Universidad de Oviedo

TEORÍA AXIOMÁTICA DE CONJUNTOS Y LÓGICA MATEMÁTICA

Malena Domínguez Sirgo

Dirigido por Consuelo Martínez López

UNIVERSIDAD DE OVIEDO

Facultad de Ciencias

Grado en Matemáticas

19/07/2024

Índice

Resumen

El objetivo de este trabajo es presentar la Teoría Axiomática de Zermelo-Fraenkel y estudiar su consistencia. Para ello, se comienza exponiendo la Lógica de Primer Orden, para comprender el lenguaje en el que se expresan los axiomas. Además, se realiza una introducción a la Teoría de Modelos, que nos permite entender el significado de los Teoremas de Completitud e Incompletitud de Gödel, los cuales serán la herramienta para analizar la consistencia.

A continuación, se presentan los axiomas de Zermelo-Fraenkel y se incluye una discusión sobre el Axioma de Elección. También se añaden conceptos que surgen de forma paralela al desarrollo de los axiomas, siendo los números naturales los que adquieren mayor relevancia.

Por último, se desarrolla el concepto de número ordinal y su aritmética. Esta última nos sitúa de nuevo en el contexto de los Teoremas de Incompletitud, por lo que el trabajo se concluye estudiando su aplicación a la Teoría Axiomática de Zermelo-Fraenkel.

Introducción

Este trabajo surge del interés por profundizar en la Teoría Axiomática de Conjuntos, tras conocer la existencia de las paradojas que surgen de la teoría convencional y que sacan a relucir la necesidad de una axiomatización.

A lo largo del Grado en Matemáticas, se han explorado numerosos conceptos, resultados y temáticas, alcanzando altos niveles de complejidad. Sin embargo, con este trabajo se ha pretendido cerrar el ciclo y volver a los orígenes de todo lo estudiado, pues, aunque su complejidad no es menor que la de aquello visto, su naturaleza pasa habitualmente más desapercibida. Aunque es lógico que en la práctica no se tenga siempre en mente, la Teoría de Conjuntos termina siendo aquello que se encuentra debajo de todo lo que podemos construir en matemáticas. Es por ello que el objetivo fundamental de este trabajo es presentar y desarrollar un fundamento sólido sobre el que poder cimentar las matemáticas.

Para llevarlo a cabo, se han consultado diversas fuentes de referencia, que aparecen especificadas en la introducción de cada capítulo. Se ha tratado de compaginar los diferentes puntos de vista presentes en los manuales, creando una notación común y unificando varias perspectivas en un único desarrollo propio. Las definiciones, aunque basadas siempre en alguna referencia, han sido en alguna ocasión adaptadas para integrarse en el hilo argumental del trabajo. En cuanto a las proposiciones, teoremas y corolarios, estos han sido obtenidos de los diferentes textos, en los que aparecían bien como un enunciado o bien como una observación que ha sido formalizada. En cuanto a las demostraciones, estas se han desarrollado de manera propia, ampliando generalmente las indicaciones dadas en los textos, aunque en algunas ocasiones se han desarrollado íntegramente de forma personal. Algunos teoremas se enuncian sin demostración, debido a que esta excede el contenido del trabajo. En ese caso, se ha citado la fuente de la que procede el enunciado, en la que se puede encontrar dicha prueba.

En cuanto al contenido del trabajo, se considera que el primer paso para conseguir una base firme es precisamente entender por qué es necesaria, y por qué no se pueden asumir como intuitivos ciertos conceptos como el de conjunto. Esto, junto con una contextualización histórica, se ha llevado a cabo en el Capítulo 0.

En el Capítulo 1 se explora la herramienta esencial para obtener nuestra base: la Lógica Matemática. Se describen los lenguajes de primer orden y los sistemas deductivos, que constituyen las piezas principales de cualquier construcción. Además, mediante la Teoría de Modelos, se dota de un significado a la lógica anterior, acercándola a nuestro propio razonamiento intuitivo. Finalmente, se presentan los Teoremas de Completitud e Incompletitud de Gödel, que permiten conocer las fortalezas y limitaciones de aquellas teorías derivadas de la lógica.

En el Capítulo 2 se desarrolla el fundamento buscado: la Teoría Axiomática de Conjuntos. Se consideran los axiomas de Zermelo-Fraenkel, aunque se mencionan otras alternativas. Se comienza explorando los primeros axiomas presentados por Zermelo y los conceptos más importantes que se derivan de ellos, con especial interés en los números naturales. Estos serán construidos desde una perspectiva conjuntista y serán a su vez el germen de conceptos de gran utilidad. A continuación, se añaden los axiomas que dan lugar a la Teoría Axiomática de Zermelo-Fraenkel. También se presenta el Axioma de Elección pues, aunque no sea imprescindible, constituye el cimiento de una gran cantidad de resultados.

Finalmente, el objetivo será evaluar la teoría que se ha presentado. Para ello, es necesaria la construcción de los números ordinales y de su aritmética, lo que constituye el contenido del Capítulo 3. Se desarrollan estos conceptos y se presentan de forma paralela nuevas herramientas que se derivan de ellos: el Teorema de Recursión Transfinita y el Principio de Inducción Transfinita. Visto esto, se termina valorando los axiomas de Zermelo-Fraenkel y conociendo sus características y alcance, que serán por extensión los de las matemáticas.

El trabajo finaliza con un cápitulo de conclusiones que recoge los resultados más importantes y sus implicaciones.

0 | Contexto y motivación.

La Teoría de Conjuntos Naive

0.1. Contexto histórico

Aunque la noción de conjunto puede parecer algo intrínseco dentro de las matemáticas, o quizás un concepto demasiado intuitivo para necesitar una definición formal, no fue hasta la segunda mitad del siglo XIX cuando, de la mano de Georg Cantor, empezó a desarollarse una teoría que explorase esta idea: la Teoría de Conjuntos. En ella se consideraba un conjunto simplemente como *"una colección de elementos"*, siendo estos elementos objetos de cualquier tipo. El principal atractivo del estudio del concepto de conjunto reside en que, tras muchos intentos por unificar las matemáticas bajo una misma rama, la teoría de conjuntos se postula a día de hoy como la mejor candidata. Como bien explica Patrick Suppes en [Suppes, 1960]:

> *"Entre las muchas ramas de las matemáticas modernas, la teoría de conjuntos ocupa un lugar único: con algunas raras excepciones, las entidades que se estudian y analizan en matemáticas pueden ser consideradas como ciertos conjuntos particulares o clases de objetos. Como consecuencia, muchas preguntas fundamentales sobre la naturaleza de las matemáticas pueden reducirse a preguntas sobre teoría de conjuntos.".*

Sin embargo, es curioso saber que el inicio de la Teoría de Conjuntos se debe a los trabajos de Cantor en el campo de las series trigonométricas, alrededor de 1870. Al estudiar la convergencia de ciertas series, descubrió propiedades de algunos conjuntos de números reales, que pronto entendió que podían ser generalizadas.

Su teoría fue percibida como revolucionaria en su época por su uso del concepto de infinito, que Cantor consideraba de manera natural, pero cuya existencia parecía plantear problemas filosóficos a sus contemporáneos. Sin embargo, para la década de 1890 gran parte de su trabajo ya estaba asentado, y antes del cambio de siglo hasta las partes más novedosas eran aceptadas por grandes personalidades, pues eran de gran utilidad en ramas como el análisis. Además, trabajos de matemáticos como Dedekind, que se desarrollaban a la par en el tiempo, ayudaron a formalizar esta teoría y a encaminarla hacia su labor de unificar las matemáticas.

Basándose en esta noción intuitiva de conjunto, comúnmente conocida en la literatura como **Teoría de Conjuntos Naive**, se ha desarrollado una teoría ampliamente extensa, que abarca relaciones, conjuntos bien ordenados o funciones entre otros aspectos. Sin embargo, entre 1895 y 1910, cuando parecía que sería finalmente reconocida y aceptada, aparecieron ciertas paradojas que pusieron en duda todo lo desarrollado hasta el momento.

0.2. Las paradojas

La primera paradoja encontrada fue publicada en 1897 por Burali-Forti, aunque había sido realmente descubierta por el propio Cantor dos años antes. Se esperaba que, haciendo pequeñas modificaciones en los conceptos teóricos, esta paradoja pudiera sortearse. Sin embargo, en 1902, Bertrand Russel encuentra una paradoja que empieza a dejar clara la necesidad de desarollar una teoría cuya definición de conjunto fuese más restrictiva y formal.

En general, las paradojas se dividen en **paradojas lógicas** y **paradojas semánticas**. Las primeras son aquellas que surgen de la propia lógica que fundamenta la construcción matemática, mientras que las segundas provienen del uso del lenguaje con el cual hablamos de los objetos matemáticos.

La más conocida dentro de las lógicas es la antes mecionada **Paradoja de Bertrand Rusell**, que este publica en su conocido Principia Mathematica [Rusell y Witehead, 1927]. Es necesario tener en cuenta que en la Teoría de Conjuntos Naive los elementos de un conjunto pueden ser objetos cualesquiera. En particular, también conjuntos. Podemos entender la Paradoja de Rusell de siguiente forma:

> Sea A el conjunto formado por todos aquellos conjuntos que no se pertenecen a sí mismos. Es decir, el conjunto dado por todos aquellos objetos que cumplen la propiedad "no pertenecerse a sí mismo". Entonces, si A es un elemento de dicho conjunto, cumple la propiedad que lo caracteriza: no pertenecerse a sí mismo, por lo que a la vez A pertenece a A y A no pertenece a A. En otro caso, si A no es un elemento del conjunto dado, tenemos que A no se pertenece a sí mismo, por lo que cumple la propiedad que caracteriza a A y de nuevo A no se pertenece a sí mismo y se pertenece simultáneamente.

Realmente, considerando el lenguaje matemático básico que se utilizó en este teoría primitiva, podemos observar la parte técnica de la contradicción de manera más sencilla. Consideramos el símbolo $\in$, que denominamos **pertenece**, de tal forma que tendremos $x \in y$ si el objeto x es un elemento del conjunto y. De la misma forma, escribimos $x \notin y$ cuando x no es un elemento de y. Además, si ϕ es una propiedad que pueden cumplir los objetos que se estén considerando, podemos definir el conjunto de todos aquellos elementos que verifican esta propiedad: $\{x \,/\, \phi\,(x)\}$. Así, lo antes expuesto se traduce en

> Sea $A = \{x \,/\, x \notin x\}$. Entonces, si $A \in A$, tenemos que A cumple por definición del conjunto que $A \notin A$, de donde $A \in A$ y $A \notin A$ simultáneamente. Por otro lado lado, si $A \notin A$, tenemos que es un elemento cumpliendo la propiedad que define al conjunto, por lo que $A \in A$, de donde obtenemos de nuevo una contradicción.

Dentro de las paradojas semánticas nos encontramos, entre otras, con la **Paradoja de Berry**, propuesta también por Rusell en 1906. Este a su vez se la atribuye a G.G.Berry, el bibliotecario jefe de la biblioteca Bodleiana de la Universidad de Oxford [Rusell y Witehead, 1927]. Dice lo siguiente:

> Teniendo en cuenta que nuestro alfabeto tiene un número finito de símbolos, si nos restrigimos a oraciones con un número máximo de palabras, podemos únicamente obtener un número finito de ellas. Así, sea S el conjunto de los números naturales que se pueden describir con menos de 20 palabras. Por el razonamiento realizado anteriormente, este es un conjunto con un número finito de elementos. En él estan por ejemplo "el primer número natural" o "el tercer natural impar". Como existe un número infinito de naturales, podemos tomar el primer número natural que no pertenezca a este conjunto: "el menor número natural que no puede ser descrito con menos de 20 palabras". De forma que este enunciado, que consta de menos de 20 palabras, determina de forma única a este natural, haciendo que este pertenezca al conjunto S.

Normalmente, cuando un razonamiento nos lleva a una contradicción de este estilo, concluimos que alguna de las premisas es incorrecta. En este caso, llegamos a la conclusión de que no siempre podremos considerar el conjunto de todos los objetos que cumplan cierta propiedad. La forma de solventar este asunto es lo que motivará las nuevas teorías que presentaremos más adelante.

0.3. La necesidad de axiomatización

De acuerdo con la aparición de estas paradojas, surge la necesidad de reevaluar la definición de conjunto dada, e incluso se empiezan a plantear por primera vez cuestiones relativas a su existencia y a la naturaleza de las propiedades que sean válidas para definirlos. Los conjuntos ya no son algo intuitivo, si no que es necesario desarollar una teoría axiomática que conserve la mayor parte de los resultados obtenidos hasta el momento, pero que evite las contradicciones.

El primero en proponer una teoría axiomática que consiguiese el objetivo propuesto fue Zermelo, en 1908. Como concluimos antes, las nuevas teorías axiomáticas deberían, de alguna forma, evitar la construcción de conjuntos de la forma "todos los elementos que cumplen cierta propiedad". Con esto en mente, Zermelo permite únicamente seleccionar objetos que cumplen una propiedad dentro de un conjunto ya existente. Es decir, si se tiene un conjunto X, se puede formar el conjunto de los elementos de X que cumple la propiedad ϕ, pero no podemos considerar el conjunto de todos los elementos verificando ϕ. De esta forma, evitamos las paradojas lógicas. Por ejemplo, ya no podemos tener en cuenta el conjunto A presentado en la paradoja de Bertrand Rusell.

Por otro lado, para evitar las paradojas semánticas, se recurre a la lógica. La propiedad ϕ no puede ser cualquiera, si no que tiene que ser correcta en el lenguaje lógico utilizado para desarrollar la teoría. Estas modificaciones son añadidas por Skolem y Fraenkel, en 1922. Con todo esto, surge la **Teoría Axiomática de Zermelo-Fraenkel** (ZF), una de las más aceptadas hoy en día y la que desarrollaremos en este trabajo.

Otro de los sistemas axiomáticos más comúnmente aceptado es la **Teoría de von Neumann-Bernays-Göedel** (NBG). En ella, todos los objetos se denominan **clases**, pero se realiza una distinción entre **elementos** y **clases propias**. Los primeros son aquellas clases que pertenecen a otras clases, mientras que los segundos son aquellas clases que no son elementos de ninguna clase. Así, podemos formar conjuntos de la forma "todas las clases que son elementos y verifican la propiedad ϕ". De esta forma, podemos considerar el conjunto $A = \{x \,/\, x \text{ es un elemento y } x \notin x\}$, pero el razonamiento de la paradoja de Rusell nos lleva a concluir que no es posible que $A \in A$ ni $A \notin A$, de donde simplemente se llega a que A es una clase propia, no a una contradicción.

Aunque no desarrollemos esta última teoría en este trabajo, ZF y NBG están íntimamente relacionadas y es sencillo entender una estando familiarizado con la otra.

1 | Introducción a la Lógica

Durante el siglo XIX, mientras Cantor exploraba la teoría de conjuntos, otros destacados matemáticos como Boole, de Morgan, Frege y Peano avanzaban en el estudio de la Lógica Matemática. Esta disciplina, mediante herramientas matemáticas, trata de estudiar la lógica ordinaria, es decir, el proceso que utilizamos los seres humanos para razonar y obtener deducciones.

Antes de adentrarnos en los Axiomas de Zermelo-Fraenkel y la teoría que se deriva de ellos, necesitamos entender cómo formalizar ideas matemáticas y qué herramientas son necesarias para ello. La respuesta a estas cuestiones vendrá de la mano de los lenguajes de primer orden. Así, se verá, en general, lo que constituye un lenguaje de primer orden. Además, clarificaremos conceptos cruciales como los axiomas y las teorías axiomáticas, demostrando cómo estas estructuras dan lugar a teorías matemáticas bien fundadas, como la que nos ocupa en este trabajo. Una vez establecidas estas bases, introduciremos los conceptos de consistencia y completitud, que nos permitirán determinar si una teoría está libre de contradicciones. Por último, terminaremos este capítulo hablando de la interpretación de los símbolos de un lenguaje y presentando los Teoremas de Completitud e Incompletitud de Gödel, que son de suma importancia a la hora de hablar de la consistencia de una teoría.

Para este capítulo se han seguido [Halbeisen y Krapf, 2020] y [San Román, 1990] como referencias. En las secciones 1.1 y 1.2 se presenta una visión propia que unifica los enfoques de ambos textos. En las secciones 1.3 y 1.4 se ha desarrollado el contenido de [Halbeisen y Krapf, 2020], aunque las demostraciones son propias.

1.1. Lenguajes de Primer Orden

La Lógica de Primer Orden es una de las diversas formas de modelar, es decir, de representar matemáticamente, la lógica ordinaria. En este enfoque se considera un **lenguaje**, también denominado **lenguaje formal**, compuesto por una colección de símbolos denominados **alfabeto**, junto con una serie de sucesiones de estos símbolos. La naturaleza de estas sucesiones se irá desarollando a lo largo de esta sección. Un lenguaje propio de la Lógica de Primer Orden se denomina **Lenguaje de Primer Orden**.

Además, lo que caracteriza a la Lógica de Primer Orden frente a otros modelos de la lógica ordinaria, como puede ser la Lógica Proposicional, es su capacidad para reflejar la estructura interna de una oración del lenguaje ordinario cuando esta se traduce al lenguaje formal, diferenciando entre el sujeto y el predicado.

Cada tipo de modelo de la lógica ordinaria tiene un lenguaje compuesto por un alfabeto distinto. Veamos cuáles son los símbolos que pueden encontrarse dentro del alfabeto de un Lenguaje de Primer Orden:

Definición 1.1. Un **alfabeto** de un Lenguaje de Primer Orden es una colección de símbolos formada por los siguientes elementos:

1. **Variables**, que denotaremos por letras como x, y, x_0, $x_1, \ldots$. Estas se utlizarán para referirse de forma genérica a los objetos que se estén considerando.

2. **Operadores lógicos**, que son $\neg$ (*no*), $\wedge$ (*y*), $\vee$ (*o*) y $\rightarrow$(*implica*).

3. **Cuantificadores lógicos**, que son el **cuantificador universal**, representado por $\forall$ y que se lee *para todo* o *para cada*, y el **cuantificador existencial**, denotado por $\exists$, y que se lee *existe*.

4. **Paréntesis**: (,).

5. **Símbolo de igualdad**: =.

6. **Constantes**, que representan objetos fijos dentro de los considerados. Suelen denotarse con las mismas letras que las variables.

7. **Símbolos de funciones**, que representan funciones fijas que toman objetos como argumentos y devuelven objetos como valores. Para cada función, definimos su **aridad** como el número de objetos que toma como argumento. Si este número es 2 decimos que tenemos una **función binaria**, si es 1, una **unaria** o **1-aria**, etcétera.

   Las funciones pueden tomar únicamente un número finito de argumentos. Para indicar que una función tiene n argumentos escribiremos una expresión de la forma $F(x_1, ..., x_n)$, donde dentro de los paréntesis encontramos variables. Si por el contrario estamos aplicando la función a objetos concretos, sustituiremos las variables por el símbolo constante que identifique a dicho elemento.

8. **Símbolos de relaciones**, también llamados **símbolos de predicado**, que representan las propiedades de los objetos. Análogamente a lo hecho con funciones, a cada relación se le asigna

su **aridad**, teniendo así **relaciones unarias, binarias**, etcétera. De la misma forma, las relaciones pueden tomar únicamente un número finito de argumentos. Si queremos expresar que cierta relación toma n argumentos escribiremos una expresión de la forma $R(x_1, ..., x_n)$, mientras que para referirnos a que ciertos objetos cumplen la relación sustituiremos las variables por las constantes correspondientes.

Los símbolos 1-5 de la definición de alfabeto forman los **símbolos lógicos**. Estos son iguales en todos los lenguajes y deben pertenecer siempre al alfabeto. Los símbolos 6-8 varían dependiendo del lenguaje y forman su **signatura**.

Nota 1. Cabe destacar que el símbolo de igualdad mencionado anteriormente es un caso particular de relación binaria, que al tener un interés especial, pues siempre ha de estar presente en un alfabeto, suele definirse por separado. Notemos también que una constante puede verse como una función con aridad 0.

Una vez que conocemos el alfabeto de un lenguaje, podemos comenzar a formar, a partir de él, lo que conoceremos como términos.

Definición 1.2. Una sucesión de símbolos del alfabeto se dice **término** si resulta de aplicar un número finito de veces las siguientes reglas:

(T0) Toda variable es un término.

(T1) Toda constante es un término.

(T2) Si $\tau_1, \ldots, \tau_n$ son términos ya construidos y F es una función con aridad n, entonces $F(\tau_1, \ldots, \tau_n)$ es también un término.

Los términos de la forma dada en (T0) y (T1) se denominan **términos atómicos**, pues a partir de ellos se construyen el resto de términos.

Nota 2. Los términos no forman parte del alfabeto. Sin embargo, sí se incluyen dentro del lenguaje. En general, para denotarlos, así como a cualquier objeto del lenguaje formal, utilizaremos letras griegas. Estas podrán referirse a un objeto en concreto o podrán ser usadas como variables, en la misma forma en la que funcionan las variables y constantes en el alfabeto.

Hasta ahora, los elementos del lenguaje que no pertenecen al alfabeto provienen de funciones de las variables y las constantes. Es decir, todavía no somos capaces de representar una relación entre dos objetos. Esto es lo que conseguimos construyendo lo que denominaremos fórmulas.

Definición 1.3. Una sucesión de símbolos se denomina **fórmula** si resulta de aplicar un número finito de veces las siguientes reglas:

(F0) Si τ_1 y τ_2 son términos, entonces $\tau_1 = \tau_2$ es una fórmula.

(F1) Si $\tau_1, \ldots, \tau_n$ son términos y R es una relación con aridad n, entonces $R(\tau_1, \ldots, \tau_n)$ es una fórmula.

(F2) Si φ y ψ son fórmulas que ya han sido construidas, entonces $\neg\varphi$, $\varphi \wedge \psi$, $\varphi \vee \psi$ y $\varphi \to \psi$ son fórmulas.

(F3) Si φ es una fórmula y x una variable arbitraria, entonces $\exists x\varphi$ y $\forall x\varphi$ son fórmulas.

Las fórmulas en la forma de (F0) y (F1) se denominan **fórmulas atómicas**. Las fórmulas también pertenecen al lenguaje.

Nota 3. Normalmente, para facilitar la lectura, se añaden paréntesis cuando se aplica la regla (F2), escribiendo por ejemplo $\varphi \to (\psi \to \varphi)$. En cuanto a la regla (F1), cuando la relación es binaria, se suele escribir xRy en vez de $R(x, y)$.

Debido a que las definiciones de término y fórmula se han realizado de manera recursiva, nos resulta mucho más sencillo demostrar que algo es cierto para todo término o función.
Si queremos probar que todos los términos cumplen una propiedad P, tenemos únicamente que probar que:

1. todos los términos atómicos verifican P.

2. si los términos $\tau_1, \ldots, \tau_n$ satisfacen P, entonces $F(\tau_1, \ldots, \tau_n)$ satisface P para cualquier símbolo de función F.

Este principio es conocido como el **Principio de Inducción para Términos**.

De la misma forma, si queremos probar que todas las fórmulas cumplen una propiedad P, tenemos únicamente que probar que:

1. toda fórmula atómica verifica P.

2. si las fórmulas φ y ψ satisfacen P y x es una variable, entonces $\neg\varphi$, $\varphi \wedge \psi$, $\varphi \vee \psi$, $\varphi \to \psi$, $\exists x\varphi$ y $\forall v\varphi$ también satisfacen P.

Este principio es conocido como el **Principio de Inducción para Fórmulas**.

Nota 4. Aunque los símbolos incluidos en el alfabeto son los utilizados habitualmente, se podría no haber incluido los símbolos $\wedge$, $\vee$ y $\exists$ y haberlos introducido como abreviaciones de la siguiente forma:

$$\varphi \wedge \psi \text{ representa } \neg(\varphi \to \neg\psi), \quad \varphi \vee \psi \text{ representa } \neg\varphi \to \psi \quad \text{y} \quad \exists x\varphi \text{ representa } \neg(\forall x \neg\varphi)$$

Se puede ver que estas definiciones dotan a los símbolos de la misma interpretación que la que se le da habitualmente (y, o y existe), donde el concepto de interpretación será formalizado más adelante. La ventaja de reducir el número de símbolos es que se reducen el número de comprobaciones necesarias en el Principio de Inducción para Fórmulas.

Nota 5. Para aligerar la notación, se pueden introducir otras abreviaciones como

$$\begin{aligned}
\varphi \leftrightarrow \beta &:\Leftrightarrow (\varphi \to \psi) \wedge (\psi \to \varphi)\\
\varphi \neq \psi &:\Leftrightarrow \neg(\varphi = \psi)\\
\exists! v\varphi &:\Leftrightarrow \exists v(\varphi(v) \wedge \forall x(\varphi(x) \to x = v)),
\end{aligned}$$

donde hemos utlizado el símbolo $:\Leftrightarrow$ para denotar simplemente que estamos definiendo un nuevo símbolo de relación, que realmente es una abreviación.
También se usarán [,] de la misma forma que los paréntesis.

Definición 1.4. Sean φ una fórmula y x una variable. Si φ no contiene un cuantificador de la forma $\exists x$ o $\forall x$, entonces se dice que x es una **variable libre**. Si ahora se construye la fórmula $\exists x\varphi$ o $\forall x\varphi$ se dirá que x es una **variable ligada**.
El conjunto de variables libres de una fórmula φ será denotado por $Free(\varphi)$. Para indicar que $Free(\varphi)$ está formado por las variables $x_1, \dots, x_n$ escribiremos $\varphi(x_1, \dots, x_n)$. Si $\tau_1, \dots, \tau_n$ son términos, podemos sustituir las variables por ellos, lo que denotaremos por $\varphi(\tau_1, \dots, \tau_n)$.

Definición 1.5. Una fórmula que no contiene variables libres se denomina **sentencia**.

Más adelante veremos que las sentencias serán aquellas fórmulas que podrán admitir una interpretación. Es decir, que podrán ser verdaderas o falsas.

1.2. Sistemas Deductivos

En el capítulo anterior se vio la importancia de establecer una teoría axiomática para formalizar la Teoría de Conjuntos Naive. En este contexto se asumió el concepto de "teoría axiomática" como intuitivo en el ámbito matemático. Sin embargo, ahora presentaremos las definiciones que nos permitirán formalizarlo y entenderlo de manera precisa.

Definición 1.6. Un **sistema deductivo** de un lenguaje esta formado por dos colecciones: una colección de fórmulas, llamadas **axiomas**, y una colección de **reglas de inferencia**, que ha de ser no vacío. La colección de axiomas está formada únicamente por los axiomas lógicos y los no lógicos, que serán definidos a continuación.

Un sistema deductivo trata de formalizar el concepto de razonamiento de la lógica ordinaria. Así, intuitivamente, entendemos que los axiomas son aquellas fórmulas que se toman como válidas. Partimos de ellas y, mediante las reglas de inferencia, derivamos fórmulas que serán válidas por provenir de las primeras.

Nota 6. Se entiende que en un sistema deductivo sólo es "válido" o "correcto" considerar cierto tipo de fórmulas. Estas serán las que esten incluidas en los axiomas o aquellas que se obtengan a través de las reglas de inferencia. El resto pueden ser escritas en el lenguaje, pero carecerán de sentido. En ocasiones, diremos que las fórmulas válidas son ciertas. También se suele decir, pues refleja la intención detrás de ellos, que los axiomas son fórmulas que siempre se cumplen.

Definición 1.7. Los **axiomas lógicos** son aquellos que se toman como válidos para un lenguaje concreto. Se entiende que son válidos universalmente, es decir, que sea cual sea la interpretación de los símbolos que intervienen en la fórmula, esta es cierta. Estos axiomas componen la agrupación más pequeña de la cual se derivan todas las sentencias universalmente válidas dentro de cualquier sistema deductivo asociado al lenguaje.

Si dentro de la fórmula del axioma lógico aparece un símbolo denotando una relación o función arbitraria, entonces diremos que se trata de un **esquema axiomático**, pues para cada función o relación concreta tenemos un axioma diferente. Cada uno de ellos se denomina una **instancia** del axioma.

Los axiomas no lógicos serán aquellos que se añaden deliberadamente a los anteriores con la intención de que pasen a ser válidos. Formalmente, tenemos la siguiente definición:

Definición 1.8. Una **teoría** es una colección de sentencias, eventualmente vacía, que se añade a los axiomas lógicos para formar la colección de axiomas de un sistema deductivo. A los elementos de la teoría se les denomina **axiomas no lógicos**.

Veamos cuáles son los axiomas lógicos dentro del lenguaje de la Lógica de Primer Orden.

Sean φ, φ_1, φ_2, φ_3 y ψ fórmulas cualesquiera, τ, $\tau_1, \ldots, \tau_n$, $\tau_1', \ldots, \tau_n'$ términos, x una variable que es libre en φ pero no en ψ, R un símbolo de relación arbitrario y F una función arbitraria. Los axiomas lógicos de un Lenguaje de Primer Orden son:

- $L_0 \,:\, \varphi \vee \neg\varphi$
- $L_1 \,:\, \varphi \to (\psi \to \varphi)$
- $L_2 \,:\, (\psi \to (\varphi_1 \to \varphi_2)) \to ((\psi \to \varphi_1) \to (\psi \to \varphi_2))$
- $L_3 \,:\, (\varphi \wedge \psi) \to \varphi$
- $L_4 \,:\, (\varphi \wedge \psi) \to \psi$
- $L_5 \,:\, \varphi \to (\psi \to (\psi \wedge \psi))$
- $L_6 \,:\, \varphi \to (\varphi \vee \psi)$
- $L_7 \,:\, \psi \to (\varphi \vee \psi)$
- $L_8 \,:\, (\varphi_1 \to \varphi_3) \to ((\varphi_2 \to \varphi_3) \to ((\varphi_1 \vee \varphi_2) \to \varphi_3))$
- $L_9 \,:\, \neg\varphi \to (\varphi \to \psi)$
- $L_{10} \,:\, \forall x \varphi(x) \to \varphi(\tau)$
- $L_{11} \,:\, \varphi(\tau) \to \exists x \varphi(x)$
- $L_{12} \,:\, \forall x(\psi \to \varphi(x)) \to (\psi \to \forall x \varphi(x))$
- $L_{13} \,:\, \forall x(\varphi(x) \to \psi) \to (\exists x \varphi(x) \to \psi)$
- $L_{14} \,:\, \tau = \tau$

- L_{15} : $(\tau_1 = \tau_i' \wedge \ldots \wedge \tau_n = \tau_n') \to (R(\tau_1, \ldots, \tau_n) \to R(\tau_1', \ldots, \tau_n'))$
- L_{16} : $(\tau_1 = \tau_i' \wedge \ldots \wedge \tau_n = \tau_n') \to (F(\tau_1, \ldots, \tau_n) \to F(\tau_1', \ldots, \tau_n'))$

En cuanto a las reglas de inferencia, las que se considerarén en este trabajo son las siguientes:

Definición 1.9. Reglas de inferencia:

- **Modus Ponens (MP)**: Si se tienen las fórmulas $\varphi \to \psi$ y φ, entonces se puede deducir ψ.
- **Generalización (G)**: Si se tiene la fórmula φ, entonces se puede deducir $\forall x \varphi$, donde x es una variable cualquiera que no aparece en ningún axioma.

Definición 1.10. En un sistema deductivo, sea Φ una colección de fórmulas. Llamamos **prueba formal** de una fórmula ψ a partir de Φ a una sucesión de fórmulas $\varphi_1, \ldots \varphi_n$ de forma que $\varphi_n \equiv \psi$ y para todo $i < n$ se cumple al menos una de las siguientes condiciones:

1. φ_i es un axioma lógico.
2. φ_i es una de las fórmulas de Φ
3. Existen $j, k < i$ tal que $\varphi_j \equiv \varphi_k \to \varphi_i$
4. Existe $j < i$ tal que $\varphi_i \equiv \forall x \varphi_j$, donde x es una variable.

Hemos utilizado el símbolo $\equiv$ como una abreviación para denotar que dos secuencias de símbolos coinciden símbolo a símbolo, es decir, son la misma. A las fórmulas de Φ se las denomina **premisas**.

Nota 7. Notar que las fórmulas que cumplan 3 serán las obtenidas mediante MP y aquella cumpliendo 4, las obtenidas mediante generalización.

Definición 1.11. Diremos que una fórmula ψ es **demostrable** de Φ si existe una prueba formal, y lo denotaremos $\Phi \vdash \psi$. Si no existe prueba formal de ψ con las fórmulas de Φ, escribiremos $\Phi \nvdash \psi$. Si la colección de fórmulas considerada es vacía, se escribe simplemente $\vdash \psi$. En este caso se dice que ψ es una **tautología**.

Definición 1.12. Se dice que dos fórmulas φ y ψ son **lógicamente equivalentes**, y se denota por $\varphi \iff \psi$, si $\varphi \leftrightarrow \psi$ es una tautología. Se cumple que $\iff$ es una relación de equivalencia.

Nota 8. En particular, en un sistema deductivo, tenemos una teoría, $\mathbf{T}$, que es un conjunto de sentencias. Por tanto, se pueden considerar las fórmulas demostrables a partir de $\mathbf{T}$. Así, cuando construimos un sistema deductivo, tenemos unas afirmaciones ciertas, los axiomas lógicos, y tomamos otras por ciertas, los axiomas no lógicos. A partir de ellos, podemos ver qué fórmulas del lenguaje derivamos mediante las reglas de inferencia. Estas serán las fórmulas demostrables a partir de $\mathbf{T}$.

Ejemplo 1.1. Sean φ y ψ dos fórmulas de un Lenguaje de Primer Orden. Una prueba formal de $\{\varphi, \psi\} \vdash \varphi \wedge \psi$ viene dada por:

φ_0:	φ	Fórmula de $\{\varphi, \psi\}$
φ_1:	ψ	Fórmula de $\{\varphi, \psi\}$
φ_2:	$\varphi \to (\psi \to (\varphi \wedge \psi))$	Instancia del axioma lógico L_5
φ_3:	$\psi \to (\varphi \wedge \psi)$	De φ_0 y φ_2 por MP
φ_4:	$\varphi \wedge \psi$	De φ_1 y φ_3 por MP

Ejemplo 1.2. De forma recíproca al ejemplo anterior, una prueba formal de $\varphi \wedge \psi \vdash \varphi$ viene dada por:

φ_0:	$\varphi \wedge \psi$	Fórmula de $\varphi \wedge \psi$
φ_1:	$(\varphi \wedge \psi) \to \varphi$	Instancia del axioma lógico L_3
φ_2:	φ	De φ_0 y φ_1 por MP

Análogamente, utilizando el axioma lógico L_4, tenemos $\varphi \wedge \psi \vdash \psi$.

Por último, definiremos dos conceptos que reflejan la idea de que una teoría esté, tal y como lo entendemos intuitivamente, libre de contradicciones.

Definición 1.13. Se dice que una teoría $\mathbf{T}$ es **completa** si para toda sentencia σ tenemos que o bien $\mathbf{T} \vdash \sigma$ o bien $\mathbf{T} \vdash \neg\sigma$. Si esto no ocurre, se dice que es **incompleta**.

Definición 1.14. Se dice que una teoría $\mathbf{T}$ es **consistente**, y se denota $Con(\mathbf{T})$, si no existe ninguna fórmula φ tal que $\mathbf{T} \vdash \varphi$ y $\mathbf{T} \vdash \neg\varphi$ o, análogamente, tal que $\mathbf{T} \vdash (\varphi \wedge \neg\varphi)$. Si existe φ cumpliendo esto, se dice que la teoría $\mathbf{T}$ es **inconsistente** o no consistente, y se denota por $\neg Con(\mathbf{T})$.

Proposición 1.1. *Si una teoría es inconsistente, entonces también es incompleta.*

1.2.1. Métodos de demostración

En los Ejemplos 1.1 y 1.2 se ilustraba cómo se llevarían a cabo pruebas formales. Sin embargo, podemos notar que, incluso para obtener resultados que pueden llegar a parecer obvios, el procedimiento puede resultar más largo de lo esperado.

En general, en las pruebas matemáticas no se lleva a cabo una demostración formal de manera exacta, si no que se realizan procedimientos, expresados en un lenguaje más intuitivo, en los que partimos de unas premisas y, a través de un razonamiento matemático, llegamos a la conclusión deseada. Sin embargo, es claro que a partir de lo desarolladoo en estas demostraciones es posible reconstruir una prueba formal. Veamos un ejemplo:

Ejemplo 1.3. Para probar $\{\neg\varphi, \neg\psi\} \vdash (\varphi \leftrightarrow \psi)$ podemos llevar a cabo el siguiente razonamiento: Aplicando el axioma lógico L_9 a las fórmulas φ y ψ tenemos que $\neg\varphi \to (\varphi \to \psi)$. Al tener $\neg\varphi$ como premisa, mediante MP, obtenemos que $\varphi \to \psi$, de la misma forma, aplicando L_9 a ψ y φ, al tener como premisa ψ, se obtiene mediante MP que $\psi \to \varphi$. Así, por el Ejemplo 1.1, al tener $\varphi \to \psi$ y $\psi \to \varphi$, se deduce $(\varphi \to \psi) \wedge (\psi \to \varphi)$. Teniendo en cuenta la definición del símbolo $\leftrightarrow$, concluimos que de $\{\neg\varphi, \neg\psi\}$ se deduce $\varphi \leftrightarrow \psi$.

Claramente, de este razonamiento se puede reconstruir una prueba formal. En primer lugar, tendríamos:

φ_0:	$\neg\varphi$	Fórmula de $\{\neg\varphi, \neg\psi\}$
φ_1:	$\neg\psi$	Fórmula de $\{\neg\varphi, \neg\psi\}$
φ_2:	$\neg\varphi \to (\varphi \to \psi)$	Instancia de L_9
φ_3:	$\varphi \to \psi$	De φ_0 y φ_2 por MP
φ_4:	$\neg\psi \to (\psi \to \varphi)$	Instancia de L_9
φ_5:	$\psi \to \varphi$	De φ_1 y φ_4 por MP

Después, replicando la prueba formal del Ejemplo 1.1, partiendo de φ_3 y φ_5, obtendríamos $(\varphi \to \psi) \wedge (\psi \to \varphi)$ y por tanto $\varphi \leftrightarrow \psi$.

Este proceso de reconstrucción de las pruebas formales puede estructurarse y llevarse a cabo para todos los métodos de demostración conocidos, como pueden ser la reducción al absurdo o una demostración por casos. Así, tenemos por ejemplo los siguientes metateoremas, donde hemos utilizado la notación $\Phi + \psi$ para denotar que a la colección de fórmulas Φ se le añade la fórmula ψ:

Teorema 1.1 (Halbeisen y Krapf, 2020, Corolario 2.5)**. *Prueba Generalizada por Casos***
Sean Φ una colección de fórmulas y $\psi_1, \ldots, \psi_n, \varphi$ fórmulas arbitrarias. Se tiene que,

$$\textit{si } \Phi \vdash \psi_1 \vee \ldots \vee \psi_n \textit{ y } \Phi + \psi_i \vdash \varphi \textit{ para todo } i \leq n, \textit{ entonces } \Phi \vdash \varphi\,.$$

Teorema 1.2 (Halbeisen y Krapf, 2020, Corolario 2.7)**. *Prueba por Contradicción***
Sean Φ una colección de fórmulas y φ una fórmula arbitraria. Si denotamos por $\#$ al imposible, se tiene que,

$$\textit{si } \Phi + \neg\varphi \vdash \#\,, \textit{ entonces } \Phi \vdash \varphi\,.$$

Nota 9. Por metateorema entendemos un teorema que arroja un resultado sobre la teoría que estamos desarrollando, en este caso la Lógica de Primer Orden. Es decir, no es un hecho dentro de la teoría, si no un resultado sobre esta.

Otro razonamiento de suma importancia en las pruebas matemáticas, presente de hecho en los teoremas anteriores, es aquel que se lleva a cabo cuando se tratan de probar implicaciones del tipo "si Φ entonces Ψ". Para ello, suele asumirse la veracidad de Φ y derivar de esta Ψ. Este razonamiento también puede ser trasladado a las pruebas formales mediente el llamado Teorema de la Deducción, que simplifica de forma notoria una gran cantiad de ellas:

Teorema 1.3 (Halbeisen y Krapf, 2020, Teorema 2.1)**. *Teorema de la Deducción***
Si Φ es una colección de fórmulas y φ y ψ son fórmulas arbitrarias, se tiene que

$$\textit{si } \Phi + \psi \vdash \varphi\,, \textit{ entonces } \Phi \vdash \psi \to \varphi\,; \textit{ y viceversa, si } \Phi \vdash \psi \to \varphi \textit{ entonces } \Phi + \psi \vdash \varphi\,.$$

En conclusión, de ahora en adelante en este trabajo, cuando se realice la prueba de algún resultado dentro de la Teoría de Conjuntos, esta se llevará a cabo mediante los métodos tradicionales. Sin embargo, no se ha de olvidar que trabajamos en un sistema deductivo y que de estas pruebas se podría obtener siempre una prueba formal.

1.3. Teoría de Modelos

Hasta ahora, hemos presentado un lenguaje a partir del cual podemos generar fórmulas. Mediante los sistemas deductivos, algunas eran tomadas como ciertas y otras resultaban válidas por ser obtenidas de las primeras a través de las reglas de inferencia. Este punto de vista se denomina **sintáctico**. Sin embargo, existe otro punto de vista, el **semántico**, según el cual se considera el significado que se pretende otorgar a los símbolos y se busca un modelo en el que todas las fórmulas sean ciertas según dicho significado. En esta sección veremos los conceptos básicos de este nuevo enfoque. Para ello, necesitaremos asumir algunas ideas de la Teoría de Conjuntos Naive, como función o n-tupla. Una vez que se formalicen en el siguiente capítulo, esta teoría podría desarrollarse de la misma manera en función de ellos.

Definición 1.15. Dado un lenguaje $\mathscr{L}$, una **$\mathscr{L}$-estructura** $\mathbb{M}$, está formada por los siguientes elementos:

- una colección de objetos no vacía M, que se denomina **dominio** de $\mathbb{M}$
- una aplicación tal que
  - asigna a cada constante c de $\mathscr{L}$ un elemento c_M de M.
  - asigna a cada símbolo de relación n-ario una colección de n-tuplas de elementos de M, que denotamos por R_M. Intuitivamente, se entiende que ciertos elementos de M están relacionados si la tupla que forman pertenece a R_M.
  - asigna a cada símbolo de función n-ario una función F_M que lleva n-tuplas de M a elementos de M.

Así, hemos vinculado a las constantes con ciertos objetos que conocemos, a los símbolos de relación con subconjuntos de estos objetos y a los símbolos de función con funciones entre ellos. Para interpretar las variables tenemos la siguiente definición.

Definición 1.16. Una **asignación** en una $\mathscr{L}$-estructura $\mathbb{M}$ es una aplicación j que asigna a cada variable de $\mathscr{L}$ un elemento del dominio M.

Con todo esto, podemos definir formalmente lo que entendemos por interpretación:

Definición 1.17. Una **interpretación** es un par $(\mathbb{M}, j)$, donde j es una asignación y $\mathbb{M}$ una $\mathscr{L}$-estructura. Además, se asocia cada término τ del lenguaje con un elemento de M que denotaremos $I(\tau)$. Teniendo en cuenta las reglas de formación de los términos (T0)-(T2) definimos:

- para una variable x, $I(x) \coloneqq j(x)$.
- para una constante c, $I(c) \coloneqq c_M$

- para un símbolo de función n-ario F y términos $\tau_1, \dots \tau_n$,

$$I\big(F(\tau_1, \dots \tau_n)\big) \coloneqq F_M\big(I(\tau_1), \dots I(\tau_n)\big).$$

Donde hemos utlizado el símbolo $\coloneqq$ para indicar que se está definiendo dicha igualdad.

Dada una interpretación, podemos querer modificarla para que a cierta variable se le asigne cierto objeto del dominio. Para ello se da la siguiente definición:

Definición 1.18. Sea una interpretación $I = (\mathbb{M}, j)$. Dadas dos variables x e y y un objeto m de M, definimos la asignación $j_{x/m}$ como

$$j_{x/m}(y) = \begin{cases} m & \text{si } y \equiv x, \\ j(y) & \text{en otro caso.} \end{cases}$$

Definimos la interpretación $I_{x/m}$ como $I_{x/m} \coloneqq (\mathbb{M}, j_{x/m})$. Intuitivamente, tomando $I_{x/m}$ estamos considerando una interpretación en la que la variable x se entiende como m.

Una vez definida la interpretación para los términos, podemos pasar a definir cuándo una fórmula es cierta bajo una interpretación.

Definición 1.19. Dada una interpretación $I = (\mathbb{M}, j)$ y una fórmula φ, decimos que φ es **cierta para** I o que **se cumple en** I si $I \models \varphi$, símbolo que definiremos a continuación. En caso contrario diremos que es **falsa** para I o que **no se cumple** en I. Teniendo en cuenta las reglas de formación de las fórmulas (F0)-(F3), definimos

$$\begin{aligned}
&I \models (\tau_1 = \tau_2) && :\Leftrightarrow I(\tau_1) \text{ es el mismo objeto que } I(\tau_2) \\
&I \models R(\tau_1, \dots, \tau_n) && :\Leftrightarrow \text{la n-tupla formada por } I(\tau_1), \dots, I(\tau_n) \text{ pertenece a } R_M \\
&I \models \neg\varphi && :\Leftrightarrow \text{ no } I \models \varphi \\
&I \models \varphi_1 \wedge \varphi_2 && :\Leftrightarrow I \models \varphi_1 \text{ y } I \models \varphi_2 \\
&I \models \varphi_1 \vee \varphi_2 && :\Leftrightarrow I \models \varphi_1 \text{ o } I \models \varphi_2 \\
&I \models \varphi_1 \to \varphi_2 && :\Leftrightarrow \text{ si } I \models \varphi_1 \text{ entonces } I \models \varphi_2 \\
&I \models \exists x\varphi && :\Leftrightarrow \text{ existe un } m \text{ en } M \text{ tal que } I_{x/m} \models \varphi \\
&I \models \forall x\varphi && :\Leftrightarrow \text{para todo } m \text{ en } M, I_{x/m} \models \varphi
\end{aligned}$$

En resumen, tenemos que una fórmula se cumple si al sustituir los símbolos por sus interpretaciones esta se cumple.

Nota 10. Es importante observar que, en el lenguaje de la lógica ordinaria, el cual usamos para interpretar las fórmulas, algo es o bien cierto o bien falso. Por tanto, siempre tendremos que o bien $I \models \varphi$ o bien $I \models \neg\varphi$, lo que no ocurría sintácticamente, pues no toda teoría es completa.

Nota 11. Dada una sentencia σ y dos interpretaciones I_1 e I_2 de una misma estructura, se tiene que $I_1 \models \sigma$ si y solo si $I_2 \models \sigma$.
Esto es debido a que, al no tener variables libres, todas ellas aparecen en la fórmula de la forma $\exists x\psi$ o $\forall x\psi$. Así, teniendo en cuenta las definiciones de $I \models \exists x\psi$ e $I \models \forall x\psi$, vemos que el hecho de que σ sea cierta en una interpretación no depende de la interpretación original. Por tanto, uniendo esto a la nota anterior, tenemos que *una sentencia σ es o bien cierta para cualquier interpretación de una estructura, o bien falsa para todas ellas.*

Como se comentaba al introducir la sección, en el enfoque semántico se buscaba encontrar una estructura en la cual todas las fórmulas de la teoría fuesen ciertas. Con este objetivo, se tienen las siguientes definiciones.

Definición 1.20. Sea $\mathbb{M}$ una $\mathscr{L}$- estructura y φ una fórmula. Escribimos $\mathbb{M} \models \varphi$ si y solo si para toda interpretación I de la estructura se tiene que $I \models \varphi$. Se dirá que φ **es cierta** o **se cumple en** $\mathbb{M}$.

Nota 12. Para una estructura $\mathbb{M}$ y para una sentencia σ, de acuerdo con la Nota 11, se tiene que o bien $\mathbb{M} \models \sigma$ o bien $\mathbb{M} \models \neg\sigma$.

Definición 1.21. Sea Φ una colección de fórmulas. Se dice que una $\mathscr{L}$- estructura $\mathbb{M}$ es un **modelo** de Φ, y escribimos $\mathbb{M} \models \Phi$ si para toda fórmula φ de Φ se tiene que $\mathbb{M} \models \varphi$. Es decir, $\mathbb{M}$ es un modelo de Φ si para cada fórmula φ de Φ y para cada interpretación I de $\mathbb{M}$, $I \models \varphi$.

Intuitivamente, $\mathbb{M}$ será un modelo de una colección de fórmulas si todas ellas son ciertas sea cual sea la interpretación de la estructura que se tome, es decir, si todas ellas se cumplen en $\mathbb{M}$.

Definición 1.22. Sean una fórmula φ y una colección de fórmualas Φ. Si para cualquier modelo $\mathbb{M}$ de Φ se cumple que $\mathbb{M} \models \varphi$, se dice que φ es **consecuencia semántica** de Φ y se escribe $\Phi \models \varphi$. Es decir, una fórmula es consecuencia semántica de una colección de fórmulas si esta es cierta para cualquier modelo y para cualquier interpretación de este.

1.4. Teoremas de Completitud e Incompletitud de Gödel

Una vez en el contexto de la Teoría de Modelos, dada Φ una colección de fórmulas, cabe plantearse la relación entre una fórmula demostrable de Φ en un sistema deductivo y una fórmula que sea consecuencia semántica de Φ. Esto se resuelve mediante el Teorema de Solidez y el Teorema de Completitud de Gödel. Por otro lado, gracias también a la Teoría de Modelos, se desarrollan dos teoremas fundamentales, los Teoremas de Incompletitud de Gödel, que arrojan importantes conclusiones sobre la consistencia y completitud de ciertas teorías. En esta sección presentaremos estos resultados y sus consecuencias más directas. Para un desarrollo de los conceptos necesarios para entender las pruebas, así como para consultar estas mismas, se puede ver [Halbeisen y Krapf, 2020, Partes I y II].

Teorema 1.4 (Halbeisen y Krapf, 2020, Teorema 3.8)**. *Teorema de Solidez***
Sea Φ una colección de fórmulas de un lenguaje $\mathscr{L}$ y $\mathbb{M}$ un modelo de Φ. Entonces, para toda fórmula φ se tiene que si $\Phi \vdash \varphi$, entonces $\mathbb{M} \models \varphi$.
Teniendo en cuenta que esto se cumple para cualquier modelo, concluimos que

$$\textit{si } \Phi \vdash \varphi \textit{ entonces } \Phi \models \varphi.$$

Este teorema afirma que una fórmula que sea demostrable a partir de una colección de fórmulas es también consecuencia semántica de estas. Además, permite obtener resultados de gran importancia.

Corolario 1.1. *Si una teoría tiene un modelo, entonces es consistente.*

Demostración. Sea $\mathbf{T}$ una teoría y $\mathbb{M}$ un modelo de esta. Supongamos que existe una fórmula φ tal que $\mathbf{T} \vdash \varphi$ y $\mathbf{T} \vdash \neg\varphi$. Entonces, por el Teorema de Solidez, tenemos que $\mathbf{T} \models \varphi$ y $\mathbf{T} \models \neg\varphi$. Es decir, para cualquier interpretación I de $\mathbb{M}$ tenemos que $I \models \varphi$ e $I \models \neg\varphi$. Sin embargo, sabemos que esto no es posible, de donde la fórmula φ no puede existir. □

Corolario 1.2. *Si $\mathbf{T}$ es una teoría que tiene un modelo, entonces existe una teoría completa $\mathbf{T}'$ que contiene a $\mathbf{T}$, en el sentido de que todas las sentencias que conforman $\mathbf{T}$ son sentencias de $\mathbf{T}'$.*

Demostración. Sea $\mathbb{M}$ un modelo de $\mathbf{T}$ y sea $\mathbf{T}'$ la colección formada por todas aquellas sentencias σ para las cuales $\mathbb{M} \models \sigma$. Es claro que en $\mathbf{T}'$ están todas las sentencias de $\mathbf{T}$, por ser $\mathbb{M}$ un modelo de esta teoría. Observamos también que $\mathbb{M}$ es un modelo de $\mathbf{T}'$. Veamos que además es completa.

Sea σ una sentencia. Es sabido que o bien $\mathbb{M} \models \sigma$ o bien $\mathbb{M} \models \neg\sigma$. Por tanto, o bien σ es parte de las sentencias de $\mathbf{T}'$ o bien lo es $\neg\sigma$. De la primera se deduciría que $\mathbf{T}' \vdash \sigma$ y de la segunda que $\mathbf{T}' \vdash \neg\sigma$.

Como por el Corolario anterior $\mathbf{T}'$ es consistente, no se pueden dar las dos opciones de forma simultánea. Así, o bien $\mathbf{T}' \vdash \sigma$ o bien $\mathbf{T}' \vdash \neg\sigma$, de donde $\mathbf{T}'$ es completa. $\square$

Pasamos ahora a enunciar el Teorema de Completitud de Gödel, que proporciona condiciones suficientes para que una teoría tenga un modelo.

Teorema 1.5 (Halbeisen y Krapf, 2020, Teorema 5.5)**.** ***Teorema de Completitud de Gödel*** *Si $\mathscr{L}$ es un lenguaje con una signatura numerable y $\mathbf{T}$ es una teoría consistente, entonces tiene un modelo. Además, si $T \nvdash \varphi$ para alguna sentencia φ, entonces $\mathbf{T} + \neg\varphi$ tiene un modelo.*

Nota 13. En el contexto de la Teoría Axiomática de Conjuntos, este teorema se puede extender a un lenguaje con una signatura arbitraria.

La primera consecuencia de este teorema es la siguiente, obtenida de sumarlo al Corolario 1.1:

Corolario 1.3. *Una teoría es consistente si y solo si tiene un modelo.*

Si por otro lado combinamos el Teorema de Completitud con el Corolario 1.2, obtenemos que:

Corolario 1.4. *Toda teoría consistente está contenida en una teoría completa.*

No obstante, las consecuencias más importantes del Teorema de Completitud vienen de combinarlo con el Teorema de Solidez. El resultado sin duda más relevante es el siguiente:

Corolario 1.5. *Sean $\mathbf{T}$ una teoría en un lenguaje con una signatura numerable y σ una sentencia. Entonces,*

$$\mathbf{T} \vdash \sigma \text{ si y solo si } \mathbf{T} \models \sigma.$$

Demostración. Ya es sabido, por el Teorema de Solidez, que si $\mathbf{T} \vdash \sigma$ entonces $\mathbf{T} \models \sigma$. Por tanto, supongamos que $\mathbf{T} \models \sigma$ y veamos que $\mathbf{T} \vdash \sigma$.

Supongamos que $\mathbf{T} \nvdash \sigma$. Entonces, por el Teorema de Completitud, $T + \neg\sigma$ tiene un modelo, que denotaremos $\mathbb{M}$. Observemos que claramente $\mathbb{M}$ es también un modelo de $\mathbf{T}$. Así, por ser $\mathbb{M}$ modelo de $T + \neg\sigma$, tenemos que $\mathbb{M} \models \neg\sigma$. Por otro lado, como $\mathbf{T} \models \sigma$ y $\mathbb{M}$ es un modelo de $\mathbf{T}$, se obtiene que $\mathbb{M} \models \sigma$. Sin embargo, ambos hechos no pueden ocurrir simultáneamente para sentencias, de donde concluimos que $\mathbf{T} \vdash \sigma$. $\square$

Esta equivalencia es de gran importancia, pues implica que es posible reemplazar, en las pruebas formales, los símbolos del lenguaje por sus interpretaciones. Esto termina de completar lo expuesto al hablar de los métodos de demostración, permitiéndonos realizar todas las pruebas matemáticas tal y como las conocemos.

A continuación, presentaremos los dos Teoremas de Incompletitud de Gödel, que nos permitirán, más adelante, sacar conclusiones sobre la consistencia y completitud de los axiomas de la Teoría de Conjuntos. Estos resultados arrojan conclusiones que involucran un sistema deductivo concreto, denominado la Aritmética de Peano, que presentaremos antes de enunciar los teoremas.

Definición 1.23. La **Aritmética de Peano** es un sistema deductivo del **Lenguaje de la Aritmética**, cuya signatura es $\mathscr{L}_{PA} = \{0, s, +, \cdot\}$, donde el 0 es una constante, s es una función de aridad uno, denominada **sucesor**, y $+, \cdot$, la **suma** y el **producto**, funciones de aridad dos. Los axiomas de la teoría, que denotaremos como PA, son

$$\begin{aligned}
\mathsf{PA}_0 &: \quad \forall x \neg(s(x) = 0) \\
\mathsf{PA}_1 &: \quad \forall x \forall y (s(x) = s(y) \rightarrow x = y) \\
\mathsf{PA}_2 &: \quad \forall x (x + 0 = x) \\
\mathsf{PA}_3 &: \quad \forall x \forall y (x + s(y) = s(x + y)) \\
\mathsf{PA}_4 &: \quad \forall x (x \cdot 0 = 0) \\
\mathsf{PA}_5 &: \quad \forall x \forall y (x \cdot s(y) = (x \cdot y) + x) \\
\mathsf{PA}_6 &: \quad \varphi(0) \wedge \forall x \big(\varphi(x) \rightarrow \varphi(s(x))\big) \rightarrow \forall x \varphi(x)
\end{aligned}$$

Donde en el último axioma φ es una fórmula y x una variable libre de φ. PA_6 recibe el nombre de **Axioma de Inducción**.

Este sistema deductivo tiene un modelo, cuyo dominio es el conjunto de los números naturales, de los que hablaremos en el siguiente capítulo, y cuyas interpretaciones van acorde a los conceptos que se conocen de la aritmética de estos.

Visto esto, podemos pasar a enunciar los teoremas. Los enunciaremos de una forma más bien intuitiva, pues los conceptos necesarios para su enunciado riguroso van más allá del propósito de este trabajo.

Teorema 1.6 (Halbeisen y Krapf, 2020, Teorema 10.12)***. Primer Teorema de Incompletitud de Gödel***
Sea $\mathscr{L}$ *un lenguaje numerable conteniendo los símbolos de* $\mathscr{L}_{PA}$ *y* $\mathbf{T}$ *una teoría en la cual se cumplen los axiomas de PA. Si* $\mathbf{T}$ *es consistente, entonces es incompleta.*

Este teorema nos dice que una teoría que sea lo suficientemente fuerte como para desarrollar en ella la Aritmética de Peano no puede ser a la vez consistente y completa. Es decir, si esta es consistente, habrá sentencias que, aunque sean ciertas a nivel semántico, ni ellas ni su negación serán demostrables en la teoría.

Para el segundo teorema, utilizaremos que, dada una teoría $\mathbf{T}$, es posible construir una sentencia, que denotaremos $con_{\mathbf{T}}$, de forma que si esta es cierta entonces la teoría es consistente.

Teorema 1.7 (Halbeisen, 2012, Teorema 3.9)***. Segundo Teorema de Incompletitud de Gödel***
Sea $\mathscr{L}$ *un lenguaje conteniendo los símbolos de* $\mathscr{L}_{PA}$ *y* $\mathbf{T}$ *una teoría en la cual se cumplen los axiomas de PA. Si* $\mathbf{T}$ *es consistente, entonces*

$$\mathbf{T} \nvdash con_{\mathbf{T}}.$$

Es decir, si una teoría en la que se cumplen los axiomas de PA *es consistente, dicha consistencia no puede ser probada únicamente a partir de la teoría.*

2 | Teoría Axiomática de Conjuntos

En 1908, Ernst Zermelo publicó en su artículo *"Investigaciones sobre los Fundamentos de la Teoría de Conjuntos I"* su primer sistema axiomático. Más adelante, en 1922, Fraenkel y Skolen mejoraron, de forma independiente, esta teoría. La versión final fue de nuevo publicada por Zermelo en 1930 y es conocida como la **Teoría Axiomática de Zermelo-Fraenkel**.
En este capítulo presentaremos los distintos axiomas que componen esta teoría, tratando de introducirlos de manera intuitiva y estudiando qué conceptos matemáticos se desprenden de cada uno de ellos. También consideraremos el conocido Axioma de Elección y discutiremos las consecuencias de incluirlo en nuestra teoría.

Se ha tomado como referencia principal para la exposición de los axiomas y el desarrollo de conceptos el texto de [Halbeisen y Krapf, 2020]. Además, se ha complementado la teoría presentada en dicho texto añadiendo la visión intuitiva de los axiomas proporcionada en [Hrbacek y Jech, 2017], así como la definición y el desarrollo de los números naturales según [Pinter, 2014] y la teoría sobre el Axioma de Elección de [Devlin, 1994]. También se desarrollaron algunas definiciones propias, como las de n-tuplas ordenadas y no ordenadas, y posteriormente se comprobó que coincidían con las de [Hrbacek y Jech, 2017]. Las demostraciones de los resultados, aunque esbozadas en los textos de referencia, se han desarrollado de forma original para integrarlas en la teoría presentada en este trabajo. Cabe destacar que el enunciado y la demostración de la Proposición 2.7, al no encontrarse en los libros consultados, fueron realizados de manera original.

2.1. Los Axiomas de Zermelo

Como se explicó en el capítulo anterior, lo primero que se debe fijar es el Lenguaje de Primer Orden que será utilizado. Una vez considerado, construiremos un sistema deductivo en el que la teoría estará compuesta por los citados Axiomas de Zermelo-Fraenkel.

Definición 2.1. El **Lenguaje de la Teoría de Conjuntos**, que denotaremos como $\mathscr{L}_{ST}$, es un Lenguaje de Primer Orden cuya signatura contiene únicamente un símbolo de relación (binaria), $\in$, denominada **relación de pertenencia**. Si $y \in x$ diremos que y **pertenece** a x o que y es un

elemento de x. Si por el contrario se tiene $\neg(y \in x)$, se dirá que y **no pertenece a** x y se denotará $y \notin x$. Además, existe un único tipo de objetos en este lenguaje, que denominaremos **conjuntos**.

De esta forma, vemos que la definición de conjunto es en gran medida distinta a aquella con la que se comenzaba este trabajo. Un conjunto ya no es una colección de elementos, si no que no es más que un objeto matemático abstracto. En palabras del propio Zermelo, [Zermelo, 1908/1967]:

"La teoría de conjuntos trata sobre un "dominio" de individuos, a los que llamamos simplemente "objetos", entre los cuales los "conjuntos" son una parte. Entre los objetos del dominio existen ciertas "relaciones fundamentales" de la forma $a \in b$. Si para dos objetos a, b se establece la relación $a \in b$, decimos que "a es un elemento del conjunto b" o "b contiene a a como elemento". Un objeto b, que contiene otro objeto a como elemento, siempre puede ser nombrado como conjunto, y solo si este es el caso, con una excepción (Axioma I)."

Donde el Axioma I, como veremos a continuación, hablará del conjunto vacío.

Así, podemos pasar a describir los axiomas. Como se veía, estos serán sentencias que se desea tomar como siempre ciertas. Sin embargo, al enunciarlos por primera vez, se trata de reflejar los conceptos que se tienen asumidos o que se ven como triviales en nuestro razonamiento. Así, trataremos de deducir los axiomas de nuestras ideas, teniendo en cuenta que deben imitar nuestro pensamiento pero que se deben hacer las menos asunciones necesarias para que la teoría sea lo mas amplia posible.

Comenzaremos con un axioma que garantice la existencia de al menos un conjunto.

I. Axioma del Conjunto Vacío

$$\exists x \forall z (z \notin x)$$

Este axioma postula la existencia de un conjunto que no tiene elementos.

Una vez dada la existencia de un conjunto, cabe plantearse cuándo dos conjuntos serán iguales, así como la unicidad del conjunto sin elementos. Esto nos lleva al siguiente axioma.

II. Axioma de Extensionalidad

$$\forall x \forall y (\forall z (z \in x \leftrightarrow z \in y) \rightarrow x = y)$$

Este axioma puede interpretarse como que dos conjuntos conteniendo los mismos elementos son iguales.

El recíproco, es decir, "dos conjuntos iguales contienen los mismos elementos", es cierto por el axioma lógico L_{15}, pues dos objetos iguales deben mantener las mismas relaciones:

$$L_{15} : (\tau_1 = \tau_i' \wedge \ldots \wedge \tau_n = \tau_n') \to (R(\tau_1, \ldots, \tau_n) \to R(\tau_1', \ldots, \tau_n')).$$

Por tanto, se tiene el siguiente resultado:

Proposición 2.1. *Dos conjuntos son iguales si y solo si contienen los mismos elementos:*

$$\forall x \forall y [\forall z (z \in x \leftrightarrow z \in y) \leftrightarrow (x = y)]$$

Demostración. Una instancia del axioma L_{15} es $(x = y \wedge z = z) \to (z \in x \to z \in y)$, donde sabemos que $z = z$ por L_{14}. De la misma forma lo es $(y = x \wedge z = z) \to (z \in y \to z \in z)$, de donde obtenemos $(x = y) \to (z \in x \leftrightarrow z \in y)$. Así, mediante generalización y aplicando L_{12}, tenemos que

$$(x = y) \to \forall z (z \in x \leftrightarrow z \in y).$$

De esta fórmula junto al Axioma de Extensionalidad obtenemos el resultado deseado. □

Observemos que esto implica también que dos conjuntos no son iguales, $x \neq y$, si y solo si existe un elemento de uno de ellos que no lo es del otro.

Visto esto, podemos probar ahora la unicidad del conjunto sin elementos dado en el Axioma del Conjunto Vacío.

Proposición 2.2. *Existe únicamente un conjunto sin elementos.*

Demostración. Supongamos que x_0 y x_1 son conjuntos sin elementos. Es decir, se cumple que

$$\forall z (z \notin x_0) \text{ y } \forall z (z \notin x_1)$$

Así, para un término a cualquiera, aplicando L_{10}, tenemos que $\neg(a \in x_0)$ y que $\neg(a \in x_1)$. Por tanto, partiendo de $\{\neg(a \in x_0), \neg(a \in x_1)\}$ y aplicando el Ejemplo 1.3, obtenemos $(a \in x_0) \leftrightarrow (a \in x_1)$. Así, mediante generalización, se concluye que

$$\forall z (z \in x_0 \leftrightarrow z \in x_1)$$

Por último, aplicando la Proposición 2.2 a x_0 y x_1, concluimos que

$$x_0 = x_1.$$

□

Este resultado nos permite definir de forma unívoca el conjunto vacío:

Definición 2.2. El único conjunto sin elementos se denomina **conjunto vacío** y se denota $\emptyset$.

Con la información descrita hasta ahora podemos definir también las siguientes relaciones binarias:

Definición 2.3. Dados x e y dos conjuntos, se define el símbolo de relación binaria $\subseteq$, denominado **subconjunto** como

$$x \subseteq y :\Leftrightarrow \forall z(z \in x \to z \in y).$$

Es decir, y será un subconjunto de x si todos los elementos de y son también elementos de x. Se dice que x **está contenido en** y o que y **contiene a** x.

De la misma forma, definimos el símbolo de relación binaria $\subsetneq$, denominado **subconjunto propio** como

$$x \subsetneq y :\Leftrightarrow (x \subseteq y) \wedge (x \neq y).$$

Es decir, y será un subconjunto de x si todos los elementos de y son también elementos de x y existe al menos un elemento de x que no lo es de y.

Nota 14. Notar que $\forall x(\emptyset \subseteq x)$.

Proposición 2.3. *Dados dos conjuntos x e y,*

$$(x \subseteq y \wedge y \subseteq x) \leftrightarrow x = y$$

Demostración. De $y \subseteq x$ tenemos que $\forall z(z \in y \to z \in x)$ y de $x \subseteq y$, que $\forall z(z \in x \to z \in y)$. Por tanto, tenemos $\forall z(z \in y \leftrightarrow z \in x)$, de donde $x = y$. Si $x = y$ es claro que se da $(x \subseteq y \wedge y \subseteq x)$.

□

Proposición 2.4. *Dados tres conjuntos x, y, z,*

$$(x \subseteq y \wedge y \subseteq z) \to x \subseteq z.$$

Demostración. Dado $w \in x$, por $x \subseteq y$, $w \in y$. Por $y \subseteq z$, como $w \in y$, $w \in z$. Así, $\forall w(w \in x \to w \in z)$. □

Hasta ahora, el único conjunto que existe con certeza es el conjunto vacío. Así, los siguientes axiomas tratarán de reflejar construcciones matemáticas habituales, que nos permitirán además definir nuevos conjuntos.

III. Axioma de Formación de Pares

$$\forall x \forall y \exists u \forall z (z \in u \leftrightarrow (z = x \vee z = y))$$

Este axioma postula que, dados dos conjuntos x e y, existe un conjunto cuyos únicos elementos son x e y.

Por el Axioma de Extensionalidad, sabemos que este conjunto es único, pues un conjunto está determinado por sus elementos. Así, tenemos la siguiente definición:

Definición 2.4. Dados dos conjuntos x e y, se define el **par no ordenado** $\{x, y\}$ como

$$\{x, y\} = u :\Leftrightarrow \forall z (z \in u \leftrightarrow (z = x \vee z = y))$$

Si $x = y$, denotaremos $\{x, x\}$ como $\{x\}$.

Nota 15. Notemos que $\{\,,\}$ es un símbolo de función binario, pues por el Axioma de Formación de Pares $\{x, y\}$ es un conjunto.

Nota 16. Supongamos que $\{x_0, y_0\} = \{x_1, y_1\}$. Los elementos del primer conjunto son x_0 e y_0. Por el Axioma de Extensionalidad, estos son elementos del segundo. Por el Axioma de Pares, $x_0 \in \{x_1, y_1\} \rightarrow (x_0 = x_1 \vee x_0 = y_1)$. Es decir, $x_0 = x_1$ o $x_0 = y_1$. De la misma forma, $y_0 = x_1$ o $y_0 = y_1$.

Gracias a este axioma, vemos que efectivamente podemos definir nuevos conjuntos distintos del conjunto vacío:
Aplicando el Axioma de Formación de Pares a los conjuntos $\emptyset$ y $\emptyset$, podemos definir el conjunto $\{\emptyset, \emptyset\} = \{\emptyset\}$. Claramente,

$$\emptyset \neq \{\emptyset\},$$

pues el primero no tiene elementos y el segundo tiene uno: el conjunto vacío. De la misma forma, aplicando el Axioma de Formación de Pares a $\emptyset$ y a $\{\emptyset\}$, construimos el conjunto $\{\emptyset, \{\emptyset\}\}$ y así de forma sucesiva.

Por otro lado, notemos que el nombre de par no ordenado está en consonancia con el hecho de que $\{x, y\} = \{y, x\}$, de acuerdo con el Axioma de Extensionalidad. Sin embargo, en ocasiones se desea que el orden de los elementos sí tenga relevancia, por lo que tenemos la siguiente definición.

Definición 2.5. Dados dos conjuntos x e y, se define el **par ordenado** $\langle x, y\rangle$ como

$$\langle x, y\rangle = u :\Leftrightarrow \forall z(z \in u \leftrightarrow (z = \{x\} \vee z = \{x, y\})).$$

Por tanto, teniendo en cuenta la definición de par no ordenado, se tiene que

$$\langle x, y\rangle = \{\{x\}, \{x, y\}\} = \{\{x, y\}, \{x\}\}.$$

Nota 17. Notemos que $\langle\, ,\rangle$ es un símbolo de función binario, pues por el Axioma de Formación de Pares $\langle x, y\rangle$ es un conjunto.

Efectivamente, esta definición nos da un par ordenado tal y como se entiende en matemáticas:

Proposición 2.5. *Dados los conjuntos x_0, y_0, x_1 e y_1, se tiene que*

$$\langle x_0, y_0\rangle = \langle x_1, y_1\rangle \leftrightarrow (x_0 = x_1 \wedge y_0 = y_1).$$

Demostración. Tenemos que

$$\langle x_0, y_0\rangle = \{\{x_0\}, \{x_0, y_0\}\} \text{ y } \langle x_1, y_1\rangle = \{\{x_1\}, \{x_1, y_1\}\}.$$

Es claro que $(x_0 = x_1 \wedge y_0 = y_1) \rightarrow \langle x_0, y_0\rangle = \langle x_1, y_1\rangle$, pues serían conjuntos con los mismos elementos. Por otro lado, si $\langle x_0, y_0\rangle = \langle x_1, y_1\rangle$, entonces tienen los mismos elementos. Tenemos varios casos:

1. Si $\{x_0\} = \{x_1\}$, lo que es análogo a que $\{x_0, x_0\} = \{x_1, x_1\}$, entonces $x_0 = x_1$ por la Nota 16. Utilizaremos el razonamiento de esta nota de ahora en adelante.

   1.1) Si $\{x_0, y_0\} = \{x_1, x_1\}$ entonces $y_0 = x_1$.

   1.1.1) Si $\{x_1, y_1\} = \{x_0, x_0\}$, entonces $y_1 = x_0 = x_1 = y_0$. Así, hemos obtenido $x_0 = x_1 \wedge y_0 = y_1$.

1.1.2) Si $\{x_1, y_1\} = \{x_0, y_0\}$, o bien $y_1 = y_0$ o bien $y_1 = x_0 = x_1 = y_0$. En ambos casos concluimos que $x_0 = x_1 \wedge y_0 = y_1$.

1.2) Si $\{x_0, y_0\} = \{x_1, y_1\}$, entonces o bien $y_0 = y_1$ en cuyo caso se obtiene lo deseado o bien $y_0 = x_1$ y se repite el razonamiento de 1.1.1 y 1.1.2.

2. Si $\{x_0, x_0\} = \{x_1, y_1\}$, entonces $x_0 = x_1$ y $x_0 = y_1$.

2.1) Si $\{x_0, y_0\} = \{x_1, x_1\}$, entonces $y_0 = x_1 = x_0 = y_1$ y tenemos $x_0 = x_1 \wedge y_0 = y_1$.

2.2) Si $\{x_0, y_0\} = \{x_1, y_1\}$, entonces o bien $y_0 = y_1$ en cuyo caso se obtiene lo deseado o bien $y_0 = x_1 = x_0 = y_1$, de donde de nuevo $x_0 = x_1 \wedge y_0 = y_1$.

$\square$

IV. Axioma de la Unión

$$\forall x \exists u \forall z (z \in u \leftrightarrow \exists w \in x (z \in w))$$

Este axioma postula que, dado un conjunto x, existe un conjunto u formado por los elementos de alguno de sus elementos.

Ejemplo 2.1. Dado el conjunto $x = \{\{a, b\}, \{b\}\}$, existe un conjunto cuyos elementos son a y b: $\{a, b\}$. Observemos que los elementos de x son $\{a, b\}$ y $\{b\}$.

De nuevo, como un conjunto viene determinado por sus elementos, este axioma define un nuevo símbolo de función 1-ario:

Definición 2.6. Definimos la **unión de un conjunto** x como:

$$\bigcup x = u :\Leftrightarrow \forall z (z \in u \leftrightarrow \exists w \in x (z \in w))$$

Ejemplo 2.2. Dado un conjunto x, se tiene que $x = \bigcup\{x\}$:
Si $z \in x$, entonces existe un elemento de $\bigcup\{x\}$, el propio x, tal que $z \in x$. De la misma forma, si $z \in \bigcup\{x\}$, existe un elemento de $\bigcup\{x\}$ conteniendo a z como elemento. Como el único elemento de la unión es x, tenemos que $z \in x$.

De la misma forma, utilizando el Axioma de Unión y el Axioma de Formación de Pares, podemos definir el siguiente símbolo de función binario:

Definición 2.7. Definimos la **unión de dos conjuntos** x e y como

$$x \cup y = u :\Leftrightarrow u = \bigcup\{x, y\}$$

Vemos que $z \in x \cup y \leftrightarrow (z \in x \vee z \in y)$.

Nota 18. Claramente, $x \subseteq x \cup y$ y análogamente $y \subseteq x \cup y$.

Proposición 2.6. *Dados tres conjuntos a, b y x*

$$(a \subseteq x \wedge b \subseteq x) \rightarrow a \cup b \subseteq x.$$

Es decir, $a \cup b$ es un subconjunto de cualquier conjunto conteniendo a a y b.

Demostración. Sea $z \in a \cup b$. Entonces, $z \in a \vee z \in b$. Como tanto a como b están contenidos en x, en ambos casos $z \in x$. □

Una consecuencia importante del Axioma de la Unión es que podemos formar la unión de una colección de conjuntos de manera recursiva:
Dados tres conjuntos x, y y z, podemos tomar $(x \cup y) \cup z$ Entonces,

$$w \in (x \cup y) \cup z \leftrightarrow (w \in x \cup y \vee w \in z)$$

Como $w \in x \cup y \leftrightarrow (w \in x \vee w \in y)$, tenemos que

$$w \in (x \cup y) \cup z \leftrightarrow (w \in x \vee w \in y \vee w \in z).$$

Por tanto, definiríamos $(x \cup y) \cup z$ y podríamos tomar la unión de este conjunto con un cuarto conjunto y así sucesivamente. Observemos que mediante este proceso sólo se puede formar la unión de un número finito de conjuntos.

Tomando conjuntos que tengan un único elemento, podemos también definir n-tuplas no ordenadas para cualquier número n:
Si queremos formar un conjunto que tenga como elementos $x_1, \ldots x_n$, podemos definir $X_1 = \{x_1\}, \ldots, X_n = \{x_n\}$ y tomar $X_1 \cup \ldots \cup X_n$. Así,

$$z \in X_1 \cup \ldots \cup X_n \leftrightarrow (z \in X_1 \vee \ldots \vee z \in X_n)$$

Sin embargo, para cualquier k, si $z \in X_k$, como $X_k = \{x_k\}$, tendremos que $z = x_k$. Así,

$$z \in X_1 \cup \ldots \cup X_n \leftrightarrow (z = x_1 \vee \ldots \vee z = x_n).$$

Es decir, de esta forma definimos de manera unívoca el conjunto formado por los elementos $x_1, \ldots, x_n$.

Definición 2.8. Dados los conjuntos $x_1, \ldots, x_n$, se define el símbolo de función n-ario $\{\,,\}$ como

$$\{x_1, \ldots, x_n\} = u :\Leftrightarrow \forall z(z \in u \leftrightarrow (z = x_1 \vee \ldots \vee z = x_n)).$$

Nota 19. Observamos que, análogamente a lo ocurrido para los pares no ordenados,

$$x_1 \cup \ldots \cup x_n = \bigcup\{x_1, \ldots, x_n\}.$$

Nota 20. Es importante observar, dado un conjunto x, la diferencia entre $\{x, \{x\}\}$ y $x \cup \{x\}$. El primero tiene dos elementos: x y $\{x\}$. El segundo, sin embargo, tiene como elementos todos los conjuntos que pertenezcan a x así como al propio x. Es decir, si por ejemplo $x = \{a, b\}$, entonces $\{x, \{x\}\} = \{\{a, b\}, \{\{a, b\}\}\}$ mientras que $x \cup \{x\} = \{a, b, x\}$.

Tal y como se ilustra en la nota anterior, con lo desarollado hasta ahora, es posible crear conjnutos de manera inductiva, por lo que se tiene la siguiente definición:

Definición 2.9. Dado un conjunto x, se dice que es **inductivo** si

$$\forall z(z \in x \to (z \cup \{z\}) \in x).$$

Se define el símbolo de relación unaria *ind* como

$$ind(x) :\Leftrightarrow \forall z(z \in x \to (z \cup \{z\}) \in x).$$

Definición 2.10. Al conjunto $x \cup \{x\}$ se llama **sucesor de** x y se denota por $x + 1$.

Nota 21. Podemos observar que el conjunto vacío es efectivamente inductivo, pues no tiene elementos.

Hasta ahora, eramos capaces de definir conjuntos que tuviesen un número finito de elementos, mediante las n-tuplas no ordenadas. Sin embargo, con los conjuntos inductivos parece que se puede tener al alcance la construcción de un conjunto infinito. Por ejemplo, se podría tomar un conjunto cuyos elementos sean $\emptyset, \emptyset \cup \{\emptyset\}, \ldots$, continuando con la unión de forma reiterada, de manera que el conjunto sea inductivo. Sin embargo, aunque parece intuitivo, no es posible probar con lo desarrollado hasta ahora que efectivamente este conjunto existe. Es por ello que se añade como un nuevo axioma.

V. Axioma del Infinito

$$\exists I(\emptyset \in I \wedge ind(I))$$

Este axioma asegura la existencia de un conjunto inductivo, no vacío, y conteniendo a $\emptyset$.

Definición 2.11. A todo conjunto verificando el Axioma del Infinito se le denomina **conjunto sucesor**. Es decir, un conjunto sucesor es un conjunto conteniendo a $\emptyset$ e inductivo.

Notemos que este axioma es realmente de gran importancia y arroja conclusiones más alla de la interpretación directa de la sentencia:
Algunos de los axiomas presentados hasta ahora, como el Axioma de Formación de Pares o el Axioma de la Unión, nos permiten formar nuevos conjuntos, pero requieren la existencia previa de algún conjunto para poder llevarlo a cabo. Sin embargo, solo dos axiomas nos garantizan la existencia de conjuntos: El Axioma del Conjunto Vacío y el presente Axioma del Infinito. Además, sin el Axioma del Infinito no es posible formar un conjunto infinito únicamente a partir de $\emptyset$. Este axioma nos asegura la esistencia de al menos un conjunto infinito, a partir del que podremos construir los demás conjuntos infinitos existentes en nuestra teoría.

Con lo visto hasta ahora, podemos comenzar a definir un concepto de suma importancia en matemáticas y sobre el que se sustentan numerosos resultados: los números naturales. Hasta ahora, ha sido utilizado de manera intuitiva, pero con los axiomas expuestos puede empezar a formalizarse. Es decir, podemos ver los números naturales como conjuntos.

Definición 2.12. Definimos los siguientes conjuntos:

$$\begin{aligned}
&0 := \emptyset \\
&1 := 0+1 = 0 \cup \{0\} = \{0\} \\
&2 := 1+1 = 1 \cup \{1\} = \{0,1\} \\
&3 := 2+1 = 2 \cup \{2\} = \{0,1,2\} \\
&\quad \dots
\end{aligned}$$

y así de forma sucesiva.

Así, el Axioma del Infinito garantiza la existencia de un conjunto I tal que $0 \in I$ e $ind(I)$, por lo que I contiene a $1, 2, 3, \ldots$. Está claro que cualquier conjunto inductivo conteniendo a $\emptyset$ como elemento (es decir, conteniendo al 0 como elemento) tiene a todos estos conjuntos como elementos. Por tanto, si existiese un conjunto cuyos únicos elementos fuesen $1, 2, 3, \ldots$, este estaría contenido en cualquier conjunto sucesor. Parece intuitivo pensar que en efecto este sería el conjunto que conocemos como los números naturales. En vista de este objetivo, se introducen los siguientes axiomas.

En primer lugar, vemos que es interesante formar el conjunto cuyos elementos son todos aquellos que satisfacen cierta propiedad, como por ejemplo el ser un conjunto inductivo. Sin embargo, ya se vio en la contextualización del trabajo que el formar este tipo de conjuntos es uno de los hechos que da lugar a paradojas. Es por ello que Zermelo decide solucionar este problema permitiendo seleccionar los conjuntos que cumplan cierta propiedad solo dentro de los elementos de un conjunto dado, dando lugar así al Esquema Axiomático de Separación.

VI. Esquema Axiomático de Separación

Dada una fórmula $\varphi(z)$ de $\mathscr{L}_{ST}$, entonces

$$\forall x \exists y \forall z (z \in y \leftrightarrow (z \in x \wedge \varphi(z))).$$

Dado un conjunto x y una fórmula φ, este axioma postula la existencia de un conjunto formado por todos los elementos de x que cumplan φ.

Definición 2.13. Dado un conjunto x y una fórmula $\varphi(z)$, se denota al conjunto dado por el Esquema Axiomático de Separación como $\{z \in x : \varphi(z)\}$, de forma que

$$\{z \in x : \varphi(z)\} = y :\Leftrightarrow \forall z (z \in y \leftrightarrow (z \in x \wedge \varphi(z))).$$

Nota 22. Efectivamente este conjunto es único, pues el axioma determina todos sus elementos.

Una primera aplicación de este axioma puede ser la formalización del concepto de intersección. Para ello, dados dos conjuntos x e y, tomamos z como variable libre y definimos $\varphi(z) \equiv z \in y$. Así, podemos considerar el conjunto $u := \{z \in x : \varphi(z)\}$. Tenemos que, por la instancia del Esquema Axiomático de Separación asociada a φ,

$$\forall z (z \in u \leftrightarrow (z \in x \wedge \varphi(z))), \text{ o equivalentemente } \forall z (z \in u \leftrightarrow (z \in x \wedge z \in y)).$$

Así, podemos definir el siguiente símbolo de función binario:

Definición 2.14. Se define la **intersección de dos conjuntos** x e y como

$$x \cap y = u :\Leftrightarrow \forall z(z \in u \leftrightarrow (z \in x \wedge z \in y).$$

Nota 23. Dados dos conjuntos x e y, $x \cap y \subseteq x$, y análogamente $x \cap y \subseteq y$.

De la misma forma, podemos definir la intersección de un único conjunto x. Este será el conjunto formado por los conjuntos que pertenezcan a todos los elementos de x. Así, considerando el conjunto $\bigcup x$ y la fórmula $\varphi(z) \equiv \forall w \in x(z \in w)$, podemos definir el símbolo de función 1-aria siguiente:

Definición 2.15. Se define la **intersección de un conjunto** x como

$$\bigcap x = u :\Leftrightarrow u = \left\{z \in \bigcup x : \varphi(z)\right\}$$

Es decir,

$$\forall z(z \in u \leftrightarrow (z \in \bigcup x \wedge \varphi(z))) \text{ o equivalentemente } \forall z(z \in u \leftrightarrow (z \in \bigcup x \wedge \forall w \in x(z \in w)))$$

Nota 24. Observemos que $\forall w(w \in x \rightarrow \bigcap x \subseteq w)$, pues $z \in \bigcap x \rightarrow \forall w \in x(z \in w)$.

Ejemplo 2.3. Si tenemos el conjunto $x = \{\{a,b\},\{b\}\}$, entonces $\bigcup x = \{a,b\}$ y $\bigcap x = \{b\}$ pues de los elementos de $\bigcup x$, b es el único que pertenece a los elementos de x, que son $\{a,b\}$ y $\{b\}$.

Otro concepto que podemos formalizar es la diferencia de dos conjuntos, que será un símbolo de función binario:

Definición 2.16. Dados dos conjuntos x e y, tomamos la fórmula $\varphi(z) \equiv z \notin y$ y definimos la **diferencia de x menos y** como

$$x \backslash y = u :\Leftrightarrow u = \{z \in x : \varphi(z)\}$$

Es decir,

$$\forall z(z \in u \leftrightarrow (z \in x \wedge \varphi(z)), \text{ o equivalentemente } \forall z(z \in u \leftrightarrow (z \in x \wedge z \notin y).$$

El Esquema Axiomático de Separación permite, dado un conjunto, seleccionar los elementos de este que cumplan cierta propiedad. Otro conjunto que parece intuitivo que exista es aquel que contenga a todos los subconjuntos de un conjunto dado, lo que comunmente conocemos como partes de un conjunto, pero cuya existencia no se puede garantizar hasta ahora. Es por ello que se introduce el siguiente axioma.

VII. Axioma de Partes de un Conjunto

$$\forall x \exists y \forall z (z \in y \leftrightarrow z \subseteq x)$$

Este axioma postula, dado un conjunto x, la existencia de otro conjunto cuyos elementos son los subconjuntos de x.

Definición 2.17. Se define el conjunto **partes de un conjunto** como

$$\mathscr{P}(x) = y :\Leftrightarrow \forall z (z \in y \leftrightarrow z \subseteq x).$$

Con este axioma, hemos visto los siete axiomas que Zermelo presentó en primer lugar. Esta teoría es conocida como la **teoría de Zermelo**, y se suele denotar como Z. Con ella, se pueden formalizar prácticamente todas las operaciones entre conjuntos que se habían desarrollado en su momento en la Teoría de Conjuntos Naive. También, como se había comentado, es posible ahora demostrar la existencia de los números naturales. Desarrollaremos a continuación los conceptos más importantes dentro de todos aquellos que se pueden derivar de estos axiomas. Para un desarrollo más extenso se puede consultar [Halbeisen y Krapf, 2020, Capítulo 13].

2.1.1. Los Números Naturales

Al introducir el Axioma del Infinito, intuíamos la naturaleza del conjunto de los números naturales. En base a esto, damos la siguiente definición.

Definición 2.18. El **conjunto de los números naturales** es un conjunto sucesor que está contenido en cualquier conjunto sucesor. Lo denotaremos como ω. Un elemento de ω se denominará **número natural**.

Teorema 2.1. *El conjunto de los números naturales existe.*

Demostración. Consideremos el conjunto sucesor I_0 cuya existencia es garantizada por el Axioma del Infinito. Por el Axioma de Partes, existe $\mathscr{P}(I_0)$. Así, si consideramos la variable libre x y la fórmula $\emptyset \in x \wedge ind(x)$, existe el conjunto

$$v = \{x \in \mathscr{P}(I_0) : \emptyset \in x \wedge ind(x)\}$$

Este es el conjunto de todos los conjuntos sucesores contenidos en I_0. Por tanto, definimos

$$\omega = \bigcap v.$$

Demostremos en pirmer lugar que es sucesor. Para ello, debemos ver que $\emptyset \in \omega$ y que $ind(\omega)$:

- Por definición de la intersección de un conjunto, $\forall z(z \in \omega \leftrightarrow (z \in \bigcup v \wedge \forall x \in v(z \in x)))$. Así, como $\emptyset \in \omega$, aplicando L_{13}, $\exists x_0 \in v(\emptyset \in x_0)$. Por tanto, tenemos que $\emptyset \in \bigcup v$. Además, $\forall x \in v(\emptyset \in x)$ por la definición de los elementos de v. De ambos hechos se concluye que $\emptyset \in \omega$.

- Sea $z \in w$. Queremos ver que $z+1 \in w$. Como $z \in \omega$, por L_{13}, $\exists x_0 \in v(z \in x_0)$. Entonces, por ser elemento de v, x_0 es inductivo, de donde $z+1 \in x_0$. Por tanto $z+1 \in \bigcup v$. Además, como $z \in \omega$, $\forall x \in v(z \in x)$. Así, como cualquier $x \in v$ es inductivo, $\forall x \in v(z+1 \in x)$. Por tanto, concluimos que $z+1 \in \omega$ y por tanto $ind(\omega)$.

Visto, esto, debemos demostrar que ω está contenido en cualquier conjunto sucesor. Tomemos I un conjunto sucesor. Vamos a probar que $\omega \cap I \in v$. Así, tendremos, por la Nota 24, que $\omega \subseteq \omega \cap I$. Además, como $\omega \cap I \subseteq \omega$, concluiríamos que $\omega = \omega \cap I$. Como $\omega \cap I \subseteq I$, se obtendría finalmente que $\omega \subseteq I$. Veamos que $\omega \cap I \in v$:

- En primer lugar, queremos ver que $\omega \cap I \in \mathscr{P}(I_0)$, lo que equivale a ver $\omega \cap I \subseteq I_0$. Tenemos que $\omega \cap I \subseteq \omega$. Además, claramente $I_0 \in v$, de donde $\omega \subseteq I_0$. Por tanto $\omega \cap I \subseteq I_0$.

- Como $\emptyset \in \omega$ y $\emptyset \in I$ por ser ambos sucesores, tenemos que $\emptyset \in \omega \cap I$.

- Dado $z \in \omega \cap I$, tenemos que $z \in \omega \wedge z \in I$. Por ser ambos inductivos, tenemos que $z+1 \in \omega \wedge z+1 \in I$, por lo que $z+1$ está en la intersección, de donde $ind(\omega \cap I)$.

□

Con este teorema hemos probado que el conjunto de los números naturales, tal y como lo hemos definido, existe. Veamos ahora que efectivamente este conjunto es aquel compuesto por $0, 1, 2, \ldots$.

Proposición 2.7. *Todo número natural es o bien el* 0 *o bien de la forma* $n+1$ *con* n *un número natural.*

Demostración. Sea m un número natural. Si $m = 0$, estaría probado. Supongamos por tanto que $m \neq 0$. Consideramos el conjunto

$$\mathscr{C} := \{z \in \omega : z \neq \emptyset \wedge \exists n \in \omega(z = n+1)\} \cup \{\emptyset\}.$$

Veamos que es sucesor. Así, por la definición de ω, tendremos $\omega \subseteq \mathscr{C}$. Por tanto, $m \in \mathscr{C}$ y como $m \neq \emptyset$, $\exists n \in w(m = n+1)$.

En primer lugar, vemos que $\emptyset \in \mathscr{C}$. Para ver que es inductivo, tomamos $z \in \mathscr{C}$ y distinguimos dos casos:

- Si $z = \emptyset$, entonces $z = 0$. Así, $0+1 \in \mathscr{C}$ pues, por ser ω inductivo y contener al 0, $0+1 \in \omega$. Además, $0+1 = \{\emptyset\} \neq \emptyset$ y claramente se cumple la útima condición.

- Si $z \neq \emptyset$, entonces $z \in \omega \wedge z \neq \emptyset \wedge \exists n \in \omega(z = n+1)$. De $z \in \omega$ tenemos que $z+1 \in \omega$. De $z \neq \emptyset$ tenemos que $z+1 \neq \emptyset$, pues $z \subseteq z+1$ y z tiene elementos. De nuevo, es claro que la última condición se cumple.

$\square$

2.1.2. Producto Cartesiano, Funciones y Relaciones

Comenzamos definiendo el producto cartesiano. Intuitivamente, dados dos conjuntos A y B, definiríamos el producto cartesiano como el conjunto de los $\langle x, y\rangle$ para los cuales $x \in A$ e $y \in B$. Sin embargo, ya sabemos que en un principio no podemos formar este conjunto, si no que, para aplicar el Axioma de Separación, necesitamos seleccionar los elementos de un conjunto dado.

Sean $x \in A$ e $y \in B$. Tenemos que $\langle x, y\rangle = \{\{x\}, \{x, y\}\}$. Así, $\{x\} \subseteq A$, por lo que $\{x\} \subseteq A \cup B$, y $\{x, y\} \subseteq A \cup B$. De esto se obtiene que $\{x\} \in \mathscr{P}(A \cup B)$ y $\{x, y\} \in \mathscr{P}(A \cup B)$. Por tanto, $\{\{x\}, \{x, y\}\} \subseteq \mathscr{P}(A \cup B)$, o equivalentemente $\{\{x\}, \{x, y\}\} \in \mathscr{P}\big(\mathscr{P}(A \cup B)\big)$. Así, podemos definir el siguente símbolo de función binario.

Definición 2.19. Se define el **producto cartesiano** de los conjuntos A y B como

$$A \times B \coloneqq \big\{z \in \mathscr{P}\big(\mathscr{P}(A \cup B)\big) : \exists x \exists y (x \in A \wedge y \in B \wedge z = \langle x, y\rangle)\big\}.$$

Para abreviar, se suele escribir

$$A \times B \coloneqq \big\{\langle x, y\rangle : x \in A \wedge y \in B\big\}.$$

Una vez definido este concepto, podemos pasar a definir las funciones entre dos conjuntos. De forma intuitiva, una función entre dos conjuntos A y B asigna a cada elemento de A un elemento de B. Esto se hace mediante subconjuntos de $A \times B$, de forma que si tenemos $\langle x, y\rangle$, la función llevaría x a y.

Definición 2.20. Dados dos conjuntos A y B, se dice que un subconjunto f de $A \times B$ es una **función** de A en B si

$$\forall x \in A \exists ! y \in B(\langle x, y\rangle \in f).$$

Al conjunto de todas las funciones de A en B se lo denota ${}^A B$. Tenemos que

$${}^A B := \{f \in \mathscr{P}(A \times B) : \forall x \in A \exists ! y \in B(\langle x, y\rangle \in f)\}.$$

Una función $f \in {}^A B$ se denota $f : A \to B$. Usualmente, se escribe $f(x) = y$ en vez de $\langle x, y\rangle \in f$. Se dice que y es la **imagen por** f **de** x. Al conjunto A se le denomina **dominio** de f, representado por $dom(f)$ y al conjunto $f[A] := \{y \in B : \exists x \in A(f(x) = y)\}$, **rango o imagen** de f.

Un caso particular de conjunto de funciones es cuando se toma ${}^n A$, con A un conjunto cualquiera y n un número natural distinto de 0 y 1. En este caso, tenemos

$${}^n A := \{f \in \mathscr{P}(n \times A) : \forall x \in n \exists ! y \in A(f(x) = y)\}.$$

Así, recordando que existe un número natural m tal que $n = m+1$, tenemos que $n = \{0, 1, 2, \ldots, m\}$. Por tanto, cada función de ${}^n A$ asignaría a cada natural desde 0 hasta m un elemento de A. Podemos identificar esta función, que es un subconjunto de $n \times A$ de la forma $\{\langle 0, x_0\rangle, \ldots, \langle m, x_m\rangle\}$, únicamente con los elementos de A a los que asocia cada natural, imitando así el concepto de par ordenado.

Definición 2.21. Dado un natural n distinto de 0 y 1, y $x_0, \ldots, x_m$ n elementos de A, con $n = m+1$, se define la **n-tupla ordenada** $\langle x_0, \ldots, x_m\rangle$ como

$$\langle x_0, \ldots, x_m\rangle := \{f \in {}^n A : f(0) = x_0 \wedge \ldots f(m) = x_m\}$$

El conjunto ${}^n A$ se denota comunmente por A^n, reflejando el hecho de que los elementos de A^n son todas las posibles n-tuplas ordenadas formadas por elementos de A.

Nota 25. Observamos que si $n = 2$ podemos identificar A^2 con el producto cartesiano $A \times A$.

Una vez definido el conjunto A^n, podemos pasar a formalizar el concepto de relación entre los elementos de un conjunto.

Definición 2.22. Dados un conjunto cualquiera A y un número natural n, cualquier subconjunto $R \subseteq A^n$ se denomina **relación n-aria** en A. Si $n = 2$, se denomina **relación binaria**, y se escribe xRy para representar que $\langle x, y\rangle \in R$.

Algunos tipos de relaciones binarias importantes son los siguientes:

Definición 2.23. Una relación binaria R en un conjunto A se dice que es **de orden** si cumple las siguientes propiedades:

- para todo $x \in A$ se tiene que xRx (**propiedad reflexiva**).
- si $x \in A$ e $y \in A$ son tales que $xRy \wedge yRx$ entonces $x = y$ (**propiedad antisimétrica**).
- si $x \in A$, $y \in A$ y $z \in A$ son tales que $xRy \wedge yRz$ entonces xRz (**propiedad transitiva**).

Se dice que (A, R) es un **conjunto ordenado**.

Definición 2.24. Una relación de orden R en un conjunto A se denomina **orden total** en A, y se dice que (A, R) es un **conjunto totalmente ordenado**, si dados dos elementos $x \in A$ e $y \in A$, se tiene que xRy o yRx. Por la propiedad antisimétrica, esta condición es equivalente a que, dados $x \in A$ y $y \in A$,

$$(x \neq y \wedge xRy) \rightarrow \neg(yRx).$$

Definición 2.25. Una relación de orden R en A se denomina **buen orden** en A si para todo conjunto no vacío $S \subseteq A$, existe $x_0 \in A$ de forma que para cualquier elemento $y \in A$ se tiene que x_0Ry. Se dice que x_0 es un **mínimo o primer elemento** del conjunto S, y que (A, R) es un **conjunto bien ordenado**.

Nota 26. Como consecuencia de que todo conjunto no vacío tenga mínimo, todo buen orden es orden total.

Nota 27. Notemos que, por ser R un orden total, el mínimo es único. Además, también por ser un orden total, se tiene que, si x_0 es el mínimo de $S \subseteq A$, con (A, R) un conjunto bien ordenado, entonces

$$\forall y(y \in S \wedge y \neq x_0 \rightarrow \neg(yRx_0)).$$

Al igual que hemos considerado el producto cartesiano de dos conjuntos, así como el conjunto A^n, que puede entenderse como el producto cartesiano de A n veces, nos planteamos la forma de poder construir el producto cartesiano de una colección arbitraria de conjuntos.

La idea intuitiva consiste en replicar lo hecho para A^n. Así, dado un conjunto I, querríamos asignar a cada elemento $i \in I$ un conjunto A_i, y definiendo A como el conjunto dado por la unión de los elementos estos A_i, definir este producto cartesiano como el conjunto de los elementos f de ${}^{I}A$ tales que para cada $i \in I,\ f(i) \in A_i$.

Sin embargo, aunque parezca intuitivo que se pueda considerar el producto cartesiano de una colección arbitraria de conjuntos, no podemos asegurar la existencia del conjunto A antes mencionado, pues este sería un conjunto de la forma $\bigcup B$ con B un conjunto cuyos elementos fueran los A_i, cuya existencia no es conocida. Esto motiva la introducción del siguiente axioma, dentro de los presentados por Fraenkel y Skolem y que completan la teoría de Zermelo.

2.2. Los Axiomas de Zermelo-Fraenkel y el Axioma de Elección

Antes de introducir el próximo axioma, necesitamos la siguiente definición.

Definición 2.26. Dada una fórmula φ con dos variables libres, se dice que esta **determina una función** si cumple que $\forall x \forall z \big(\exists y \varphi(x,y) \wedge (\varphi(x,z) \rightarrow y = z)\big)$.

VIII. Esquema Axiomático de Reemplazamiento

Dada una fórmula φ que determina una función, se tiene que

$$\forall A \exists B \forall y (y \in B \leftrightarrow \exists x \in A \varphi(x,y)).$$

Este axioma postula que si tenemos una fórmula φ tal que para cualquier conjunto x existe un único conjunto y tal que $\varphi(x,y)$, entonces, para cualquier conjunto A, existe un conjunto consistiendo en aquellos y tales que $\varphi(x,y)$ para algún $x \in A$. Intuitivamente, si tenemos un conjunto A y una fórmula que relaciona cada elemento x de A con otro elemento y, existe un conjunto obtenido de reemplazar cada $x \in A$ por su y correspondiente.

Veamos que, en efecto, este axioma nos permite definir el producto cartesiano deseado.

Definición 2.27. Dada una fórmula φ que determina una función, definimos el símbolo de función 1-ario

$$F(x) = y :\Leftrightarrow \varphi(x,y),$$

que se denomina **función de clase**. Entonces, escribimos el conjunto dado por el Esquema Axiomático de reemplazamiento aplicado a φ como $F[A]$ o como $\{F(x) : x \in A\}$, de forma que

$$F[A] = u :\Leftrightarrow \forall y (y \in u \leftrightarrow \exists x \in A (F(x) = y)).$$

Nota 28. Observemos que si tenemos un conjunto A y una fórmula tal que para todo $x \in A$ se cumple que $\exists y \varphi(x,y) \wedge \big(\forall z(\varphi(x,z) \to y = z)\big)$, podemos definir la fórmula

$$\psi \equiv (x \in A \wedge \varphi(x,y)) \vee (x \notin A \wedge y = x),$$

que claramente define una función. Así, podemos aplicar el Esquema Axiomático de Reemplazamiento a ψ y a A, obteniendo el conjunto $\{F(x) : x \in A\}$ donde F es la función de clase asociada a ψ. Como para cualquier elemento de A ψ coincide con φ, se pude suponer que $F(x) = y \Leftrightarrow \varphi(x,y)$. Así, para aplicar este esquema axiomático, solo es necesario realmente que para todo $x \in A$ exista un único elemento y cumpliendo $\varphi(x,y)$. En este caso diremos que F es una **función de clase definida en o sobre** A.

Visto esto, podemos definir el producto cartesiano de una colección arbitraria de conjuntos.

Definición 2.28. Sea F una función de clase e I un conjunto arbitrario. Para cada $i \in I$ definimos $A_i \coloneqq F(i)$ y sea $A \coloneqq \bigcup F[I]$, donde tenemos que $F[I] = \{A_i : i \in I\}$. Se define el **producto cartesiano de los conjuntos** A_i**,** $i \in I$ como

$$\bigtimes_{i \in I} A_i \coloneqq \big\{f \in {}^{I}A : \forall i \in I(f(i) \in A_i)\big\}.$$

Nota 29. Notar que tomamos las funciones en ${}^{I}A$ pues estas llevan a cada elemento de i a un elemento de A. Como los elementos de A son precisamente los elementos de alguno de los A_i, una función $f \in \bigtimes_{i \in I} A_i$ llevará a cada elemento de I, que se puede entender como un índice, a un elemento de A_i. Esto generaliza los conceptos vistos hasta ahora.

Observamos que una consecuencia importante de este axioma es que, aunque no se nos de una función como tal, a posteriori podemos definir una función entre I y $F[I]$.

Nota 30. Un hecho a tener en cuenta es que con lo desarrollado hasta ahora no se puede probar que el producto cartesiano de conjuntos no vacios sea no vacío.

El último axioma dentro de la teoría de Zermelo-Fraenkel, aunque quizás menos intuitivo, se introduce para evitar ciertas paradojas. Más adelante se verán otras utilidades, como que otorga a la relación $\in$ propiedades de gran importancia.

IX. Axioma de Fundamento

$$\forall x(x \neq \emptyset \to \exists y \in x(x \cap y = \emptyset))$$

Este axioma postula que para un conjunto x, existe un elemento y de x tal que su intersección con x es el conjunto vacío.

Algunas de las implicaciones de este axioma son las siguientes:

Proposición 2.8. *Para cualquier conjunto x, $x \notin x$.*

Demostración. Dado un conjunto x, consideramos el conjunto $\{x\}$, cuyo único elemento es x. Por tanto, por el Axioma de Fundamento, $x \cap \{x\} = \emptyset$. Es decir, este conjunto no tiene elementos, por lo que x no puede pertenecer a él. Así, no se pueden dar simultáneamente $x \in x$ y $x \in \{x\}$. Como lo segundo es cierto, se obtiene que $x \notin x$. □

Proposición 2.9. *No pueden existir conjuntos $x_0, \ldots, x_n$ cumpliendo que $x_0 \in x_1 \in \ldots \in x_n \in x_0$.*

Demostración. Supongamos que existen dichos conjuntos. Entonces podemos considerar el conjunto $\{x_0, \ldots, x_n\}$. Por el Axioma de Fundamento,

$$\exists z \in \{x_0, \ldots, x_n\}\big(z \cap \{x_0, \ldots, x_n\} = \emptyset\big).$$

Como $z \in \{x_0, \ldots, x_n\}$, z ha de ser igual a uno de estos conjuntos. Por hipótesis, existe $x_k \in \{x_0, \ldots, x_n\}$ de forma que $x_k \in z$. Así, $x_k \in z \cap \{x_0, \ldots, x_n\}$, lo que sería una contradicción con el axioma. Por tanto, concluimos que no pueden existir dichos conjuntos. □

Una consecuencia inmediata de esta proposición es la siguiente:

Proposición 2.10. *Dados dos conjuntos x e y, no puede ocurrir que $x \in y$ e $y \in x$.*

Proposición 2.11. *No existe una sucesión infinita de la forma $x_0 \ni x_1 \ni x_2 \ni \ldots$.*

Demostración. Supongamos que existe una sucesión de esta forma, es decir, exsiste un conjunto x_0 e, inductivamente, $\forall n \in \omega \exists (y \in x_n)$, que denotamos x_{n+1}.

Si existiese un conjunto x cuyos elementos fuesen los x_n para todo $n \in \omega$, este incumpliría el Axioma de Fundamento, pues para todo conjunto x_n, $x_{n+1} \in x_n \cap x$, de donde $x_n \cap x \neq \emptyset$. Es decir, no existiría un elemento de x de forma que su intersección con x fuera vacía. Veamos que dicho conjunto existe.

Consideramos la fórmula $\varphi(n, y) \equiv y = x_n$, que vemos que determina una función sobre ω, pues para cada número natural, existe un único conjunto que la cumple. Aplicando el Esquema Axiomático de Reemplazamiento a φ y al conjunto ω, existe un conjunto x tal que

$$\forall y(y \in x \leftrightarrow \exists n \in \omega(y = x_n)).$$

Vemos que efectivamente este conjunto es el buscado, pues sus elementos son los x_n para todo $n \in \omega$. $\square$

Con estos 9 axiomas se concluye la **teoría de Zermelo-Fraenkel**, que denotaremos ZF. Sin embargo a estos axiomas se suele añadir el denominado Axioma de Elección, que presentaremos a continuación. Este ayuda a resolver cuestiones como la planteada en la Nota 30, aunque el hecho de su inclusión o no es debatido por algunos matemáticos. A la teoría que surge de añadir el Axioma de Elección a ZF la denotaremos por ZFC.

X. Axioma de Elección

$$\forall \mathcal{F} \left(\mathcal{F} \neq \emptyset \wedge \emptyset \notin \mathcal{F} \rightarrow \exists f \in {}^{\mathcal{F}}\!\left(\bigcup \mathcal{F}\right) (\forall x \in \mathcal{F} \, (f(x) \in x)) \right)$$

Este axioma postula que, dado un conjunto $\mathcal{F}$ no vacío y cuyos elementos son no vacíos, existe una función de $\mathcal{F}$ a $\bigcup \mathcal{F}$ tal que lleva a cada $x \in \mathcal{F}$ a un elemento de x. Es decir, existe una función que a cada $x \in \mathcal{F}$ le asigna uno de sus elementos. Como cada x solo tiene una imagen, intuitivamente este axioma postula que dado un conjunto $\mathcal{F}$, se puede elegir un único elemento dentro de cada uno de sus elementos y formar un conjunto con ellos.

Definición 2.29. Sea $\mathcal{F}$ un conjunto cuyos elementos son no vacíos. Se denomina **función de elección de** $\mathcal{F}$ a una función $f : \mathcal{F} \rightarrow \bigcup \mathcal{F}$ tal que $f(x) \in x$ para todo $x \in \mathcal{F}$.

Vista esta definición, vemos que el Axioma de Elección se puede reformular de la siguiente manera:

Todo conjunto cuyos elementos son no vacíos posee una función de elección.

Además de este enunciado, existe otra forma equivalente de presentar el Axioma de Elección:

Sea $\mathcal{F}$ un conjunto de conjuntos no vacíos y disjuntos dos a dos. Entonces existe un conjunto M que consiste en precisamente un elemento de cada miembro de $\mathcal{F}$. El conjunto M se llama ***conjunto de elección para*** *$\mathcal{F}$.*

Veamos que, en efecto, son equivalentes:

Teorema 2.2. *Los siguientes enunciados son equivalentes:*

(i) Todo conjunto cuyos elementos son no vacíos posee una función de elección.

(ii) Todo conjunto cuyos elementos son conjuntos no vacíos y disjuntos dos a dos posee un conjunto de elección.

Demostración. Probemos en primer lugar que, si suponemos cierta (i), obtenemos (ii).
Sea $\mathcal{F}$ un conjunto cuyos elementos son conjuntos no vacíos y disjuntos. Por (i), al ser los elementos de $\mathcal{F}$ no vacíos, existe una función de elección f. Sea $M := f[\mathcal{F}] = \{f(x) : x \in \mathcal{F}\}$. Como $f(x) \in x$ para todo $x \in \mathcal{F}$, tenemos que M es un conjunto de elección para $\mathcal{F}$.

Supongamos ahora que (ii) es cierta y veamos que se da entonces (i).
Para ello, tomamos $\mathcal{F}$ un conjunto cuyos elementos son no vacíos. Para cada $x \in \mathcal{F}$, definimos $x^* = x \times \{x\}$. Como la fórmula $y = x^*$ es tal que para cada $x \in \mathcal{F}$ existe un único y cumpliéndola, por el Esquema Axiomático de Reemplazamiento y por la Nota 28, podemos tomar el conjunto

$$\mathcal{F}^* := \{x^* : x \in \mathcal{F}\}.$$

Es claro que los elementos de $\mathcal{F}^*$ son no vacíos, por ser no vacíos los elementos de $\mathcal{F}$: dado $x^* \in \mathcal{F}^*$, existe $x \in \mathcal{F}$ tal que $x^* = x \times \{x\}$. Como x es no vacío, $\exists y \in x$. Entonces, tenemos que $\langle y, x\rangle \in x \times \{x\} = x^*$.
Además, se tiene también que los elementos de este conjunto son disjuntos dos a dos: sean x_1^* y x_2^* elementos de $\mathcal{F}^*$ de forma que $x_1^* \neq x_2^*$. Entonces, existen x_1 y x_2 elementos de $\mathcal{F}$ tales que

$$x_1^* = x_1 \times \{x_1\} \text{ y } x_2^* = x_2 \times \{x_2\}.$$

Es claro que $x_1 \neq x_2$, pues si fuesen iguales también lo serían x_1^* y x_2^*.
Así, si suponemos que $x_1^* \cap x_2^* \neq \emptyset$, tenemos que $\exists z \in x_1^* \cap x_2^*$. Por tanto, tenemos por un lado que $z \in x_1^* = x_1 \times \{x_1\}$, de donde $z = \langle a, x_1\rangle$, con $a \in x_1$. Por el otro lado, de $z \in x_2^* = x_2 \times \{x_2\}$ obtenemos que $z = \langle b, x_1\rangle$, con $b \in x_2$. Por tanto, $\langle a, x_1\rangle = \langle b, x_1\rangle$, lo que implica que $x_1 = x_2$. Por tanto, llegamos a una contradicción al suponer que $x_1^* \cap x_2^* \neq \emptyset$, de donde $x_1^* \cap x_2^* = \emptyset$.

Así, podemos aplicar (ii) al conjunto $\mathcal{F}^*$, por lo que ese posee un conjunto de elección $\mathcal{M}$. Este conjunto estará compuesto por uno y solo un elemento de cada $x^* \in \mathcal{F}^*$. Así, dado $x \in \mathcal{F}$, existe un único elemento en $\mathcal{M}$ perteneciendo a $x^* = x \times \{x\}$. Este elemento será de la forma $\langle m_x, x\rangle$, con $m_x \in x$. Observemos que entonces $m_x \in \bigcup \mathcal{F}$. Por tanto, aplicando el Esquema Axiomático de Reemplazamiento al conjunto $\mathcal{F}$ y a la fórmula $y = \langle x, m_x\rangle$, podemos considerar el conjunto

$$f := \{\langle x, m_x\rangle : x \in \mathcal{F}\},$$

que es un subconjunto de $\mathcal{F} \times \bigcup \mathcal{F}$ cumpliendo que cada $x \in \mathcal{F}$ tiene una única imágen. Por consiguiente, f es una función de $\mathcal{F}$ en $\bigcup \mathcal{F}$. Además, está claro que $\forall x \in \mathcal{F},\ f(x) \in x$, de donde conluimos que f es una función de elección de $\mathcal{F}$. $\square$

En la introducción de este axioma, se comentó que una de sus utilidades era el resolver la cuestión sobre el producto cartesiano de conjuntos no vacíos. Veamos que en efecto esto es cierto:

Proposición 2.12. *En ZFC, el producto cartesiano de conjuntos no vacíos es un conjunto no vacío.*

Demostración. Sean I un conjunto no vacío cualquiera, F una función de clase cualquiera y A_i y A los conjuntos dados en la Definición 2.28. Supongamos que los A_i son no vacíos. Entonces, aplicando el Axioma de Elección a $F[I]$, que es no vacío y contiene elementos no vacíos,

$$\exists f : F[I] \to \bigcup F[I] \text{ cumpliendo que } \forall A_i \in F[I](f(A_i) \in A_i).$$

Aplicando el Esquema Axiomático de Reemplazamiento al conjunto I y a la fórmula $y = \langle i, f(A_i)\rangle$, existe un conjunto formado por los pares ordenados $\langle i, f(A_i)\rangle$ para todo $i \in I$. Lo denotamos $\{\langle i, f(A_i)\rangle : i \in I\}$. Vemos que efectivamente podemos aplicar este axioma, pues para cada $i \in I$ existe un único y cumpliendo la fórmula.

Tenemos entonces que

$$\{\langle i, f(A_i)\rangle : i \in I\} \in \bigtimes_{i \in I} A_i,$$

pues es un conjunto de ${}^I A$ cumpliendo que para todo $i \in I$ la imagen de i está en A_i. $\square$

2.2.1. La Consistencia del Axioma de Elección

El Axioma de Elección, aunque de apariencia intuitiva, tiene una naturaleza muy distinta al resto de axiomas. Aunque garantiza la existencia de una función de elección, no da en ningún momento herramientas para construirla. Es decir, la existencia de una función de elección es un hecho distinto de la posiblidad real de que una puede ser construida. Es por esto que la inclusión de este axioma es debatida por algunos matemáticos.

En 1940, Gödel demostró el siguiente resultado de suma importancia:

Teorema 2.3. *Si $Con(ZF)$, entonces $Con(ZFC)$.*

Es decir, supuesta la consistencia de la Teoría de Zermelo-Fraenkel, asumir el Axioma de Elección no produce nuevas contradicciones. La consistencia de una teoría habiendo asumido como consistente otra se denomina **consistencia relativa**.

En 1959, el matemático Paul Cohen demostró la consistencia relativa de la negación del Axioma de Elección, demostrando así que este axioma es independiente del resto de axiomas de ZF. Por tanto, la decisión de incluir o no este axioma no es más que de caracter personal.

Sin embargo, se conocen numerosos resultados de gran importancia cuya prueba no sería posible sin el Axioma de Elección. Debido a su caracter intuitivo y a la forma en la que facilita la teoría matemática, este axioma es generalmente aceptado como cierto.

Algunos de los enunciados equivalentes al Axioma de Elección, además del enunciado alternativo visto anteriormente, son los siguientes:

- Teorema de Buen Orden: *Todo conjunto puede ser bien ordenado.*
- Lema de Zorn: *Sea un conjunto parcialmente ordenado en el que toda cadena tiene cota superior. Entonces este conjunto contiene al menos un elemento maximal.*
- *Todo espacio vectorial posee una base.*

También es imprescindible para demostrar resultados como que *todo anillo tiene ideal maximal* o que *todo producto de una familia de espacios topológicos compactos es compacto*, conocido como Teorema de Tychonoff.

Para una demostración de la equivalencia de estos resultados, así como para conocer otros enunciado equivalentes, puede consultarse [Devlin, 1994, Sección 2.7].

3 | Los Números Ordinales y la Completitud y Consistencia de ZF

Dentro de su trabajo en Teoría de Conjuntos, Cantor desarrolla un concepto de suma importancia: los números ordinales. Estos surgen de generalizar y abstraer los números naturales, reflejando la capacidad que estos tienen para enumerar u ordenar conjuntos y extendiendo este orden. Aunque inicialmente presentados dentro de la Teoría de Conjuntos Naive, fueron trasladados por Von Neumann a la Teoría Axiomática en su artículo *"Sobre la introducción de los números transfinitos"*, publicado en 1923.

En este capítulo desarrollaremos los conceptos básicos sobre los números ordinales y presentaremos su aritmética. Además, veremos como esta nos permite obtener conclusiones acerca de si la Teoría de Zermelo-Fraenkel es completa o consistente. Para ello se verá que ω es un modelo de la Aritmética de Peano y se aplicarán los Teoremas de Incompletitud de Gödel. A la par, se verán resultados de gran relevancia y alcance: la Recurrencia y la Inducción Transfinitas.

En la Sección 3.1 se ha desarrollado una teoría que unifica la visión dada en [Devlin, 1994] y [Halbeisen, 2012]. Se ha elaborado un razonamiento propio que motiva la definición de ordinal dada en el primer texto y se ha derivado de esta la dada en el segundo, presentando después resultados seleccionados de ambos, junto con algunos expuestos en [Suppes, 1960]. Las demostraciones, aunque algunas de ellas esbozadas en las referencias, se han realizado de forma original, presentando en algunos casos alternativas a las propuestas, con el objetivo de enmarcar las pruebas en el contexto del trabajo. La Sección 3.2 es una aplicación directa de los Teoremas de Incompletitud, aunque la demostración de que ω consituye un modelo de PA es propia.

3.1. Los Números Ordinales

Como se comentaba en la introducción a este capítulo, hasta ahora hemos construido los números naturales: $0, 1, 2, 3, \ldots$. Para ello, hemos utilizado el concepto de conjunto sucesor, de forma que tenemos 0, $0 + 1 = 1$, $1 + 1 = 2$, continuando de forma consecutiva. Con todos estos conjuntos formamos ω. Sin embargo, ahora podríamos proseguir, realizando $\omega + 1, (\omega + 1) + 1, \ldots$, extendiendo así los números naturales y generando conjuntos que se suceden de forma ordenada. Parece intuitivo pensar entonces que, al igual que podemos definir los naturales, exista un concepto que pueda englobar a todos estos conjuntos. Para entenderlo, pensaremos qué propiedades caracterizan a los naturales y trataremos de abstraerlas.

Por construcción, tenemos que

$$0 \in 1 \quad 0,1 \in 2 \quad 0,1,2 \in 3 \quad \ldots$$

Es decir, de alguna forma la relación de pertenencia "ordena" a los números naturales de la siguiente forma:

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

Además, si tomamos un $n \in \omega$,

$$n = \{0, 1, \ldots, n-1\},$$

donde a partir de ahora $n-1$ denotará al elemento del cual n es sucesor. Es decir, n es el conjunto formado por aquellos ordinales anteriores a n en esta ordenación que nos proporcionaría la relación de pertenencia.

Para abstraer esto, podemos tomar conjuntos ordenados en los que cada elemento esté formado por aquellos conjuntos relacionados con él. Además, hay otras propiedades características de ω que harían que tomar únicamente un conjunto ordenado no se asemejase del todo a lo buscado. En primer lugar, vemos que no solo hay una ordenación de los naturales, sino que todos están relacionados entre sí. Más aún, encontramos un elemento, el 0, que pertenece a todos los demás conjuntos. Todas estas características son las que motivan las siguientes definiciones.

Definición 3.1. Sea un conjunto ordenado $(X, \leq)$. Dados b y c elementos de X, escribamos $b < c :\Leftrightarrow b \leq c \wedge b \neq c$. Entonces, dado $a \in X$, se define el **segmento** X_a como

$$X_a \coloneqq \{x \in X : x < a\}.$$

Definición 3.2. Un **ordinal**, también llamado **número ordinal**, es un conjunto bien ordenado tal que $\forall a \in X, a = X_a$.

Consideremos entonces un ordinal $(X, \leq)$ con $X \neq \emptyset$. ¿Cómo son sus elementos? Como $X \subseteq X$ y este es no vacío, existe un mínimo, x_0, que cumple por definición que $\forall y \in X (x_0 \leq y)$. Además, si

$x \neq y$, por ser $\leq$ un orden total tenemos que $\neg(y \leq x_0)$, de donde $\neg(y < x_0)$. Claramente, tampoco se tiene que $x_0 < x_0$. Es decir, no existe ningún elemento y de X tal que $y < x_0$, de donde

$$x_0 = X_{x_0} = \{y \in X : y < x_0\} = \emptyset = 0.$$

Así, hemos visto que el mínimo de un ordinal ha de ser siempre el 0. De la misma forma, por ser $(X, \leq)$ totalmente ordenado, existirá un elemento x_1 de forma que $x_0 < x_1$ y $x_1 < y$ para todo $y \in X$ con $y \neq x_0$ e $y \neq x_1$. Así,

$$x_1 = X_{x_1} = \{y \in X : y < x_1\} = \{x_0\} = \{0\} = 1.$$

De forma consecutiva, dado un natural n, si suponemos que X tiene más de $n+1$ elementos, el (n+1)-ésimo elemento de X, será

$$x_n = X_{x_n} = \{y \in X : y < x_n\} = \{0, 1, \ldots, n-1\} = n.$$

De esta forma, hemos visto que los primeros elementos de un ordinal no pueden ser conjuntos cualesquiera, sino que son naturales. Pasamos ahora a preguntarnos cómo ha de ser la relación $\leq$ y si esta tiene también alguna restricción.

Proposición 3.1. *Sean $(X, \leq)$ un ordinal y a, b dos elementos de X. Entonces*

$$a \leq b \leftrightarrow X_a \subseteq X_b \text{ o equivalentemente } a \leq b \leftrightarrow a \subseteq b.$$

Demostración. Supongamos en primer lugar que $a \leq b$ y sea $y \in X_a$. Entonces, tenemos que $y \in X_a \leftrightarrow (y \in X \land y < a)$. Así, como $a \leq b$, tendríamos o bien $a = b$ en cuyo caso $X_a = X_b$ y por tanto $X_a \subseteq X_b$, o bien $a < b$. En este caso, por la propiedad transitiva de $\leq$ tenemos que $y \leq b$. Además, no se puede dar $y = b$, pues esto implicaría que se tienen a la vez $a \leq b$ y $b \leq a$, de donde, por la propiedad antisimétrica, $a = b$, lo que hemos supuesto que no ocurre. Por tanto, se tiene que $y < b$, de donde $y \in X_a \to (y \in X \land y < b)$ y por tanto $y \in X_b$, de donde se concluye también que $X_a \subseteq X_b$.

Supongamos ahora que $X_a \subseteq X_b$. Esto quiere decir que $\forall y(y \in X_a \to y \in X_b)$. Para probar que $a \leq b$, veremos que no se da $b < a$, por lo que, al ser un orden total, tendrá que darse $a \leq b$. Así, si $b < a$, como $b \in X$, tendríamos que $b \in X_a$ y por tanto $b \in X_b$, de donde $b < b$. Pero esto implicaría que $b \neq b$, lo que es una contradicción.

La segunda equivalencia se deduce de $X_a = a$ y $X_b = b$. $\square$

Así, hemos visto que la relación de orden no puede ser cualquiera, si no que es precisamente el contenido. Sin embargo, teniendo en cuenta el comportamiento de los naturales, resulta interesante estudiar el papel de $\in$ en esta relación de orden. Es por ello que introducimos la siguiente definición.

Definición 3.3. Sea X un conjunto. Definimos la relación binaria $R_\in$ como

$$R_\in \coloneqq \big\{\langle a, b\rangle \in X \times X : a \in b \vee a = b\big\}.$$

Nota 31. Es necesario definir $R_\in$, pues $\in$ por si misma no podría ser nunca una relación de orden, ya que por el Axioma de Fundamento no se cumpliría la propiedad reflexiva.

Veamos que, como cabía esperar, para los ordinales el contenido es equivalente a $R_\in$:

Proposición 3.2. *Sean* $(X, \subseteq)$ *un ordinal y* a, b *elementos de* X*. Entonces*

$$a \subseteq b \leftrightarrow aR_\in b.$$

Demostración. Comencemos suponiendo que $a \subseteq b$, donde $b = X_b = \{y \in X : y \subseteq b \wedge y \neq b\}$. Entonces, tenemos que, o bien $a = b$, en cuyo caso $aR_\in b$ o bien $a \neq b$, lo que junto a $a \subseteq b$ y $a \in X$ hace que $a \in b$ y por tanto $aR_\in b$.
Supongamos ahora que $aR_\in b$. Entonces o bien $a = b$, de donde $a \subseteq b$, o bien $a \in b$, lo que por la definición de b implica que $a \subseteq b$. □

En conclusión, hemos visto que un ordinal X es un conjunto bien ordenado por la relación $R_\in$, lo que comunmente se expresa como bien ordenado por $\in$, y tal que cada elemento a es igual al segmento X_a. Notemos que podemos referirnos al ordinal simplemente como X pues la relación de orden viene especificada implícitamente. Sin embargo, ahora que hemos visto la estrecha relación que hay entre la pertenencia y el contenido en este tipo de conjuntos, la condición de los segmentos suele ser reemplazada por la siguiente propiedad:

Definición 3.4. Un conjunto x se dice **transitivo** si cumple que

$$\forall y(y \in x \to y \subseteq x) \text{ o equivalentemente } \forall y \forall z(z \in y \wedge y \in x \to z \in x).$$

Vemos que el hecho de que un conjunto sea transitivo refleja bien el comportamiento derivado de que los elementos sean segmentos. Por ejemplo, si tomamos el número natural $6 = \{0, 1, 2, 3, 4, 5\}$, vemos que $5 \in 6$. Pero $5 = \{0, 1, 2, 3, 4\}$, por lo que también se tiene que $5 \subseteq 6$.

Vamos a comprobar que efectivamente estas dos definiciones de ordinal son equivalentes. Antes de ello, veamos dos propiedades sobre la relación $R_\in$, que vienen dadas por el Axioma de Fundamento, y que nos facilitarán las demostraciones posteriores.

Proposición 3.3. *Sea $(X, R_\in)$ un conjunto totalmente ordenado. Entonces está bien ordenado.*

Demostración. Sea S un subconjunto no vacío de X y supongamos que no tiene mínimo. Como es no vacío, podemos tomar un elemento $x_0 \in S$. Si para todo $y \in S$ se tiene que $x_0 R_\in y$, entonces x_0 sería mínimo. Por tanto, existe $x_1 \in S$ que no cumple esto, por lo que, al ser totalmente ordenado, $x_1 R_\in x_0$ y además $x_1 \neq x_0$, de donde $x_1 \in x_0$. Aplicando el mismo razonamiento a x_1, encontraríamos $x_2 \in S$ tal que $x_2 \in x_1 \in x_0$. Así, podríamos construir una sucesión de la forma $x_0 \ni x_1 \ni x_2 \ni \ldots$, lo que sabemos que es imposible por el Axioma de Fundamento. □

Proposición 3.4. *Un conjunto X está bien ordenado por $\in$ si y solo si para cualesquiera a, b elementos de X, o bien $a \in b$ o bien $b \in a$ o bien $a = b$.*

Demostración. Supongamos que $(X, R_\in)$ está bien ordenado. Entonces, dados a, b elementos de X, se tiene que $a R_\in b$ o $b R_\in a$. Esto es equivalente a que o bien $a \in b$ o bien $b \in a$ o bien $a = b$, pues por el Axioma de Fundamento los tres casos son excluyentes.

Supongamos ahora que se da que para cualesquiera a, b elementos de X, o bien $a \in b$ o bien $b \in a$ o bien $a = b$. Demostraremos que $R_\in$ es una relación de orden en X. Visto esto, claramente será de orden total, pues la hipótesis es equivalente a que se de $a R_\in b$ o $b R_\in a$. Al ser de orden total, tendremos que X está bien ordenado por $\in$ por la proposición anterior. Así, veamos que se cumplen las 3 propiedades necesarias:

- sea $a \in X$. Como $a = a$, claramente $a R_\in a$, de donde obtenemos la propiedad reflexiva.
- sean a y b elementos de X de forma que $a R_\in b$ y $b R_\in a$. Si se tuviese que $a \neq b$, tendríamos que $a \in b \in a$, lo que no es posible por el Axioma de Fundamento. Por tanto, $a = b$, asegurando así la propiedad antisimétrica.
- sean a, b y c elementos de X cumpliendo que $a R_\in b$ y $b R_\in c$. Si $a = b$ o $b = c$ es claro que $a R_\in c$. Si no se da ninguna de las igualdades, tenemos que $a \in b \in c$. Por hipótesis, tenemos que o bien $a \in c$ o bien $a = c$ o bien $c \in a$. De la segunda opción se obtendría que $a \in b \in a$ y de la tercera que $c \in a \in b \in c$, imposibles por el Axioma de Fundamento. Por tanto, se tiene que $a \in c$ y se demuestra así la propiedad transitiva.

□

Demostremos ahora la equivalencia deseada:

Proposición 3.5. *Un conjunto X es un ordinal si y solo sí es transitivo y está totalmente ordenado por $\in$.*

Demostración. Sea X un ordinal. Entonces, como la relación de orden es equivalente a $R_\in$, sabemos que $(X, R_\in)$ está bien ordenado, por lo que está totalmente ordenado. Veamos que es transitivo: Sea $a \in X$. Sabemos que $a = X_a$. Además, las relaciones $R_\in$ y $\subseteq$ son equivalentes en este conjunto, de donde $a = X_a = \{y \in X : y \subseteq a \wedge y \neq a\}$. Vemos por tanto que todos los elementos de a son elementos de X, por lo que $a \subseteq X$.

Recíprocamente, supongamos que X está totalmente ordenado por $\in$ y es transitivo. Por la proposición anterior, $(X, R_\in)$ está bien ordenado. Veamos entonces que, dado $a \in X$, $a = X_a$. Sea en primer lugar $y \in a$. Entonces por ser X transitivo, como $a \in X$, tenemos que $y \in X$. Así, $y \in \{z \in X : z \in a\} = \{z \in X : zR_\in a \wedge z \neq a\} = X_a$. Por otro lado, si $y \in X_a$, tenemos que $yR_\in a \wedge y \neq a$, por lo que $y \in a$. Así, concluimos que $a = X_a$. □

Nota 32. Notemos que decir que X está totalmente ordenado por $\in$ es análogo a decir que está bien ordenado por $\in$.

Así, en muchos textos un ordinal se define como un conjunto transitivo y bien ordenado por $\in$. A partir de ahora usaremos las dos definiciones indistintamente. Además, denotaremos a los ordinales con letras griegas, pues es lo más común en la literatura.

Visto esto, pasaremos a demostrar una serie de propiedades que nos servirán para comprender mejor el comportamiento de estos conjuntos.

Proposición 3.6. *Sean α y β dos ordinales. Se cumple que:*

1. *O bien $\alpha = \emptyset$ o bien $\emptyset \in \alpha$.*
2. *Si $\alpha \neq \beta$, entonces $\alpha \in \beta \leftrightarrow \alpha \subseteq \beta$*

Demostración.

1. Esto se demostró en el desarrollo posterior a la Definición 3.2.
2. Supongamos que $\alpha \in \beta$. Entonces, por ser β un ordinal, es transitivo, de donde $\alpha \subseteq \beta$.

   Supongamos ahora que $\alpha \subseteq \beta$. Como $\alpha \neq \beta$, tenemos que $\beta \backslash \alpha \neq \emptyset$. Por tanto, al ser un subconjunto no vacío de β, tiene mínimo, que llamaremos γ. Veamos que $\gamma = \alpha$. Entonces tendremos que $\alpha \in \beta$, pues por definición $\gamma \in \beta \backslash \alpha$ y por tanto $\gamma \in \beta$.

Comenzamos tomando $x \in \gamma$. Entonces $x \notin \beta \backslash \alpha$, pues lo contrario contradice la definición de mínimo. Como $x \in \gamma$, $\gamma \in \beta$ y β es transitivo, tenemos que $x \in \beta$. Por tanto, para que $x \notin \beta \backslash \alpha$, tiene que darse que $x \in \alpha$. Así, vemos que $\gamma \subseteq \alpha$.

Supongamos que $\gamma \neq \alpha$. Entonces $\exists \delta \in \alpha \backslash \gamma$, de donde δ es tal que $\delta \in \alpha$ y $\delta \notin \gamma$. De $\delta \in \alpha$, como $\alpha \subseteq \beta$, obtenemos que $\delta \in \beta$. Así, tanto γ como δ son elementos de β, que está bien ordenado por $R_\in$, por lo que se da $\delta R_\in \gamma$ o $\gamma R_\in \delta$. Como $\delta \notin \gamma$, lo que implica que no se da $\delta R_\in \gamma$, tenemos que $\gamma R_\in \delta$. Es decir, o bien $\gamma = \delta$ o bien $\gamma \in \delta$. Sin embargo, la primera opción no es posible pues $\delta \in \alpha$ y $\gamma \notin \alpha$. Por tanto, se tendría que $\gamma \in \delta$, pero $\delta \in \alpha$ y α transitivo implicaría que $\gamma \in \alpha$. En conclusión, hemos llegado a una contradicción al suponer que $\gamma \neq \alpha$, por lo que se concluye que $\gamma = \alpha$.

□

Llegados a este punto, algo que cabe preguntarse es si la colección de todos los números ordinales es un conjunto. Con el objetivo de responder esta cuestión tenemos la siguiente proposición.

Proposición 3.7. *Se cumple que:*

1. *Dados α y β ordinales, o bien $\alpha \in \beta$ o bien $\alpha = \beta$ o bien $\beta \in \alpha$.*
2. *Si β es un ordinal y $\alpha \in \beta$ entonces α es un ordinal*
3. *En toda colección no vacía formada por ordinales existe un elemento α tal que $\alpha \in \beta$ para todo β distinto de α en la colección.*

Demostración. 1. Notemos que, como consecuencia del Axioma de Fundamento, los tres casos son efectivamente excluyentes.

Ahora, si $\alpha = \beta$ ya estaría. Supongamos entonces que $\alpha \neq \beta$ y probaremos que se tiene $\alpha \subseteq \beta$ o $\beta \subseteq \alpha$. Así, al tener $\alpha \neq \beta$, por el punto *2* de la proposición anterior se deduce que $\alpha \in \beta$ o $\beta \in \alpha$ respectivamente. Para obtener estos contenidos, veremos que se da que $\alpha \cap \beta = \alpha$, lo que implicaría $\alpha \subseteq \beta$, o que $\alpha \cap \beta = \beta$, lo que implicaría $\beta \subseteq \alpha$.

Supongamos que se tiene que $\alpha \cap \beta \neq \alpha$. Entonces, como $\alpha \cap \beta \subseteq \alpha$, por el punto *2* de la proposición anterior $\alpha \cap \beta \in \alpha$. Si además $\alpha \cap \beta \neq \beta$, como también se tiene $\alpha \cap \beta \subseteq \beta$, se obtendría por el mismo motivo que $\alpha \cap \beta \in \beta$. Así, se deduciría que $\alpha \cap \beta \in \alpha \cap \beta$, lo que sabemos que no es posible por el Axioma de Fundamento. Así, no pueden darse simultáneamente ambas condiciones, por lo que se tendrá que $\alpha \cap \beta = \alpha$ o que $\alpha \cap \beta = \beta$.

2. Queremos probar que α está bien ordenado por $\in$ (o equivalentemente que está totalmente ordenado) y que es transitivo. Como $\alpha \in \beta$ y β transitivo, tenemos que $\alpha \subseteq \beta$. Por tanto, $R_\in$ será un orden total en α, pues todos los elementos de α lo serán de β, lo que implica que se cumplirán las propiedades necesarias al estar β totalmente ordenado por $R_\in$.

   Para ver que es transitivo, sean $\gamma \in \alpha$ y $\delta \in \gamma$. Queremos ver que $\delta \in \alpha$. Por un lado, como $\alpha \in \beta$ y β es transitivo, se tiene que $\gamma \in \beta$. De esto junto a $\delta \in \gamma$, de nuevo por la transitividad de β, $\delta \in \beta$. Así, como $\alpha \in \beta$, $\delta \in \beta$ y β es un conjunto totalmente ordenado por $R_\in$, tenemos que $\delta R_\in \alpha$ o $\alpha R_\in \delta$, lo que es equivalente a que se de o bien $\alpha = \delta$ o bien $\alpha \in \delta$ o bien $\delta \in \alpha$.

   Si $\alpha = \delta$, como $\delta \in \gamma$, se tiene que $\alpha \in \gamma$. Pero por hipótesis $\gamma \in \alpha$, por lo que se entra en contradicción con las consecuencias del Axioma de Fundamento. De la misma forma, si $\alpha \in \delta$, se tendría que $\alpha \in \delta \in \gamma \in \alpha$, lo que también contradice las consecuencias del Axioma de Fundamento. Así, se obtiene que $\delta \in \alpha$, probando lo deseado.

3. Supongamos que no existe dicho elemento. Entonces, como la colección es no vacía, podemos tomar un elemento α_0 en ella. Si este fuera el único elemento cumpliría el enunciado, por lo que existen más elementos en la colección. Como no puede darse que $\alpha_0 \in \beta$ para todo $\beta \neq \alpha_0$ en la colección, existe otro conjunto α_1 tal que $\alpha_0 \notin \alpha_1$ y $\alpha_0 \neq \alpha_1$. Así, por *1*, se tiene que $\alpha_1 \in \alpha_0$. Aplicando ahora el mismo razonamiento a α_1, existe α_2 en la colección de forma que $\alpha_2 \in \alpha_1 \in \alpha_0$. De esta forma, podríamos construir una sucesión de la forma $\alpha_0 \ni \alpha_1 \ni \alpha_2 \ni \ldots$, lo que no es posible por el Axioma de Fundamento.

□

Una vez visto esto, podemos demostrar que la colección de todos los ordinales, denotada comunmente por Ω, no es un conjunto. Para ello sólo necesitamos los dos primeros puntos de la proposición anterior, pero se ha incluido el tercero para ilustrar más propiedades interesantes.

Proposición 3.8. Ω *no es un conjunto.*

Demostración. Si Ω es un conjunto, podemos considerar en él la relación $R_\in$. Así, el punto *1* de la proposición anterior junto con la Proposición 3.4 aseguran que $(\Omega, R_\in)$ está bien ordenado. Además, el punto *2* implica que Ω es transitivo. Así, tendríamos que Ω es un ordinal, por lo que $\Omega \in \Omega$, contradiciendo el Axioma de Fundamento. □

Nota 33. A una colección de conjuntos cumpliendo cierta propiedad se la denomina **clase**. Así, todo conjunto es una clase pero no toda clase es un conjunto, como es el caso de Ω.

En el contexto de la Teoría de Conjuntos Naive, todas las propiedades anteriores también podían ser demostradas. Además, aunque se carecía del Axioma de Fundamento, era posible probar que $\alpha \notin \alpha$ para cualquier ordinal. Así, por un lado, Ω era un conjunto, pues no existían restricciones sobre estos, y por otro lado no era posible que $\Omega \in \Omega$ pues, razonando como en la última proposición, Ω era un ordinal. Esta es la razón de que este resultado sea conocido como la **Paradoja de Burali-Forti**, pues en el contexto en el que apareció se trataba en efecto de una contradicción.

Otra cuestión importante a tratar es si, en efecto, los números naturales son ordinales, así como si ω lo es. En vista de este objetivo se tienen las siguientes proposiciones.

Proposición 3.9. *Si α es un ordinal, entonces $\alpha + 1$ también lo es.*

Demostración. Veamos en primer lugar que es transitivo. Para ello, sea $\beta \in \alpha + 1$. Como $\alpha + 1 = \alpha \cup \{\alpha\}$, se tiene que $\beta \in \alpha$ o $\beta \in \{\alpha\}$, donde la última opción es equivalente a $\beta = \alpha$. En el primer caso, como α es transitivo, se tiene que $\beta \subseteq \alpha \subseteq \alpha \cup \{\alpha\}$, por lo que $\beta \subseteq \alpha \cup \{\alpha\}$. En el segundo, como $\alpha \subseteq \alpha \cup \{\alpha\}$, también se obtiene que $\beta \subseteq \alpha \cup \{\alpha\}$. Así, $\beta \subseteq \alpha + 1$.

Para ver que $\alpha + 1$ está bien ordenado por $\in$, basta observar que los elementos de este conjunto son o bien elementos de α o bien el propio α. Es decir, son todos ordinales. Entonces, por el punto *1* de la Proposición 3.7 junto con la Proposición 3.4, tenemos que efectivamente está bien ordenado por $\in$.

□

Proposición 3.10. *Todo número natural es un ordinal.*

Nota 34. Para demostrar este resultado, trataremos de probar que el conjunto de todos los naturales tales que son ordinales coincide con ω. Para poder tomar dicho conjunto, usaremos el Esquema Axiomático de Separación, para lo cual necesitamos expresar el hecho de ser ordinal mediante una fórmula. Hemos visto que ser ordinal se caracteriza por ser un conjunto bien ordenado por $\in$ y transitivo. A la vez, la primera condición se puede sustituir por la dada en la Proposición 3.4. Así, consideramos las siguientes definiciones:

$$\begin{aligned} trans(x) &:\Leftrightarrow \forall y(y \in x \to y \subseteq x) \\ word(x) &:\Leftrightarrow (\forall y \in x)(\forall z \in x)(z \in y \lor y = z \lor y \in z) \\ ord(x) &:\Leftrightarrow trans(x) \land word(x) \end{aligned}$$

Observamos que, efectivamente, $ord(x)$ implica que x es un ordinal. Visto esto, demostremos el resultado:

Demostración. Consideramos el siguiente conjunto:

$$\mathscr{C} := \big\{n \in \omega : ord(n)\big\}.$$

Por definición, $\mathscr{C} \subseteq \omega$. Así, si vemos que $\mathscr{C}$ es un conjunto sucesor, por la definición de ω tendremos $\omega \subseteq \mathscr{C}$, de donde $\omega = \mathscr{C}$ y por tanto todos los elementos de ω, los naturales, serán ordinales.

En primer lugar, observamos que $\emptyset$ es un ordinal, como se vio en la Proposición 3.6. Como además $\emptyset \in \omega$, tenemos que $\emptyset \in \mathscr{C}$. Ahora, sea $n \in \mathscr{C}$, lo que implica que $n \in \omega$ y que n es un ordinal. De lo primero se deduce que $n+1 \in \omega$ y de lo segundo, por la proposición anterior, que $n+1$ es un ordinal. Así, concluimos que $n+1 \in \mathscr{C}$. De esta forma hemos visto que $\mathscr{C}$ es un conjunto sucesor, probando así lo deseado. □

Proposición 3.11. *Si A es un conjunto cuyos elementos son ordinales, entonces $\bigcup A$ es un ordinal.*

Demostración. Veamos en primer lugar que $\bigcup A$ es transitivo. Sean x, y de forma que $y \in x \in \bigcup A$. Como $x \in \bigcup A$, $\exists \alpha \in A(x \in \alpha)$. Además, como α es un ordinal, es transitivo, por lo que $y \in \alpha$. Así, por definición de la unión de un conjunto, $y \in \bigcup A$, por lo que este es transitivo.

Para comprobar que está bien ordenado por $\in$, tomemos x, y elementos de $\bigcup A$ y veamos que o bien $x \in y$ o bien $x = y$ o bien $y \in x$. De $x \in \bigcup A$ obtenemos que $\exists \alpha \in A(x \in \alpha)$ y de $y \in \bigcup A$, que $\exists \beta \in A(y \in \beta)$. Por ser α y β ordinales, se tiene que o bien $\alpha \in \beta$ o bien $\alpha = \beta$ o bien $\beta \in \alpha$. Si $\alpha = \beta$, entonces $x \in \alpha$ e $y \in \alpha$, y por estar α bien ordenado por $\in$ se tiene que o bien $x \in y$ o bien $x = y$ o bien $y \in x$. Si $\alpha \neq \beta$, supongamos que $\alpha \in \beta$, de donde $\alpha \subseteq \beta$. Así, tanto x como y serán elementos de β, que está bien ordenado, por lo que se obtiene la misma conclusión. En el caso $\beta \in \alpha$ el razonamiento es análogo. □

Proposición 3.12. *El conjunto ω es un ordinal.*

Demostración. Sabemos que todos los elementos de ω son ordinales. Por tanto, si demostramos que $\omega = \bigcup w$, por la proposición anterior tendremos lo deseado.

Sea $n \in \omega$. Entonces $n+1 \in \omega$ y $n \in n+1$, por lo que $n \in \bigcup \omega$. Sea ahora $n \in \bigcup \omega$. Entonces $\exists m \in \omega(n \in m)$. Como m está compuesto por todos los naturales anteriores a él, tenemos que $n \in \omega$. Así, se demuestra la igualdad deseada. □

En resumen, hemos comprobado que en efecto todos los naturales son ordinales, así como que lo es ω. Además, como sabemos que el conjunto sucesor de cualquier ordinal también lo es, los conjuntos $\omega+1, (\omega+1)+1, \ldots$ son también ordinales. Parece que hemos obtenido por tanto lo que se buscaba: un concepto que generalizase a los números naturales y que extendiese su ordenación. Como hemos visto que Ω no es un conjunto, no podemos definir una relación de orden propiamente dicha. Sin embargo, dados α y β ordinales, comúnmente se utiliza la siguiente notación:

$$\alpha < \beta :\Leftrightarrow \alpha \in \beta \qquad \alpha \leq \beta :\Leftrightarrow \alpha \in \beta \vee \alpha = \beta,$$

y se dice que α es **menor** que β o **menor o igual** que β respectivamente, o que β es **mayor** o **mayor o igual** que α.

Así, presentamos una última proposición que nos ayudará a determinar finalmente la ordenación y el comportamiento de los ordinales:

Proposición 3.13. *Se cumple que:*

1. *Si α y β son ordinales de forma que $\alpha < \beta$, entonces $\alpha + 1 \leq \beta$. Es decir, $\alpha + 1$ es el menor ordinal mayor que α.*

2. *Si α es un ordinal, entonces o bien $\alpha = \bigcup \alpha$ o bien existe un ordinal β de forma que $\alpha = \beta + 1$.*

Demostración. 1. Tenemos que $\alpha < \beta$, es decir, $\alpha \in \beta$. Queremos ver que $\alpha + 1 \subseteq \beta$, pues esto es equivalente a que $\alpha + 1 = \beta \vee \alpha + 1 \in \beta$, es decir, a $\alpha + 1 \leq \beta$. Como $\alpha \in \beta$, $\{\alpha\} \subseteq \beta$. Además, como β es transitivo, $\alpha \subseteq \beta$. Por tanto, $\alpha \cup \{\alpha\} \subseteq \beta$, obteniendo lo deseado.

2. En primer lugar, notamos que, $\bigcup \alpha \subseteq \alpha$, pues dado $x \in \bigcup \alpha$, $\exists \gamma \in \alpha (x \in \gamma)$, y como α es transitivo, tenemos $x \in \alpha$. Así, exsten dos posibilidades: que $\alpha \backslash \bigcup \alpha = \emptyset$, de donde se tendría $\alpha = \bigcup \alpha$, o que $\alpha \backslash \bigcup \alpha \neq \emptyset$. En el segundo caso, por ser un subconjunto no vacío de α, existe $\beta \in \alpha \backslash \bigcup \alpha$ mínimo de este conjunto. Probemos que $\beta + 1 = \alpha$.

   Por su definición, $\beta \in \alpha$ y $\beta \notin \bigcup \alpha$. Del primer hecho obtenemos que β es un ordinal, por lo que $\beta + 1$ también lo es. Así, por el punto *1* y el hecho de que $\beta \in \alpha$, $\beta + 1 \subseteq \alpha$. Por tanto, o bien $\beta + 1 = \alpha$ o bien $\beta + 1 \in \alpha$. En el segundo caso, se tendría $\beta \in \beta + 1 \in \alpha$, lo que implica que $\beta \in \bigcup \alpha$, pero sabemos que $\beta \notin \bigcup \alpha$. En conclusión, $\alpha = \beta + 1$.

□

Nota 35. En el caso de que $\alpha = \beta + 1$, se suele denotar a β como $\alpha - 1$.

Este resultado motiva la siguiente definición:

Definición 3.5. Un ordinal α se denomina **ordinal sucesor** si existe otro ordinal β tal que $\alpha = \beta + 1$, y se denomina **ordinal límite** si $\alpha = \bigcup \alpha$.

Así, como se esperaba en la motivación intuitiva, se comienzan los ordinales en el 0 y se van añadiendo de forma reiterada los conjuntos sucesores, obteniendo los naturales. Así, todos los naturales son ordinales sucesores. Tomando la unión de todos ellos, se forma el conjunto ω, el primer ordinal límite no vacío. A continuación, se prosigue tomando los sucesores de ω, para luego formar un nuevo ordinal límite con la unión de todos ellos. Este proceso continua hasta el infinito.
Llegados a esta punto, podemos comenzar a visualizar la ordenación de los ordinales:

$$0 < 1 < 2 < \ldots < n < \ldots < \omega < \omega + 1 < (\omega + 1) + 1 < \ldots$$

Hasta ahora, conocemos bien los números naturales y ω. El siguiente objetivo será estudiar los conjuntos posteriores. Sabemos que $\omega + 1$ es el sucesor de ω. A su vez, $(\omega + 1) + 1$ es el sucesor de $\omega + 1$. Así, parece intuitivo pensar que este último conjunto se pueda escribir como $\omega + 2$, en el sentido de que se toma dos veces el conjunto sucesor. En general, ¿qué significaría $\alpha+\beta$ cuando estos son dos ordinales? Teniendo estas preguntas como motivación, pasamos a estudiar la aritmética de los ordinales.

3.1.1. La Aritmética de los Ordinales

Para definir las distintas operaciones entre ordinales, aprovecharemos su principal característica: su orden. Así, definiremos la suma, la multiplicación y la exponenciación de ordinales de manera recursiva. Es decir, se determinará el resultado de una operación entre dos ordinales en base a los resultados de dicha operación en los ordinales previos. Para ello, haremos uso del Teorema de Recursión Transfinita, que permite definir una función de clase cuyo valor en un ordinal depende de su valor en los anteriores. Existe también una definición alternativa que no hace uso de este teorema, sino que se basa en la concepción de un ordinal como conjunto ordenado, aunque es equivalente. Esta puede encontrarse en [Devlin, 1994, Capítulo 3].

Teorema 3.1 (Halbeisen, 2012, Teorema 3.19)**. *Teorema de Recursión Transfinita***
Sea F una función de clase. Entonces existe una única función de clase G, definida en los ordinales, de forma que para cada ordinal α se tiene que

$$G(\alpha) = F\big(G|_\alpha\big) \quad \text{donde } G|_\alpha \coloneqq \big\{\langle \beta, G(\beta)\rangle : \beta \in \alpha\big\}.$$

Este teorema afirma que, dada una función de clase F, existe una única forma de definir una función de clase G sobre los ordinales, de forma que para cada ordinal α, $G(\alpha) = F(G|_\alpha)$. Es decir, el valor de G en α se obtendría aplicando F a la función que asigna $G(\beta)$ a cada ordinal $\beta < \alpha$. Esto nos permite construir G de manera recursiva, teniendo en cuenta los resultados anteriores.

Con este teorema como herramienta, pasemos a definir las operaciones con ordinales. Es importante recordar que, como Ω no es un conjunto, no se puede definir en él una operación binaria interna propiamente dicha. Sin embargo, hablaremos de operación en el sentido de que asignaremos, para cada operación y par de ordinales, un conjunto único que sea también un ordinal. Daremos primeramente su definición y posteriormente comprobaremos que estan efectivamente bien definidas.

Comencemos definiendo la suma. Sean α y β ordinales. Teniendo en cuenta la ordenación de los ordinales, al igual que en el ejemplo de $(\omega+1)+1$, intuitivamente podemos entender que $\alpha+\beta$ sería el resultado de comenzar en α y desplazarse β posiciones hacia delante, es decir, tantas posiciones como elementos tenga β. Además, como haremos uso del Teorema de Recursión Transfinita, podemos suponer que $\alpha + \gamma$ está definido para cualquier $\gamma < \beta$. Siguiendo esta idea intuitiva, si $\beta = 0$ definiremos $\alpha + \beta$ como α. Si β fuese un ordinal sucesor, como conoceríamos $\alpha + (\beta - 1)$, bastaría con tomar el sucesor de este ordinal. Para entender la definición cuando β es un ordinal límite, veamos el siguiente ejemplo:

Supongamos que queremos calcular $\omega + \omega$. Este se obtendría comenzando en ω y desplazándose tantas veces como elementos tiene ω. Entonces, teniendo en cuenta que un ordinal es el conjunto formado por todos los ordinales anteriores al él, tendríamos que:

$$\omega + \omega = \{0, 1, 2, \ldots, \omega, \omega + 1, \omega + 2, \ldots\} = \bigcup\{\omega + \gamma : \gamma \in \omega\}.$$

Así, tenemos la siguiente definición:

Definición 3.6. Suma de Ordinales. Para un ordinal arbitrario α se define:

- $\alpha + 0 \coloneqq \alpha$.
- $\alpha + \beta \coloneqq \big(\alpha + (\beta - 1)\big) + 1$ si β es un ordinal sucesor.
- $\alpha + \beta \coloneqq \bigcup\{\alpha + \gamma : \gamma \in \beta\}$ si β es un ordinal límite no vacío.

Nota 36. Observamos que ahora $\alpha + 1$ denota tanto la suma de un ordinal α con el ordinal 1 como la operación de tomar el conjunto sucesor de α. Sin embargo, esto no genera ningún conflicto pues, por la definición de la suma, el sucesor de un ordinal α es el mismo que el sucesor de $\alpha + 0$. Así, como $1 - 1 = 0$, se tiene de nuevo por definición que $\alpha + 1$ es el sucesor de $\alpha + (1 - 1) = \alpha + 0$, es decir, el sucesor de α.

Es importante notar que para definir la suma, al igual que ocurrirá con el resto de operaciones, fijamos un ordinal α y definimos el resultado de sumarle todos los ordinales posibles. Se lleva a cabo de esta forma, como se verá más adelante, para poner en práctica el Teorema de Recursión Transfinita, pero acarrea algunas consecuencias, como que la suma no es conmutativa. Por ejemplo:

$$1 + \omega = \bigcup \{1 + n : n \in \omega\} = \omega \neq \omega + 1.$$

Gracias a la suma de ordinales, podemos hacernos una mayor idea de la ordenación:

$$0 < 1 < 2 < \ldots < n < \ldots < \omega < \omega + 1 < \omega + 2 < \ldots < \omega + n < \ldots < \omega + \omega <$$
$$\omega + \omega + 1 < \omega + \omega + 2 < \ldots < \omega + \omega + n < \ldots < \omega + \omega + \omega < \omega + \omega + \omega + 1 < \ldots$$

Observamos que los ordinales límite no vacíos serán ω, $\omega + \omega$, $\omega + \omega + \omega$, y así de forma sucesiva.

Introducimos ahora la multiplicación de ordinales. Esta surge para expresar términos en los que el mismo ordinal aparece sumado varias veces, como por ejemplo $\omega + \omega + \omega$, que sería $\omega \cdot 3$. Así, la idea intuitiva de $\alpha \cdot \beta$ será sumar α tantas veces como elementos tenga β. Teniendo en cuenta que queremos aplicar el Teorema de Recursión transfinita, si $\beta = 0$, $\alpha \cdot \beta$ será 0, si β es sucesor, será sumar α a $\alpha \cdot (\beta - 1)$ y si es ordinal límite, la unión de $\alpha \cdot \gamma$ para todos los $\gamma < \beta$.

Definición 3.7. Multiplicación de Ordinales. Para un ordinal arbitrario α se define:

- $\alpha \cdot 0 \coloneqq 0$.
- $\alpha \cdot \beta \coloneqq \big(\alpha \cdot (\beta - 1)\big) + \alpha$ si β es un ordinal sucesor.
- $\alpha \cdot \beta \coloneqq \bigcup \{\alpha \cdot \gamma : \gamma \in \beta\}$ si β es un ordinal límite no vacío.

Observamos que la multiplicación tampoco es conmutativa, pues por ejemplo,

$$2 \cdot \omega = \bigcup \{2 \cdot n : n \in \omega\} = \omega \neq \omega + \omega = \omega \cdot 2.$$

De nuevo, gracias a la multiplicación tenemos una visión aún más amplia de los ordinales:

$$0 < 1 < 2 < \ldots < n < \ldots < \omega < \omega + 1 < \omega + 2 < \ldots < \omega + n < \ldots < \omega \cdot 2 <$$
$$\omega \cdot 2 + 1 < \omega \cdot 2 + 2 < \ldots < \omega \cdot 2 + n < \ldots < \omega \cdot 3 < \ldots <$$
$$\omega \cdot \omega < \omega \cdot \omega + 1 < \omega \cdot \omega + 2 < \ldots < \omega \cdot \omega \cdot n < \ldots < \omega \cdot \omega \cdot \omega < \ldots$$

Así, observamos que los ordinales límite no vacios serán ordinales de la forma $\omega \cdot \alpha$ para algún ordinal α.

Por último, la exponenciación de ordinales surge para expresar el mismo ordinal multiplicado varias veces, como puede ser el caso de $\omega \cdot \omega \cdot \omega$, que sería ω^3. Aplicando el mismo razonamiento intuitivo que para las operaciones anteriores, tenemos la siguiente definición.

Definición 3.8. Exponenciación de Ordinales. Para un ordinal arbitrario α se define:

- $\alpha^0 \coloneqq 1$.
- $\alpha^\beta \coloneqq \alpha^{\beta-1} \cdot \alpha$ si β es un ordinal sucesor.
- $\alpha^\beta \coloneqq \bigcup\{\alpha^\gamma : \gamma \in \beta\}$ si β es un ordinal límite no vacío.

De nuevo, la exponenciación de ordinales nos permite extender aún más la visión de nuestra ordenación:

$$
\begin{aligned}
&0 < 1 < 2 < \ldots < \omega < \omega + 1 < \ldots < \omega \cdot 2 < \ldots < \omega \cdot 3 < \ldots < \omega \cdot \omega < \ldots \\
&\omega^3 < \omega^3 + 1 < \ldots < \omega^3 + \omega < \ldots < \omega^3 + \omega^2 < \ldots < \omega^4 < \omega^4 + 1 < \ldots \\
&\omega^\omega < \omega^\omega + 1 < \ldots < \omega^{\omega \cdot 2} < \ldots < \omega^{\omega \cdot 3} < \ldots < \omega^{\omega \cdot \omega} < \ldots < \omega^{\omega^3} < \ldots < \omega^{\omega^\omega} < \ldots
\end{aligned}
$$

En conclusión, gracias a la aritmética de los ordinales somos capaces de visualizar una gran parte de estos conjuntos. Sin embargo, es importante no perder de vista que incluso el número ω^{ω^ω} es pequeño en comparación con la mayoría de ordinales.

Visto esto, pasaremos ahora a demostrar que efectivamente estas tres operaciones están bien definidas en el sentido de que tienen un único resultado, que es a su vez un ordinal.

Teorema 3.2. *La suma, multiplicación y exponenciación de ordinales está bien definida.*

Demostración. Comencemos probándolo para la suma. Sea α un ordinal cualquiera pero fijo. Vamos a demostrar, mediante el Teorema de Recursión Transfinita, que existe una función de clase G_α definida sobre los ordinales de forma que para cada ordinal β, $G_\alpha(\beta) = \alpha + \beta$. Para ello, definimos una función de clase F_α de forma que $F_\alpha(x) \coloneqq 0$ si x no es una función y toma el siguiente valor si

x es una función:

$$F_\alpha(x) \coloneqq \begin{cases} \alpha & \text{si } x = 0 \\ x(\beta - 1) + 1 & \text{si } dom(x) = \beta \text{ y } \beta \text{ es un ordinal sucesor} \\ \bigcup\{x(\gamma) : \gamma \in \beta\} & \text{si } dom(x) = \beta \text{ y } \beta \text{ es un ordinal límite no vacío} \\ 0 & \text{en otro caso.} \end{cases}$$

Por el Teorema de Recursión Transfinita, existe una única función de clase G definida sobre los ordinales de forma que, para cada ordinal β,

$$G(\beta) = F_\alpha(G|_\beta) \quad \text{donde } G|_\beta = \big\{\langle\gamma, G(\gamma)\rangle : \gamma \in \beta\big\}.$$

Así, dado un ordinal β,

- si $\beta = 0$, $G|_\beta = \emptyset = 0$, por lo que $F_\alpha(G|_\beta) = \alpha = \alpha + \beta$.
- si β es un ordinal sucesor, entonces $G|_\beta = \big\{\langle\gamma, G(\gamma)\rangle : \gamma \in \beta\big\}$ es una función tal que $dom\big(G|_\beta\big) = \beta$. Así, suponiendo que ya se ha probado que para cualquier ordinal menor que β la función de clase G define la suma con α,

$$G(\beta) = F_\alpha(G|_\beta) = G(\beta - 1) + 1 = \big(\alpha + (\beta - 1)\big) + 1 = \alpha + \beta.$$

- si β es un ordinal límite, entonces $G|_\beta$ tiene la misma expresión que en el caso anterior, y el dominio de esta función también es β. Así, de nuevo suponiendo que para cualquier ordinal menor que β la función de clase G define la suma con α, $G(\beta) = F_\alpha(G|_\beta) = \bigcup\{G(\gamma) : \gamma \in \beta\} = \bigcup\{\alpha + \gamma : \gamma \in \beta\} = \alpha + \beta$.

Por tanto, tomando $G_\alpha = G$, encontramos la función de clase buscada. Observamos también que el resultado de esta operación es siempre un ordinal, pues se obtiene tomando el sucesor de un ordinal o la unión de un conjunto de ordinales.

Para la multiplicación y la exponenciación se repite el mismo proceso con una variación adecuada de la función de clase F_α tomada anteriormente. En el primer caso el valor de F_α para una función x sería

$$F_\alpha(x) \coloneqq \begin{cases} 0 & \text{si } x = 0 \\ x(\beta - 1) + \alpha & \text{si } dom(x) = \beta \text{ y } \beta \text{ es un ordinal sucesor} \\ \bigcup\{x(\gamma) : \gamma \in \beta\} & \text{si } dom(x) = \beta \text{ y } \beta \text{ es un ordinal límite no vacío} \\ 0 & \text{en otro caso,} \end{cases}$$

y en el segundo,

$$F_\alpha(x) := \begin{cases} 1 & \text{si } x = 0 \\ x(\beta - 1) \cdot \alpha & \text{si } dom(x) = \beta \text{ y } \beta \text{ es un ordinal sucesor} \\ \bigcup\{x(\gamma) : \gamma \in \beta\} & \text{si } dom(x) = \beta \text{ y } \beta \text{ es un ordinal límite no vacío} \\ 0 & \text{en otro caso.} \end{cases}$$

□

3.1.2. Inducción Transfinita

Una de las propiedades más importantes de los números naturales es el Principio de Inducción, una herramienta fundamental que permite demostrar que cierta propiedad se cumple para todos estos conjuntos. Este principio es ampliamente utilizado en las matemáticas, pero en esta sección lo demostraremos en el marco de la Teoría Axiomática de Conjuntos. Además, dado que los ordinales amplian el conjunto de los números naturales, es razonable esperar que el Principio de Inducción también pueda extenderse a ellos. Esta extensión es conocida como el Principio de Inducción Transfinita, que presentaremos también a continuación.

Teorema 3.3. ***Principio de Inducción***
Sea C un conjunto de números naturales cumpliendo que:

$$\textit{(i)}\ 0 \in C \qquad \textit{(ii)}\ \forall n \in \omega(n \in C \to n+1 \in C).$$

Entonces $C = \omega$.

Demostración. Basta ver que, por hipótesis, $C \subseteq \omega$ y C es un conjunto sucesor. Entonces, por la definición de ω, $\omega \subseteq C$ y por tanto $\omega = C$. □

Una reformulación de este resultado es la siguiente:

Corolario 3.1. *Dada una fórmula φ, si se cumple que*

$$\varphi(0) \wedge \forall n \in \omega\big(\varphi(n) \to \varphi(n+1)\big),$$

entonces $\forall n \in w\, \varphi(n)$.

Demostración. Basta tomar $C := \{n \in \omega : \varphi(n)\}$ y aplicar el Principio de Inducción. □

Veamos ahora la generalización de este principio para los ordinales:

Teorema 3.4. ***Principio de Inducción Transfinita***
Sea C una clase de ordinales tal que

(i) si un ordinal α pertenece a C, entonces $\alpha + 1$ pertenece a C.

(ii) si α es un ordinal límite y para cualquier $\beta \in \alpha$, β pertenece a C, entonces α pertenece a C.

Entonces C es la clase de todos los ordinales.

Demostración. Notemos en primer lugar que, por *(ii)*, tenemos que 0 está en C, por lo que es una clase no vacía. Así, si suponemos que C no es la clase de todos los ordinales, podemos tomar la clase formada por todos aquellos ordinales que no estén en C. La denotaremos por $\Omega \backslash C$. Así, por el punto *(iii)* de 3.7, tenemos que existe $\alpha_0 \in C$ tal que $\alpha_0 < \beta$ para todo β en la colección. Como α_0 es un ordinal, puede o bien ser sucesor, o bien límite.
Por un lado, en el primer caso, se tiene que $\alpha_0 - 1$ pertenece a C, pues si no sería un conjunto en la colección $\Omega \backslash C$ menor que α_0, lo que no es posible. Pero en este caso, por *(i)*, α_0 pertenecería también a C. Por el otro lado, si α_0 fuese un ordinal límite, todos sus elementos deberían pertenecer a C por el razonamiento anterior, pero por *(ii)* se tendría que α_0 pertenecería a C.
Así, llegamos a una contradicción proveniente de suponer que C no es la clase de todos los ordinales, de donde decucimos que en efecto sí lo es. $\square$

Una formulación equivalente de este teorema, expresada en función de fórmulas, es la presentada en el siguiente corolario. Para demostrarlo basta tomar, de forma análoga a lo realizado en los naturales, la clase de los ordinales cumpliendo dicha fórmula y aplicar el Principio de Inducción Transfinita.

Corolario 3.2. *Dada una fórmula φ, si se cumple que*

$$\forall\alpha \left(ord(\alpha) \wedge \forall\beta \in \alpha\big(\varphi(\beta)\big) \rightarrow \varphi(\alpha)\right),$$

entonces $\forall\alpha\big(ord(\alpha) \rightarrow \varphi(\alpha)\big)$.

Como ejemplos de la aplicación de este principio, demostraremos una serie de propiedades sobre la suma de ordinales. Las análogas se dan también para la multiplicación.

Proposición 3.14. *Dado un ordinal α, $0 + \alpha = \alpha$.*

Demostración. Consideramos la clase de los ordinales cumpliendo esta propiedad. En primer lugar, si α es un ordinal en ella, entonces $0 + (\alpha + 1) = (0 + \alpha) + 1 = \alpha + 1$, por lo que $\alpha + 1$ pertenece a la clase, donde en la segunda igualdad se utiliza la hipótesis de inducción.

Supongamos ahora que α es un ordinal límite tal que todos sus elementos pertenecen a la clase. Si $\alpha = 0$, entonces $0 + \alpha = 0 + 0 = 0$ y por tanto α está en la colección. En caso contrario,

$$0 + \alpha = \bigcup\{0 + \beta : \beta \in \alpha\} = \bigcup\{\beta : \beta \in \alpha\} = \bigcup \alpha = \alpha,$$

donde en la segunda igualdad se ha utilizado la hipótesis de inducción. Así, α estaría en la clase.

En conclusión, aplicando el Principio de Recursión Transfinita tenemos que la clase considerada es la compuesta por todos los ordinales, de donde todos cumplen esta propiedad. □

Nota 37. Vemos que el proceso de una prueba por inducción transfinita sobre α consiste en probar que el 0 cumple la propiedad deseada, que si lo hace α lo hace $\alpha + 1$, y que si α es un ordinal límite tal que todos sus elementos cumplen la propiedad, entonces α la cumple. Procederemos siguiendo este esquema a partir de ahora.

Proposición 3.15. *Sean α, β y γ ordinales. Entonces,*

$$\beta < \gamma \to \alpha + \beta < \alpha + \gamma.$$

Demostración. Fijemos α y β y realicemos la prueba por inducción sobre γ. Observemos que, por L_9, cuando $\gamma \leq \beta$, lo que es análogo a $\neg(\beta < \gamma)$, podremos deducir directamente que se cumple la propiedad deseada.

- Si $\gamma = 0$, $\gamma \leq \beta$, de donde el 0 cumple la propiedad.

- Sea γ un ordinal tal que $\beta < \gamma \to \alpha + \beta < \alpha + \gamma$. Si $\gamma < \beta$, entonces $\gamma + 1 \leq \beta$, de donde $\gamma + 1$ cumple la propiedad.

  Si $\gamma = \beta$, entonces $\alpha + \beta = \alpha + \gamma \in (\alpha + \gamma) + 1 = \alpha + (\gamma + 1)$. Así, $\alpha + \beta \in \alpha + (\gamma + 1)$, lo que es equivalente a $\alpha + \beta < \alpha + (\gamma + 1)$.

  Si $\beta < \gamma$, entonces se cumple que $\alpha + \beta < \alpha + \gamma$. Así, tenemos $\alpha + \beta \in \alpha + \gamma \in \alpha + (\gamma + 1)$. Como $\alpha + (\gamma + 1)$ es transitivo por ser un ordinal, se cumple que $\alpha + \beta \in \alpha + (\gamma + 1)$ y por tanto $\alpha + \beta < \alpha + (\gamma + 1)$.

Aplicando el Principio de Recursión Transfinita se tendría el resultado. □

Probaremos ahora que la suma de ordinales es asociativa. Para ello, veamos el siguiente resultado previo:

Proposición 3.16. *Sean α y β ordinales, siendo β un ordinal límite. Entonces $\alpha+\beta$ es un ordinal límite.*

Demostración. Supongamos que $\alpha+\beta$ no es un ordinal límite, de donde es un ordinal sucesor. Así, existe un ordinal γ tal que $\alpha+\beta=\gamma+1$. Como $\gamma \in \gamma+1$, tenemos que $\gamma \in \alpha+\beta$.
Por definición, $\alpha+\beta=\bigcup\{\alpha+\delta : \delta \in \beta\}$. Así, existe $\delta_0 \in \beta$ de forma que $\gamma \in \alpha+\delta_0$ o, equivalentemente, $\gamma < \alpha+\delta_0$, de donde $\gamma+1 \leq \alpha+\delta_0$. Por tanto, o bien $\gamma+1 \in \alpha+\delta_0$, lo que implicaría que $\gamma+1 \in \alpha+\beta=\gamma+1$, lo que no es posible, o bien $\gamma+1=\alpha+\delta_0$. En este último caso, como $\delta_0<\beta$, $\delta_0+1 \leq \beta$. Además, la igualdad no se puede dar porque si no β sería sucesor. Así, $\delta_0+1 \in \beta$. Con esto, se tiene que

$$\gamma+1=\alpha+\delta_0 \in (\alpha+\delta_0)+1=\alpha+(\delta_0+1),$$

con $\delta_0+1 \in \beta$, de donde $\gamma+1 \in \alpha+\beta=\gamma+1$, lo que de nuevo es una contradicción.
Así, concluimos que $\alpha+\beta$ es un ordinal límite. □

Proposición 3.17. *La suma de ordinales es asociativa. Es decir, dados tres ordinales α, β y γ, se cumple que*

$$(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma).$$

Demostración. Fijemos α y β y realicemos la prueba por inducción sobre γ.

- si $\gamma=0$, por definición de la suma con el 0 se tiene que $(\alpha+\beta)+0=\alpha+\beta=\alpha+(\beta+0)$.
- sea γ un ordinal cumpliendo que $(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$. Entonces,

  $$(\alpha+\beta)+(\gamma+1)=\big((\alpha+\beta)+\gamma\big)+1=\big(\alpha+(\beta+\gamma)\big)+1=\alpha+\big((\beta+\gamma)+1\big)=\alpha+\big(\beta+(\gamma+1)\big),$$

  donde en la segunda igualdad se ha aplicado la hipótesis de inducción y en el resto, la definición de suma.
- sea ahora γ un oridnal límite tal que para todo $\delta \in \gamma$ se da la asociatividad. Por un lado, tenemos que

  $$(\alpha+\beta)+\gamma=\bigcup\{(\alpha+\beta)+\delta : \delta \in \gamma\}=\bigcup\{\alpha+(\beta+\delta) : \delta \in \gamma\},$$

donde el la primera igualdad se ha aplicado la definición de suma de un ordinal y en la segunda la hipótesis de inducción.

Por otro lado, teniendo en cuenta, por el resultado anterior, que $\beta + \gamma$ es un ordinal límite,

$$\alpha + (\beta + \gamma) = \bigcup\{\alpha + \eta : \eta \in \beta + \gamma\}.$$

Así, para ver que γ cumple la propiedad asociativa, tenemos que probar que

$$\bigcup\{\alpha + (\beta + \delta) : \delta \in \gamma\} = \bigcup\{\alpha + \eta : \eta \in \beta + \gamma\}.$$

Denotemos, por simplicidad, A al primer conjunto y B al segundo.

Si $x \in A$, entonces existe $\delta_0 \in \gamma$ tal que $x \in \alpha + (\beta + \delta_0)$. Si vemos que $(\beta + \delta_0) \in \beta + \gamma$ tendremos que $x \in B$. Por definición, $\beta + \gamma = \bigcup\{\beta + \delta : \delta \in \gamma\}$. Además, como $\delta_0 \in \gamma$, $\delta_0 + 1 \leq \gamma$ y al ser γ un ordinal límite, $\delta_0 + 1 \in \gamma$. Así, $\beta + \delta_0 \in (\beta + \delta_0) + 1 = \beta + (\delta_0 + 1)$ con $\delta_0 + 1 \in \gamma$. Por tanto, $\beta + \delta_0 \in \beta + \gamma$.

Ahora, si $x \in B$, tenemos que existe $\eta_0 \in \beta + \gamma$ tal que $x \in \alpha + \eta_0$. Por definición de $\beta + \gamma$, existe $\delta_1 \in \gamma$ tal que $\eta_0 \in \beta + \delta_1$. Así, tenemos que $\eta_0 < \beta + \delta_1$, por lo que por la proposición anterior se deduce que $\alpha + \eta_0 < \alpha + (\beta + \delta_1)$. Por la transitividad de este último conjunto, $x \in \alpha + (\beta + \delta_1)$, con $\delta_1 \in \gamma$. Así, se concluye que $x \in A$.

Aplicando ahora el Principio de Inducción Transfinita se obtiene el resultado. □

Como acabamos de ver, el Principio de Inducción Transfinita permite probar numerosas propiedades sobre los ordinales de una forma bastante sencilla. Sin embargo, su alcance va mucho más alla de este campo, pues es una herramienta muy poderosa en todas las ramas de las matemáticas. Una de sus principales aplicaciones es la construcción de la **Jerarquía de Borel**, que ordena los conjuntos de la σ- álgebra de Borel asignando a cada uno de ellos un ordinal. Se puede consultar esta construcción en [Jech, 2003, Capítulo 11]. Esta jerarquía, a su vez, permite probar resultados importantes en Teoría de la Medida o sobre la clasificación de funciones.

La gran relevancia de este principio reside realmente en que la inducción ya no solo se restringe a elementos que puedan ser ordenados mediante los naturales, si no que se amplía a un orden cuya extensión es inimaginablemente mayor.

3.2. Consistencia y Completitud de ZF

En este trabajo, hemos presentado los axiomas de la Teoría de Zermelo-Fraenkel. Sin embargo, no se ha hablado aún de su completitud o de su consistencia. Tras haber estudiado la aritmética de los ordinales, contamos ya con las herramientas para relacionar los Teoremas de Incompletitud de Gödel con los axiomas de ZF. Para ello, veremos que los números naturales cumplen los axiomas de la Aritmética de Peano, lo que nos permitirá aplicar los teoremas.

Proposición 3.18. *ω cumple los axiomas de la Aritmética de Peano (1.23). Es decir, $\omega \models PA$.*

Demostración. En primer lugar, recordemos que el Lenguaje de la Aritmética de Peano está compuesto por la signatura $\mathscr{L}_{PA} = \{0, s, +, \cdot\}$. Así, vamos a formar una $\mathscr{L}_{PA}$-estructura con dominio ω y vamos a ver que efectivamente es un modelo de PA. Observemos que como ω es un conjunto, la operación conjunto sucesor es una función 1-aria en ω y la suma y la multipicación de ordinales son funciones binarias dentro de este conjunto, pues como los naturales son ordinales sucesores, la suma de cualquiera de ellos seguirá siendo un sucesor menor que ω.

Así, tomamos ω como dominio y asignamos a la constante 0 el elemento 0 de ω, al símbolo de función s la función 1-aria sucesor, y a los símbolos de función $+$ y $\cdot$, las funciones binarias suma y producto de naturales.

Veamos entonces que los axiomas de la teoría de la Aritmética de Peano se cumplen según la interpretación dada en 1.19:

- Los Axiomas de PA_2 a PA_5 se cumplen por definición de las operaciones entre ordinales.
- PA_6 se trata del Principio de Inducción, por lo que se cumple.
- Dado $n \in \omega$, si $n + 1 = 0$ entonces $n \in n + 1 = \emptyset$, lo que no es posible. Por tanto, para todo $n \in \omega$ no se puede dar que $n + 1 = 0$. De esto se obtiene PA_0
- Sean n y m naturales. Si $n + 1 = m + 1$, veamos que $n = m$. Supongamos que $n \neq m$. Como $n \in n + 1 = m + 1 = m \cup \{m\}$, tenemos que o bien $n \in m$ o bien $n \in \{m\}$. Como la segunda opción implica que $n = m$, lo que suponemos que no ocurre, tenemos que $n \in m$. Pero aplicando el mismo razonamiento a $m \in m + 1$ obtenemos $m \in n$. Sabemos que esto no es posible por el Axioma de Fundamento, de donde se concluye que $n = m$ y que por tanto se cumple PA_1.

□

Así, tenemos que el Lenguaje de la Teoría de Conjuntos es un lenguaje numerable en el que se pueden definir los símbolos de $\mathscr{L}_{PA}$ y tanto ZF como ZFC son teorías en las que se cumplen los axiomas de PA. Así, aplicando los Teoremas de Incompletitud de Gödel obtenemos el siguiente resultado.

Teorema 3.5. ***Teoremas de Incompletitud de Gödel para ZF y ZFC***
Si ZF es consistente, entonces es incompleta y $ZF \nvdash con_{ZF}$. Lo mismo ocurre con ZFC.

Así, se concluye que cualquier prueba de la consistencia de ZF no podría llevarse a cabo dentro de ZF, si no que debería realizarse en el marco de una teoría más amplia, cuya consistencia habría que asumir. Por otro lado, si se asume la consistencia de ZF, esta teoría sería incompleta, por lo que existirían sentencias que no se podrían probar en ZF. Es decir, tanto si asumimos la consistencia de ZF como si tratamos de probarla en una estructura mayor, llegará un punto en el que será necesario asumir cierta teoría como consistente, y dentro de esta siempre habrá hechos que, aunque puedan ser ciertos, no podrán llegar a ser demostrados.

Conclusiones

En este trabajo se propuso estudiar la Teoría Axiomática de Conjuntos a través de la Lógica Matemática, con el fin de presentar los fundamentos sobre los que se basan las distintas ramas de las matemáticas. A lo largo de estas páginas se han obtenido numerosas conclusiones, que se recopilan a continuación.

En primer lugar, se ha entendido la necesidad de una axiomatización de la Teoría de Conjuntos, evitando así las paradojas derivadas de la concepción puramente intuitiva. Con esto en mente, se han presentado las herramientas lógicas que permiten su construcción, explicando el lenguaje en el que se expresan todos los conceptos y las reglas que permiten obtener conclusiones válidas. En este contexto, se ha podido definir rigurosamente el concepto de axioma, haciendo ver que son aquellas afirmaciones que se toman como verdaderas y a partir de las cuales se deducen los resultados. Además, mediante los modelos, los Teoremas de Solidez y Completitud y los métodos de demostración, se ha asegurado que la forma de trabajar en mátemáticas, utilizando las interpretaciones intuitivas de los símbolos, está sustentada por la lógica.

Visto esto, se ha pasado a desarrollar la Teoría Axiomática de Zermelo-Fraenkel. Se ha visto que un conjunto no es más que un objeto matemático cumpliendo una serie de axiomas, y que al contrario de lo que se podría esperar, no toda colección de elementos constituye un conjunto. Se han presentado además los nueve axiomas que conforman la teoría. Únicamente asumiendo estos enunciados, se ha podido demostrar que conceptos habituales como el contenido, la unión o las partes de un conjunto existen y se comportan tal y como se conoce. También se han demsotrado hechos que en un principio se asumen como intuitivos, como la existencia y unicidad del conjunto vacío o la existencia de al menos un conjunto infinito. Además, se ha visto que conceptos en los que se basan gran parte de las matemáticas, como las funciones o las relaciones, se pueden construir en base a conjuntos.

Una de las conclusiones más interesantes que se deriva de estos axiomas es la construcción de los números naturales y su esencia. Se demuestra la existencia de estos números, así como la del conjunto formado por todos ellos y se observa que los naturales son también conjuntos, cuya característica principal es que contienen a sus anteriores.

También se ha visto que el Axioma de Elección es independiente de la teoría de Zermelo-Fraenkel.

Es decir, tanto añadirlo como negarlo no crea contradicciones. Así, por mucho que la naturaleza del enunciado pueda ser cuestionada, no existen razones puramente lógicas para decantarse por una de estas opciones.

Una vez determinada la Teoría Axiomática, se ha demostrado la existencia de un nuevo sistema numérico, los ordinales, que constituyen una extensión de los naturales y de su orden. Se ha presentado su aritmética, que engloba la conocida para los naturales, y se ha visto que su naturaleza se basa en continuar el proceso de construcción de los naturales y ω. Además, se ha visto cómo construir funciones de clase de forma recursiva, se ha demostrado el conocido Principio de Inducción y se ha visto la forma de extenderlo a todos los ordinales mediante el Principio de Inducción Transfinita, que constituye una herramienta de gran potencia.

Para terminar, se ha demostrado que el conjunto ω constituye un modelo de la Aritmética de Peano. Esto, a través de los Teoremas de Incompletitud de Gödel, implica que la consistencia de la Teoría de Zermelo-Fraenkel nunca podrá ser probado dentro de sí misma. Así, se manifiesta la necesidad de asumir, en algún punto, la consistencia de una teoría axiomática, aunque, aún así, siempre habrá ciertos hechos que no podrán ser probados, lo que refleja las limitaciones que tiene una teoría matemática. Sin embargo, el recuento de todo lo desarrollado es claramente positivo: asumiendo únicamenente nueve axiomas de caracter profundamente intuitivo, es posible construir prácticamente todos los conceptos matemáticos conocidos. Es decir, aunque puedan existir limitaciones, con este pequeño número de asunciones, que no dejan de replicar nuestra propia intuición, se está en posesión de los materiales necesarios y suficientes para constuir matemáticas que abarquen todos los conceptos deseados.

Referencias

Devlin, K. (1994). *The Joy of Sets: Fundamentals of Contemporary Set Theory*. Springer.

Halbeisen, L., & Krapf, R. (2020). *Gödel's Theorems and Zermelo's Axioms. A Firm Foundation of Mathematics*. Birkhäuser.

Halbeisen, L. (2012). *Combinatorial Set Theory*. Springer.

Hrbacek, K., & Jech, T. (2017). *Introduction to Set Theory. Revised and Expanded*. Crc Press.

Jech, T. (2003). *Set theory: The third millennium edition, revised and expanded*. Springer.

Pinter, C. C. (2014). *A book of set theory*. Dover Publications, INC.

Rusell, B., & Witehead, A. N. (1927). *Principia Mathematica*. Cambridge University Press.

San Román, J. S. (1990). *Lógica Matemática y Computabilidad*. Díaz de Santos, S. A.

Suppes, P. (1960). *Axiomatic Set Theory*. D. Van Nostrand Company, Inc.

Zermelo, E. (1967). Investigations in the Foundation of Set Theory I. En J. van Heijenoort (Ed. & Trad.), *From Frege to Gödel: A Source Book in Mathematical Logic, 1879-1931* (J. van Heijenoort, Ed. & Trad.; pp. 200-215). Harvard University Press. (Fecha inicial de publicación 1908)