha propiedad[9].

Observación 1.28. — Las definiciones anteriores se generalizan naturalmente a las relaciones binarias definidas sobre una clase. Recordemos que en teoría de conjuntos, una *clase A* es definida por una fórmula $A(x)$ que depende de una variable x, mientras una *relación binaria R* es definida por una fórmula $x R y \equiv R(x, y)$ que depende de dos variables x e y. (Tales fórmulas $A(x)$ y $R(x, y)$ pueden depender de otras variables, que tienen un papel de parámetros.) En tal contexto, diremos por ejemplo que la relación R es transitiva sobre la clase A cuando

$$(\forall x, y, z : A)\, (x R y \land y R z \Rightarrow x R z)\,,$$

[7] A diferencia de las relaciones no homogéneas $R \subseteq A \times B$, donde $A \neq B$.

[8] Véase la nota 6 p. 26 acerca de las dos nociones de relación total.

[9] ¡Cuidado! Algunas propiedades de las relaciones no se mantienen a través de la operación de restricción. Por ejemplo, si $(<) \subseteq A \times A$ es un orden estricto *denso* sobre A (es decir: una relación irreflexiva, transitiva y tal que $(\forall x, y \in A)(x < y \Rightarrow (\exists z \in A)\, (x < z \land z < y))$), entonces la relación inducida $(<)_{\restriction X} \subseteq X \times X$ todavía es un orden estricto sobre X, pero no cumple la propiedad de densidad sobre X en general. El mismo problema ocurre típicamente con las propiedades de completitud (existencia de ínfimo/supremo), y más generalmente con todas las propiedades que involucran cuantificaciones existenciales.

utilizando la abreviatura $(\forall x : A)\, \phi(x) \ \equiv\ \forall x\, (A(x) \ \Rightarrow\ \phi(x))$. (Las otras propiedades de las relaciones binarias se generalizan del mismo modo.) Una gran parte de la terminología y de los resultados sobre las relaciones de equivalencia y las relaciones de orden (véase Secciones 1.9.2 y 1.9.3 más abajo) se generaliza a este marco, teniendo en cuenta que las clases *no son* objetos matemáticos (es decir: objetos del universo $\mathscr{U}$), lo que implica que:

(1) Una clase nunca puede pertenecer a un conjunto, ni siquiera a otra clase.

(2) Nunca se puede cuantificar sobre todas las clases, aunque se pueda *parametrizar* un resultado y su demostración por una o múltiples clases, utilizando *esquemas* de teoremas y de demostraciones (siguiendo el ejemplo de los esquemas de axiomas).

1.9.2. Relaciones de equivalencia

Sea A un conjunto. Se recuerda que una *relación de equivalencia* sobre A es una relación binaria $R \subseteq A \times A$ reflexiva, simétrica y transitiva:

$$\begin{array}{rcll} R \text{ equivalencia sobre } A & \equiv & R \subseteq A \times A & \wedge \\ & & (\forall x \in A)\, x\, R\, x & \wedge \\ & & (\forall x, y \in A)\, (x\, R\, y \Rightarrow y\, R\, x) & \wedge \\ & & (\forall x, y, z \in A)\, (x\, R\, y \wedge y\, R\, z \Rightarrow x\, R\, z) & \end{array}$$

En lo siguiente, se utilizarán símbolos tales que $\sim$, $\simeq$, $\cong$, etc. (posiblemente afectados con sub/superíndices) para indicar las relaciones de equivalencia.

Clases de equivalencia y conjunto cociente Sea $\sim$ una relación de equivalencia sobre A. Dado un elemento $x \in A$, se llama la *clase de equivalencia de* x (respecto a la relación $\sim$) y se escribe $[x]_\sim$ el subconjunto de A definido por

$$[x]_\sim \ :=\ \{y \in A : x \sim y\}\,.$$

Se observa que para todos $x, y \in A$, se cumplen las siguientes equivalencias lógicas:

$$x \sim y \ \Leftrightarrow\ y \in [x]_\sim \ \Leftrightarrow\ x \in [y]_\sim \ \Leftrightarrow\ [x]_\sim \cap [y]_\sim \neq \varnothing \ \Leftrightarrow\ [x]_\sim = [y]_\sim\,.$$

Más generalmente, se llama una *clase de equivalencia* de la relación $\sim$ cualquier conjunto de la forma $[x]_\sim$ para algún $x \in A$. Se llama el *conjunto cociente* de A por la relación $\sim$ y se escribe $A/{\sim}$ el conjunto de todas las clases de equivalencias de la relación $\sim$:

$$A/{\sim} \ :=\ \{[x]_\sim : x \in A\} \qquad\qquad (\subseteq \mathfrak{P}(A))$$

La función $\pi_\sim : A \to (A/{\sim})$ definida por $\pi_\sim(x) = [x]_\sim$ para todo $x \in A$ es sobreyectiva; se llama la *sobreyección canónica* asociada a la relación de equivalencia $\sim$.

Particiones de un conjunto La noción de relación de equivalencia (sobre un conjunto A) es fuertemente relacionada con la noción de *partición* (del mismo conjunto A). Formalmente, se llama una *partición* de A todo conjunto de partes $P \subseteq \mathfrak{P}(A)$ tal que:

(1) Para todo $C \in P$: $\quad C \neq \varnothing$.

(2) Para todos $C, C' \in P$: $\quad C \neq C'$ implica $C \cap C' = \varnothing$.

(3) $A = \bigcup P$.

La siguiente proposición establece que las particiones de A son exactamente los conjuntos cocientes de A por todas las relaciones de equivalencia posibles sobre A:

Proposición 1.29. — *Sea A un conjunto cualquiera.*

(1) Para toda relación de equivalencia $\sim$ sobre A, el cociente $A/{\sim}$ es una partición de A, cuyos elementos son las clases de equivalencia de la relación $\sim$.

(2) Recíprocamente, para toda partición P de A, la relación $(\sim) \subseteq A \times A$ definida por

$$x \sim y \;\equiv\; (\exists C \in P)\,(x \in C \wedge y \in C) \qquad (\textit{para todos } x, y \in A)$$

es una relación de equivalencia sobre A, cuyas clases de equivalencia son los elementos de P, de tal modo que $(A/{\sim}) = P$.

(Se deja la demostración en ejercicio al lector.)

Funciones compatibles Sea $\sim$ una relación de equivalencia sobre A y X un conjunto cualquiera. Se dice que una función $f : A \to X$ es *compatible* con la relación de equivalencia $\sim$ cuando $x \sim y$ implica $f(x) = f(y)$ para todos $x, y \in A$.

El conjunto cociente $A/{\sim}$ cumple la siguiente propiedad universal:

Proposición 1.30 (Propiedad universal). — *Sea A un conjunto equipado con una relación de equivalencia $\sim$. Para todo conjunto X dado con una función $f : A \to X$ compatible con la relación de equivalencia $\sim$, existe una única función $\tilde{f} : (A/{\sim}) \to X$ tal que $f = \tilde{f} \circ \pi_{\sim}$:*

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d, "\pi_\sim"'] & X \\ A/{\sim} \arrow[ur, dotted, "\tilde{f}"'] & \end{tikzcd}$$

(Se deja la demostración en ejercicio al lector.)

Más generalmente, si A y B son dos conjuntos dados con relaciones de equivalencia $(\sim_A) \subseteq A \times A$ y $(\sim_B) \subseteq B \times B$, se dice que una función $f : A \to B$ es *compatible* con las relaciones $\sim_A$ y $\sim_B$ cuando para todos $x, y \in A$, $x \sim_A y$ implica $f(x) \sim_B f(y)$. Se deduce de la propiedad anterior que existe una única función $\hat{f} : (A/{\sim_A}) \to (B/{\sim_B})$ tal que $\hat{f} \circ \pi_{\sim_A} = \pi_{\sim_B} \circ \hat{f}$:

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d, "\pi_{\sim_A}"'] & B \arrow[d, "\pi_{\sim_B}"] \\ A/{\sim_A} \arrow[r, dotted, "\hat{f}"'] & B/{\sim_B} \end{tikzcd}$$

(la función $\hat{f}$ siendo definida por $\hat{f} = (\pi_{\sim_B} \circ f)^{\sim_A}$).

1.9.3. Relaciones de orden

Sea A un conjunto. Se recuerda que una *relación de orden* sobre A es una relación binaria $R \subseteq A \times A$ reflexiva, antisimétrica y transitiva:

$$\begin{array}{rcll} R \text{ orden sobre } A & \equiv & R \subseteq A \times A & \land \\ & & (\forall x \in A)\, x\, R\, x & \land \\ & & (\forall x, y \in A)\,(x\, R\, y \land y\, R\, x \Rightarrow x = y) & \land \\ & & (\forall x, y, z \in A)\,(x\, R\, y \land y\, R\, z \Rightarrow x\, R\, z) & \end{array}$$

Precisemos que dicha noción no presupone que la relación R sea total, y es la razón para la cual algunos autores llaman un *orden parcial* lo que se llama aquí un *orden*. Un *conjunto ordenado* es un par (A, R) formado por un conjunto A equipado con una relación de orden R sobre A.

Orden inverso Dado una relación de orden R sobre A, se observa que la relación inversa

$$R^{-1} := \{(y, x) \in A \times A : x\, R\, y\}$$

también es una relación de orden sobre A; ésta se llama el *orden inverso*, el *orden dual* o el *orden opuesto* del orden R sobre A.

En lo siguiente, se utilizarán símbolos tales que $\leq$, $\preceq$, $\preccurlyeq$, etc. (posiblemente afectados con sub/superíndices) para indicar las relaciones de orden, así como los símbolos "en espejo" $\geq$, $\succeq$, $\succcurlyeq$, etc. para indicar las relaciones de orden opuestas correspondientes.

Elementos notables Sea $(A, \leq)$ un conjunto ordenado. Dado un elemento $x \in A$, se dice que:

- x es un *elemento minimal* de A si: $(\forall y \in A)\,(y \leq x \Rightarrow x = y)$;
- x es un *elemento maximal* de A si: $(\forall y \in A)\,(y \geq x \Rightarrow x = y)$;
- x es el *mínimo* de A si: $(\forall y \in A)\,(x \leq y)$;
- x es el *máximo* de A si: $(\forall y \in A)\,(x \geq y)$.

Cuando existe, el mínimo (resp. el máximo) de A es único, y se escribe $\text{mín}(A)$ (resp. $\text{máx}(A)$); también es el único elemento minimal (resp. el único elemento maximal) de A. Al contrario, el conjunto A puede tener cero, uno o múltiples elementos minimales (resp. elementos maximales) sin tener mínimo (resp. máximo).

Además, dado un elemento $x \in A$ y un subconjunto $S \subseteq A$, se dice que:

- x es una *cota inferior* de S si: $(\forall y \in S)\,(x \leq y)$;
- x es una *cota superior* de S si: $(\forall y \in S)\,(x \geq y)$.

Un subconjunto $S \subseteq A$ puede tener cero, una o múltiples cotas inferiores (resp. cotas superiores) en A. Cuando una cota inferior (resp. una cota superior) de S en A pertenece a S si mismo, es única (en S); y ésta es el mínimo (resp. el máximo) de S para el orden inducido.

Se llama el *ínfimo* (resp. *supremo*) de S el máximo (resp. el mínimo) del conjunto de las cotas inferiores (de las cotas superiores) de S en A, cuando existe:

$$\begin{array}{rcl} \text{ínf}(S) & := & \text{máx}\{x \in A : (\forall y \in S)\,(x \leq y)\} \\ \sup(S) & := & \text{mín}\{x \in A : (\forall y \in S)\,(x \geq y)\} \end{array}$$

(Insistamos en que tales elementos no siempre existen.) Se observa que, cuando existe, el ínfimo (resp. el supremo) de S pertenece a S si y sólo si S tiene mínimo (resp. máximo); y en este caso, tenemos que $\text{ínf}(S) = \text{mín}(S)$ (resp. $\sup(S) = \text{máx}(S)$). También se observa que $\text{ínf}(\varnothing) = \sup(A) = \text{máx}(A)$ (si existe) y $\sup(\varnothing) = \text{ínf}(A) = \text{mín}(A)$ (si existe).

Funciones monótonas Sean $(A, \leq_A)$, $(B, \leq_B)$, $(C, \leq_C)$ conjuntos ordenados. Se dice que una función $f : A \to B$ es *monótona* (respecto a los órdenes $\leq_A$ y $\leq_B$) cuando $x \leq_A y$ implica $f(x) \leq_B f(y)$ para todos $x, y \in A$:

$$f : A \to B \text{ monótona } \equiv (\forall x, y \in A)\, (x \leq_A y \Rightarrow f(x) \leq_B f(y))\,.$$

Se verifica inmediatamente que la función identidad $\mathrm{id}_A : A \to A$ es monótona y que la compuesta $g \circ f : A \to C$ de dos funciones monótonas $f : A \to B$ y $g : B \to C$ también es una función monótona (respecto a los órdenes correspondientes). Sin embargo, la función inversa $f^{-1} : B \to A$ de una función $f : A \to B$ monótona y biyectiva no es monótona en general.

Se llama un *isomorfismo* de $(A, \leq_A)$ hasta $(B, \leq_B)$ toda biyección $f : A \to B$ tal que ambas funciones $f : A \to B$ y $f^{-1} : B \to A$ son monótonas (respecto a los órdenes correspondientes). Se verifica sin dificultad que una biyección $f : A \to B$ es un isomorfismo de $(A, \leq_A)$ hasta $(B, \leq_B)$ si y sólo si: $(\forall x, y \in A)\, (x \leq_A y \Leftrightarrow f(x) \leq_B f(y))$.

Observación 1.31 (Estructura categórica)**.** — La clase de los conjuntos ordenados equipados con las funcionas monótonas forma naturalmente una *categoría*, que se llama la *categoría de los órdenes*, y se escribe **Ord**. En esta categoría:

- Los *objetos* son los conjuntos ordenados $\mathcal{A} = (A, \leq)$.
- Las *flechas* entre dos objetos $\mathcal{A} = (A, \leq_A)$ y $\mathcal{B} = (B, \leq_B)$ son las funciones monótonas f de $\mathcal{A}$ hasta $\mathcal{B}$ (notación: $f : \mathcal{A} \to \mathcal{B}$).
- La *compuesta de dos flechas* $f : \mathcal{A} \to \mathcal{B}$ y $g : \mathcal{B} \to C$ es la función $g \circ f : \mathcal{A} \to C$.
- La *flecha identidad* en un objeto $\mathcal{A}$ es la función monótona $\mathrm{id}_{\mathcal{A}} : \mathcal{A} \to \mathcal{A}$.

Esta estructura cumple obviamente los dos axiomas de las categorías:

(Asociatividad)	$(h \circ g) \circ f = h \circ (g \circ f)$	(si $f : \mathcal{A} \to \mathcal{B}$, $g : \mathcal{B} \to C$, $h : C \to \mathcal{D}$)
(Identidad)	$f \circ \mathrm{id}_{\mathcal{A}} = \mathrm{id}_{\mathcal{B}} \circ f = f$	(si $f : \mathcal{A} \to \mathcal{B}$)

Preórdenes Sea A un conjunto. Una *relación de preorden* sobre A es una relación binaria reflexiva y transitiva sobre A. (Se observa que las relaciones de orden como las relaciones de equivalencia son casos particulares de preórdenes.) Un *conjunto preordenado* es un par $(A, \leq)$ formado por un conjunto A equipado con un preorden $\leq$ sobre A.

Cada preorden $\leq$ sobre A induce una relación de equivalencia $\sim$ sobre A, definida por

$$x \sim y \equiv x \leq y \wedge y \leq x \qquad \text{(para todos } x, y \in A)$$

(Ésta es la relación de igualdad cuando la relación $\leq$ ya es una relación de orden.) A través de la sobreyección canónica $\pi_\sim : A \to (A/{\sim})$ asociada a la equivalencia $\sim$, el preorden $\leq$ sobre A induce una relación de orden $\leq$ sobre el conjunto cociente $A/{\sim}$, la cual es definida por:

$$c \leq c' \equiv (\exists x \in c)\, (\exists x' \in c')\, (x \leq x') \qquad \text{(para todos } c, c' \in A/{\sim})$$

(La definición no depende de los representantes x y x' elegidos en las clases c y c'.)

En otros términos, el mecanismo anterior permite transformar cualquier conjunto preordenado en un conjunto ordenado, remplazando el conjunto de base por su cociente a través la relación de equivalencia inducida por el preorden.

Más generalmente, la terminología asociada a los conjuntos ordenados se generaliza naturalmente a los conjuntos preordenados, remplazando en cada definición la relación de igualdad $x = y$ por la relación de equivalencia $x \sim y$ asociada al preorden. Por ejemplo, si $x \in A$ es un elemento de un conjunto preordenado $(A, \preceq)$, se dice que:

- x es un *elemento minimal* de A si: $\quad (\forall y \in A)\,(y \preceq x \Rightarrow x \sim y)$;
- x es un *elemento maximal* de A si: $\quad (\forall y \in A)\,(y \succeq x \Rightarrow x \sim y)$.

En otro lado, las nociones de máximo/mínimo, de cota inferior/superior y de ínfimo/supremo se definen exactamente del mismo modo que en los conjuntos ordenados, teniendo en cuenta que en un conjunto preordenado $(A, \preceq)$, el mínimo/máximo de A y el ínfimo/supremo de un subconjunto $S \subseteq A$ (cuando existen) no son únicos en general.

Asimismo, las nociones de *función monótona* y de *isomorfismo* entre dos conjuntos preordenados se definen con las mismas fórmulas que para los conjuntos ordenados. Como la clase de los conjuntos ordenados, la clase de los conjuntos preordenados equipados con las funciones monótonas (generalizadas a los preórdenes) tiene una estructura de categoría, que se llama la *categoría de los preórdenes*, y se escribe **Preord**.

Órdenes estrictos La noción de *orden* que definimos al comienzo de la Sección 1.9.3 es la noción de orden *en el sentido amplio*, que incluye el caso de igualdad (por reflexividad):

$$(\forall x, y \in A)(x = y \Rightarrow x \leq y) \qquad \text{(si } \leq \text{ es un orden sobre } A\text{)}$$

Cuando se desea excluir el caso de igualdad, se utiliza la noción de *orden estricto* en lugar de la noción de orden (en el sentido amplio). Formalmente, dado un conjunto A, se llama una relación de *orden estricto* sobre A toda relación binaria $R \subseteq A \times A$ irreflexiva y transitiva:

$$\begin{aligned} R \text{ orden estricto sobre } A \;\equiv\;\; & R \subseteq A \times A && \land \\ & (\forall x \in A)\,\neg(x\,R\,x) && \land \\ & (\forall x, y, z \in A)\,(x\,R\,y \land y\,R\,z \Rightarrow x\,R\,z) \end{aligned}$$

Insistamos en que las relaciones de orden estricto *no son* casos particulares de las relaciones de orden (en el sentido amplio), ambas nociones siendo disjunta cuando $A \neq \varnothing$. (En efecto, la relación vacía sobre el conjunto vacío es la única relación que es a la vez un orden estricto y un orden en el sentido amplio.) Sin embargo, se demuestra sin dificultad que:

Proposición 1.32. — *Dado un conjunto A:*

(1) Para toda relación de orden estricto $<$ sobre A, la relación $\leq$ definida por

$$x \leq y \;\equiv\; x < y \lor x = y \qquad (\textit{para todos } x, y \in A)$$

(es decir: $(\leq) := (<) \cup \mathrm{id}_A$) es una relación de orden sobre A (en el sentido amplio).

(2) Para toda relación de orden $\leq$ sobre A (en el sentido amplio), la relación $<$ definida por

$$x < y \;\equiv\; x \leq y \land x \neq y \qquad (\textit{para todos } x, y \in A)$$

(es decir: $(<) := (\leq) - \mathrm{id}_A$) es una relación de orden estricto sobre A.

(3) Las dos correspondencias anteriores son recíprocas la una de la otra, y definen biyecciones inversas entre el conjunto de las relaciones de orden estricto sobre A y el conjunto de las relaciones de orden (en el sentido amplio) sobre A.

1.10. Ejercicios

1.10.1. Equipotencia

Se recuerda que dos conjuntos A y B son *equipotentes* cuando existe una biyección $f : A \xrightarrow{\sim} B$. La relación de equipotencia es una relación de equivalencia sobre el universo $\mathscr{U}$.

Ejercicio 1.1. — Para cada uno de los pares de conjuntos siguientes, construir una biyección "natural" entre los dos conjuntos:

(1) $\mathfrak{P}(A)$ y 2^A (con $2 = \{0, 1\}$)
(2) $A \times (B + C)$ y $(A \times B) + (A \times C)$
(3) $A \times \sum_{i \in I} B_i$ y $\sum_{i \in I}(A \times B_i)$
(4) A^{B+C} y $A^B \times A^C$
(5) $A^{\left(\sum_{i \in I} B_i\right)}$ y $\prod_{i \in I} A^{B_i}$
(6) $(A \times B)^C$ y $A^C \times B^C$
(7) $(\prod_{i \in I} A_i)^B$ y $\prod_{i \in I} A_i^B$
(8) $A^{B \times C}$ y $(A^B)^C$
(9) $(A + B)^C$ y $\sum_{D \in \mathfrak{P}(C)}(A^D \times B^{C-D})$ (fórmula del binomio)

Ejercicio 1.2 (Cantor)**.** — Sea A un conjunto.

(1) Demostrar que no existe ninguna sobreyección $f : A \to \mathfrak{P}(A)$.
(*Sugerencia*: considerar el conjunto $\{x \in A : x \notin f(x)\}$.)
(2) Deducir de (1) que no existe ninguna inyección $g : \mathfrak{P}(A) \to A$.
(3) Deducir de lo anterior que para todo conjunto a, tenemos que $\mathfrak{P}(\bigcup a) \notin a$.

Ejercicio 1.3 (Teorema de Cantor-Bernstein-Schröder)**.** — Sea X un conjunto dado con una inyección $h : X \hookrightarrow X$ (que define así una biyección entre X y $h(X)$), e Y un conjunto tal que $h(X) \subseteq Y \subseteq X$. Se trata de demostrar que Y es equipotente a X.

(1) Demostrar que existe un subconjunto *minimal* $Z \subseteq X$ tal que $(X - Y) \subseteq Z$ y $h(Z) \subseteq Z$.
(2) Demostrar que $h(Z) = Z \cap Y$.
(3) Utilizando la inyección $h : X \hookrightarrow X$ y el subconjunto $Z \subseteq X$ definido en (1), construir una biyección $h' : X \xrightarrow{\sim} Y$. (*Sugerencia*: definir $h'(x)$ por casos según que $x \in Z$ o no.)
(4) Deducir de lo anterior el siguiente teorema:

Teorema (Cantor-Bernstein-Schröder). — *Si A y B son dos conjuntos tales que existen inyecciones $f : A \hookrightarrow B$ y $g : B \hookrightarrow A$, entonces A y B son equipotentes.*

Conjuntos numerables Se dice que un conjunto A es *numerable*[10] cuando existe una biyección $f : \mathbb{N} \xrightarrow{\sim} A$. En el siguiente ejercicio, se supone conocidas las definiciones y propiedades básicas de las operaciones aritméticas $+$, $\times$ y del orden $\leq$ en $\mathbb{N}$.

Ejercicio 1.4 (Conjuntos numerables)**.** — Demostrar las siguientes propiedades, utilizando el teorema de Cantor-Bernstein-Schröder (Ejercicio 1.3) cuando se necesita.

[10] Algunos autores utilizan la palabra *numerable* en el sentido de «numerable o finito». En este curso, se utiliza la palabra *numerable* con su sentido estricto, es decir: «infinito numerable».

(1) La siguiente función $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ es biyectiva:

$$f(n,m) := (n+m)(n+m+1)/2 + n \qquad \text{(para todo } (n,m) \in \mathbb{N} \times \mathbb{N})$$

(2) Si dos conjuntos A y B son numerables, entonces $A \times B$ es numerable igualmente.

(3) Los conjuntos $\mathbb{Z}$ (enteros relativos) y $\mathbb{Q}$ (números racionales) son numerables.

(4) Si $(A_i)_{i\in I}$ es una familia de conjuntos numerables indizada por un conjunto numerable I, entonces ambos conjuntos $\sum_{i\in I} A_i$ y $\bigcup_{i\in I} A_i$ son numerables igualmente.

(5) Si A es un conjunto numerable, entonces el conjunto $\mathfrak{P}_{\text{fin}}(A)$ $(\subset \mathfrak{P}(A))$ formado por todos los subconjuntos finitos de A es numerable igualmente.

(6) Si A es un anillo conmutativo numerable (por ejemplo: $\mathbb{Z}$), entonces el anillo $A[X]$ de los polinomios con coeficientes en A es numerable igualmente.

(7) Deducir de lo anterior que el conjunto de los *números algebraicos*

$$\mathbb{A} = \{x \in \mathbb{R} : (\exists p \in \mathbb{Z}[X])\, p(x) = 0\}$$

es numerable igualmente.

El conjunto $\mathbb{R}$ de los números reales En lo siguiente, se suponen conocidas la definición y las propiedades de los números reales, y se escribe $\mathbb{R}$ el conjunto correspondiente. Para todos $a, b \in \mathbb{R}$ tales que $a < b$, se definen los intervalos $(a,b), (a,b], [a,b), [a,b] \subseteq \mathbb{R}$ por:

$$\begin{array}{llll} (a,b) := \{x \in \mathbb{R} : a < x < b\} & & (a,b] := \{x \in \mathbb{R} : a < x \le b\} \\ [a,b) := \{x \in \mathbb{R} : a \le x < b\} & & [a,b] := \{x \in \mathbb{R} : a \le x \le b\} \end{array}$$

Ejercicio 1.5 (Intervalos de $\mathbb{R}$)**.**

(1) Verificar que la función $f : \mathbb{R} \to (-1,1)$ definida por $f(x) = x/(1+|x|)$ es biyectiva.

(2) Deducir que (a,b) es equipotente con $\mathbb{R}$ para todos $a, b \in \mathbb{R}$ tales que $a < b$.

(3) Utilizando el teorema de Cantor-Bernstein-Schröder (Ejercicio 1.3 (4)), deducir que los intervalos (a,b), $(a,b]$, $[a,b)$ y $[a,b]$ son equipotentes con $\mathbb{R}$ para todos $a < b$.

Ejercicio 1.6 (Desarrollo de un número real en base b)**.** — Sea $b \in \mathbb{N}$, $b \ge 2$ una base de numeración. Se escribe igualmente $b = \{0, \dots, b-1\}$ el conjunto de los *dígitos en base* b[11]. Para todo $x \in [0,1)$, el *desarrollo de x en base b* es la sucesión $(x \#_b n)_{n\ge1} \in b^{\mathbb{N}^*}$ definida por:

$$x \#_b n := \lfloor b^n x \rfloor \bmod b \qquad \text{(para todo } n \ge 1)$$

donde $\lfloor y \rfloor$ indica la parte entera inferior del número real $y \in \mathbb{R}$, y donde $k \bmod b$ indica el resto de la división euclidiana del entero $k \in \mathbb{N}$ por b.

(1) Verificar que para todo $x \in [0,1)$: $\quad x = \sum_{n\ge1} (x \#_b n) b^{-n}$.

(2) Demostrar que la función $f : [0,1) \to b^{\mathbb{N}}$ dada por $f(x) = (x \#_b (n+1))_{n\in\mathbb{N}}$ para todo $x \in [0,1)$ es bien definida e inyectiva.

[11] La notación $b = \{0, \dots, b-1\}$ es consistente con la representación de los enteros naturales como ordinales que introduciremos y adoptaremos en el Capítulo 2.

(3) Demostrar que la función $f' : b^{\mathbb{N}} \to [0,1]$ dada por $f'((d_n)_{n\in\mathbb{N}}) = \sum_{n\in\mathbb{N}} d_n b^{-n-1}$ para toda sucesión de dígitos $(d_n)_{n\in\mathbb{N}} \in b^{\mathbb{N}}$ es bien definida y sobreyectiva, pero no inyectiva. ¿Cuáles son los números $x \in [0,1]$ que tienen más de un antecedente por f'?
(4) Demostrar que la función $f'' : b^{\mathbb{N}} \to [0,1]$ dada por $f''((d_n)_{n\in\mathbb{N}}) = \sum_{n\in\mathbb{N}} d_n (b+1)^{-n-1}$ para toda sucesión de dígitos $(d_n)_{n\in\mathbb{N}} \in b^{\mathbb{N}}$ es bien definida e inyectiva.
(5) Utilizando el teorema de Cantor-Bernstein-Schröder (Ejercicio 1.3 (4)), deducir que los conjuntos $\mathbb{R}$, $[0,1)$ y $b^{\mathbb{N}}$ son equipotentes para toda base $b \geq 2$.
(6) Deducir de lo anterior que $\mathbb{R}$ es equipotente con $\mathfrak{P}(\mathbb{N})$, y que no es equipotente con $\mathbb{N}$.

Ejercicio 1.7. — Utilizando los resultados del ejercicio anterior, demostrar que:

(1) $\mathbb{R}^2 := \mathbb{R} \times \mathbb{R}$ es equipotente con $\mathbb{R}$.
(2) $\mathbb{R}^n$ es equipotente con $\mathbb{R}$ para todo $n \geq 1$.
(3) $\mathbb{R}^{\mathbb{N}}$ es equipotente con $\mathbb{R}$. (*Sugerencia*: observar que $\mathbb{R}^{\mathbb{N}}$ es equipotente con $(2^{\mathbb{N}})^{\mathbb{N}}$.)
(4) $\mathbb{R}^{\mathbb{R}}$ es equipotente con $\mathfrak{P}(\mathbb{R})$, y no es equipotente con $\mathbb{R}$.

1.10.2. Relaciones

Ejercicio 1.8 (Clausuras de relaciones)**.** — A cada relación binaria $R \subseteq A \times A$ se asocian las tres relaciones $R^=, R^\leftrightarrow, R^+ \subseteq A \times A$ definidas por:

$$\begin{array}{rcll} R^= & := & R \cup \mathrm{id}_A & (\text{donde } \mathrm{id}_A = \{(x,y) \in A \times A \ : \ x = y\}) \\ R^\leftrightarrow & := & R \cup R^{-1} & (\text{donde } R^{-1} = \{(x,y) \in A \times A \ : \ y\,R\,x\}) \\ R^+ & := & \multicolumn{2}{l}{\bigcap \{S \in \mathfrak{P}(A \times A) \ : \ S \text{ transitiva } \wedge \ R \subseteq S\}.} \end{array}$$

Las tres relaciones $R^=$, $R^\leftrightarrow$ y R^+ se llaman respectivamente la *clausura reflexiva*, la *clausura simétrica* y la *clausura transitiva* de la relación R.

(1) Demostrar que $R^=$ es la relación reflexiva más pequeña que contiene R.
(2) Demostrar que $R^\leftrightarrow$ es la relación simétrica más pequeña que contiene R.
(3) Demostrar que R^+ es la relación transitiva más pequeña que contiene R.
(4) Deducir de lo anterior que $(R^=)^= = R^=$, $(R^\leftrightarrow)^\leftrightarrow = R^\leftrightarrow$ y $(R^+)^+ = R^+$.
(5) Deducir de (3) que para todos $x, y \in A$:

$$\begin{array}{rcl} x\,R^+\,y & \Leftrightarrow & x\,R\,y \vee (\exists z \in A)\,(x\,R\,z \wedge z\,R^+\,y) \\ & \Leftrightarrow & x\,R\,y \vee (\exists z \in A)\,(x\,R^+\,z \wedge z\,R\,y) \end{array}$$

Se define la *clausura reflexiva-transitiva* de R por: $R^* := (R^+)^=$.

(6) Demostrar que R^* es el preorden más pequeño que contiene R.
(7) Demostrar que: $R^* = (R^+)^= = (R^=)^+ = (R^=)^* = (R^*)^= = (R^+)^* = (R^*)^+ = (R^*)^*$.

Se define la *clausura reflexiva-simétrica-transitiva* de R por: $R^\sim := (R^\leftrightarrow)^*$.

(8) Demostrar que $R^\sim$ es la relación de equivalencia más pequeña que contiene R.
(9) Demostrar que $(R^*)^\leftrightarrow \subseteq (R^\leftrightarrow)^* = R^\sim$, y dar un contraejemplo a la inclusión recíproca.
(10) Explicar por qué no se puede definir una noción de clausura antisimétrica.

Relaciones bien fundadas Dado una relación binaria $R \subseteq A \times A$, se llama el *principio de inducción bien fundada* (respecto a la relación R) la siguiente fórmula:

$$(\forall X \subseteq A)\,[(\forall x \in A)\,((\forall y \in A)\,(y\,R\,x \Rightarrow y \in X) \;\Rightarrow\; x \in X) \;\Rightarrow\; X = A]\,.$$

Cuando esta fórmula se cumple, se dice que la relación R es *bien fundada* sobre A. Es importante observar que la propiedad de buena fundación no implica la propiedad de transitividad; por ejemplo, la relación $R \subseteq \mathbb{N} \times \mathbb{N}$ («relación sucesor») definida por

$$n\,R\,m \;\equiv\; s(n) = m \qquad \text{(para todos } n, m \in \mathbb{N}\text{)}$$

es bien fundada (véase Ejercicio 1.17 más abajo), pero no es transitiva.

Ejercicio 1.9 (Caracterización de las relaciones bien fundadas)**.**

(1) Demostrar que una relación $R \subseteq A \times A$ es bien fundada sobre A si y sólo si todo subconjunto no vacío $X \subseteq A$ tiene un elemento R-minimal, es decir:
$$(\forall X \subseteq A)\,[X \neq \varnothing \;\Rightarrow\; (\exists x \in X)\,(\forall y \in X)\,\neg(y\,R\,x)]$$
(2) Deducir que toda relación bien fundada es irreflexiva.
(3) Demostrar que si una relación $R \subseteq A \times A$ es bien fundada sobre A, entonces no existe ninguna sucesión $(x_n)_{n\in\mathbb{N}} \in A^{\mathbb{N}}$ tal que $x_{n+1}\,R\,x_n$ para todo $n \in \mathbb{N}$.
(4) Utilizando el axioma de elección (véase Capítulo 2), demostrar la recíproca de (3).

Ejercicio 1.10 (Clausura transitiva de una relación bien fundada)**.** — Sea R una relación binaria sobre un conjunto A, y R^+ su clausura transitiva (véase Ejercicio 1.8). A cada subconjunto $X \subseteq A$, se asocia el subconjunto $X^+ \subseteq A$ definido por

$$X^+ \;:=\; X \cup \{z \in A : (\exists x, y \in X)\,(x\,R^+\,z \wedge z\,R^+\,y)\}$$

Se llama un *elemento R-minimal de X* todo elemento $x \in X$ tal que $(\forall y \in X)\,\neg(y\,R\,x)$; la noción de *elemento R^+-minimal* se define de modo similar.

(1) Demostrar que si x es un elemento R-minimal de X^+, entonces $x \in X$ y x es un elemento R^+-minimal de X. (*Sugerencia:* utilizar las equivalencias del Ejercicio 1.8 (5).)
(2) Deducir que si la relación R es bien fundada, entonces su clausura transitiva R^+ es bien fundada igualmente. (*Sugerencia:* utilizar la caracterización del Ejercicio 1.9 (1).)
(3) Suponiendo la relación R bien fundada sobre A, ¿qué se puede decir en general sobre su clausura reflexiva $R^=$ y su clausura simétrica $R^{\leftrightarrow}$?

1.10.3. Relaciones de orden

Ejercicio 1.11 (Orden producto)**.** — Sea $(\mathcal{A}_i)_{i\in I} = (A_i, \leq_i)_{i\in I}$ una familia de conjuntos ordenados indizada por un conjunto I cualquiera. Se considera la relación binaria $(\leq_P)$ definida sobre el producto cartesiano (generalizado) $P = \prod_{i\in I} A_i$ por:

$$(a_i)_{i\in I} \leq_P (a'_i)_{i\in I} \;\equiv\; (\forall i \in I)(a_i \leq_i a'_i) \qquad \text{(para todos } (a_i)_{i\in I}, (a'_i)_{i\in I} \in P\text{)}$$

(1) Demostrar que $(\leq_P)$ es una relación de orden sobre P. El orden $\leq_P$ se llama el *orden producto* sobre $P = \prod_{i\in I} A_i$ respecto a la familia de órdenes $(\leq_i)_{i\in I}$.

(2) Demostrar que para todo $i \in I$, la proyección $\pi_i : P \to A_i$ definida por $\pi_i((a_i)_{i\in I}) = a_i$ (para todo $(a_i)_{i\in I} \in P$) es una función monótona respecto a los ordenes $\leq_P$ y $\leq_i$.

(3) Demostrar que, en general, la relación de orden $\leq_P$ no es total sobre P, incluso cuando todos los órdenes $\leq_i$ son totales sobre los conjuntos A_i ($i \in I$).

(4) Demostrar que en la categoría de los conjuntos ordenados (véase Obs. 1.31 p. 33), el conjunto ordenado $(P, \leq_P)$ satisface una propiedad universal similar a la de la Prop. 1.14 p. 22, remplazando en el diagrama todos los conjuntos involucrados por conjuntos ordenados, y sólo considerando funciones monótonas.

Ejercicio 1.12 (Orden suma). — Sea $(\mathcal{A}_i)_{i\in I} = (A_i, \leq_i)_{i\in I}$ una familia de conjuntos ordenados indizada por un conjunto I cualquiera. Se considera la relación binaria $(\leq_S)$ definida sobre la suma directa $S = \sum_{i\in I} A_i$ por:

$$(i, a) \leq (i', a') \ \equiv \ i = i' \wedge a \leq_i a' \qquad \text{(para todos } (i, a), (i', a') \in S)$$

(1) Demostrar que $(\leq_S)$ es una relación de orden sobre S. El orden $\leq_S$ se llama el *orden suma* sobre $S = \sum_{i\in I} A_i$ respecto a la familia de órdenes $(\leq_i)_{i\in I}$.

(2) Demostrar que para todo $i \in I$, la inyección $\sigma_i : A_i \to S$ definida por $\sigma_i(a) = (i, a)$ (para todo $a \in A_i$) es una función monótona respecto a los ordenes $\leq_i$ y $\leq_S$.

(3) Demostrar que, en general, la relación de orden $\leq_S$ no es total sobre S, incluso cuando todos los órdenes $\leq_i$ son totales sobre los conjuntos A_i ($i \in I$).

(4) Demostrar que en la categoría de los conjuntos ordenados (véase Obs. 1.31 p. 33), el conjunto ordenado $(S, \leq_S)$ satisface una propiedad universal similar a la de la Prop. 1.16 p. 23, remplazando en el diagrama todos los conjuntos involucrados por conjuntos ordenados, y sólo considerando funciones monótonas.

Ejercicio 1.13 (Orden lexicográfico sobre la suma directa). — Sea $(\mathcal{A}_i)_{i\in I} = (A_i, \leq_i)_{i\in I}$ una familia de conjuntos ordenados indizada por un conjunto ordenado $\mathcal{I} = (I, \leq_I)$. Se considera la relación binaria $(\leq_{\text{lex}})$ definida sobre la suma directa $S = \sum_{i\in I} A_i$ por:

$$(i, a) \leq (i', a') \ \equiv \ i <_I i' \vee (i = i' \wedge a \leq_i a') \qquad \text{(para todos } (i, a), (i', a') \in S)$$

(1) Demostrar que $(\leq_{\text{lex}})$ es una relación de orden sobre S. El orden $\leq_{\text{lex}}$ se llama el *orden lexicográfico* sobre la suma $S = \sum_{i\in I} A_i$ respecto a los órdenes $\leq_I$ y $\leq_i$ (para todo $i \in I$).

(2) Demostrar que para todo $i \in I$, la inyección $\sigma_i : A_i \to S$ definida por $\sigma_i(a) = (i, a)$ (para todo $a \in A_i$) es una función monótona respecto a los ordenes $\leq_i$ y $\leq_{\text{S}}$.

(3) Considerando el orden $\leq_S$ definido en el ejercicio anterior (Ejercicio 1.12), demostrar que la función $\mathrm{id}_S : (S, \leq_S) \to (S, \leq_{\text{lex}})$ es monótona, pero en general no un isomorfismo.

(4) Demostrar que si los órdenes $\leq_I$ (sobre I) y $\leq_i$ (sobre A_i para todo $i \in I$) son totales, entonces el orden lexicográfico $\leq_{\text{lex}}$ es total igualmente sobre S.

Observación. En el caso particular donde $\mathcal{A}_i = \mathcal{A} = (A, \leq_A)$ para todo $i \in I$ (es decir: cuando la familia $(\mathcal{A}_i)_{i\in I}$ es constante), vimos que la suma directa $S = \sum_{i\in I} A_i$ se reduce a un producto cartesiano binario: $S = I \times A$. En este caso, el conjunto ordenado $(S, \leq_{\text{lex}})$ se escribe $\mathcal{I} \times_{\text{lex}} \mathcal{A}$, y se llama el *producto lexicográfico* de los conjuntos ordenados $\mathcal{I}$ y $\mathcal{A}$.

Ejercicio 1.14 (Orden lexicográfico sobre las palabras)**.** — Dado un conjunto ordenado $(A, \leq)$, se escribe A^* el conjunto de todas las sucesiones finitas de la forma $(x_i)_{i\in[1..n]}$, donde $n \in \mathbb{N}$ y $x_i \in A$ para todo $i \in [1..n]$. Los elementos $(x_i)_{i\in[1..n]} \in A^*$ también se llaman las *palabras* sobre el alfabeto A, y se utiliza la notación sugestiva $x_1 \cdots x_n := (x_i)_{i\in[1..n]}$ para representarlas. (La palabra vacía se escribe ε.) Se define la relación binaria $\leq^*_{\text{lex}}$ sobre A^* por:

$$\begin{array}{rcl}(x_i)_{i\in[1..n]} \leq^*_{\text{lex}} (y_i)_{i\in[1..m]} & \equiv & (n \leq m \ \wedge\ (\forall i \leq n)\, x_i = y_i) \quad \vee \\ & & (\exists k \leq \text{mín}(n, m))\, ((\forall i < k)\, x_i = y_i \ \wedge\ x_k < y_k)\end{array}$$

(En la definición anterior, los enteros naturales i y k son supuestos ≥ 1.)

(1) Demostrar que $\leq^*_{\text{lex}}$ es una relación de orden sobre el conjunto A^*, cuyo mínimo es la palabra vacía ε. El orden $\leq^*_{\text{lex}}$ se llama el *orden lexicográfico* sobre el conjunto A^*.

(2) Demostrar que si el orden $\leq$ es total sobre A, entonces el orden $\leq^*_{\text{lex}}$ es total sobre A^*.

Buenos ordenes Un *buen orden* sobre un conjunto A es una relación de orden $(\leq) \subseteq A \times A$ tal que todo subconjunto no vacío de A tenga un mínimo:

$$\begin{array}{rcl}(\leq) \text{ buen orden sobre } A & \equiv & (\leq) \text{ orden (amplio) sobre } A \ \wedge \\ & & (\forall X \subseteq A)\, (X \neq \varnothing \Rightarrow (\exists x \in X)(\forall y \in X)\, x \leq y)\end{array}$$

Un conjunto *bien ordenado* es un conjunto ordenado $(A, \leq)$ cuyo orden es un buen orden.

Ejercicio 1.15 (Orden lexicográfico y buen orden)**.** — Sea $(\mathcal{A}_i)_{i\in I} = (A_i, \leq_i)_{i\in I}$ una familia de conjuntos bien ordenados indizada por un conjunto bien ordenado $\mathcal{I} = (I, \leq_I)$, y $S = \sum_{i\in I} A_i$ la suma directa de los conjuntos subyacentes.

(1) Demostrar que el orden lexicográfico $\leq_{\text{lex}}$ sobre S (Ejercicio 1.13) es un buen orden.

Sean $(A, \leq)$ un conjunto bien ordenado, y A^* el conjunto de las palabras sobre el alfabeto A (Ejercicio 1.14). Para toda palabra $u = (x_i)_{i\in[1..n]} \in A^*$, se escribe $|u| = n$ su *longitud*.

(2) Mostrar con un contraejemplo que el orden lexicográfico $\leq^*_{\text{lex}}$ sobre A^* (Ejercicio 1.14) no es un buen orden en general.

(3) Demostrar que la relación binaria $(\leq^{**}_{\text{lex}}) \subseteq A^* \times A^*$ definida por

$$u \leq^{**}_{\text{lex}} v \ \equiv\ |u| < |v| \ \vee\ (|u| = |v| \wedge u \leq^*_{\text{lex}} v)$$

(para todas palabras $u, v \in A^*$) es un buen orden sobre A^*.

En lo siguiente, se llama un *buen orden estricto* todo orden estricto sobre un conjunto A cuya relación de orden asociada (en el sentido amplio) es un buen orden sobre A:

$$\begin{array}{rcl}(<) \text{ buen orden estricto sobre } A & \equiv & (<) \text{ orden estricto sobre } A \ \wedge \\ & & (\forall X \subseteq A)\, (X \neq \varnothing \Rightarrow (\exists x \in X)(\forall y \in X)\, (x = y \vee x < y))\,.\end{array}$$

Ejercicio 1.16 (Caracterización de los buenos órdenes estrictos)**.** — Se dice que una relación binaria $R \subseteq A \times A$ es *conexa* cuando: $(\forall x, y \in A)\, (x\, R\, y \vee x = y \vee y\, R\, x)$.

(1) Demostrar que todo buen orden estricto sobre A es una relación conexa.

(2) Demostrar que todo buen orden estricto sobre A es una relación bien fundada.

(3) Recíprocamente, demostrar que toda relación conexa y bien fundada sobre A es un buen orden estricto sobre A. (No se supone que dicha relación es un orden estricto.)

Ejercicio 1.17 (Buen orden sobre una estructura aritmética). — En este ejercicio, se fija una estructura aritmética (N, o, s) cualquiera (Def. 1.20 p. 24). Se considera la relación «sucesor» $S \subseteq N \times N$ definida para todos $x, y \in N$ por $x\,S\,y \equiv s(x) = y$, y se escriben respectivamente S^+ y S^* su clausura transitiva y su clausura reflexiva-transitiva (véase Ejercicio 1.8).

(1) Demostrar que la relación S es bien fundada. Deducir (con el resultado del Ejercicio 1.10) que su clausura transitiva S^+ es bien fundada igualmente.

(2) Deducir de lo anterior que la relación S^+ es un orden estricto, mientras la relación S^* es el orden asociado (en el sentido amplio).

En lo siguiente, se escriben $(<) := S^+$ y $(\leq) := S^*$.

(3) Demostrar por inducción que: $(\forall x \in N)\,(x = o \lor o < x)$.

(4) Demostrar que la relación $(<) = S^+$ es conexa: $(\forall x, y \in N)\,(x < y \lor x = y \lor y < x)$. (*Sugerencia:* se efectúa la demostración por una inducción doble sobre x e y, utilizando las equivalencias del Ejercicio 1.8 (5).)

(5) Utilizando la caracterización del Ejercicio 1.16, deducir que la relación $(\leq) \subseteq N \times N$ definida por $(\leq) := S^*$ (clausura reflexiva-transitiva de S) es un buen orden sobre N.

(6) Verificar que el elemento $o \in N$ y la función sucesor $s : N \to N$ se pueden caracterizar a partir del buen orden $(\leq)$ sobre N definido en (5) por:

$$o = \text{mín}_{\leq}(X) \qquad \text{y} \qquad s(x) = \text{mín}_{\leq}\{y \in N : x < y\} \qquad \text{(para todo } x \in N)$$