La Axiomática de la Teoría de Conjuntos

Carlos Ivorra

(http://www.uv.es/=ivorra)

1 Introducción

Durante el siglo XIX se llevó a cabo un proceso de fundamentación de la matemática en virtud del cual se fueron precisando paulatinamente todos los conceptos básicos, desde el concepto de límite hasta el de número natural. Finalmente, Frege presentó lo que debería haber sido la culminación de este proceso: una *teoría axiomática de conjuntos,* es decir, un sistema de axiomas a partir de los cuales podían demostrarse rigurosamente todos los resultados básicos aceptados por los matemáticos y, a partir de ellos, todos los teoremas matemáticos. Desgraciadamente, Bertrand Russell descubrió que la axiomática de Frege era contradictoria. En efecto, uno de los axiomas básicos de Frege afirmaba lo siguiente:

> *Para toda propiedad $\phi(X)$ definible en la teoría, existe un conjunto Y cuyos elementos son exactamente los conjuntos X que cumplen $\phi(X)$.*

En otros términos, Frege postulaba la existencia del conjunto

$$Y = \{X \mid \phi(X)\}.$$

Lo que Russell observó fue que esto podía aplicarse a $\phi(X) \equiv X \notin X$, que era una propiedad trivialmente definida en la teoría de Frege, de modo que debía existir un conjunto

$$R = \{X \mid X \notin X\},$$

que claramente nos lleva a la contradicción $R \in R \leftrightarrow R \notin R$.

A partir de aquí, la minuciosa lógica de Frege permitía probar con el mismo rigor que $2+2 = 4$ y que $2+2 = 5$, por lo que su teoría se volvía inservible. El mismo Russell, junto con A. N. Whitehead, presentó un tiempo después otra teoría axiomática que, al menos en apariencia, estaba exenta de contradicciones, si bien era tan inútil como la de Frege, esta vez no por contradictoria sino por complicada. Se trata de los *Principia Mathematica.*

La primera teoría axiomática construida por un matemático a gusto de los matemáticos fue la de Zermelo. La forma en que Zermelo evitó la paradoja de Russell fue debilitar el axioma de formación de conjuntos de Frege, reduciéndolo a:

Para toda propiedad $\phi(X)$ definible en la teoría y todo conjunto U, existe un conjunto Y cuyos elementos son exactamente los elementos $X \in U$ que cumplen $\phi(X)$.

Así, lo que Zermelo postulaba era la existencia de

$$Y = \{X \in U \mid \phi(X)\}.$$

Ahora bien, este axioma sólo permite definir conjuntos a partir de otros conjuntos, por lo que Zermelo tuvo que añadir otros axiomas que garantizaran la existencia de aquellos conjuntos necesarios que no podían obtenerse como subconjuntos de otros conjuntos dados. Enseguida describiremos con detalle la axiomática de Zermelo, pero antes daremos algunas indicaciones sobre la lógica matemática que subyace a toda teoría de conjuntos moderna.

2 La lógica de la teoría de conjuntos

El punto de partida de la teoría de conjuntos moderna consiste en admitir que no podemos dar ninguna definición operativa de "conjunto". El paso siguiente es darse cuenta de que no necesitamos hacerlo. Consideremos el silogismo siguiente:

Toda palabra properispómena es barítona,
δῶρον es una palabra properispómena,
luego δῶρον es una palabra barítona.

Si consultamos una gramática griega y un diccionario veremos que todas estas palabras son de verdad, pero lo maravilloso del caso es que no necesitamos saber lo que significan para concluir que el razonamiento es correcto: Si sabemos que toda palabra properispómena (sea esto lo que sea) es barítona (sea esto lo que sea), así como que "δῶρον" (sea lo que sea) es una palabra properispómena (sea lo que sea), podemos afirmar sin miedo a equivocarnos que "δῶρον" (sea lo que sea) es una palabra barítona (sea lo que sea).

Técnicamente, hacer matemáticas es esto mismo: hablar con absoluto rigor lógico sin preocuparse del significado de los términos empleados. Más concretamente, todo teorema matemático podría formularse así:

Si admitimos que los conjuntos (sean lo que sean), junto con la relación de pertenencia (sea esto lo que sea), cumplen unos axiomas dados, entonces tal afirmación es cierta.

Esto no significa que la palabra "conjunto" carezca de significado (al fin y al cabo "properispómeno" es una palabra más rara y sí que significa algo muy concreto). Existen muchas opiniones al respecto, desde los formalistas radicales que niegan todo significado a los conceptos matemáticos hasta los platonistas radicales que creen que los conjuntos existen de forma objetiva en algún sentido de la palabra. En el término medio estarían las posturas finitistas y similares, según las cuales podemos atribuir un significado concreto a ciertos conjuntos (como mínimo a todos los conjuntos finitos, tal vez también a los numerables o a algunos de ellos, etc.), pero no a otros.

Sea como sea, la lógica matemática permite que estas cuestiones no afecten a la fundamentación de la matemática: tanto si los conjuntos son algo como si no son nada, podemos dar unos axiomas y deducir cosas de ellos con todo rigor.

El primer paso es indicar explícitamente todos los signos del lenguaje matemático. Existen varias alternativas, pero una muy habitual es considerar que el lenguaje de la teoría de conjuntos consta de los 12 signos siguientes:

$$(,\,),\neg,\rightarrow,\vee,\wedge,\leftrightarrow,\forall,\exists,|,=,\in,$$

más una lista potencialmente infinita de signos llamados *variables:*

$$x, y, z, \alpha, \beta, x_1, x_2, \ldots$$

Con estos signos podemos formar cadenas de signos, tales como

$$\forall = (\rightarrow xy, \qquad \text{o} \qquad ((x = y) \rightarrow (y = x)).$$

Las cadenas de signos se dividen en dos tipos: *expresiones* y *cadenas no expresivas.* Informalmente, una expresión es una cadena con sentido (como el segundo ejemplo), mientras que una cadena no expresiva es un galimatías como el primer ejemplo. Ahora bien, no podemos permitirnos el lujo de aludir al posible significado de las cadenas de signos, ya que no está claro que ninguna de ellas vaya a tener alguno. Formalmente, la distinción entre expresiones y cadenas no expresivas se hace a través de reglas sintácticas de construcción. Antes de verlo necesitamos introducir una distinción adicional: una expresión es un *término* si nombra a un objeto y una *fórmula* si afirma algo. De nuevo esta "definición" informal no nos sirve porque alude a un presunto significado tabú. La definición rigurosa de término, fórmula y, por consiguiente, expresión, está contenida en las reglas siguientes:

1. Toda variable es un término.

2. Si t_1 y t_2 son términos, entonces $(t_1 = t_2)$ y $(t_2 \in t_2)$ son fórmulas.

3. Si α y β son fórmulas, entonces $\neg\alpha$, $(\alpha \rightarrow \beta)$, $(\alpha \vee \beta)$, $(\alpha \wedge \beta)$, $(\alpha \leftrightarrow \beta)$ son fórmulas.

4. Si α es una fórmula y X es una variable entonces $\forall X\alpha$ y $\exists X\alpha$ son fórmulas.

5. Si α es una fórmula y X es una variable, entonces $X \mid \alpha$ es un término.

Estas reglas permiten determinar sin ambigüedad alguna si una cadena de signos es una expresión (un término o una fórmula) o si, por el contrario, es no expresiva. Por ejemplo, tenemos que las variables x e y son términos por 1), luego $(x = y)$ e $(y = x)$ son fórmulas por 2), luego $((x = y) \rightarrow (y = x))$ es una fórmula por 3) (y no porque significa algo, como decíamos antes). En cambio, la cadena $\forall = (\rightarrow xy$ no es una expresión, pues la regla 4) exige que detrás de un $\forall$ vaya una variable, y nunca un $=$.

Así pues, podemos afirmar que la cadena

$$\forall X((X \in A) \rightarrow (X \in B))$$

tiene significado: afirma que todos los conjuntos (sea esto lo que sea) que pertenecen (sea esto lo que sea) al conjunto (slqs) A pertenecen (slqs) también al conjunto (slqs) B. Y si algún malicioso nos cuestiona que podamos atribuir un significado a una afirmación sin atribuir un significado concreto a sus signos, tenemos a nuestra disposición una respuesta técnicamente impecable: cuando decimos que tiene significado sólo queremos decir que es una fórmula en el sentido siguiente: X, A y B son términos por la regla 1), $(X \in A)$ y $(X \in B)$ son fórmulas por la regla 2), $((X \in A) \rightarrow (X \in B))$ es una fórmula por la regla 3) y $\forall X((X \in A) \rightarrow (X \in B))$ es una fórmula por la regla 4). Lo cual es absolutamente objetivo y no admite discusión.

Yendo un poco más lejos podríamos atribuir un significado a algunos signos, y establecer que $\wedge$ significa "y", $\neg$ significa "no", etc., es decir, podríamos convertir esto en afirmaciones precisas y rigurosas. Es cuestionable si tiene sentido decir que $=$ significa "igual" o si esto es no decir nada. En cualquier caso, los signos siguientes se resisten a toda precisión operativa:

$$\forall, \quad \exists, \quad \in .$$

Como no disponemos de ninguna definición explícita de conjunto, no podemos explicar qué significa "para todo conjunto se cumple que" ni "existe un conjunto tal que". En esencia, la posibilidad de un uso puramente formal de estos tres signos nos dispensa de definir lo que es un conjunto y lo que es la pertenencia entre conjuntos.

Ahora podríamos definir formalmente lo que es un razonamiento lógico aceptable, tal y como hizo (bien) el propio Frege, pero no merece la pena, pues los problemas de fundamentación de la matemática no están ahí. Después de precisar meticulosamente qué se puede deducir y qué no de unas premisas dadas llegaríamos a la noción intuitiva de deducción lógica que todos tenemos.

A partir de aquí somos libres de introducir tantas definiciones como consideremos oportuno. Técnicamente, una definición no es más que una abreviatura. Por ejemplo, podemos definir (antes incluso de dar ningún axioma) la noción de *inclusión* de conjuntos:

$$A \subset B \equiv \forall X((X \in A) \rightarrow (X \in B)).$$

Las tres rayas $\equiv$ indican que el miembro izquierdo es una abreviatura del miembro derecho (no una mera equivalencia lógica). Teóricamente, cada vez que un matemático escribe $A \subset B$ podría sustituir esto por el miembro derecho. Técnicamente $\subset$ no es un signo del lenguaje de la teoría de conjuntos, sino una mera abreviatura de una fórmula que (como todas las fórmulas) no contiene más que los signos admisibles que hemos indicado al principio.

Para terminar comentaremos que en la práctica relajaremos ligeramente la sintaxis impuesta por la definición de término y fórmula, siempre que esto no lleve a confusión. Por ejemplo, podemos escribir $\forall X \in A(X = U \vee X = V)$, entendiendo que con ello nos referimos a la fórmula $\forall X((X \in A) \rightarrow ((X = U) \vee (X = V)))$.

3 La Axiomática de Zermelo

El primer axioma de la teoría de conjuntos de Zermelo es el *axioma de extensionalidad:*

$$\forall XY(\forall U(U \in X \leftrightarrow U \in Y) \rightarrow X = Y).$$

Afirma que si dos conjuntos tienen los mismos elementos entonces son iguales (el recíproco es un caso particular de un principio lógico: si $X = Y$ entonces todo lo que vale para X vale para Y).

Según hemos comentado en la introducción, el problema que presenta la fundamentación de la teoría de conjuntos es que no podemos permitirnos el lujo de postular que toda propiedad define un conjunto. En su lugar, la teoría de Zermelo postula la existencia de conjuntos definidos por ciertas propiedades inofensivas (no como $X \notin X$). Tal vez el conjunto más inofensivo de todos sea el que nos da el *axioma del conjunto vacío:*

$$\exists X \forall U \; U \notin X.$$

Este axioma afirma la existencia de un conjunto que no tiene elementos. Dicho conjunto es único, pues dos conjuntos sin elementos tendrían los mismos elementos. Esto nos permite definir el término

$$\varnothing \equiv X \mid \forall U \; U \notin X.$$

Según comentábamos en la sección anterior, $\varnothing$ no es un signo del lenguaje de la teoría de conjuntos, sino una abreviatura de un término que puede eliminarse de cualquier expresión sin más que sustituirla por el miembro derecho de la definición. No volveremos a insistir en ello.

En particular tenemos que existen conjuntos. El *axioma del par* implica entre otras cosas que existen infinitos conjuntos:

$$\forall XY \exists Z \forall U(U \in Z \leftrightarrow (U = X \vee U = Y)).$$

Este axioma afirma que, dados dos conjuntos X e Y, existe un tercer conjunto Z cuyos elementos son exactamente X e Y. Dicho Z es único, pues si Z y Z' cumplen lo mismo entonces ambos tienen los mismos elementos, luego son iguales. Esto nos permite definir el término

$$\{X, Y\} \equiv Z \mid \forall U(U \in Z \leftrightarrow (U = X \vee U = Y)).$$

Nunca hemos dicho que X e Y deban ser distintos. Abreviaremos $\{X\} \equiv \{X, X\}$, de modo que, para todo conjunto X tenemos justificada la existencia de un conjunto

$\{X\}$ cuyo único elemento es X. En particular todo conjunto pertenece a otro conjunto. También vemos que existen infinitos conjuntos, como

$$\varnothing, \quad \{\varnothing\}, \quad \{\varnothing, \{\varnothing\}\}, \quad \{\{\varnothing\}\}, \quad \text{etc.}$$

Notemos que los axiomas vistos hasta aquí **no** garantizan que, dados X, Y, Z, exista un conjunto $\{X, Y, Z\}$ que los tenga a ellos solos como elementos. Para ello necesitamos el *axioma de la unión:*

$$\forall X \exists Y \forall U (U \in Y \leftrightarrow \exists V (U \in V \wedge V \in X)).$$

Este axioma afirma que, dado un conjunto X, existe un conjunto Y cuyos elementos son los elementos de los elementos de X. De nuevo es único, pues dos conjuntos que cumplieran lo mismo tendrían los mismos elementos. Esto nos permite definir el término

$$\bigcup_{V \in X} V \equiv Y \mid \forall U (U \in Y \leftrightarrow \exists V (U \in V \wedge V \in X)).$$

Ahora podemos demostrar el teorema siguiente:

Teorema $\forall XY \exists Z \forall U (U \in Z \leftrightarrow U \in X \vee U \in Y)$.

DEMOSTRACIÓN: Hemos de probar que, dados dos conjuntos X e Y, existe un conjunto Z que contiene a los elementos de ambos. Basta tomar

$$Z = \bigcup_{V \in \{X, Y\}} V.$$

■

El conjunto Z dado por el teorema anterior es único por el axioma de extensionalidad, luego podemos definir el término

$$X \cup Y \equiv Z \mid \forall U (U \in Z \leftrightarrow U \in X \vee U \in Y).$$

Se comprueba inmediatamente que la unión es asociativa, por lo que podemos definir uniones $X_1 \cup \cdots \cup X_n$ sin necesidad de indicar los paréntesis. En particular podemos definir

$$\{X_1, \ldots, X_n\} \equiv \{X_1\} \cup \cdots \cup \{X_n\},$$

y es inmediato que el conjunto así definido tiene como elementos a $X_1, \ldots, X_n$.

Otro conjunto cuya existencia necesita ser postulada es el conjunto de partes, dado, naturalmente, por el *axioma de partes:*

$$\forall X \exists Y \forall U (U \in Y \leftrightarrow U \subset X).$$

Recordemos que la fórmula $U \subset X$ está definida en la sección anterior (esta definición no requiere ningún axioma). Nuevamente, el axioma de extensionalidad garantiza que el conjunto Y cuya existencia postula este axioma es único, por lo que podemos definir el término

$$\mathcal{P}X = Y \mid \forall U(U \in Y \leftrightarrow U \subset X).$$

La teoría de conjuntos de Zermelo tiene infinitos axiomas. Ello se debe a que el axioma siguiente no es un axioma, sino un *esquema axiomático,* es decir, una regla que determina infinitas fórmulas que la teoría acepta como axiomas (si bien en cada demostración en concreto sólo se podrá usar un número finito de casos particulares de dicha regla.) Se trata del *esquema axiomático de especificación:*

> Para cada fórmula $\phi(U)$ del lenguaje de la teoría de conjuntos (tal vez con más variables libres, además de X), la fórmula siguiente es un axioma:
>
> $$\forall X \exists Y \forall U(U \in Y \leftrightarrow U \in X \wedge \phi(U)).$$

Una vez más, X es único, luego podemos definir

$$\{Y \in X \mid \phi(U)\} \equiv Y \mid \forall U(U \in Y \leftrightarrow U \in X \wedge \phi(U)).$$

Observemos cómo la noción de "propiedad" queda completamente precisada al sustituirla por la de "fórmula". Por ejemplo, si aplicamos el esquema de especificación a un conjunto X y a la fórmula $\phi(U) \equiv U \in Y$ obtenemos la existencia del conjunto

$$X \cap Y \equiv \{U \in X \mid U \in Y\},$$

que tiene la propiedad de contener exactamente a los elementos comunes de X e Y.

Si aplicamos el esquema de especificación a un conjunto X y a $\phi(U) \equiv U \notin Y$ obtenemos el conjunto

$$X \setminus Y \equiv \{U \in X \mid U \notin Y\}.$$

En resumen, hasta aquí tenemos probada la existencia de

$$\varnothing, \quad X \cup Y, \quad X \cap Y, \quad X \setminus Y, \quad X \subset Y.$$

Supondremos conocidas las numerosas relaciones entre estos conceptos, pues todas se demuestran ahora fácilmente.

Definimos $(X, Y) \equiv \{\{X\}, \{X, Y\}\}$. Una simple rutina justifica el teorema siguiente:

Teorema $\forall XYZW((X, Y) = (Z, W) \leftrightarrow X = Z \wedge Y = W)$.

Vamos a demostrar que, dados dos conjuntos A y B, existe un conjunto C cuyos elementos son los pares ordenados con primera componente en A y segunda componente en B. Formalmente:

Teorema $\forall AB \exists C \forall U(U \in C \leftrightarrow \exists X \in A \exists Y \in B\ U = (X, Y))$.

DEMOSTRACIÓN: Notemos que no podemos aplicar el esquema de especificación con $\phi(U) \equiv \exists X \in A \exists Y \in B\ U = (X, Y)$ y considerar

$$C = \{U \mid \exists X \in A \exists Y \in B\ U = (X, Y)\},$$

pues necesitamos exigir que U pertenezca a un conjunto dado de antemano. El problema es encontrar un conjunto Z de modo que al exigir $U \in Z$ no nos dejemos fuera ningún par. Para ello observamos que si $X \in A$ e $Y \in B$ entonces $\{X\} \subset A \subset A \cup B$ y $\{X, Y\} \subset A \cup B$, luego $\{X\}, \{X, Y\} \in \mathcal{P}(A \cup B)$, luego

$$(X, Y) = \{\{X\}, \{X, Y\}\} \in \mathcal{P}\mathcal{P}(A \cup B).$$

Por consiguiente concluimos que el conjunto

$$C = \{U \in \mathcal{P}\mathcal{P}(A \cup B) \mid \exists X \in A \exists Y \in B\ U = (X, Y)\}$$

cumple el teorema. ■

El axioma de extensionalidad garantiza que el conjunto dado por el teorema anterior es único, luego podemos definir

$$A \times B \equiv C \mid \forall U(U \in C \leftrightarrow \exists X \in A \exists Y \in B\ U = (X, Y)).$$

A su vez esto nos permite definir, para cada fórmula $\phi(X, Y)$, el término

$$\{(X, Y) \in A \times B \mid \phi(X, Y)\} \equiv \{U \in A \times B \mid \exists X \in A \exists Y \in B(U = (X, Y) \wedge \phi(X, Y))\}.$$

A partir de aquí podemos definir y demostrar fácilmente todos los conceptos y teoremas relacionados con relaciones y funciones. Los supondremos conocidos.

El último axioma de la teoría de Zermelo postula la existencia de un conjunto infinito (sabemos que existen infinitos conjuntos, pero no conjuntos infinitos). Para introducirlo adecuadamente conviene definir $X' \equiv X \cup \{X\}$. El *axioma de infinitud* afirma que

$$\exists Y(\varnothing \in Y \wedge \forall X \in Y\ X' \in Y).$$

Ahora conviene introducir una notación alternativa: $0 \equiv \varnothing$, es decir, usaremos indistintamente 0 o $\varnothing$ como abreviaturas para el conjunto vacío. Así, un conjunto Y en las condiciones que postula el axioma de infinitud contiene necesariamente a 0, pero entonces contiene también a $1 \equiv 0' \equiv \varnothing \cup \{\varnothing\} = \{0\}$, pero entonces contiene también a $2 \equiv 1' \equiv 1 \cup \{1\} = \{0, 1\}$, y también a $3 \equiv 2' = \{0, 1, 2\}$, etc. En suma, Y es un conjunto infinito.

Podemos precisar aún más. Definamos

$$Y \text{ inductivo} \equiv \varnothing \in Y \wedge \forall X \in Y\ X' \in Y.$$

En estos términos, el axioma de infinitud afirma que existe un conjunto inductivo Y. Podemos construir entonces el conjunto

$$N = \{U \in Y \mid \forall Z \subset Y(Z \text{ inductivo } \rightarrow U \in Z)\},$$

es decir, N es el conjunto de los elementos de Y que pertenecen a todos los subconjuntos inductivos de Y. Vamos a probar que N es inductivo y que está contenido en todo conjunto inductivo.

Para probar que N es inductivo demostramos en primer lugar que $0 \in N$. En efecto, $0 \in Y$ y si Z es un subconjunto inductivo de Y entonces $0 \in Z$ por definición de conjunto inductivo. Esto prueba que $0 \in N$.

Supongamos ahora que $X \in N$, y hemos de probar que $X' \in N$. Por definición de N tenemos que $X \in Y$. Como Y es inductivo, $X' \in Y$. Ahora hemos de ver que si $Z \subset Y$ es inductivo, entonces $X' \in Z$. Ahora bien, por definición de N tenemos que $X \in Z$, y por definición de conjunto inductivo $X' \in Z$ como queríamos probar.

Ahora supongamos que I es un conjunto inductivo cualquiera (no necesariamente contenido en Y), y vamos a probar que $N \subset I$. En efecto, se demuestra inmediatamente que $Z = I \cap Y$ es un conjunto inductivo contenido en Y, luego por definición de N tenemos que $N \subset I \cap Y \subset I$. Con esto hemos demostrado el teorema siguiente:

Teorema $\exists N(N \text{ inductivo} \wedge \forall Y(Y \text{ inductivo} \rightarrow N \subset Y))$.

Este conjunto N es claramente único, pues si N y N' cumplen lo mismo, entonces $N \subset N'$ (porque N' es inductivo) y $N' \subset N$ (porque N lo es). Esto nos permite definir

$$\mathbb{N} \equiv N \mid (N \text{ inductivo} \wedge \forall Y(Y \text{ inductivo} \rightarrow N \subset Y)).$$

En otras palabras, el conjunto de los *números naturales* se define como el menor conjunto inductivo. A partir de aquí es fácil demostrar los axiomas de Peano:

Teorema (Axiomas de Peano)

1. $0 \in \mathbb{N}$,
2. $\forall n \in \mathbb{N}\ n' \in \mathbb{N}$,
3. $\neg\exists n \in \mathbb{N}\ 0 = n'$,
4. $\forall mn \in \mathbb{N}(m' = n' \rightarrow m = n)$,
5. $\forall X \subset \mathbb{N}((0 \in X \wedge \forall m \in X\ m' \in X) \rightarrow X = \mathbb{N})$.

No entraremos en los detalles de la definición de la suma y el producto de números naturales. A partir de aquí pueden definirse de la forma habitual todos los conceptos matemáticos (números enteros, racionales, reales, complejos, sucesiones, n-tuplas, etc.

4 Variantes sobre la teoría de Zermelo

La teoría de Zermelo presenta ciertas lagunas a la hora de obtener resultados más profundos en teoría de conjuntos, concretamente al trabajar con ordinales y cardinales infinitos. Fraenkel arregló el problema añadiendo el *esquema axiomático del reemplazo:*

> *Para toda fórmula $\phi(X,Y)$, tal vez con más variables libres, la fórmula siguiente es un axioma:*

$$\forall A(\forall X \in A \forall Y Z(\phi(X,Y) \land \phi(X,Z) \rightarrow Y = Z) \rightarrow \exists B \forall Y(Y \in B \leftrightarrow \exists X \in A \phi(X,Y)).$$

Este esquema afirma que si la fórmula $\phi(X,Y)$ define una función sobre un subconjunto de A, es decir, si a cada $X \in A$ le podemos asociar a lo sumo un conjunto Y que cumpla $\phi(X,Y)$, entonces existe un conjunto B que contiene a todos los conjuntos Y que podemos asignar de este modo a los elementos de A.

En otros términos, si llamamos $C = \{X \in A \mid \exists Y \phi(X,Y)\}$ y llamamos

$$F = \{(X,Y) \mid X \in A \land \phi(X,Y)\},$$

entonces F es una aplicación y el conjunto B cuya existencia postula el esquema de reemplazo es simplemente $B = F[C]$. El problema es que la existencia de F no puede probarse a partir de los axiomas de Zermelo. Una vez tenemos asegurada la existencia del conjunto B, entonces sí podemos definir

$$F = \{(X,Y) \in A \times B \mid \phi(X,Y)\},$$

pero sin tener a B la definición anterior no es correcta, pues no indicamos ningún conjunto del cual F sea un subconjunto, y no podemos aplicar el esquema de especificación.

A partir del esquema del reemplazo puede demostrarse el esquema de especificación, por lo que la teoría de conjuntos de Zermelo-Fraenkel (ZF) es la que resulta de sustituir la especificación por el reemplazo.

A los axiomas de ZF puede añadirse el *axioma de regularidad* de von Neumann:

$$\forall X \exists Y(Y \in X \land Y \cap X = \varnothing).$$

Este axioma no es necesario para demostrar ningún teorema importante, pero prohíbe la existencia de "monstruos" tales como un conjunto $X = \{X\}$, o un par de conjuntos $X = \{Y\}$, $Y = \{X\}$ o cosas peores. Garantiza, por el contrario, que todo conjunto puede construirse a partir del conjunto vacío en una cantidad finita o infinita de pasos (obsérvese, por ejemplo, cómo hemos construido el conjunto de los números naturales a partir del conjunto vacío).

Por último, podemos añadir el *axioma de elección:*

$$\forall X \exists F(F : X \longrightarrow \bigcup_{V \in X} V \land \forall U \in X(U \neq \varnothing \rightarrow F(U) \in U)).$$

Este axioma afirma que para toda familia X de conjuntos existe una "función de elección" F que asigna a cada uno de ellos U uno de sus elementos (siempre que $U \neq \varnothing$). Al contrario que el axioma de regularidad, este axioma es crucial para la demostración de muchos resultados importantes, como la existencia de bases en espacios vectoriales, la existencia de clausura algebraica, el teorema de Tychonoff sobre la compacidad de un producto de espacios topológicos, el teorema de Hann-Banach sobre extensión de funcionales lineales, etc.).

5 Clases

La axiomática de Zermelo (o de Zermelo-Fraenkel) evita las paradojas de la teoría de conjuntos negando la existencia de los conjuntos que las provocan. Por ejemplo:

Teorema $\neg\exists R \forall X (X \in R \leftrightarrow X \notin X)$.

La prueba de este teorema es trivial: si existiera un R en estas condiciones tendríamos la consabida contradicción. La diferencia con la teoría de Frege es que ésta postulaba la existencia de R, mientras que los axiomas de ZF no lo hacen.

Como consecuencia tenemos lo siguiente:

Teorema $\neg\exists V \forall X\ X \in V$.

Es decir, no existe ningún conjunto que contenga a todos los conjuntos. La demostración es la siguiente: Si existiera V podríamos definir

$$R = \{X \in V \mid X \notin X\},$$

por el esquema de especificación. Pero este conjunto contradice al teorema anterior. ■

Más en general, no podemos definir el complementario de un conjunto X, es decir, no existe un conjunto $\overline{X}$ cuyos elementos sean los conjuntos que no pertenecen a X. Si existiera, entonces el conjunto $V = X \cup \overline{X}$ contendría a todos los conjuntos, en contradicción con el teorema anterior.

Georg Cantor desarrolló una teoría de cardinales infinitos que extiende a la bien conocida teoría de cardinales finitos. Según esta teoría, la sucesión de los cardinales continúa más allá de los números naturales:

$$0, \quad 1, \quad 2, \quad \ldots \quad \aleph_0, \quad \aleph_1, \quad \aleph_2, \quad \ldots$$

de modo que a cada conjunto X se le puede asignar un único cardinal $|X|$ en esta sucesión con la propiedad de que $|X| = |Y|$ si y sólo si existe una aplicación $F : X \longrightarrow Y$ biyectiva. Otro ejemplo clásico de "conjunto que no existe" es el conjunto de todos los cardinales, es decir, si suponemos la existencia de tal conjunto llegamos a una contradicción.

De estos hechos se desprende una consecuencia filosófica muy interesante: si estamos dispuestos a admitir que la palabra "conjunto" significa algo, hemos de tener bien claro que no significa "colección de objetos". En efecto, admitamos que los axiomas de la teoría de conjuntos (y, con ellos, sus teoremas) hablan de unos objetos llamados conjuntos. Entonces, la colección de todos esos conjuntos es sin duda una colección de objetos, pero no se corresponde con ninguno de esos conjuntos. Así pues, podríamos admitir que existen unos objetos llamados conjuntos que satisfacen los axiomas de ZF, y que cada uno de ellos puede verse como una colección de objetos, a saber, aquellos otros conjuntos que le pertenecen según el (presunto) significado de la relación de pertenencia; pero lo que es inadmisible es postular el recíproco, es decir, que toda colección de conjuntos constituya la extensión de un conjunto, es decir, que haya un conjunto que los tenga a ellos por elementos y sólo a ellos. Por ejemplo, si en algún sentido existen los conjuntos, la colección de todos aquellos conjuntos que no cumplen la definición de número natural es ciertamente una colección de conjuntos, pero no hay ningún conjunto que los tenga a ellos y sólo a ellos por elementos (sería el complementario de $\mathbb{N}$, y ya hemos demostrado que no existe).

Sin embargo, el hecho de que no exista el conjunto de todos los conjuntos, o el conjunto de todos los cardinales, o el complementario del conjunto de los números naturales no es obstáculo para que podamos hablar de ellos. El único arte que ello requiere es que tengamos bien claro lo que hacemos. Concretamente, podemos hacer lo siguiente:

Para cada fórmula $\phi(X)$, convendremos en que la notación

$$C = \{X \mid \phi(X)\}$$

(que leeremos "C es la clase de todos los conjuntos que cumplen $\phi(X)$") significará simplemente que $X \in C$ es una forma alternativa de escribir $\phi(X)$.

De este modo, una clase es esencialmente lo mismo que una fórmula. Por ejemplo, podemos definir $V = \{X \mid X = X\}$, de modo que V es la clase de todos los conjuntos. Sabemos que V no existe, en el sentido de que no existe ningún conjunto que contenga a todos los conjuntos, pero podemos usar V en el sentido limitado en virtud del cual $X \in V$ es una forma cómoda de expresar que $X = X$, es decir, que X es un conjunto cualquiera. Igualmente, podemos definir $V \times V = \{X \mid \exists YZ\ X = (Y, Z)\}$, de modo que $V \times V$ es la clase de todos los pares ordenados. Es fácil probar que no existe ningún conjunto que contenga a todos los pares ordenados (ejercicio), pero no hay ningún problema en escribir $X \in V \times V$ como abreviatura de $\exists YZ\ X = (Y, Z)$.

En realidad podemos usar clases en un contexto más general que las meras fórmulas de tipo $X \in C$. Por ejemplo, pese a que ni existe V ni existe $V \times V$, podemos interpretar la afirmación $V \times V \subset V$ como una fórmula válida en ZF. Basta recordar la definición de la inclusión, según la cual $V \times V \subset V$ es equivalente a

$$\forall X(X \in V \times V \to X \in V).$$

Puesto que tanto $X \in V \times V$ como $X \in V$ tienen un significado preciso, lo mismo sucede con toda la fórmula. En definitiva, $V \times V \subset V$ significa que todo par ordenado es un conjunto, lo cual es cierto.

En definitiva, si en una fórmula aparecen clases, pero únicamente en la forma $X \in C$, la fórmula tiene sentido como afirmación en ZF, pese a que las clases en sí no existan, pues las clases que aparecen pueden sustituirse por las fórmulas que las definen.

Más aún, no necesitamos exigir que las clases aparezcan en la forma $X \in C$. Por ejemplo, imaginemos que en una fórmula aparece una igualdad de clases $C = E$, o una igualdad de la forma $X = C$, donde X es una variable. El axioma de extensionalidad nos permite darles una interpretación. Por ejemplo, convenimos en que $C = E$ es una abreviatura de la fórmula

$$\forall X(X \in C \leftrightarrow X \in E),$$

donde a su vez podemos sustituir C y E por las fórmulas que las definen.

Todo esto es consistente mientras recordemos que $\forall X$ y $\exists X$ hacen referencia a conjuntos. Por ejemplo, se cumple que $\neg\exists X\ X = V$, pues esta fórmula significa, según los convenios que hemos establecido, a que

$$\neg\exists X \forall Y(Y \in X \leftrightarrow Y \in V)$$

o, equivalentemente, que no existe ningún conjunto que contenga a todos los conjuntos. Sería un error pensar que, como $V = V$, entonces $\exists X\ X = V$. Esta última fórmula afirma que existe un conjunto igual a V, lo cual es falso.

Por último, también podemos dar sentido a cualquier fórmula en la que una clase propia aparezca en la forma $C \in X$. Convenimos en que esto significa

$$\exists Y(Y = C \wedge Y \in X).$$

Para que esto pueda suceder es necesario que la clase C se corresponda en realidad con un conjunto, es decir, que sí que exista un conjunto cuyos elementos coincidan con los de C. En resumen:

> *Cada fórmula de ZF nos permite definir una clase C, de tal modo que toda fórmula de ZF en la que aparezcan clases puede interpretarse como la abreviatura de una fórmula que no las contenga sin más que eliminar las subfórmulas $X \in C$, $X = C$, $C \in X$, etc. del modo que hemos indicado.*

A partir de aquí, las operaciones conjuntistas pueden aplicarse a las clases con total libertad. Por ejemplo, podemos definir la intersección de dos clases C y E como la clase

$$C \cap E = \{X \mid X \in C \wedge X \in E\}.$$

En otras palabras, estamos conviniendo en que $X \in C \cap E$ es simplemente una forma de abreviar $X \in C \wedge X \in E$, que a su vez tendrá un significado concreto en función de las fórmulas que definan a C y E.

Otro ejemplo, podemos definir $R = \{(X,Y) \mid X \subset Y\}$ y afirmar que la clase R es una relación de orden parcial en la clase universal V. Con un poco de paciencia, usando la definición de relación de orden, podríamos reescribir esta afirmación en términos de conjuntos únicamente.

Terminamos la sección observando que el hecho de que podamos hablar de clases como si fueran conjuntos no significa que su comportamiento sea el mismo que el de los conjuntos. Por ejemplo, podemos definir en la clase universal V la relación de equivalencia

$$X\,R\,Y \leftrightarrow |X| = |Y|.$$

(Con más precisión, $R = \{(X,Y) \mid |X| = |Y|\}$). Ahora podríamos pensar en formar la clase cociente, y esperar encontrarnos con tantos elementos como cardinales hay, pero no es así, la clase cociente resulta ser $V/R = \{\{\varnothing\}\}$. En efecto, la clase de equivalencia del conjunto vacío es $\{\varnothing\}$, ya que $\varnothing$ es el único conjunto de cardinal 0, pero puede probarse que, para cualquier otro cardinal $\kappa > 0$ (finito o infinito) no existe ningún conjunto que contenga a todos los conjuntos de cardinal κ, de modo que la clase de equivalencia $C = \{X \mid |X| = \kappa\}$ no es un conjunto, luego no puede pertenecer a la clase V/R ni a ninguna otra clase, ya que hemos convenido en que $C \in V/R$ ha de entenderse como

$$\exists X (X = C \land X \in V/R),$$

y esto es falso. Por lo tanto V/R sólo contiene la clase de equivalencia de $\varnothing$.

Fenómenos como éste hacen que las clases sean útiles como meros auxiliares para tratar con conjuntos, pero no podemos centrar en ellas la teoría.

6 Axiomatización de las clases

La discusión de la sección anterior puede parecer escurridiza y poco rigurosa, pero no es así. De hecho, partiendo de una axiomática exótica de von Neumann, Bernays diseñó otra en la que había dos tipos de objetos básicos: los conjuntos y las clases. Los conjuntos se comportaban igual que los de ZF y las clases se comportaban como hemos descrito en la sección anterior. La axiomática de Bernays era un tanto complicada, pero Gödel la simplificó drásticamente, hasta hacerla tan sencilla o más que la de ZF. Su versión se conoce como teoría de conjuntos de von Neumann-Bernays-Gödel (NBG). Aquí veremos una versión ligeramente distinta (más sencilla aún) llamada teoría de Morse-Kelley (MK). Luego comentaremos las diferencias respecto a NBG.

Si el planteamiento de ZF es "vamos a hablar de unos objetos llamados conjuntos y una relación de pertenencia que no definimos, pero que cumplen los axiomas siguientes", el planteamiento de MK es "vamos a hablar de unos objetos llamados clases y una relación de pertenencia que no definimos, pero que cumplen los axiomas siguientes". En otras palabras, ahora $\forall X$ y $\exists X$ se leen "para toda clase X" y "existe una clase X".

Antes de dar ningún axioma, damos una definición: X es un conjunto $\equiv \exists Y\ X \in Y$. Es decir, llamaremos conjuntos a las clases que pertenecen a alguna clase. Adoptaremos también el convenio de usar letras minúsculas para referirnos a conjuntos, de modo que

$$\forall x \phi(x) \equiv \forall X(X \text{ es un conjunto} \rightarrow \phi(X)),$$

$$\exists x \phi(x) \equiv \exists X(X \text{ es un conjunto} \wedge \phi(X)).$$

La inclusión de clases se define como

$$X \subset Y \equiv \forall u(u \in X \rightarrow u \in Y).$$

El primer axioma es el de *extensionalidad:*

$$\forall XY(\forall u(u \in X \leftrightarrow u \in Y) \rightarrow X = Y).$$

Este axioma afirma que si dos clases tienen los mismos conjuntos como elementos entonces son iguales.

La diferencia fundamental de MK (o NBG) frente a ZF es que nos permite dar un *esquema de formación de clases* mucho más parecido al original de Frege, y, desde luego, mucho menos restrictivo que el esquema de especificación:

Para toda fórmula $\phi(X)$, la fórmula siguiente es un axioma de MK:

$$\exists Y \forall x(x \in Y \leftrightarrow \phi(x)).$$

En otras palabras, para toda propiedad, existe una clase cuyos elementos son los conjuntos que cumplen la propiedad. Esta clase Y es única, pues dos clases que cumplan esto tienen los mismos elementos. Por lo tanto podemos definir el término

$$\{x \mid \phi(x)\} \equiv Y \mid \forall x(x \in Y \leftrightarrow \phi(x)).$$

Es crucial observar que $\{x \mid \phi(x)\}$ es la *clase* de todos los *conjuntos* que cumplen $\phi(x)$, es decir, para que un X esté en esta clase *no basta* con que cumpla $\phi(X)$, sino que además ha de ser un conjunto.

A partir de aquí podemos definir:

1. $\varnothing \equiv \{x \mid x \neq x\}$ (la clase vacía),

2. $V \equiv \{x \mid x = x\}$ (la clase universal),

3. $X \cap Y \equiv \{u \mid u \in X \wedge u \in Y\}$ (intersección de clases),

4. $X \cup Y \equiv \{u \mid u \in X \vee u \in Y\}$ (unión de clases),

5. $\overline{X} \equiv \{u \mid u \notin X\}$ (clase complementaria),

6. $X \setminus Y \equiv \{u \mid u \in X \wedge u \notin Y\}$ (diferencia de clases),

7. $\mathcal{P}X \equiv \{u \mid u \subset X\}$ (clase de partes),

8. $\{X_1, \ldots, X_n\} \equiv \{u \mid u = X_1 \vee \cdots \vee u = X_n\}$.

Es fácil probar que estos conceptos tienen las propiedades habituales, salvo los problemas que pueden ocasionarse porque una clase dada no sea un conjunto. Para entender esto consideremos la clase

$$R = \{x \mid x \notin x\}.$$

Esta clase no es un conjunto, pues si R fuera un conjunto y $R \in R$, entonces debería ser $R \notin R$, y viceversa. Puesto que no es un conjunto, se cumple que $R \notin R$, y esto no obliga a que $R \in R$, pues para que algo pertenezca a R no basta con que cumpla $x \notin x$, sino que además ha de ser un conjunto.

Esto hace que $\{R\} = \varnothing$, pues para pertenecer a $\{R\}$ se han de cumplir dos condiciones: ser igual a R y ser un conjunto, y no pueden darse las dos a la vez.

Otro ejemplo: obviamente $R \subset V$, pero $R \notin \mathcal{P}V$, ya que para pertenecer a $\mathcal{P}V$ no basta estar contenido en V, sino que hace falta ser un conjunto.

Ahora necesitamos axiomas que garanticen que existen conjuntos. Con lo que sabemos hasta ahora podría ocurrir que $V = \varnothing$. Para el axioma de la unión definimos la clase

$$\bigcup_{v \in X} v \equiv \{u \mid \exists v \in X \; u \in v\}.$$

1. $\varnothing$ es un conjunto. (Axioma del conjunto vacío.)
2. $\forall xy \; \{x, y\}$ es un conjunto. (Axioma del par.)
3. $\forall x \bigcup\limits_{v \in x} v$ es un conjunto. (Axioma de la unión.)
4. $\forall x \; \mathcal{P}x$ es un conjunto. (Axioma de partes.)

El axioma del par afirma que el par $\{x, y\}$ formado por dos conjuntos es de nuevo un conjunto, lo que permite definir el par ordenado $(x, y) = \{\{x\}, \{x, y\}\}$ y demostrar el teorema $\forall xyzw((x, y) = (z, w) \leftrightarrow x = z \wedge y = w)$. Esto no vale para pares ordenados con componentes que no sean conjuntos. Por ejemplo,

$$(R, R) = \{\{R\}, \{R, R\}\} = \{\varnothing, \varnothing, \} = \{\varnothing\}$$

y, en general, siempre que X e Y no sean conjuntos se cumple que $(X, Y) = \{\varnothing\}$.

Definimos el producto cartesiano de clases como

$$X \times Y \equiv \{u \mid \exists vw(v \in X \wedge w \in Y \wedge u = (v, w))\}.$$

A su vez definimos

$$F : A \longrightarrow B \equiv F \subset A \times B \wedge \forall x \in A \exists! y \in B (x, y) \in F.$$

Así, tiene sentido hablar de que una clase F sea una aplicación entre dos clases. Podemos definir los conceptos usuales de dominio, rango, imagen, aplicación inyectiva, suprayectiva, etc., todos ellos para clases arbitrarias. Esto nos permite enunciar el *axioma del reemplazo* (que aquí es un axioma, no un esquema axiomático) de forma especialmente simple:

$$\forall FXY(F : X \longrightarrow Y \text{ suprayectiva} \land X \text{ es un conjunto} \to Y \text{ es un conjunto.})$$

De aquí deducimos el último teorema importante de formación de conjuntos:

Teorema Toda subclase de un conjunto es un conjunto.

DEMOSTRACIÓN: Sea x un conjunto y supongamos que una clase Y cumple $Y \subset x$. Distinguimos dos casos: si $Y = \varnothing$, entonces Y es un conjunto por el axioma del conjunto vacío. Si $Y \neq \varnothing$ tomamos un $u \in Y$ y definimos la aplicación $F : X \longrightarrow Y$ dada por

$$F(t) = \begin{cases} t & \text{si } t \in Y, \\ u & \text{si } t \notin Y. \end{cases}$$

Claramente F es suprayectiva, luego por el axioma del reemplazo Y es un conjunto. ∎

En particular, si x e y son conjuntos también lo es $x \cap y$, pues $x \cap y \subset x$. Para la unión observamos que $\{x, y\}$ es un conjunto por el axioma del par y entonces

$$x \cup y = \bigcup_{v \in \{x,y\}} v$$

es un conjunto por el axioma de la unión.

Teorema $\forall xy\ x \times y$ es un conjunto.

DEMOSTRACIÓN: Basta observar que $x \times y \subset \mathcal{PP}(x \cup y)$ (ver la introducción del producto cartesiano en la teoría de Zermelo). Hemos visto que $x \cup y$ es un conjunto. Por el axioma de partes $\mathcal{PP}(x \cup y)$ es un conjunto y por el teorema anterior $x \times y$ también lo es. ∎

A partir de aquí ya es fácil probar que cualquier cosa que debería ser un conjunto lo es realmente. Por el contrario, la clase universal V no es un conjunto, ya que $R \subset V$ y R no es un conjunto. Si x es un conjunto, la clase complementaria $\overline{x}$ no es un conjunto, pues si lo fuera $V = x \cup \overline{x}$ también sería un conjunto.

Podemos definir

$$\mathbb{N} \equiv \{x \mid \forall Y(Y \text{ es una clase inductiva} \to x \in Y)\}.$$

Es fácil probar que $\mathbb{N}$ es una clase inductiva contenida en cualquier otra clase inductiva, así como que cumple los axiomas de Peano. No obstante, no podemos demostrar que $\mathbb{N}$ sea un conjunto. Esto es precisamente lo que afirma el *axioma de infinitud.* Con esto completamos la axiomática de Morse-Kelley (en realidad faltarían los axiomas de regularidad y elección, pero éstos no presentan diferencias frente a ZF).

La única diferencia de MK respecto a la teoría NBG de von Neumann-Bernays-Gödel es que ésta última restringe el esquema de formación de clases a fórmulas $\phi(X)$ en las que todos los cuantificadores $\forall X$, $\exists X$ estén restringidos a conjuntos, es decir, sean de hecho de la forma $\forall x$, $\exists x$. Esto hace que NBG sea equivalente a MK en el sentido siguiente:

> *Todo teorema de ZF puede probarse en NBG y, recíprocamente, todo teorema de MK en cuyo enunciado sólo aparezcan conjuntos puede demostrarse en ZF (aunque la demostración haga referencia a clases que no sean conjuntos).*

Por el contrario, en MK pueden demostrarse teoremas sobre conjuntos que no pueden demostrarse ni en NBG ni en ZF. Otra característica de NBG que no tiene MK es que es finitamente axiomatizable, es decir, el esquema de formación de clases (restringido en la forma indicada) puede sustituirse por un número finito de casos particulares, a partir de los cuales pueden demostrarse todos los demás (esto lo probó Gödel).

7 Observaciones finales

Las teorías axiomáticas que hemos visto son intentos de evitar las paradojas de la teoría de conjuntos, pero no existe ninguna garantía de que lo consigan. Los teoremas de incompletitud de Gödel tienen como consecuencia que si cualquiera de estas teorías es consistente no es posible dar un argumento que lo pruebe.

Debería estar claro que los axiomas de la teoría de conjuntos no son de ningún modo "verdades básicas" de la matemática. Los matemáticos aceptan en sus razonamientos decenas de hechos igualmente básicos. Los axiomas de la teoría de conjuntos son simplemente una selección arbitraria de unos pocos de estos hechos básicos que bastan para demostrar todos los demás. Por ejemplo, la existencia de la unión es tan básica como la existencia de la intersección, pero en ZF la primera es un axioma y la segunda un teorema. En cambio, en NBG ambas son consecuencias de un mismo axioma general, pero luego el hecho de que la unión de conjuntos es un conjunto es un axioma y el hecho análogo para la intersección es consecuencia del axioma del reemplazo. Todo esto son convenios útiles que en absoluto pueden interpretarse como que la unión sea más o menos básica que la intersección.