UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE ESTUDIOS SUPERIORES
ACATLÁN

ALGUNAS APLICACIONES, EQUIVALENCIAS Y DEMOSTRACIONES RELACIONADAS CON EL AXIOMA DE ELECCIÓN

T E S I S

QUE PARA OBTENER EL TITULO DE
ACTUARIO

PRESENTA
OSVALDO ANGTUNCIO HERNÁNDEZ

Asesor: MIGUEL ANGEL SÁNCHEZ BARQUÍN

FEBRERO DE 2011

AGRADECIMIENTOS

A cada uno de los involucrados en la realización y mejora de esta tesis.

Índice general

LISTA DE ABREVIACIONES Y SÍMBOLOS

$\vee$	Abreviación para «o» (en el sentido de «los dos o alguno»)
$\wedge$	Abreviación para «y»
$\exists$	Abreviación para «para algún» (o «existe»)
$\forall$	Abreviación para «para todo»
$\Rightarrow$	Abreviación para «si – entonces» (o «implica»)
ssi, $\Leftrightarrow$	Abreviación para «si y solo si»
i.e.	Abreviación para «es decir»
$\|$, $:$	Tal que, con, se tiene que, con la propiedad
$p \vee q$	(léase: p o q) denota la disjunción de p y q
$p \wedge q$	(léase: p y q) denota la conjunción de p y q
$\neg\, q$	(léase: no q) denota la negación de q
$p \Rightarrow q$	(léase: p implica q). Denota «$(\neg p) \vee q$»
$p \Leftrightarrow q$	(léase p es el equivalente lógico a q, o también, p ssi q). Denota «$(p \Rightarrow q) \wedge (q \Rightarrow p)$»
$\mathbb{N}$	El conjunto de los números naturales
$\mathbb{R}$	El conjunto de los números reales
$\mathbb{R}^n$	El conjunto de las n-adas de números reales
$A \times B$	El producto cartesiano de A y B
$\aleph_0$	El cardinal de $\mathbb{N}$
$2^{\aleph_0}$	La potencia del continuo $\mathbb{R}$
ω	El menor de los ordinales infinitos
Ω	El menor de los ordinales incontables, también conocido como ω_1
$\|A\|$	El número cardinal del conjunto A
$\bigcup_{x \in B} A_x$	La unión de todos los conjuntos A_x, tales que $x \in B$
$\mathscr{P}(A)$	El conjunto potencia de A
$m(A)$	La medida del conjunto A
γ	Función de elección, o Zermelo´s covering
$W = On$	La clase de todos los ordinales
$L(S)$	El espacio generado por el conjunto S
G_n	El grupo de todas las isometrías de $\mathbb{R}^n$
$(\beta(X), p)$	Compactificación de Stone-Čech de X
$\overline{\mathbb{R}}$	Reales extendidos
$I(\mathbb{R})$	Conjunto de todos los subconjuntos integrables de $\mathbb{R}$
$B_d(x, r)$	d-bola de radio r y centro x
$A \cup B$	Unión de A y B, es $\{x \in V \| x \textit{ pertenece por lo menos a alguno de los conjuntos } A \text{ y } B\}$

$A \cap B$	Intersección de A y B, es $\{x \in V \vert x$ *pertenece a ambos conjuntos* A y $B\}$
$A - B$	Diferencia de A con B, es $\{x \in A \vert x \notin B\}$
$\mathscr{C}_A B$	Si $A \subset B$, es la operación *complemento* de B respecto de A, que es $A - B$
$A \Delta B$	La *diferencia simétrica* (o *suma Booleana*) de A y B, y es $(A-B) \cup (B-A)$
$\bigcup\limits_{\alpha} A_\alpha$	La unión de la familia $\{A_\alpha \vert \alpha \in \mathscr{A}\}$, y es $\{x \in V \vert \exists \alpha \in \mathscr{A} : x \in A_\alpha\}$
$\bigcap\limits_{\alpha} A_\alpha$	La intersección de la familia $\{A_\alpha \vert \alpha \in \mathscr{A}\}$, y es $\{x \in V \vert \forall \alpha \in \mathscr{A} : x \in A_\alpha\}$
$\mathscr{P}(A)$	El *conjunto potencia* de un conjunto A, es $\{B \vert B \subset A\}$
aRb	Si $(a,b) \in R$ se dice que a se relaciona o esta en R-relación, con b
Dom R	*Dominio* de una relación, es $\{x : \textit{para algún } y,\ (xRy)\}$
Ran R	*Rango* (también llamado *contradominio* o *dominio converso*) de una relación, es $\{y : \textit{para algún } x,\ (xRy)\}$
R^{-1}	La *inversa* de la relación R, es la relación $\{(y,x) \vert xRy\}$
$R\vert_A$	La *relación restringida* o *restricción* de la relación R en A y es $R \cap (A \times \mathit{Ran}\ R)$
$R(A)$	La *imagen* de un conjunto A bajo R, se define como $\mathit{Ran}\ (R\vert_A)$
$[a]$	Se llama la *clase de equivalencia de a* (con respecto a R) y es $\{b \in A \vert bRa\}$
$f(A)$	La *imagen* de A bajo f, con f una función, es $\{f(x) \vert x \in A\}$[1]
$f^{-1}(B)$	La *imagen inversa* de B bajo f, con f una función, es $\{x \in X \vert f(x) \in B\}$

Con V un conjunto dado, y $\{A_\alpha | \alpha \in \mathscr{A}\}$ una familia de subconjuntos de V (i.e. $\forall \alpha \in \mathscr{A},\ A_\alpha \subset V$). $\bigcap\limits_{\alpha} A_\alpha = \bigcap\{A_\alpha | \alpha \in \mathscr{A}\}$ y de la misma manera $\bigcup\limits_{\alpha} A_\alpha = \bigcup\{A_\alpha | \alpha \in \mathscr{A}\}$.

También se le llama *cuantificador universal* a $\forall$ y significa «para cada», y el *cuantificador existencial* también es el símbolo $\exists$.

INTRODUCCIÓN

El principal objetivo de esta tesis es de ser una referencia informativa. Se pretende que el lector adquiera conocimientos sobre los diversos temas matemáticos e históricos en donde se aplica el Axioma de Elección, encontrando demostraciones detalladas y entendiendo la historia del tema.

De un solo vistazo al índice de la tesis, uno se puede dar cuenta la vasta gama de temas donde esta presente el Axioma, dando pinceladas de conocimientos matemáticos en diferentes ramas como son: Matemáticas Generales, Álgebra, Topología, Análisis Matemático, Teoría de Cardinales, Teoría de Ordinales, Lógica Matemática y Teoría de Conjuntos, principalmente. Aclaro que solo se dan pinceladas, debido a que me enfoco a demostrar resultados importantes relacionados al Axioma en cada sección del capítulo 3, dando solamente los teoremas y definiciones que se requerirán para demostrar dicho resultado. Por otro lado, en la parte histórica y de comprensión del Axioma, doy información sobre el uso de éste, las formas en que se puede encontrar, las aplicaciones que surgen, las contradicciones, los puntos a favor y en contra, la justificación matemática y la humana.

En el capítulo 1, se enuncia la maquinería básica para la comprensión de los capítulos subsecuentes, se definen conceptos como familia de conjuntos, función, relación, número cardinal y se enuncian los sistemas axiomáticos más comunes ZFC y NBG. Además, se incluyen teoremas necesarios para demostraciones del capítulo 3, cuya necesidad es recurrente en varias secciones.

El capítulo 2 introduce al lector en el contexto histórico del Axioma, relatando desde que los filósofos adentraban sus pensamientos al concepto de infinito, hasta la culminación de la búsqueda de preguntas como ¿es posible crear un sistema matemático que sea capaz de deducir las verdades del universo?, ¿cuántos sistemas así son posibles?. Se muestran resultados que hacían imperante la aplicación del Axioma a la matemática de la época, debido a las consecuencias de dichos resultados. Con esto, se logra entender la necesidad de crear un "axioma" que ayude al matemático a sobreponerse a las dificultades del infinito, se entiende cómo y por qué se propuso por primera vez el Axioma, por qué fue tan rechazado y debatido, y cómo nos ayuda hoy en día. Se va conduciendo al lector por diferentes épocas, denotando lugares y personajes involucrados en la evolución del Axioma, escribiendo qué aportaron a la historia y sobre todo las dificultades que encontraron en el camino.

En el capítulo 3 se proporcionan algunas de las principales demostraciones relacionadas con el Axioma, mostrando al lector la influencia de éste en varias ramas de la matemática, ligando los teoremas de la rama al Axioma para crear resultados imprescindibles en dicha área. En cada sección los teoremas ayudan a profundizar el conocimiento de la rama, proporcionando en cada caso las definiciones y resultados pertinentes; se ayuda a esclarecer el lugar del teorema donde se usa una forma del Axioma; se ayuda a entender el uso frecuente de alguna forma del Axioma en las matemáticas modernas; y principalmente, se resuelve el objetivo de enseñar al lector, demostraciones en temas importantes para las matemáticas, mediante explicaciones detalladas, así como justificaciones sobre los teoremas y corolarios usados para la resolución de una afirmación.

Esta tesis le puede servir al lector, para darse una idea del contenido de temas como Álgebra, Topología y Teoría de Conjuntos; puede descubrir el gusto por algún tema en especial que no es profundizado o visto en la carrera. También el lector puede, aprender de demostraciones que aquí se dan, para aplicarlas él mismo de manera similar, o para incrementar su pensamiento y abstracción matemáticas.

PREFACIO

Como estudiante de la carrera, en materias como Álgebra Lineal y Análisis Matemático escuché hablar a mis profesores del Axioma de Elección, en Lineal aprendí que el Axioma de Elección enunciaba la existencia de una función que elegía elementos en conjuntos. La única demostración sobre el Axioma de Elección que vi en la carrera fue el teorema 3.1.2 (página 91) por mi asesor Barquín. Siempre tuve dudas sobre las aplicaciones y el significado del Axioma de Elección, para mi era un tema intrigante, porque nadie lo mencionaba con conocimientos mayores a los que se incluían en el libro de texto, pero sobre todo, porque sabía que aplicaba a conjuntos muy generales: para cualquier espacio vectorial podemos encontrar una base, el producto cartesiano no vacío de conjuntos no vacíos es no vacío. Lo anterior, aunado a mi curiosidad, me hicieron investigar y encontrar rápidamente temas interesantes y diversos en internet, leí que el Axioma de Elección implica la Paradoja de Banach-Tarski, que tiene más de 300 equivalencias, que está relacionado con la Hipótesis de Continuo, con la posibilidad de contar más allá de infinito, con teoremas sobre Ideales y Ultrafiltros, teoremas sobre Análisis Real y Funcionales Lineales, sobre Teoría de Anillos, Teoría de Demostraciones, sobre Lógica Matemática, teoremas de Categoría de Baire, con la existencia de conjuntos que no son Lebesgue Medibles, la existencia de un buen orden en conjuntos, el Lema de Zorn, y sobre todo, su relación con la teoría axiomática en la que descansan las matemáticas modernas.

Las justificaciones de la elección de este tema son su vasta aplicación y su importancia. El campo donde se ocupa actualmente el Axioma es enorme, reducido únicamente por la postura del lector (el constructivista no acepta el Axioma de Elección). La importancia que tiene el tema es descrita en varios lados de la tesis, es considerado el segundo problema más importante y controversial de las matemáticas, rebasado solo por el quinto postulado de Euclides, y esa, para mi, es razón suficiente para querer investigar del tema.

Creo que es necesaria dar toda esta información, porque justifica varios procedimientos, de varias materias impartidas en la carrera: en Cálculo Diferencial e Integral nos enseñan que la unión de conjuntos contables es contable, que $\mathbb{R}^2$ es isomorfo a $\mathbb{R}$; en Álgebra Superior nos enseñan que las clases de equivalencia de un conjunto tienen representantes, los cuales pueden ser infinitos; en Álgebra Lineal nos advierten que existe una diferencia importante entre espacios vectoriales de dimensión finita e infinita, incluso a veces, nos dicen que se puede encontrar una base para cualquier espacio vectorial; en procesos estocásticos, utilizamos Cadenas de Markov, las cuales resultan de sucesiones infinitas que tienen términos dependientes unos de los anteriores; y principalmente, en análisis matemático, ocupamos sin saber, o sin darnos cuenta, muchísimas veces al Axioma, en diferentes formas, en métodos variados, en demostraciones donde se encuentra escondido para el que no está acostumbrado o desconoce del Axioma. Aquél que tiene la curiosidad de saber qué hay detrás de él, porqué se ocupa tanto y en tantos lugares, esta tesis absuelve el misterio detrás del Axioma de Elección.

En esta tesis doy a conocer al lector el Axioma de Elección desde un punto de vista general, enseñando al lector los hechos desde una perspectiva externa, sin inclinación a una postura filosófica, relatando la historia del mismo para entender su necesidad y controversia, remarcando cuándo es necesario y cuándo se puede evitar.

Debido a la amplia diversificación del tema, en la mayoría de los resultados importantes se citará el texto donde la información fue obtenida para futuras referencias y la página donde se encuentra, esto se hará entre corchetes. Por ejemplo, para citar la definición que da Gregory H. Moore en su libro «Zermelo´s Axiom of Coice, its Origins, Development, and Influence» de *punto límite de A* que aparece en la página 17, se denotará como (véase [Moo, p. 17]), y la bibliografía completa del libro aparecerá en la parte de referencias de esta tesis (p. 165).

Se enunciarán en cada tema las definiciones necesarias para la comprensión del mismo, suponiendo solo un conocimiento general en las materias enseñadas en la carrera de Actuaría. La bibliografía principalmente se compone de libros descargables gratuitamente en internet, los cuales se consiguen en formato .pdf o en formato .djvu, y se pueden abrir usando el programa Adobe Reader® o DjVu Viewer® respectivamente. Algunos libros son muy difíciles de conseguir, pero la mayoría los bajé de internet, además, anexo páginas de internet donde conseguí libros de consulta necesaria para el lector interesado en el tema (véase página 165).

Debido a mi formación como Actuario, mi posición sobre los temas será general, las demostraciones no irán inclinadas a algún tema en particular y se expondrán las que más fáciles y relacionadas con los temas vistos en la carrera haya encontrado, esto debido a que existen autores que su especialidad se basa en Lógica Matemática o Teoría Axiomática de Conjuntos por ejemplo, y ellos basan sus demostraciones en sus conocimientos respecto a sus temas de interés.

Al dar una definición, se citará la palabra en inglés si la fuente está en ese idioma, para que, cuando el lector consulte fuentes de información del tema, no encuentre ambigüedades. Se usará la palabra en inglés o en su idioma original cuando no conozca traducción o haya encontrado varias palabras para definir un concepto, por ejemplo, en su libro, Gregory H. Moore (véase [Moo, p. 17]), define un *limit point of a set* A como un punto p en el que toda vecindad[2] de él contiene un punto q en $A-\{p\}$, pero en su libro «Topology», Dugunji (véase [Dug, p. 70]) define un *cluster point of a set* A como un punto p tal que cada vecindad de éste contenga al menos un punto de A distinto de p. Además, en la p. 210, Dugunji define a un *limit point* p de una secuencia[3] $\{y_n\}$ si para toda vecindad de p, existe un N, tal que para toda $n>N$, y_n esta en la vecindad; y define a un *accumulation point* p si para toda vecindad de p y para toda N, existe un n tal que $n>N$ y además y_n esta en la vecindad. Dugunji menciona que los conceptos de *limit point, accumulation point and cluster point* deben ser delineados de cada uno, debido a que *p is a limit point* $\Rightarrow$ *p is an accumulation point*, y en espacios de Hausdorff, *p is a cluster point* $\Rightarrow$ *p is an accumulation point*, pero ninguna de esas implicaciones es reversible.

Se ocupará un término que ocupa W. Sierpinski en [Sie, p. 29]. Afirma que para demostrar que dos conjuntos son equivalentes es necesario y suficiente demostrar que existe una función biyectiva que tenga de dominio a uno y contradominio al otro. Una demostración indirecta es posible, o deducir su existencia de teoremas antes demostrados también es válido, y la definición de equivalencia no pide encontrar a dicha función biyectiva y dar un ejemplo de ella, o de establecerla, solo demostrar que existe. Si se establece una función biyectiva entre los dos conjuntos (se da dicha función o se encuentra un ejemplo de alguna) se dice que los conjuntos son efectivamente equivalentes. Dicho uso de efectividad se corresponde a las matemáticas, y se ocupa no solo para definiciones de equivalencia, sino para definiciones en las que un objeto matemático existe y no se ha encontrado un ejemplo de ello, o se ha encontrado tal. Una *forma del Axioma* puede ser una equivalencia del Axioma (la forma implica al Axioma y viceversa), una implicación del Axioma (el Axioma implica dicho teorema pero no viceversa, y dicho teorema no puede ser deducido de los demás Axiomas), o una generalización del Axioma (dicha forma es una axioma e

[2]Véase la definición 3.6.3.

[3]**Secuencia o Sucesión=** Una secuencia en un espacio Y es una función $\phi:\mathbb{N}\mapsto Y$, también denotada $\{\phi_n\}$.

implica al Axioma de Elección pero no viceversa).

El Axioma de Elección será aquí llamado el Axioma. Para definiciones muy utilizadas en el texto, se usan abreviaciones que se señalan después de haber definido el concepto, el lector puede encontrar en el Índice Alfabético (p. 169) la mayoría de los conceptos importantes.

ABSTRACT

En esta tesis se da un panorama general de la historia, importancia, necesidad, utilidad y desventajas que ofrece el Axioma de Elección desde varias perspectivas. Se abordará la historia del Axioma de Elección, desde sus inicios hasta los diferentes puntos de vista que surgieron al haberse comprobado que el Axioma es independiente de los demás axiomas de ZF. Se demostrarán resultados importantes de diversas ramas de las matemáticas como La Paradoja de Banach-Tarski en Topología, El Principio del Buen Orden y el Lema de Zorn en Teoría de Conjuntos, La Tricotomía de los Cardinales en Teoría de Cardinales, la existencia de una base para todo Espacio Vectorial en Álgebra, la existencia de conjuntos no Lebesgue Medibles en Análisis. Se comprenderá la utilidad del Axioma en la carrera.

Capítulo 1

PRELIMINARES MATEMÁTICAS

Este capítulo introduce los conceptos, notación y resultados básicos matemáticos para comprender mejor los subsecuentes capítulos. Se utiliza notación actual y cuando en la tesis se utilice una diferente para respetar la notación del autor original, se dirá explícitamente. La notación de lógica simbólica se empleará para hacer más breves los contenidos después de la introducción de la tesis. Se usará mucha redacción al principio para ir acorde con los contenidos (más históricos que matemáticos) de la segunda parte. Estos preliminares son solo informales, la aproximación axiomática se verá en la sección 1.8. La sección 1.1 introduce el concepto fundamental, recurrente y vago de conjunto, se sigue con la forma en que manipulamos los conjuntos y los denotamos (1.2 y 1.3), con definiciones y teoremas sobre relaciones y funciones (1.4, 1.5 y 1.6). En las siguientes secciones se habla de los fundamentos del Axioma, de cómo puede ser tratado y usado junto con otros axiomas en las matemáticas modernas, mostrando donde descansan las reglas matemáticas que la mayoría usamos, hablo de las diferentes teorías de conjuntos (1.8), pero dichas teorías se fundamentan en una forma lógica y formal de escribir y deducir proposiciones, basándose en los lenguajes de la lógica (1.7). Al final del capítulo se introducen nuevos temas para la carrera: los Cardinales (1.9) necesarios para contar mucho más allá del número de reales, para agrupar los conjuntos; y la Inducción Transfinita (1.10) principal para hacer generalizaciones como las que hace la inducción matemática pero en conjuntos más grandes.

1.1. CONJUNTOS

Si quisiéramos definir la palabra *conjunto* utilizaríamos sinónimos para hacerlo: colección de objetos, elementos agrupados por tal característica, etc., cayendo en un círculo debido a que no tenemos palabras precisas para definirlo. Los conjuntos, como son identificados, tienen elementos, intuitivamente, son una colección[1] de cosas u objetos que satisfacen cierta propiedad, e incluso, sus elementos pueden ser conjuntos como el conjunto que contiene al conjunto que tiene como elemento al número 8. Lo que es sorprendente no es que un conjunto pueda tener como elementos a conjuntos, sino que para propósitos matemáticos ¡ningún otro objeto es considerado!; se estudian conjuntos, conjuntos de conjuntos, y anidaciones espeluznantes e increíblemente complejas, pero nada mas, solo conjuntos[2].

Formalmente, la palabra conjunto no tiene definición en matemáticas. Lo que hacemos es dar ciertas propiedades que creemos claro debería de tener un conjunto, y con ellas desarrollamos la teoría sobre el mismo. Algunos se preguntarán: ¿cómo trabajar con algo que ni siquiera esta definido?, pues esto ha sido

[1] La palabra «*familia*» es un sinónimo común de conjunto, como también «*colección*».

[2] Claro, también existen teorías en las que se introducen nuevos elementos, por ejemplo, *átomos*, *constantes* y las famosas *clases* (véanse las páginas 30 y 30).

así desde hace mucho tiempo. Debido a la ambigüedad que podría generar, utilizamos definiciones para reconocer objetos, pero no siempre es así, hay veces solo decimos qué debería hacer o qué hace dicho objeto, y así se sigue manejando. Un ejemplo clásico es en la Geometría Elemental: tampoco existe una definición de punto, ni de línea, pero decimos qué hacen éstos, y que podemos nosotros crear con esas cosas, sean lo que sean. Halmos dice acerca de esto (véase [Hal, p. 1]): «*The semi-axiomatic point of view adopted here assumes that the reader has de ordinary, human, intuitive (and frequently erroneous) understanding of what sets are; the purpose of the exposition is to delineate some of the many things that one can correctly do with them*».

¿Y cuáles son dichas propiedades de lo que sea que es un «conjunto»?, una importante es que entre sus elementos no hay orden, en el conjunto no nos importa que elemento esta adelante de quién, quién es el primero, ni si están numerados. Otra pregunta relevante, ¿todo objeto puede ser un «conjunto»?: entendiéndose por objeto a cualquier cosa, incluso no imaginable, no física, no pensable. Aquí hay que detenernos un momento, es posible pensar en ¿el conjunto de todas las cosas inimaginables?, ¿tiene esto sentido?, una vez expresó Hardy: «*Even if we knew that it was impossible ever to define a single member of a class, it would not of course follow that members of the class did not exist*» [Her]; y Schröder investigando sobre el *Ordering Principle* sospechaba que sería imposible ordenar el dominio de *todo lo pensable*, el *absolute Denkbereich* (véase [Moo, p. 48]); también Dedekind en 1888 usó el dominio de todo lo pensable para justificar la existencia de un conjunto infinito. Un matemático prominente como Russell, escribió algo al respecto, como una razón para aceptar el Axioma multiplicativo de la página 51:

> *It is obvious* a priori *that there are existent (non-empty) classes of which no member can be specified: e.g., the class of indefinable ordinals. If one could generalize the multiplicative axiom, ... Thus we could do with a general axiom giving us members of classes in cases where we cannot specify any members. Such an axiom might well be legitimate.*[3]

Imagine lector el conjunto (si es que lo fuese) de todas las cosas «impensables», sus elementos serían cosas en las que no podemos pensar, y no podríamos definir alguno solo, si lo hiciéramos en algún momento, el elemento ¡dejaría de pertenecer al conjunto!. Entonces, para responder la pregunta (con un ¡no!), necesitamos conceptos importantes que restringirán la clase de todas las cosas que pueden ser conjuntos.

Este tipo de incongruencias fueron las que llevaron a los matemáticos a axiomatizar su ciencia, y a encontrarse con problemas como el Axioma de Elección, al no saber si era axioma o teorema.

Primero se define un universo de discurso, llamémoslo V, es como el conjunto de donde vamos a deducir propiedades, de donde sacaremos los elementos y los subconjuntos. Éste se tiene que definir porque el hecho de no hacerlo puede llevar a paradojas o a incongruencias (véase 24). El concepto primitivo, elemental e indefinido con el que se empieza es el de *pertenencia*. Se escribe «$x \in A$» y se dice que x pertenece, es un elemento, un punto o un miembro de A. Dentro de las siguientes anotaciones, se tomará la convención de que A, B y C serán subconjuntos de V. La relación de pertenencia se denota por $\in$, los elementos de un conjunto generalmente serán letras minúsculas, como: a, b, $c,\ldots$, y los conjuntos por mayúsculas A, B, C, $\ldots$. Para diferenciar la jerarquía entre elementos, conjuntos que los contienen a éstos, y conjuntos que los contienen a éstos (conjuntos de conjuntos) a veces se usan letras más complicadas, por ejemplo, $x \in A$, $A \in X$ y $X \in \mathscr{C}$. La igualdad de conjuntos se demuestra haciendo ver que un conjunto está contenido en el otro y viceversa, esto es por el siguiente axioma:

Axioma 1.1.1 (Axioma de Extensión). *Dos conjuntos son iguales si y solo si tienen los mismos elementos.*

[3]Sacado en [Moo, p. 131], escrito en un artículo de 1906 en las páginas 4-7 llamado «Multiplicative Axiom», no publicado, pero guardado en *Russell Archives*.

De esta forma, aún sin tener la definición de lo que es un conjunto, de este axioma tenemos tres propiedades que debe cumplir sea lo que sea, $\{a,b\} = \{b,a\}$ ya que el axioma anterior nos lo afirma al tener elementos iguales, la otra es que $\{a\} = \{a,a\}$ (que de la misma manera, tienen los mismos elementos), y la otra es que solo nos importa qué elementos tiene, no cómo vienen ordenados ni siquiera si están ordenados, ya que $\{a,b\} = \{b,a\}$.

Daremos una forma básica de definir lo que significa *sentencia*. Hay dos tipos básicos de sentencias, afirmaciones de pertenencia $x \in A$ y de igualdad $A = B$; todas las demás sentencias son obtenidas mediante tales sentencias «*atómicas*» por aplicaciones repetidas de los *operadores lógicos usuales* como los operadores para todo o negación.

Una expresión $p(x)$ que se convierte en proposición siempre que los valores de un dominio especificado sean sustituidos por x, es llamada función proposicional, o condición en x; y p es llamada propiedad o predicado. La afirmación «y tiene la propiedad p» significa que «$p(y)$» es verdadera.

La mayoría de los axiomas son para formar conjuntos a partir de otros. Pero para formar conjuntos de un conjunto dado, hay dos formas: el Axioma de Elección y el Axioma de Especificación. El segundo afirma que podemos sacar un subconjunto de uno dado mediante la aplicación de una propiedad, en otras palabras, para todo conjunto A, y para toda condición $S(x)$, existe el conjunto B que consta de los elementos de A para los cuales la sentencia $S(x)$ se cumple, considerando que cada elemento de A tiene o no tiene la propiedad. Este subconjunto esta únicamente determinado por S y viene dado por $B = \{x \in A | S(x)\}$ (léase: todos los x en A tal que $S(x)$ es verdadera). Formalmente:

Axioma 1.1.2 (Axioma de Especificación)**.** *Se puede obtener el subconjunto $B = \{x \in A | S(x)\}$ de todo conjunto A y toda condición dada $S(x)$.*

La «condición» es una sentencia, la letra «x» está libre en la sentencia $S(x)$, lo que significa que ocurre por lo menos una vez en $S(x)$ y no está encajada en alguna frase «para algún x» o «para todo x».

Con esto en mente, tenemos varias formas de definir un conjunto. Si A_α son los elementos de un conjunto A y α varía sobre un conjunto de índices $\mathscr{A}$, se escribirá $A = \{A_\alpha\}_{\alpha \in \mathscr{A}}$ (véase 1.3). Si A esta formado por elementos que cumplen cierta propiedad p, se escribirá $A = \{x \in V | p(x)\}$, que es el conjunto de todos los elementos que tienen la propiedad p.

1.2. ÁLGEBRA BOOLEANA

Se mencionará un poco de la estructura en la que se basa la Teoría de Conjuntos, esta definición es tomada del libro «A survey in Modern Algebra» (véase [BirMac, p. 361]).

Definición 1.2.1. *Un Álgebra Booleana B es un conjunto compuesto de dos operaciones binarias $\wedge$ (llamada en inglés wedge) y $\vee$ (llamada en inglés vee), y una operación unaria $'$ de complementación, que satisfacen los siguientes axiomas, donde a, b, y c varían sobre todos los elementos de B:*

1. *$\wedge$ y $\vee$ satisfacen las leyes:*

   idempotentes
$$a \wedge a = a \vee a = a,$$
   conmutativas, i.e.
$$a \wedge b = b \wedge a \qquad \text{y también} \qquad a \vee b = b \vee a,$$
   asociativas
$$a \wedge (b \wedge c) = (a \wedge b) \wedge c \qquad \text{y} \qquad a \vee (b \vee c) = (a \vee b) \vee c.$$

2. *$\wedge$ y $\vee$ satisfacen las leyes de absorción*

$$a \wedge (a \vee b) = a \vee (a \wedge b) = a.$$

3. *Esas dos operaciones son mutuamente distributivas:*

$$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c), \qquad a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c).$$

4. *Existen elementos únicos en B denotados por O y I, tales que:*

$$a \vee O = a, \qquad a \wedge O = O, \qquad a \wedge I = a, \qquad a \vee I = I.$$

5. *La operación unaria ´ cumple*

$$a \wedge a' = O, \qquad y \qquad a \vee a' = I.$$

La colección de todos los subconjuntos de un conjunto fijo V, con $\vee$, $\wedge$, ´ , O, I interpretados como $\cup$, $\cap$, $\mathscr{C}_V$ (o escrito también como el superíndice ^{c} cuando V se sobreentiende), $\emptyset$, V, respectivamente, forman un Álgebra Booleana.

1.3. FAMILIAS DE CONJUNTOS

Si para cada elemento α de algún conjunto $\mathscr{A} \neq \emptyset$ corresponde un conjunto A_α, entonces la colección de conjuntos $\{A_\alpha | \alpha \in \mathscr{A}\}$ es llamada una *familia de conjuntos* y $\mathscr{A}$ es llamado un *conjunto de índices* para la familia[4]. Cualquier conjunto $\mathscr{F}$ puede ser convertido a una familia de conjuntos, simplemente «auto-indexamos»; usas a $\mathscr{F}$ como el conjunto de índices y asignas a cada conjunto de la colección el elemento que representa, i.e., si $A \in \mathscr{F}$, renombramos $A_A = A$, y así la familia de conjuntos sería $\{A_A | A \in \mathscr{F}\}$.

Se llamará a $\{A_\alpha | \alpha \in \mathscr{A}\}$ una *familia no vacía* si el conjunto de índices cumple con $\mathscr{A} \neq \emptyset$, y claro, una *familia de conjuntos no vacíos* será tal que $\forall \alpha \in \mathscr{A}$, $A_\alpha \neq \emptyset$.

Una propiedad de las familias de conjuntos importante. Si tenemos $\{A_i\}$ con $i \in \{1, 2, \ldots, n\}$ y $n > 1$, tal que $\bigcup A_i = X$, siempre se puede encontrar una familia $\{B_i\}$ de conjuntos disjuntos que tiene la misma unión que la familia $\{A_i\}$, a saber, si se define

$$B_1 = A_1 \text{ y}$$

$$B_i = A_i - \bigcup_{k=1}^{i-1} A_k \text{ para } 2 \leq i \leq n,$$

la familia $\{B_i\}$ con $i \in \{1, 2, \ldots, n\}$, cumplirá estas tres propiedades:

1. $B_i \subset A_i$,
2. $B_i \cap B_j = \emptyset$, con $1 \leq i < j \leq n$, es decir, para todos los índices distintos, y por último
3. $\bigcup_{k=1}^{n} A_k = \bigcup_{k=1}^{n} B_k$.

[4]Como se mencionó en la sección 1.1 p. 11 familia es sinónimo de conjunto, y así familia de conjuntos también significa *conjunto de conjuntos*.

1.4. RELACIONES

Por *relación* es lo que intuitivamente relaciona a dos o varias cosas, i.e., una proposición tal que para parejas ordenadas (a,b) uno puede decir si a se relaciona con b o no. A veces a una relación se le llama *relación binaria* dando a entender que requiere de dos argumentos.

Definición 1.4.1. *Una relación R en un conjunto A es un subconjunto $R \subset A \times A$. Si $(a,b) \in R$ se escribe aRb.*

Si $R = X \times Y$, entonces $Dom\ R = X$ y $Ran\ R = Y$. Si R es una relación incluida en el producto cartesiano $X \times Y$ (i.e., $Dom\ R \subset X$ y $Ran\ R \subset Y$), se dice que R es una relación *de X a Y*, si $X = Y$ se dice que R es una relación *en X*.

Un tipo de relación muy importante, y usada en la tesis es:

Definición 1.4.2. *Una relación R en A es una* relación de equivalencia *si:*

1. $\forall a \in A : aRa$ *(reflexiva).*

2. $(aRb) \Rightarrow (bRa)$ *(simétrica).*

3. $(aRb) \wedge (bRc) \Rightarrow (aRc)$ *(transitiva).*

Si aRb se dice que a y b son equivalentes *y se denota por $a \sim b$.*

Una *partición* de un conjunto X es una colección disjunta $\mathscr{C}$ de conjuntos no vacíos de X cuya unión es X. Más matemáticamente, si A es una familia no vacía, la familia de subconjuntos de A, $\{A_\alpha | \alpha \in \mathscr{A}\}$ es una partición de A si:

$$\begin{array}{rcll} A_\alpha & \neq & \emptyset & \text{para cada } \alpha \in \mathscr{A}, \\ \bigcup\limits_{\alpha \in \mathscr{A}} A_\alpha & = & A, & \\ A_\alpha \cap A_\beta & = & \emptyset & \text{para todo } \alpha \neq \beta \in \mathscr{A} \end{array}$$

La tercera condición es la que nos dice qué es una *colección disjunta* de conjuntos.

Teorema 1.4.1. *Sea R una relación de equivalencia de A. La colección de clases de equivalencia distintas particiona a A en conjuntos mutuamente excluyentes, llamados* clases de R-equivalencia*, que cumplen con que dos elementos de A pertenecen a una clase de R-equivalencia común ssi ellos son equivalentes.*

Todo elemento a de una clase de R-equivalencia A_α es llamado *representativo* o *representante* de A_α, y $a \in A_\alpha \Leftrightarrow A_\alpha = [a]$.

Por último, una definición de un concepto muy ocupado en varias ramas de la matemática:

Definición 1.4.3. *Una* estructura *es una pareja (A,R) consistiendo de un conjunto A y una relación binaria R en A (i.e. $R \subset A \times A$).*

1.5. FUNCIONES

Este concepto es sin duda de los mas importantes en matemáticas, tiene aplicaciones tan diversas y fundamentales, que se debe poner especial atención a los conceptos relacionados con él. Desde el siglo XVIII, la generalización y clarificación del concepto de función han tenido mucha fama; pero ha tenido sus baches, la representación de Fourier de funciones «arbitrarias» por series trigonométricas encontró oposición, y cuando Weierstrass y Riemann dieron ejemplos de funciones continuas pero derivables en ningún punto, muchos se rehusaron a considerarlas seriamente.

La función *identidad* sobre X es $x \mapsto x$, y se denota por $1 : X \mapsto X$ o 1_X. Si a la función identidad, cambiamos el dominio por $A \subset X$, la función resultante es $i : A \mapsto X$, llamada función *inclusión* o *encaje* de A en X. Si $f : Y \mapsto Z$ y $X \subset Y$, se construye la función g con $f(x) = g(x), \forall x \in X$; así, a g se le llama *restricción* de f en X, y f es llamada la *extensión* de g en Y. Si $A \subset X$, la *función característica* de A es:

$$\chi_A(x) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \in A^c \end{cases}$$

«Y^X» denotará el conjunto de todas las funciones de X a Y, o en términos serios, $Y^X = \{f : f \; es \; función \wedge f : X \mapsto Y\}$. A veces, se escribe al conjunto de todas las funciones de X a Y como ${}^XY$.

Para dos conjuntos A y B en los que existe una función biyectiva entre ellos $f : A \mapsto B$, se dice que son *equivalentes* o *equipotentes* o que tienen la misma *cardinalidad*, y se escribe $A \sim B$[5]. Los conjuntos se dicen *comparables* si uno es equipotente con un subconjunto del otro, es decir, $A \sim C$ o $B \sim D$ con $C \subset B$ y $D \subset A$, esto se escribirá $A \preceq B$ o $B \preceq A$ respectivamente.

Es lo mismo decir que si $A \preceq B$, entonces existe una función inyectiva $f : A \mapsto B$. Se escribe $A \prec B$ significando que $A \preceq B$ pero no $A \sim B$, o lo que es lo mismo, existe una función biyectiva de A a un subconjunto de B pero no existe función biyectiva alguna de B a A. Una nota importante es aclarar que $A \prec B$ no es lo mismo que decir que existe una función biyectiva de A a un subconjunto propio de B.

La imagen inversa cumple que:

Teorema 1.5.1. *Sea $f : X \mapsto Y$, la inducida $f^{-1} : \mathscr{P}(Y) \mapsto \mathscr{P}(X)$ preserva operaciones elementales sobre conjuntos:*

$$\begin{aligned} f^{-1}(\bigcup_\alpha B_\alpha) &= \bigcup_\alpha f^{-1}(B_\alpha) \\ f^{-1}(\bigcap_\alpha B_\alpha) &= \bigcap_\alpha f^{-1}(B_\alpha) \\ f^{-1}(A-B) &= f^{-1}(A) - f^{-1}(B) \\ f^{-1}(A^c) &= (f^{-1}(A))^c \end{aligned}$$

Teorema 1.5.2. *Con las mismas condiciones que el teorema anterior pero tomamos a $f : \mathscr{P}(X) \mapsto \mathscr{P}(Y)$:*

$$\begin{aligned} f(\bigcup_\alpha B_\alpha) &= \bigcup_\alpha f(B_\alpha) \\ f(\bigcap_\alpha B_\alpha) &\subset \bigcap_\alpha f(B_\alpha) \end{aligned}$$

[5]Otras notaciones son escribir $\overline{\overline{A}} = \overline{\overline{B}}$ o también una más reciente $|A| = |B|$.

Llamaremos a dos funciones f y g compatibles, si se cumple que $f(x) = g(x)$ para toda $x \in dom\ f \cap dom\ g$. Un sistema de funciones compatibles será un conjunto de funciones las cuales cualesquiera dos son compatibles. Se cumple la proposición de que dos funciones son compatibles ssi su unión es compatible.

Teorema 1.5.3. *Si F es un sistema de funciones compatibles, entonces $\bigcup F$ es una función con dominio $dom\ (\bigcup F) = \bigcup\{dom\ f | f \in F\}$.*

1.6. RELACIONES DE ORDEN

Las definiciones básicas de la Teoría del Orden son simples e intuitivas, solo hay que tener en cuenta que las nociones vienen de las propiedades familiares de «menor que», «igual» y «mayor que». Un *orden* es cierto tipo de relación sobre un conjunto, generalmente se denota por «$\leq$» o por «$\preceq$». Empezamos con una relación R sobre un conjunto A:

Definición 1.6.1. *R es* reflexiva *en $A \Leftrightarrow \forall x \in A \Rightarrow\ xRx$.*

Definición 1.6.2. *R es* antirreflexiva *en $A \Leftrightarrow \forall x \in A \Rightarrow \neg(xRx)$.*

Definición 1.6.3. *R es* simétrica *en $A \Leftrightarrow (\forall x,\ y \in A)\ \wedge\ (xRy) \Rightarrow\ yRx$.*

Definición 1.6.4. *R es* asimétrica *en $A \Leftrightarrow (\forall x,\ y \in A)\ \wedge\ (xRy) \Rightarrow\ \neg(yRx)$.*

Definición 1.6.5. *R es* antisimétrica *en $A \Leftrightarrow (\forall x,\ y \in A)\ \wedge\ (xRy)\ \wedge\ (yRx) \Rightarrow\ y = x$.*

Definición 1.6.6. *R es* transitiva *en $A \Leftrightarrow (\forall x,\ y,\ z \in A)\ \wedge\ (xRy)\ \wedge\ (yRz) \Rightarrow xRz$.*

Definición 1.6.7. *R es* conexa *en $A \Leftrightarrow (\forall x,\ y \in A)\ \wedge\ (x \neq y)\ \Rightarrow (xRy)\ \vee\ (yRx)$.*

Definición 1.6.8. *R es* fuertemente conexa[6] *en $A \Leftrightarrow (\forall x,\ y \in A)\ \Rightarrow\ (xRy)\ \vee\ (yRx)$.*

Con estas definiciones se introducen 5 tipos de ordenes:

Definición 1.6.9. *R es un* cuasi-orden[7] *de $A \Leftrightarrow R$ es reflexivo y transitivo en A.*

Definición 1.6.10. *R es un* orden parcial[8] *de $A \Leftrightarrow R$ es reflexivo, antisimétrico y transitivo en A.*

Definición 1.6.11. *R es un* orden simple *de $A \Leftrightarrow R$ es antisimétrico, transitivo y fuertemente conexo en A.*

Definición 1.6.12. *R es un* orden parcial estricto *de $A \Leftrightarrow R$ es asimétrico y transitivo en A.*

Definición 1.6.13. *R es un* orden simple estricto[9] *de $A \Leftrightarrow R$ es asimétrico, transitivo y conexo en A.*

---

[6] También llamada *cadena*. *Anticadena* se significa que $x \leq y$ implique que $x = y$ para todos $x, y \in A$.

[7] También llamado *preorden*.

[8] Algunos autores le llaman a este orden con las características mencionadas, simplemente orden y al conjunto (A,R) le llaman un conjunto ordenado, sin embargo, puede confundirse con creer que el conjunto tiene un orden de otro tipo, por lo tanto utilizaremos el término orden parcial.

[9] Sierpinski en [Sie, p. 188] define un conjunto ordenado de ésta forma, lo utilizaremos en la sección de equivalencias.

En el estudio de ordenes parciales, resulta que éstos tienen muchas propiedades parecidas como las que tienen los símbolos $<$, $\leq$, $>$ y $\geq$ definidos como comúnmente los ocupamos para los números. No hay convenio sobre ocupar ordenes débiles $\leq$ u ordenes estrictos $<$, para un orden parcial el primero pide reflexividad, mientras que para un orden parcial el segundo pediría irreflexividad. La diferencia entre definiciones radica en qué tipo de estrictez se quiere trabajar, y se debe tener cuidado en cómo se definen estos conceptos en los libros.

De la definición (1.6.10) es importante decir que un *conjunto parcialmente ordenado* es un conjunto junto con un orden parcial en él. La formulación precisa es: una conjunto parcialmente ordenado es un par ordenado $(X, \leq)$, donde X es un conjunto y $\leq$ un orden parcial en X. Sin embargo, esta definición es redundante, debido a que dado un orden parcial, se puede encontrar su dominio; desgraciadamente, en materia de lenguaje y notación, casi siempre la costumbre gana sobre la razón pura. Esto es así en muchas partes de las matemáticas, se dan parejas ordenadas que se refieren a una estructura basada en un conjunto y a una operación, una función e incluso otro conjunto en él, las estructuras como se definen en matemáticas están en la página 15; tenemos varios ejemplos: en análisis se usan los espacios topológicos (X, τ) o espacios métricos (X, d), en álgebra espacios vectoriales con producto interno $(V, <\ ,\ >)$, en probabilidad espacios de probabilidad $(X, \mathscr{P})$. Debido a esto, seguiremos la costumbre y diremos «sea X un conjunto parcialmente ordenado» cuando en realidad queremos decir: «sea X el dominio de un orden parcial».

Un conjunto con un preorden definido se llama un *conjunto preordenado*. Si el orden es «$\preceq$», y consideramos a $\mathscr{P}(X)$, con $A \preceq B$ ssi $A \subset B$, entonces se tiene un preorden, y cualquier familia de conjuntos ordenada de esta manera es llamada *ordenada por inclusión*.

Un *conjunto totalmente ordenado* o también llamado un *conjunto linealmente ordenado* es un conjunto con un orden parcial, junto con una cadena, es decir: un conjunto con una relación reflexiva, antisimétrica, transitiva y en que cualesquier elementos se relacionan.

La teoría usa definiciones cuya explicación técnica es muy intuitiva y casi la misma palabra se explica a si misma. Los ordenes parciales generalmente se denotan por $\leq$ en vez de R. Si tenemos $(A, \leq)$ (o (A,R)) y a, $b \in A$ y además $a \leq b$, se dice que «a es menor o igual que b» o «b es mayor o igual que a» en el orden $\leq$. Se dirá que a y b son *comparables*[10] en el orden $\leq$ si se tiene que $a \leq b$ o que $b \leq a$, serán *incomparables* si se tiene que no son comparables (ni $a \leq b$ y ni $b \leq a$). Al orden fuertemente conexo o cadena también se define de otra forma. Si se tiene un $B \subset A$ y A es un conjunto parcialmente ordenado, B será una *cadena* de A si cualquier pareja de elementos de B es comparable (claro, respecto al orden de A).

Si X es un conjunto parcialmente ordenado y $x, y \in X$, «$y \geq x$» significará «$x \leq y$», en otras palabras «$\geq$» es la relación inversa de «$\leq$». Si $x \leq y$ y $x \neq y$ se escribe $x < y$ y se dice q x es mas *pequeño* o *menor* que y, o que x es un *predecesor* de y. Análogamente, $y > x$ quiere decir que y es *mayor* o mas *grande* que x, o que y es un *sucesor* de x. Casi siempre, *estricto* y *débil* se refieren a $<$ y a $\leq$ respectivamente.

Si X es un conjunto parcialmente ordenado, y si a es un elemento de X, el conjunto $s(a) = \{x \in X | x < a\}$ es el *segmento inicial* o *segmento inicial estricto* determinado por a, y el *segmento inicial débil* será $\overline{s(a)} = \{x \in X | x \leq a\}$. Entonces, el segmento inicial determinado por a es descrito por el conjunto de los *predecesores estrictos* de a, mientras que el segmento inicial débil determinado por a es el conjunto de todos los *predecesores débiles* de a. Un conjunto $A \subset X$ será un *segmento inicial* de X si A es subconjunto propio[11] de X, y si para todo $a \in A$, todos los x´s que cumplan con $x < a$ son también elementos de A,

[10]Enderton en [End] define a un orden parcial como uno que es irreflexivo y transitivo. La diferencia con el que hemos definido radica en que la definición de Enderton, se basa en hacer la semejanza del orden con el símbolo que conocemos $<$ y nuestra nuestra definición ocupa $\leq$. De esta manera, el autor dice que comparabilidad es tricotomía, porque se tiene que pasan $a < b$ o $a = b$ o $a > b$.

[11]Autores como [RubRub, p. xxii] no piden que sea un subconjunto propio, esta definición la ocuparemos en la sección de

i.e., $(\forall a \in A)(\forall x \in X)(xRa \Rightarrow x \in A)$. Si $x \leq y$ y $y \leq z$ decimos que y esta *entre* x y z; si $x < y$ y $y < z$ se dice que y está *estrictamente entre* x y z. Si $x < y$ y no existe un elemento estrictamente entre ellos, se dice que x es un *predecesor inmediato* de y, o que y es un *sucesor inmediato* de x.

Dado un conjunto parcialmente ordenado X, si existe un $a \in X$ tal que $a \leq x$, $\forall x \in X$ se le llama el *menor, primero* o *más pequeño* elemento de X[12]. Por la antisimetría, si X tiene un elemento menor, éste es único. Similarmente, si X tiene un $a \in X$ tal que $x \leq a$, $\forall x \in X$ se le llama el *mayor, último* o *más grande* elemento de X (también único)[13]. Un elemento $a \in X$ se llama elemento *minimal* de X si no existe un elemento en X estrictamente menor que a, formalmente: a es minimal si $x \leq a \Rightarrow x = a$. Casi de la misma manera, un elemento *maximal* de X es a tal que si $a \leq x$, entonces $x = a$.

Resumiendo, un elemento menor precede a cualquier otro elemento, pero el elemento minimal no tiene predecesores; de esta manera, todo primer elemento es un elemento minimal, pero no al contrario (análogo para mayor y supremo).

Una *cota inferior* (o en inglés, *lower bound*) de un subconjunto E de un conjunto X (aún parcialmente ordenado) es un $a \in X$ que cumple con $a \leq x$ para todo $x \in E$; y una *cota superior* (o en inglés, *upper bound*) de E si $x \leq a$, $\forall x \in E$. Una aclaración importante es que el subconjunto puede tener muchas cotas, o ninguna, y éstas pueden o no pertenecer al subconjunto. Si llamamos a E^* el conjunto de todas las cotas superiores de E y a E_* el de las inferiores, estos pueden o no ser vacíos. Si $E^* \cap E$ no es vacío, entonces es un único $\{a\}$, que es el mayor elemento de E (similarmente para E_*). Si pasara que E_* tuviera un menor elemento a (único necesariamente), se le llamaría la *menor de las cotas superiores* (l.u.b por sus siglas en inglés) o *supremo* de E denotado por $\sup E$. El *ínfimo* o *mayor de las cotas inferiores* de E es el mayor elemento de E_* (si tal existe); para resumir la aseveración anterior, $\inf E$ es el único elemento en X que es una cota inferior de E, y que es mayor que cualquier otra cota inferior de E. El ínfimo y el supremo pueden o no pertenecer a E, al conjunto dado[14].

Esta enumeración contiene las propiedades de unas de las definiciones que acabamos de dar, con (X,R) un orden parcial y $E \subset X$:

1. x es un primer elemento de X ssi $x \in X \wedge (\forall y : (y \in X \wedge x \neq y \Rightarrow xRy))$.
2. x es un último elemento de X ssi $x \in X \wedge (\forall y : (y \in X \wedge x \neq y \Rightarrow yRx))$.
3. x es un elemento minimal de X ssi $x \in X \wedge (\forall y : (y \in X \Rightarrow \neg(yRx)))$.
4. x es un elemento maximal de X ssi $x \in X \wedge (\forall y : (y \in X \Rightarrow \neg(xRy)))$.
5. x es una cota inferior de E ssi $\forall y : (y \in E \Rightarrow xRy)$.
6. x es un ínfimo de E ssi x es una cota inferior de E y $\forall y :$ (y es una cota inferior de E $\Rightarrow yRx$).
7. x es una cota superior de E ssi $\forall y : (y \in E \Rightarrow yRx)$.
8. x es un supremo de E ssi x es una cota superior de E y $\forall y :$ (y es una cota superior de E $\Rightarrow xRy$).

Ahora introduciremos el concepto de conjunto bien ordenado, o conjunto con un buen-orden, concepto muy importante para esta tesis.

---

equivalencias.

[12]En inglés, *first, smallest, least element*.

[13]En inglés, *greatest, largest, last element*.

[14]En todas estas definiciones, que esos elementos cumplan con alguna, depende de la relación, así, si el conjunto es E con relación R, y se requiere no tener confusión sobre la relación, se escribe por ejemplo, elemento R-minimal o R-sucesor.

Definición 1.6.14. *Un conjunto parcialmente ordenado A es un conjunto* bien ordenado *(y su orden es llamado un* buen orden*) si todo subconjunto de A no vacío tiene un elemento menor.*

De esta definición se sigue esta interesante propiedad:

Proposición 1.6.1. *Todo conjunto bien ordenado A es un conjunto totalmente ordenado.*

Creo que el ejemplo más común pero importante de conjunto bien ordenado es el conjunto de los números enteros no negativos, junto con el orden habitual $\leq$. Este conjunto bien ordenado se denota por $\omega = (\mathbb{N}, \leq)$.

Una propiedad que se ocupará mucho, es que todo subconjunto (A, R) de un conjunto bien ordenado (B, R), es bien ordenado, claro, con el orden que hereda del conjunto B.

Se dirá que un conjunto bien ordenado B es una *continuación* de un conjunto bien ordenado A si $A \subset B$, A es un segmento inicial de B y además el orden de los elementos de A es el mismo que el orden en B.

1.7. LÓGICA MATEMÁTICA

La *Lógica* es la ciencia del razonamiento, y aunque algunos autores plantean que esto es muy vago debido a la vasta diversidad que éste ofrece, es un buen punto de partida. La Lógica antigua se basa en gran parte en las ideas Aristotélicas, mientras la actual en los trabajos realizados por Frege y Peano. En nuestro contexto, la Lógica Matemática es la ciencia de las deducciones. Esta abarca áreas como Teoría de Demostraciones (estudia la estructura y limitaciones de ellas), Teoría de Modelos (estudia la conexión entre sintaxis y semántica[15] mediante las estructuras algebraicas que interpretan los lenguajes formales), Teoría de la Recursión (estudia los procesos algorítmicos) y Teoría de Conjuntos.

La introducción (muy breve) que se da aquí de Lógica Matemática (sólo llamada Lógica ya que no se discutirá alguna otra) se debe a dos principales razones que ahora expongo. Es necesario aclarar al lector sobre los diferentes tipos de modelos que existen para diferentes teorías formales (se explicarán estos conceptos a continuación); dependiendo de cómo uno define los axiomas y el modelo que acepta, es como resultará la teoría que deduzca; citando a Carlos Ivorra, sobre esto dice: «*De esta forma la lógica ha probado ser indispensable a la hora de trabajar en teoría de conjuntos, hasta el punto de que es inconcebible el estudio de ésta sin un buen conocimiento de aquella.*» [Ivo, p. 5]. La otra razón, dar una idea intuitiva para comprender referencias posteriores sobre la equivalencia o no equivalencia de una sentencia con el axioma.

Es posible construir modelos de la teoría de conjuntos, encontrando objetos a los que, si los llamamos «conjuntos» cumplen los axiomas que se postulan sobre los conjuntos, y así el Axioma de Elección, interpretado como una afirmación sobre estos objetos, es verdadero, mientras que se puede hacer lo mismo con otra interpretación de «conjunto» resultando en que el Axioma es falso. Esto quiere decir que, interpretando de formas distintas la palabra «conjunto» (esto es posible dado que los axiomas no determinan completamente el concepto), podemos construir 2 objetos que sin ambigüedad llamaríamos «el conjunto de los conjuntos construíbles»[16] para las dos teorías, de tal forma que en una sea cierto el Axioma y en el otro sea falso.

El problema es decidir cuál de las 2 alternativas posibles tomar, i.e., si se tiene una afirmación como el Axioma, cuál es la teoría que tiene a nuestra interpretación encajada, y será un buen modelo de lo que

[15] Sintaxis se refiere a conceptos como fórmula, demostración y teorema por ejemplo, mientras que semántica se refiere a nociones de modelo, consecuencia lógica y verdad, por citar algunas.

[16] Véase p. 77, donde se definen los conjuntos mediante jerarquías, donde un conjunto es construíble si está en una jerarquía.

consideramos «conjunto» y responderá si es o no verdad el Axioma. De esta manera, debemos de tener presente que nuestros axiomas determinarán el mundo con el que querremos trabajar.

El objetivo es *formalizar* las reglas de razonamiento en general y desarrollar sus propiedades, de tal manera que se tengan mejores herramientas como para responder a: «¿son el Axioma de Elección o su negación demostrables a partir de los axiomas de Teoría de Conjuntos?» o «si se acepta el Axioma, ¿se pueden llegar a contradicciones?».

1.7.1. Lenguajes de Primer Orden

Una forma simple de describir a una *Teoría Matemática formalizada* es decir que es un conjunto de estos conjuntos:

1. Un conjunto de *símbolos básicos* o *símbolos primitivos* llamado $\mathscr{V}$, usado para construir *secuencias simbólicas* (también llamadas *cadenas, expresiones* o *palabras*) sobre $\mathscr{V}$.
2. Un conjunto de *cadenas* sobre $\mathscr{V}$ denotado **Wff**, que es el conjunto de las *fórmulas* de la teoría.
3. Un subconjunto de **Wff**, denotado por **Thm**, el conjunto de *teoremas* de la teoría.

Una *cadena de símbolos* de $\mathscr{V}$ es una sucesión finita de símbolos de $\mathscr{V}$, repetidos o no, pero con orden.

Hay dos partes en cada alfabeto de primer orden, la primera es la colección de *símbolos lógicos* (como objetos, conectores Booleanos, cuantificadores, relator de igualdad) y la otra de *símbolos no lógicos* (como constantes, relatores, funciones o funtores).

El conjunto Term de términos, se define como el conjunto más pequeño[17] de cadenas sobre $\mathscr{V}$ que satisface:

1. Todas las constantes y las variables individuales están incluidas.
2. Si f es una función de rango n (para cualquier $n > 0$) y $t_1, t_2, \ldots, t_n$ están en Term, entonces también lo está la cadena «$ft_1t_2\ldots t_n$»[18].

El conjunto de *fórmulas atómicas* **Af** $\subset$ **Wff** contiene:

1. Las cadenas $t = s$ para cada elección posible de términos t y s.
2. Las cadenas $Pt_1t_2\ldots t_n$ para toda posible elección de relatores P de rango n para todas las elecciones de n y todas las posibles elecciones de términos $t_1, t_2, \ldots, t_n$[19].

Entonces, las expresiones son secuencias finitas de los 5 tipos de símbolos del vocabulario. Toda fórmula empieza por un relator, un cuantificador o el implicador. Todo término empieza por una variable, una constante o una función. Analizando una cadena de símbolos de izquierda a derecha puede comprobarse mediante un número finito de pasos si es una expresión o no.

El conjunto de *fórmulas bien formadas* (o *well-formed formulas*), es **Wff**, es el conjunto más pequeño de cadenas o expresiones sobre el alfabeto $\mathscr{V}$ con las propiedades:

---

[17] Por más pequeño nos referimos a la intersección de todos los conjuntos no vacíos que cumplan con las propiedades, dicho en inglés «least inclusive set».

[18] En vez de esta notación, a veces se escribe $f(t_1, \ldots, t_n)$, otras son más específicas dependiendo de la teoría en particular.

[19] También aquí se escribe $P(t_1, \ldots, t_n)$.

1. **Af**$\subset$**Wff**

2. Si $A,\ B \in$ **Wff** son cadenas sobre $\mathscr{V}$, también están en **Wff** $A \vee B$ y $\neg A$

3. Si A es una cadena que esta incluida en **Wff** y x es cualquier variable (que puede o no ocurrir como subcadena en A), entonces también está incluida la cadena $((\exists x)A)$ (léase: existe un x en A).

Estas reglas de sintaxis dadas no permiten escribir cosas como $\exists f$ o $\exists P$ con P y f relator y función resp., esa restricción de que la cuantificación actúe sólo sobre variables objeto y no sobre funciones o relatores hace al lenguaje de *primer orden*[20].

En resumen, una teoría consiste en un alfabeto de símbolos primitivos, algunas reglas usadas para generar el «lenguaje de la teoría» de esos símbolos, y algunas reglas adicionales usadas para generar los teoremas. De las reglas que hablamos son reglas para la manipulación de cadenas, es decir, funciones que reciben cadenas y dan como resultados cadenas. El *lenguaje formal de primer orden* denotado aquí por «L»[21], donde la teoría es «hablada», es una terna $(\mathscr{V}, \mathbf{Term}, \mathbf{Wff})$, donde $\mathscr{V}$ es el *alfabeto* o *vocabulario* del lenguaje y lo usamos para formar *expresiones* que son *términos* si están en **Term** o *fórmulas* si están en **Wff**. Los términos nombrarán «objetos» y las fórmulas harán «declaraciones» acerca de los objetos.

Razonando en la teoría significará el proceso de descubrir «*declaraciones verdaderas*»[22] acerca de objetos, es decir, teoremas. Este proceso, comienza por fórmulas que dan declaraciones que tomamos por ciertas (i.e., se aceptan sin demostración como «verdades básicas»). A estas fórmulas se les llama *axiomas*.

Hay dos tipos de axiomas, los *axiomas especiales o no lógicos* describen aspectos específicos de cualquier teoría específica que se construye (un ejemplo es «$n+1 \neq 0$» como axioma para la Teoría de Números); los *axiomas lógicos* se encuentran en todas las teorías, son «universalmente válidos», de ahí su nombre (por ejemplo «x=x»). Los axiomas se pueden pensar como reglas que no reciben argumento y dan un resultado.

Por último, se necesitan *reglas* para razonar, éstas son las *reglas de inferencia*. Éstas nos ayudan a deducir afirmaciones de otras afirmaciones ya establecidas ciertas. Las reglas de inferencia deben ser elegidas por su obviedad en significado.

En la sección 1.8 se verá que sólo se trabajan con conjuntos sin siquiera haberlos definido en nuestra teoría formal, esto se justifica con este ejemplo:

Ejemplo 1.7.1.

$$\forall xy(\forall u(u \in x \Leftrightarrow u \in y) \Leftrightarrow x = y)$$

Esta expresión tiene sentido para una matemático ya que le dice que dos conjuntos son iguales ssi tienen los mismos elementos. Para el que estudia lógica matemática, es una expresión ya que tiene un algoritmo basado simplemente en los símbolos de la cadena para verificar que en realidad cumple con la definición de expresión*, y para verificarlo no necesita saber qué es un conjunto. Y de ésta manera, él puede hablar de conjuntos sin haberlos definido siquiera.*

Discutiremos ahora los conceptos de variables ligadas (o acotadas) y libres. Una variable x está *libre* en un término t o fórmula atómica A ssi ocurre en t o A como una subcadena. Ocurre libre en $A \vee B$ ssi ocurre libre en al menos alguna. Ocurre libre en $(\exists y)A$ ssi ocurre libre en A y $x \neq y$. Las variables que

[20]En un lenguaje de segundo orden se pueden cuantificar esos dos.

[21]Véase la p. 22 para darse una idea de porqué se dice de *primer orden*.

[22]Verdaderas significa *deducibles*, que se refiere a llegar a ellas mediante el uso de las reglas de inferencia y los axiomas establecidos.

no están cuantificadas están libres, así, aunque y esté libre en A, no lo está en $(\exists y)A$, y decimos que está *ligada* (o en inglés *bound*).

Una expresión esta *cerrada* ssi no hay variables libres en ella. Una fórmula cerrada se llama *sentencia*. Una fórmula esta *abierta* ssi no contiene cuantificadores.

Acabando de definir el concepto de sentencia, podremos ocuparlo más adelante en la tesis, éste es importante debido a que el Esquema Axiomático de Separación (o el Áxioma de Separación), se basa en sentencias y en lo llamado propiedad definida, y tanto paradojas como discusiones se dieron en torno a los comienzos del Axioma de Elección en la historia de las matemáticas[23].

Dado un lenguaje $\mathscr{L}$, una estructura (modelo) $\mathscr{M} = (U, \mathscr{I})$ apropiada para $\mathscr{L}$, viene determinado por:

1. Una colección de objetos no vacía U llamada universo de M.

2. Un criterio que asocie a cada constante $c \in \mathscr{L}$ a $M(c) \in U$.

3. Un criterio que asocie a cada relator R que sea n-ádico tal que $R \in \mathscr{L}$, una relación R_1 tambié n-ádica con $R_1 \in U$.

4. Un criterio que asocie a cada funtor f que es n-ádico tal que $f \in \mathscr{L}$ una función f_1 al igual n-ádica con $f_1 \in U$.

Dicho criterio es la función $\mathscr{I}$, que es la interpretación.

Una colección de fórmulas de un lenguaje formal $\mathscr{L}$ es contradictoria si existe una fórmula la cuál sea deducible, y su negación también. En caso contrario se dice que es consistente. Una teoría axiomática es contradictoria si existe dicha fórmula deducida ella y su negación mediante los axiomas dados.

1.8. TEORÍA DE CONJUNTOS

Halmos comienza el prefacio de su libro diciendo que todo matemático esta de acuerdo en que todo matemático debe conocer algo sobre la Teoría de Conjuntos, el problema es delimitar «cuánto» es ese algo [Hal]. Daremos aquí las herramientas que nos permitirán comprender conceptos que ocuparemos más adelante.

La primera axiomatización para evitar las paradojas de la Teoría de Conjuntos fue creada por Whitehead y Russell en el libro *Principia Mathematica* entre 1910 y 1913[24]. Aunque ésta es demasiado axiomática y no es más provechosa que las subsecuentes teorías.

La importancia de este tema no solo es por ser fundamental para todo matemático, sino que el Axioma esta tan ligado a él, que estudiarlo sin dar una recopilación de los conceptos básicos que recaen en cuanto a Teoría de Conjuntos me refiero, tendría muy poco sentido. Además, entran temas no vistos en la carrera que cuentan con un número grandísimo de resultados, como Ordinales y Cardinales (véase 3.3), Inducción Transfinita (véase 1.10), Forcing, etc.

Admitiendo que los axiomas que se estudian en ésta teoría son consistentes, el Teorema de Completitud dice que para la palabra «conjunto» existe una interpretación[25], y se puede demostrar que entonces tiene infinitas. Se tiene que existen infinitos modelos no iguales, refiriéndonos a que si tomamos dos de

[23] Véase la p. 61.
[24] Fue entre los 3 volúmenes del libro, cuyas referencias están en [Moo, p. 368].
[25] Entiéndase por esto lo que se dice en la p. 23.

ellos, uno tiene una sentencia cierta que en el otro es falsa[26]. También, como ya se dijo en el ejemplo (1.7.1), los axiomas no definen completamente qué es un conjunto, solo dan nociones y propiedades que éstos cumplen, y así, las nociones que metamos en nuestra teoría (axiomas) nos darán diferentes conceptos de conjunto.

Daremos primero la Teoría Intuitiva de Conjuntos, que básicamente es la que se define en las primeras secciones de este capítulo, la segunda teoría a considerar es la de Zermelo-Fraenkel, y la última es la que desarrollaron von Neumann y Gödel, como una teoría que buscaba mejorar la de ZF. Von Neumann afirmó que su motivación de hacerlo fue porque la noción de «propiedad definida» no era clara, el Axioma de Reemplazo era necesario para la teoría de números ordinales y los conjuntos «muy grandes»[27] tenían que ser admitidos para formular el Axioma de Reemplazo.

Una *paradoja* o *antinomía* ocurre cuando en un sistema axiomático una proposición y su negación son deducibles de los axiomas. Esto implica (por un ejercicio de lógica elemental) que «toda» declaración en el sistema es verdadera, según Carlos Ivorra.

Ésta fue razón por la que esas teorías se fueron modificando, por las inconsistencias que se encontraban en ellas. El ejemplo más típico es el de que no todo lo debemos de considerar como conjunto, el «conjunto» de todos los conjuntos no es un conjunto. Analicemos a un objeto B que suponemos es un subconjunto de nuestro universo de discurso V, y que sus elementos cumplen con la propiedad de que $x \in B \Leftrightarrow (x \notin x) \wedge (x \in V)$, i.e., $B = \{x \in V | x \notin x\}$[28]. La pregunta es ¿$B \in V$?, pues la penúltima expresión nos dice que $B \in B \Leftrightarrow (B \notin B) \wedge (B \in V)$, entonces, si $B \in V$ se llega a una contradicción, por lo tanto se debe de tener que $B \notin V$, pero V es un conjunto arbitrario y lo que se acaba de demostrar es que **nada contiene todo**, o lo que es lo mismo, no hay un conjunto que contenga a todos los conjuntos. Ésta se conoce como la *paradoja de Russell*. Esta paradoja justifica el hecho de mejorar las teorías, y de tener presente que no solo uno dice arbitrariamente qué es un conjunto, o que sus elementos cumplen con cierta «propiedad», uno debe ser cauteloso en estos casos.

Una gran ventaja de NBG es que es finitamente axiomatizable. En el sistema ZFC de primer orden, se necesitan los esquemas axiomáticos de Separación y Reemplazo, generando denumerables axiomas por cada esquema, y por esto no se puede axiomatizar finitamente, sin embargo, NBG tiene finitos axiomas ya que las clases juegan el papel de «propiedad definida» o condiciones de primer orden en ZFC, y las clases se derivan de solo un número finito de axiomas. Esto nos dice que las propiedades que hacían esquemas a los axiomas en ZFC se convierten en tipos de clases, y para hacer éstas (construirlas recursivamente, visto en la p. 77.) solo necesitamos un número finito de axiomas.

Para las siguientes subsecciones se enunciarán los axiomas que dan cuerpo a las teorías, sin embargo, en ocasiones hay varias formulaciones de éstos, así que se dará la más común y cuando se repita se intentará dar otra equivalencia. Se darán sus diferencias y breves introducciones históricas, el lector interesado en la profundización de esos temas se puede referir al capítulo VIII «Los Axiomas de la Teoría de Conjuntos» de [Ivo] y al capítulo XI «Consolidation of Set Theory» de [Fer] respectivamente.

1.8.1. Teoría de Conjuntos Intuitiva

Se le llama intuitiva debido a que el lenguaje y la notación son informales, y este término fue propuesto por von Neumann. Se ocupan en este teoría palabras de uso común en muchos resultados, además, la Teoría de Conjuntos la toma como una serie de hechos, mientras que el punto de vista axiomático puro, muy ortodoxo, se fija en las relaciones lógicas entre los axiomas. Un geómetra es puramente intu-

[26]Véase p. 21 para recordar porqué nuestros resultados que creemos ciertos cambian el modelo a utilizar.

[27]Como los que llevan a paradojas, en la p. 30 se menciona porqué se hace alusión de esta manera.

[28]Estamos ocupando el Axioma de Abstracción de la p. 13.

itivo si solo usa papel y lápiz junto con su intuición de geometría, y sería axiomático, si los axiomas de la geometría no euclidiana los estudia tanto como de la geometría de Euclides.

El problema de esta teoría, es que llevó a contradicciones cerca de 1900, donde se encontraron paradojas y conjuntos con las propiedades definidas.

Una de ellas es la de Burali-Forti. En la Teoría Intuitiva de Conjuntos, todo conjunto bien ordenado tiene un número ordinal [29]. Además, el conjunto de todos los ordinales es bien ordenado, y por eso tiene a un ordinal $\mathfrak{n}$, pero si añadimos al conjunto de todos los ordinales cualquier ordinal, también será bien ordenado, y además su cardinal será $\mathfrak{n}+1$, que es mayor que $\mathfrak{n}$, por eso $\mathfrak{n}$ no es el número ordinal de todos los cardinales[30].

La paradoja de Cantor, hecha en 1899, crea el mayor número cardinal. Considera el cardinal $\mathfrak{n}$ del conjunto V de todos los conjuntos, éste debe ser el mayor cardinal posible, pero si se considera el conjunto potencia de V y su cardinal respectivo cardinal de éste $\mathfrak{p}$, se llega a una contradicción ya que $\mathfrak{p} > \mathfrak{n}$ (ésto por el teorema 1.9.1).

Los axiomas básicos se sumarizan en 3, y estos se sacan no de un trabajo o escrito de alguien que los haya definido, sino de los postulados en que se basaron autores que no formalizaban sus tratados, y sólo daban por hecho nociones básicas sobre los conjuntos. Pero principalmente los axiomas se basan en los primeros trabajos de Cantor y son: el Axioma de Extensionalidad de conjuntos, el Axioma de Abstracción y el Axioma de Elección[31]. Hay que tener en cuenta que la definición Cantoriana de conjunto es: cualquier totalidad de objetos bien distinguidos de nuestro pensamiento.

La formulación explícita del Axioma de Abstracción la dio Frege en 1893. El problema de la teoría intuitiva radica en el Axioma de Abstracción, debido a que en 1901 Russell descubrió que se puede llegar a una contradicción a partir de éste axioma (véase la p. 24). La reacción de Frege respecto a esto es clara: «*Difícilmente puede sobrevenirle a un escritor científico algo más infortunado que ver vacilar uno de los fundamentos de su edificio, después de que su trabajo está terminado.*» [32].

Se definirán los axiomas de esta teoría:

Axioma 1.8.1. Axioma de Extensión. *Dos conjuntos son iguales ssi tienen los mismos elementos.*

Axioma 1.8.2. Axioma de Abstracción. *A cada condición $S(x)$, le corresponde un conjunto B cuyos elementos son exactamente esos x que hacen $S(x)$ verdadera, i.e., $B = \{x|S(x)\}$.*

Aquí, *condición* significa *sentencia*, vista ya en la p. 23. La variable x está libre (véase p. 22) en $S(x)$, y dicho B es único por el axioma anterior.

Otra cosa que debemos notar, es que el Axioma de Abstracción es en verdad una infinidad de axiomas, ya que cuando reemplazamos la expresión $S(x)$ en el axioma (1.8.2) por cualquier fórmula que no tenga a y como variable libre, obtenemos un nuevo axioma. Un axioma que permite la substitución de fórmulas de esta forma se llama *esquema axiomático*.

Axioma 1.8.3. Axioma de Elección. *Existe una función f tal que $f(x) \in x$ para todo subconjunto no vacío de un conjunto no vacío dado.*

Los demás axiomas están implícitos en esta teoría y son los más comunes, sin embargo, son los que menciona [Hal] en su libro «Naive Set Theory» y no tiene que ver tanto con la historia de la teoría intuitiva de Cantor, sino más bien con los fundamentos esenciales e intuitivos de un primer curso de teoría axiomática de conjuntos. Son:

[29] Véase la p. 32 para la definición de este término.
[30] Véase la p. 35 para la definición de este concepto.
[31] Véase las secciones 2.1.1 y 2.1.2 para saber porqué y como utilizó intuitivamente Cantor el Axioma de Elección.
[32] Publicado en el apéndice del segundo volumen de *Grundgesetze der Arithmetik, 1903*.

Axioma 1.8.4. Axioma de Par. *Para cualesquier dos conjuntos, existe un conjunto al cuál ellos pertenecen.*

Axioma 1.8.5. Axioma de Unión. *Para toda colección de conjuntos existe un conjunto que contiene a todos los elementos que pertenezcan al menos a un conjunto de la colección.*

Axioma 1.8.6. Axioma del Conjunto Potencia. *Para todo conjunto existe la colección de conjuntos que contiene entre sus elementos a los subconjuntos del conjunto dado.*

Axioma 1.8.7. Axioma de Infinitud. *Existe un conjunto que contiene al 0 y que contiene al sucesor de cada uno de sus elementos.*

Este axioma está muy relacionado con la definición de números naturales y el conjunto de éstos[33], pudiendo definir al *sucesor de un conjunto x* como:

$$x^+ = x \cup \{x\},$$

también se le denota por x'. De esta manera formamos a los números naturales como conjuntos, pero el conjunto de todos ellos lo forma el axioma.

De aquí se desprende el *Principio de Inducción Matemática* que dice[34]:

(1.10.1.1) Si $S \subset \omega$, $0 \in S$ y $n^+ \in S$ siempre que $n \in S$, se tiene entonces que $S = \omega$.

1.8.2. Teoría de Conjuntos de Zermelo-Fraenkel

La mayoría de estos axiomas son los de Zermelo. Pero para la teoría de inducción transfinita y aritmética ordinal se necesita un esquema axiomático mas fuerte, llamado el esquema axiomático de reemplazo o esquema axiomático de sustitución, que se atribuye a Fraenkel. Por esta razón se llama Teoría de Conjuntos de Zermelo-Fraenkel (abreviada ZF) aunque, más apropiadamente, debería llamarse Teoría de Conjuntos de Zermelo-Fraenkel-Skolem, ya que el mismo axioma lo propuso independientemente y al mismo tiempo Skolem.

La intención de Zermelo al axiomatizar la teoría Cantoriana, era hacerla más clara y comprensible, pero consideraba que no se necesitaba lógica formal. Luego, Skolem añadió al modelo de Zermelo cálculo de predicados de primer orden con igualador.

Los axiomas son:

Axioma 1.8.8. Axioma de Extensionalidad. *Si X y Y tienen los mismos elementos, entonces $X = Y$.*

Axioma 1.8.9. Axioma de Par. *Para todo a y b existe el conjunto $\{a, b\}$ que contiene exactamente a dichos a y b.*

Axioma 1.8.10. Esquema Axiomático de Separación. *Si P es una propiedad con parámetro p, entonces para todo X y todo p existe un conjunto $Y = \{u \in X | P(u, p)\}$ que contiene todos esos $u \in X$ que tienen la propiedad P.*

Axioma 1.8.11. Axioma de Unión. *Para todo X existe el conjunto $Y = \bigcup X$, la unión de todos los elementos de X.*

[33] Véase la p. (1.8.2), en ella hay una forma de enunciar el Axioma de Infinitud sin siquiera definir los números naturales, y de ésta manera seguir trabajando con la teoría de conjuntos hasta ahora hecha.

[34] Véase la p. (1.6) donde viene la definición de ω.

Axioma 1.8.12. Axioma de Partes o del Conjunto Potencia. *Para todo X existe un conjunto* $Y = \mathscr{P}(X)$, *el conjunto de todos los subconjuntos de* X.

Axioma 1.8.13. Axioma de Infinitud. *Existe un conjunto infinito.*

Axioma 1.8.14. Esquema Axiomático de Reemplazo. *Si una clase* F *es una función, entonces para todo* X *existe un conjunto* $Y = F(X) = \{F(x) | x \in X\}$.

Axioma 1.8.15. Axioma de Regularidad. *Todo conjunto no vacío tiene un elemento* $\in$*-minimal*[35].

Axioma 1.8.16. Axioma de Elección. *Para toda familia no vacía de conjuntos no vacíos el* Producto Cartesiano[36] *de esa familia es no vacío, es decir, si* $\mathscr{A} \neq \emptyset$ *y* $X_\alpha \neq \emptyset \ \forall \alpha \in \mathscr{A}$, *se cumplirá:*

$$\prod_{\alpha \in \mathscr{A}} X_\alpha \neq \emptyset.$$

Un esquema axiomático que se utiliza es el de Zermelo, llamado el *esquema axiomático de separación* (o *Aussonderung Axiom*), que nos permite separar elementos de un conjunto «ya dado» que satisfacen la propiedad, y formar el nuevo conjunto. Este dice:

$$(\exists y)(\forall x)(x \in y \Leftrightarrow x \in z \vee S(x))$$

El cambio con respecto al Axioma de Abstracción es que se pide que exista un z a donde x pertenezca. De esta manera, no se llega a una contradicción como la de Russell.

Para entender el Axioma de Infinitud se necesita definir primero lo que es un conjunto finito, y para ello se necesitan introducir los número naturales, pero hay una manera de burlar esto. Se cumple que este axioma y el dado en la teoría intuitiva son equivalentes, la liga es definir un *conjunto inductivo* como el que cumple la condición:

$$\exists S(\emptyset \in S \wedge (\forall x \in S) x \cup \{x\} \in S).$$

Esta teoría descarta paradojas de una manera restrictiva: no hay axiomas que indiquen la existencia «general» de conjuntos. En esta teoría no existe un conjunto universal que contenga a todos los demás.

Los *conjuntos no bien fundados* (o *non-well-founded*) se caracterizan por tener sucesiones descendientes de la forma $\ldots \in x_2 \in x_1 \in x$. Lo que un conjunto no bien fundado dice es que es un x, que tiene un elemento, y que éste tiene un elemento, y que éste tiene un elemento, y que éste tiene un elemento, ... En realidad éstos conjuntos fueron los que dieron lugar al *Axioma de Fundación* o Regularidad.

Un axioma muy importante para formar conjuntos es el de reemplazo. El contenido esencial del axioma de remplazo dice:

(1.10.2.1) Si $F : A \mapsto B$ es una aplicación suprayectiva y A un conjunto, entonces también lo será B.

Con él se puede demostrar que muchas construcciones de clases son conjuntos, como los subconjuntos, el producto cartesiano, las relaciones sobre conjuntos, las funciones con dominio conjunto y demás.

En realidad Zermelo tomó el esquema axiomático de especificación, y fue Fraenkel el que lo sustituyó por el esquema axiomático de reemplazo.

Hay tres caminos diferentes para obtener números cardinales. Uno es introducir una noción primitiva y un axioma para números cardinales. Otro es definirlos como ciertos número ordinales, ésta definición requiere del Axioma de Elección para mostrar que todo conjunto tiene un número cardinal. La última es operar con los axiomas actuales y con el Axioma del Infinito, mediante la noción del *rango* de un conjunto[37], pero esta construcción es muy complicada.

[35] Véase la p. 19 y la p. 27 para la definición y anotaciones sobre elementos minimales en relaciones binarias.
[36] Esta definición se verá en la p. 91 en la subsección 3.1.1.
[37] Véase [Ivo, p. 339] para una exposición sobre este concepto.

1.8.3. Teoría de Conjuntos de von Neuman-Bernays-Gödel

Esta subsección se trata sobre describir el universo de todos los conjuntos «reales» y objetos de nuestra intuición. El sentido de esta teoría axiomática, es tener una teoría «formal» que derive declaraciones verdaderas acerca de los conjuntos de von Neumann, permitiéndonos conocer la estructura y naturaleza de este universo; y para esto los axiomas elegidos deben tener cierta «verdad» en el universo que describimos. Para describir este universo, se proponen los axiomas que parecen seguirse de la formación de conjuntos por etapas, o de un principio que mantenga al sistema sin paradojas. Sobre este aspecto Tourlakis escribe: «*...the reader is constantly made aware that we are building a meaningful set theory that has the relevance to mathematical intuition and expectations (the "real" mathematics), and is not just an artificial choice of a contradiction-free set of axioms followed by the mechanical derivation of a few theorems.*» [Tou, p. xiii].

Originalmente formulada por von Neumann en una serie de escritos en 1925, 1929 y 1938. Su idea original difiere considerablemente de la teoría de Zermelo debido a que se le da más importancia a la noción de función que a la de clase o conjunto. En una serie de escritos para la «Journal of Symbolic Logic», Bernays modificó la teoría de von Neumann para acercarla más al sistema original de Zermelo, introduciendo 2 relaciones de pertenencia (dos relatores de rango 1), una entre conjuntos y la otra entre conjuntos y clases. Para 1940 Gödel simplificó más la teoría, sus nociones primitivas eran las de conjuntos, clases y pertenencia, aunque hoy sabemos que con la pertenencia es suficiente.

Tiene dos diferencias esenciales respecto al sistema ZF. La teoría de von Neumann-Bernays-Gödel (abreviada NBG) puede ser axiomatizada finitamente. No se requiere esquema axiomático de separación para construir, basta con unas construcciones finitas y específicas de conjuntos y clases (se menciona porqué en la p. 24). Además, en NBG, la noción básica es la de clase. De esta manera, burla las dificultades que traen las *multiplicidades inconsistentes*[38] de Cantor, restringiendo los axiomas oportunos a solamente los conjuntos.

En principio se creía que eran muy diferentes ZFC y NBG, pero luego se demostró que NBG sin Axioma de Elección es una *extensión conservativa* de ZF, queriendo decir esto que si un teorema de conjuntos puede ser probado en NBG, una proposición correspondiente se puede demostrar en ZF.

Si una sentencia de conjuntos es demostrable en ZFC entonces lo es en NBG. Para el lado contrario, un teorema de Shoenfield que utiliza métodos de teoría de demostraciones, afirma que si una sentencia que solo incluye variables de conjuntos es demostrable en NBG son el Axioma, entonces lo es en ZF. Para extenderlo a NBG y ZFC se utiliza el método de *forcing*.

Enunciamos los axiomas de la teoría NBG, mediante su forma primitiva:

Axioma 1.8.17 (NBG-1)**. Existencionalidad.**

$$\forall XY\left(\forall u(u \in X \Leftrightarrow u \in Y) \Rightarrow X = Y\right)$$

Axioma 1.8.18 (NBG-2)**. Intersección.**

$$\forall XY\exists Z\forall u(u \in Z \Leftrightarrow u \in X \wedge u \in Y)$$

Axioma 1.8.19 (NBG-3)**. Complemento.**

$$\forall X\exists Y\forall u(u \in Y \Leftrightarrow u \notin X)$$

Axioma 1.8.20 (NBG-4)**. Par.**

$$\forall uv\exists y\forall x(x \in y \Leftrightarrow x = u \vee x = v)$$

[38] Véase p. 47.

Axioma 1.8.21 (NBG-5)**. Pertenencia.**

$$\exists A \forall xy((x,y) \in A \Leftrightarrow x \in y)$$

Axioma 1.8.22 (NBG-6)**. Dominio.**

$$\forall A \exists B \forall x(x \in B \Leftrightarrow \exists y(x,y) \in A)$$

Axioma 1.8.23 (NBG-7)**. Producto Cartesiano.**

$$\forall A \exists B \forall xy((x,y) \in B \Leftrightarrow x \in A)$$

Axioma 1.8.24 (NBG-8)**. Relación Inversa.**

$$\forall A \exists B \forall xy((x,y) \in B \Leftrightarrow (y,x) \in A)$$

Axioma 1.8.25 (NBG-9)**. Permutación.**

$$\forall A \exists B \forall xyz((x,y,z) \in B \Leftrightarrow (y,z,x) \in A)$$

Axioma 1.8.26 (NBG-10)**. Permutación.**

$$\forall A \exists B \forall xyz((x,y,z) \in B \Leftrightarrow (x,z,y) \in A)$$

Axioma 1.8.27 (NBG-11)**. Conjunto Vacío.**

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

Axioma 1.8.28 (NBG-12)**. Unión.**

$$\forall x \exists y \forall uv(u \in v \wedge v \in x \Rightarrow u \in y)$$

Axioma 1.8.29 (NBG-13)**. Reemplazo.**

$$\forall xA(Un\, A \Rightarrow \exists y \forall u(u \in y \Leftrightarrow \exists v \in x(v,u) \in A))^{39}$$

Axioma 1.8.30 (NBG-14)**. Infinitud.**

$$\exists fx(f : x \mapsto x \ \ \textit{inyectiva y no suprayectiva})$$

Axioma 1.8.31 (NBG-15)**. Partes.**

$$\forall x \exists y \forall u(u \subset y \Rightarrow u \in y)$$

Axioma 1.8.32 (NBG-16)**. Regularidad.**

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

Axioma 1.8.33 (NBG-17)**. Axioma de Elección.**

$$\forall x \exists f(f \textit{ es una función } \wedge Dom\ f = x \wedge \forall u \in x(u \neq \emptyset \Rightarrow f(u) \in u))$$

---

[39] Donde $Un\, X \equiv \forall uvw((u,v) \in X \wedge (u,w) \in X \Rightarrow v = w)$.

A la función del axioma NBG-17, que «elige» un elemento de cada conjunto no vacío de x, se le llama *función de elección* sobre x, y también se puede decir que el axioma garantiza la existencia de una función de elección sobre todo conjunto.

En esta teoría, todo conjunto es una clase, a las clases que no son conjuntos se les llaman *clases propias* (llamadas así, por Kurt Gödel, y son las que Cantor llamó multitudes inconsistentes o absolutamente infinitas, hablamos de éstas en la p. 47), y se distinguen por no pertenecer a alguna clase, otra forma de decirlo, es que un conjunto es una clase que pertenece por lo menos a otra clase. De esta manera, las paradojas de Burali-Forti, Cantor y Russell no pueden ser construidas, debido a que existen, pero no son conjuntos.

De cierta manera, se definen las clases a partir de fórmulas[40]:

Definición 1.8.1. *Si $P(\ x,\ p_1,\ \ldots,\ p_n)$ es una fórmula, se llama* clase *a:*

$$C = \{x | P(\ x,\ p_1,\ \ldots,\ p_n)\},$$

sus elementos son los conjuntos x que satisfacen $P(\ x,\ p_1,\ \ldots,\ p_n)$.

Se dice que C es *definible* de $p_1,\ \ldots,\ p_n$, si $P(x)$ no tiene parámetros p_i, la clase C es definible.

La *clase universal* o *universo* (usualmente denotada por V) es la clase de todos los conjuntos:

$$V = \{x | x = x\}$$

Von Neumann en realidad estableció que una clase propia era equivalente al conjunto de todos los objetos (c es clase propia $\Leftrightarrow \exists$ función (clase propia) f tal que $\forall x \exists y \in c$ y tal que $f(y) = x$), de aquí se desprende que se mencionen las clases propias como «muy grandes». Esta fue la aproximación que da la importancia a las funciones. Este axioma que introdujo von Neumann, es tan fuerte que implica los axiomas de Separación, Reemplazo, el Axioma de Elección, y hasta el Axioma de Elección Global[41]: existe una relación sencilla (es decir una clase y no un conjunto) que selecciona simultáneamente un elemento de cada conjunto del universo. La demostración informal es: la clase de todos lo ordinales (que tiene un buen orden) lleva a la paradoja de Burali-Forti (véase p. 25), por lo tanto por el axioma de von Neumann, es «muy grande» y equivalente a la clase de todas las cosas, por el buen orden se obtiene un buen orden de la clase universal, y por esto, todo conjunto o clase puede ser bien ordenado, y la última afirmación sabemos es equivalente al Axioma de Elección (véase p. 96).

El problema de este axioma, es que va más allá y afirma que todas las clases cuya cardinalidad sea menor que la del conjunto universal son conjuntos. Por eso, von Neumann restringió su sistema axiomático sustituyendo su axioma por el de Reemplazo y el de Elección.

Aunque no debe de sorprendernos, NBG también cuenta con sus propias multiplicidades inconsistentes, por ejemplo, la clase de todas las clases, existe una demostración de que ésta no existe; una posible forma de evitarlo sería componer de «niveles» a las clases, pero, eso sería dotar de importancia a las clases, tomándolas como objeto de estudio, y no a los conjuntos en sí.

Átomos

El tema de los átomos como otro tipo de elementos matemáticos a considerar es muy importante. En una primera etapa de axiomatización y de pruebas de independencia de varios axiomas, sirvieron como herramienta fundamental.

[40]Una clase no es una fórmula, sino una fórmula con una variable destacada.

[41]Término acuñado por Azriel Levy en 1973, en 1961 le llamó Axioma Universal de Elección.

La teoría de conjuntos que hacen varios de los autores actuales, permite los llamados *átomos* (o *urelemente*), pero no insiste en que haya alguno, esto claro, tiene sus ventajas y desventajas. Algunos matemáticos dicen que la Teoría de Conjuntos sin átomos, solo habla de conjuntos y nada más, de ningún otro objeto existente. Pero en realidad, es sabido que la teoría es capaz de construir otros objetos «artificiales». De esta manera, una teoría con átomos es más general, y además permite dar ejemplos y objetos de naturaleza distinta a solo conjuntos.

El resultado más importante de Fraenkel fue la independencia del Axioma de los otros postulados. Esto lo logró mediante un modelo M de la teoría de conjuntos junto con *átomos* (urelements), suponiendo que en el modelo hay infinitos contables de éstos. Un átomo es un objeto que pertenece a un conjunto, que no contiene elementos y sin embargo es distinto del conjunto vacío.

También von Neumann probó la independencia del Axioma de Elección, que no es posible demostrar el Axioma a partir de ZF con átomos. Encontró contraejemplos consistentes al Axioma, pero ésos eran átomos. Hoy se sabe que tales contraejemplos también son consistentes en ZF, así que la existencia de átomos no es necesaria.

1.9. CARDINALES Y ORDINALES

Daremos una breve introducción a estos dos temas, debido a que aparecerán en la historia y son básicos como temas de la teoría de conjuntos. La presentación de la teoría de cardinales puede basarse en la teoría de ordinales, pero también puede no basarse en ella, claro, las dos formas tienen sus ventajas y pesares. Trataré de dar un compendio de lo más común.

A dos conjuntos parcialmente ordenados los llamaremos *similares* o *semejantes*[42] si existe una función biyectiva que preserve el orden entre ellas. Escribiéndolo explícitamente, diremos que $(X,R) \cong (Y,S)$ si existe una función $f:(X,R) \mapsto (Y,S)$ biyectiva tal que si a, $b \in X$, entonces $aRb \Leftrightarrow f(a)Sf(b)$, generalmente se suele escribir a la relación por $\leq$. A tal f le llamaremos *similaridad* o *semejanza*[43]. Se dirá que dos clases ordenadas son similares si existe una similaridad entre ellas. A veces, se define un *monomorfismo* como una función inyectiva $f:(X,R) \mapsto (Y,S)$ que preserva el orden; y un *isomorfismo* como un monomorfismo biyectivo (que es lo que definimos como similaridad), también se dice que un isomorfismo es una función biyectiva $f:(X,R) \mapsto (Y,S)$ y que f y f^{-1} preservan el orden. Cuando $X = Y$, se dice que f es un *automorfismo*.

El conjunto sucesor de un conjunto X se definió en la p. 26 como $x^+ = x \cup \{x\}$, y de ahí se formaron los números naturales y ω como el conjunto más pequeño que contiene a todos los sucesores de sus elementos y al vacío. Para seguir contando se necesita formar el conjunto que tenga no solo a los naturales, sino al sucesor de ω, y al sucesor de éste, y seguir así *ad infinitum*. Diremos que una función f es una función *ω-sucesora* si $Dom f = n$ con $n \neq 0$ (i.e., el conjunto de todos los predecesores del número n), y además $f(0) = \omega$ y $f(m^+) = f(m)^+$ siempre que $m^+ < n$. Sea $S(n,x)$ la sentencia que nos dice: $n \in \mathbb{N}$ y $x \in Ran f$ con $Dom f = n$ y f una función ω-sucesora. Lo que buscamos es el conjunto I tal que $x \in I$ ssi $S(n,x)$ es verdadera para algún n. Ese conjunto es el análogo a la distancia entre 0 y ω pero para ω e I, y para formarlo necesitamos una función F tal que $(n,x) \in F \Leftrightarrow x \in F(n) = \{y : S(n,y)\}$.

El axioma que nos permite la formación del conjunto I es el de reemplazo (sin éste no se puede demostrar la existencia de dicho conjunto), que aquí enunciaremos de otra forma equivalente:

---

[42]También se dice en inglés que tienen el mismo *order type*.

[43]Algunos autores, como Sierpinski en [Sie] definen similaridad en otros términos, piden un orden que sea conexo, transitivo e irreflexivo, y que f solo cumpla con ser biyectiva y $a < b \Rightarrow f(a) < f(b)$. La diferencia lo dice la relación, radica en un «menor estricto» como lo entendemos y no un menor igual.

Axioma 1.9.1 (Axioma de Reemplazo)**.** *Si $S(a,b)$ es una sentencia que forma un conjunto $\{b : S(a,b)\}$ para toda $a \in A$, entonces existe una función F con $DomF = A$ que cumple $F(a) = \{b : S(a,b)\}$ para toda $a \in A$.*

Una aplicación del axioma es contar mucho más lejos después de los números naturales. Estos cumplen que si $m,\ n \in \mathbb{N}$, $m < n$ es lo mismo que $m \in n$[44]. Recordando que el segmento inicial de un conjunto x en un conjunto bien ordenado es $s(x) = \{y|y < x\}$[45], se cumple que para todo natural $s(n) = \{m \in \omega : m < n\} = n$.

De esta manera, definimos a un *número ordinal* como un conjunto bien ordenado α con $s(\beta) = \beta$ para todo $\beta \in \alpha$. También se puede definir como un conjunto α que cumple con las dos propiedades siguientes: si $\eta \in \beta$ y también $\beta \in \alpha$ se sigue que $\eta \in \alpha$, y está bien ordenado por la relación $\in$ en el conjunto dado[46].

Con las funciones ω-sucesoras cuya existencia depende del Axioma de Reemplazo, podemos formar el conjunto ω^+, con una función de dominio $n = 2$ por ejemplo, ya que $F(1) = F(0)^+ = \omega^+$, este conjunto está ordenado de la manera de ω, y éste orden es un buen orden que cumple las condiciones de los números ordinales. El conjunto ω es un número ordinal (que no es un número natural) ya que contiene a todos sus predecesores, los naturales. Si $\beta \in \omega^+$, entonces se debe cumplir que $\beta \in \omega$ o $\beta = \omega$; en el primer caso, β sería un natural, y por lo tanto un ordinal, que cumpliría con $s(\beta) = \beta$, si $\beta = \omega$, se cumple que ω por ser un número ordinal contiene a todos sus antecesores, de esta forma $s(\beta) = \omega = \beta$.

Este argumento un poco general se puede aplicar para cualquier ordinal, demostrando que el sucesor de cualquier ordinal, vuelve a ser un ordinal. Un orden parcial o total en un conjunto X está determinado únicamente por sus segmentos iniciales; si se tienen dos ordenes R y S en X, y si para cualquier $x \in X$ el conjunto de R-predecesores es igual al conjunto de S-predecesores, necesariamente $R = S$, ya que aRx ssi aSx (por ser iguales los segmentos iniciales). Lo que esto nos dice, es que si un conjunto lo podemos bien ordenar para hacerlo un número ordinal, hay solo una manera de hacerlo, ya que si tenemos dos ordenes en el número ordinal, los segmentos iniciales coincidirían por la definición de número ordinal, y por lo tanto las relaciones serían iguales.

Un hecho elemental sobre números ordinales, es que cada elemento de ellos es también un subconjunto. Si $\beta \in \alpha$, como $s(\beta) = \beta$ entonces $\beta \subset \alpha$ porque todos los elementos de β son predecesores de β en α, y por lo tanto, también serán predecesores de α. De esta manera, los predecesores de α son los elementos de α y por lo tanto la relación en cuestión es la de pertenencia. Esto dice que los números ordinales también son *conjuntos transitivos*, un conjunto transitivo X cumple con que si $a \in x$ y además $x \in X$ se sigue que $a \in X$, otra definición análoga es que todo elemento de X es a la vez subconjunto. Como los elementos de los números ordinales son subconjuntos también, entonces serán conjuntos bien ordenados con el orden heredado del número ordinal dado[47], y de aquí se sigue que si $\beta \in \alpha$ con éste último un número ordinal, también será el primero, debido a que si $\gamma \in \beta$, los segmentos iniciales de γ en β y en α serán los mismos ya que $\gamma \in \alpha$, con esto, los segmentos iniciales de los elementos de β, son iguales a dichos elementos y por lo tanto, β es un número ordinal.

Un conjunto X se dice *bien ordenado por epsilon* ssi la relación

$$\in_X = \{(x,y) \in X \times X | x \in y\}$$

es un buen orden en X.

[44]Véase la p. 26.

[45]Véase la p. 18.

[46]Esto justifica que no sea ambiguo escribir $\eta \in \beta \in \alpha$.

[47]Véase la página 20.

Un número ordinal, de esta manera, es un conjunto transitivo bien ordenado por epsilon, y además la implicación contraria también es cierta.

Regresando a lo que habíamos requerido, preguntando cómo contamos después de ω, el Axioma de Sustitución o Reemplazo nos dice que existe una función F en ω tal que $F(0) = \omega$ y $F(n^+) = F(n)^+$, esto $\forall n \in \mathbb{N}$. Si escribimos $F(n) = \omega + n$ podemos escribir al conjunto $\omega 2$ como el conjunto que consiste de toda n y toda $\omega + n$ con $n \in \omega$. De esta manera empezaremos a contar (la suma de números ordinales se dará en la sec. 3.3):

Primero comenzamos con 0, 1, 2, ..., después de ellos viene $\omega, \omega+1, \omega+2, \ldots$, luego está $\omega 2$, después $\omega 2+1, \omega 2+2, \ldots$, luego $\omega 3$. Así, se encuentran $\omega, \omega 2, \omega 3, \ldots$, después de todos ellos una aplicación del Axioma de Reemplazo nos trae a ω^2. Y de ahí se comienza a contar de nuevo, pasando por ejemplo por $\omega^\omega, \omega^{\omega^\omega}$, después de ahí siguen muchos ordinales más.

Otra propiedad de los números ordinales, es que si dos de ellos son similares, entonces ellos son iguales.

Si se tienen dos números ordinales β y α, entonces ellos son conjuntos bien ordenados, y por lo tanto son similares o uno de ellos será similar a un segmento inicial del otro, todo conjunto de números ordinales está totalmente ordenado.

Un hecho importante es que todo conjunto no vacío de números ordinales esta bien ordenado. El hecho de que un número ordinal α tenga un elemento menor β, significa que no existe $\gamma \in \alpha$ tal que $\gamma \in \beta$. La demostración dice así: si $\alpha \in X$ un conjunto de números ordinales y si α es el menor elemento de X estamos bien, porque $\alpha \cap X = \emptyset$ y si $\beta \in X \Rightarrow \beta \notin \alpha$, si no fuera α el elemento menor, entonces existe un $\beta \in X$ (ya que $\alpha \cap X \neq \emptyset$) que será el menor elemento de $\alpha \cap X$ ya que este conjunto está contenido en α y por lo tanto es un número ordinal y en particular un conjunto bien ordenado. Ese β será un menor elemento de A en el orden <, ya que para cualquier $\gamma \in X$ distinto de β, si

$$\gamma \in \alpha \Rightarrow \gamma \in \alpha \cap X \Rightarrow \gamma \notin \beta$$

$$\gamma \notin \alpha \Rightarrow \alpha \in \gamma \Rightarrow \alpha \subset \gamma \Rightarrow \beta \in \gamma$$

Como son conjuntos transitivos los números ordinales, serán equivalentes siempre que ellos son diferentes[48]:

$$\beta \in \alpha$$

$$\beta \subset \alpha$$

$$\alpha \text{ es una continuación de } \beta,$$

si alguna de ellas sucede, se dirá que $\beta < \alpha$.

Esta relación que acabamos de definir tiene todas las propiedades de buen orden. En realidad, se cumple que todo conjunto de ordinales está bien ordenado.

Los números ordinales que son *finitos* son los números naturales, los elementos de ω. Los demás son llamados *transfinitos* o simplemente *infinitos*, ω es el número ordinal transfinito más pequeño. Todo número ordinal distinto de 0 tiene un predecesor inmediato, si un número ordinal transfinito α tiene un predecesor inmediato β, entonces $\alpha = \beta^+$. Un número *ordinal límite* es el que no tiene predecesores inmediatos.

Notemos que si α es un número ordinal, $\alpha = \{\beta | \beta \text{ ordinal y } \beta < \alpha\}$. Si α es un sucesor, digamos $\beta + 1$, tendrá un elemento mayor, que será β. Si es un ordinal límite, no tendrá elemento mayor y será $\alpha = \sup\{\beta | \beta < \alpha\}$.

[48]Continuación se ve en la p. 20.

El propósito de contar los elementos de un conjunto es comparar el tamaño entre ellos, el método que ocupamos con los números ordinales es acomodar los elementos en un orden apropiado. El Principio del Buen Orden de la p. 46 nos dice que todo conjunto puede ser bien ordenado, y como acabamos de ver, para este tipo de conjuntos parece una buena medida de su tamaño su número ordinal. Si denotamos el número ordinal de un conjunto X por ord X, no nos basta con encontrar uno de sus «buenos ordenes» de cada conjunto X e Y y comparar ord X con ord Y; el problema es comparar el tamaño de los conjuntos cuando no parece haber algo entre ellos, otro problema es que un conjunto puede ser bien ordenado de varias formas[49]. En realidad, ord X mide más el buen orden de X que su tamaño. Suponga que se tiene a ω el conjunto de todos los números naturales, si definimos un orden en ω poniendo a 0 enfrente de cualquier otro elemento ($m \neq 0 \Rightarrow m < 0$) y para los demás elementos se deja el orden habitual, se encuentra un buen orden de ω y este buen orden tiene como número ordinal a $\omega + 1$ ya que su ordenación es la misma que la de los números transfinitos menores que $\omega + 1$.

Ya que ord X<ord Y ssi X es similar a un segmento inicial de Y, para comparar el tamaño de los números ordinales necesitamos saber el orden de los conjuntos, o también el concepto de similaridad. El análogo que se encuentra en la teoría para conjuntos no ordenados (o conjuntos donde no nos importa comparar el orden) es el de equipotencia o equivalencia[50]. Se dirá que Y *domina* a X si $X \preceq Y$ (que significa que existe una función biyectiva entre X y un subconjunto de Y o que existe una función inyectiva de X a Y), y también se tiene una relación de orden parcial entre todos los pares de subconjuntos de un conjunto I dado, que tengan la relación de dominancia. Cualquier conjunto X es equipotente a un subconjunto de él mismo (en realidad, a él mismo), mediante la función identidad. Si f es una biyección de $X \preceq Y$, y g también entre $Y \preceq Z$, entonces la función compuesta y restringida $g|_{Ran f} \circ f$ es una biyección entre X y un subconjunto de Z, ya que es una composición de funciones biyectivas. La antisimetría de la relación se verá en la p. 104 ya que es un resultado importante llamado Teorema de Schröder-Bernstein.

En realidad la relación $\preceq$ es un orden total entre conjuntos, que nos dice que para cualesquiera X y Y se tiene que $X \preceq Y$ o que $Y \preceq X$. Si Y domina a X pero X no domina a Y, se dirá que Y *domina estrictamente* a X, denotándolo por $X \prec Y$.

Dos teoremas importantes y muy utilizados serán:

Teorema 1.9.1 (Cantor)**.** *Todo conjunto X está estrictamente dominado por su conjunto potencia $\mathscr{P}(X)$.*

Teorema 1.9.2. *Para todo conjunto X, se cumple que su conjunto potencia es equivalente a el conjunto de todas las funciones que van de X a $\{0,1\}$, i.e.: $\mathscr{P}(X) \sim 2^X$.*

Para comparar tamaños de conjuntos usaremos el concepto de *número cardinal* y asociaremos a cada conjunto X un número cardinal, denotándolo por card X, o de la forma definida en la p. 16 se utilizará card $X = |X|$. Se busca un orden también denotado por $\leq$ y se busca que a cada número cardinal $\mathfrak{n}$ se le pueda encontrar un conjunto X tal que card $X = \mathfrak{n}$. La conexión con lo que hemos visto en párrafos anteriores será que card $X =$ card Y ssi $X \sim Y$, además card $X <$ card Y[51] ssi $X \prec Y$ (i.e., que exista una función inyectiva de X a un subconjunto de Y y no de lado contrario, o que X sea equipotente a un subconjunto de Y y no de manera contraria).

Es importante notar que todo conjunto es equivalente a un número ordinal.

Así, podemos decir que el número cardinal de un conjunto es la propiedad que éste tiene en común con todos los conjuntos equivalentes a él. Pero para cualquier conjunto X, hay muchos conjuntos equivalentes a éste, aunque sabemos que todo conjunto es equivalente a un número ordinal, pero como dijimos,

[49] Claro, dijimos que si se logra bien ordenar para hacerlo un número ordinal, será con una única relación, pero de todas formas, pueden haber varios ordenes que den varios números ordinales.

[50] Véase la p. 16

[51] Como siempre si $\mathfrak{n}$ y $\mathfrak{m}$ son números cardinales, $\mathfrak{n} < \mathfrak{m}$ significará que $\mathfrak{n} \leq \mathfrak{m}$ y $\mathfrak{n} \neq \mathfrak{m}$.

dependiendo del buen orden dado o encontrado, el conjunto es equivalente a muchos números ordinales. Pero hay un punto bueno en esto, en realidad es un conjunto la clase de números ordinales equivalentes a un conjunto X. Para demostrarlo, primero debemos saber que es posible encontrar un número ordinal mayor estrictamente a todos los números ordinales equivalentes a X; si β es un ordinal equivalente a $\mathscr{P}(X)$ y α es un ordinal equivalente a X, por el teorema 1.9.1 $\alpha < \beta$ y eso implica que card $\alpha <$ card β. Pero $\alpha < \beta$ es lo mismo que $\alpha \in \beta$ por ser números ordinales, y así hemos exhibido un conjunto β que contiene a todo número ordinal equivalente a X, diciendo que éstos forman un conjunto. ¿Cuál de todos los números ordinales se escoge por el cardinal de X?, pues como todo conjunto de números ordinales está bien ordenado por el Principio del Buen Orden de la p. 46, el primer elemento será el elegido.

De esta manera relacionamos a los dos conceptos de ordinalidad y cardinalidad. Cantor en su trabajo «Contribuciones a la fundamentación de la teoría de conjuntos transfinita» fue cuando definió con rigor estos dos conceptos. El «ordinal» de un conjunto parcialmente lo idealiza como el concepto al que se llega cuando se hace una abstracción de sus elementos y se conserva su orden, diciendo que dos conjuntos tienen el mismo ordinal ssi se pueden poner en correspondencia biyectiva conservando el orden. El «cardinal» de un conjunto es lo que él especifica como el concepto al que se llega cuando se hace una abstracción de sus elementos, así como de cualquier orden en ellos.

Un número ordinal α será llamado *ordinal inicial* si no es equipotente a algún $\beta < \alpha$.

Un *número cardinal* es un número ordinal α tal que si β es un número ordinal equivalente a él, lo que nos dice que card $\alpha =$ card β, entonces $\alpha \leq \beta$. Si tenemos un conjunto X, card X significará el número cardinal de X o la *potencia* de X o el menor número ordinal equivalente a X. También es lo mismo decir que card X es el único ordinal inicial equipotente a X.

Una conclusión que se saca de lo anterior, es que como todo conjunto es equivalente a card X ($X \sim$ card X), se tiene que si card X= card Y entonces $X \sim Y$; recíprocamente, si $X \sim Y$, entonces card $X \sim$ card Y, pero como card X es el número ordinal menor equivalente a X, se cumple que card $X \leq$ card Y, simétricamente para Y se concluye que card X = card Y. Este si y solo si era lo que se buscaba al definir los números cardinales.

Un número cardinal será *finito* si es el número cardinal de un conjunto finito, de otra manera, será *infinito*. Un número ordinal finito (un número natural) solo es equivalente a ese número natural, y por lo tanto, el número cardinal será el mismo que el ordinal.

Ya vimos lo que significa para números ordinales que $\alpha < \beta$. De esta manera, los números cardinales también vienen equipados con un orden. Si tenemos que card $X <$ card Y, entonces card $X \subset$ card Y, y por eso $X \preceq Y$, es decir, X es equivalente a un subconjunto de Y. No puede pasar que $X \sim Y$ debido a que se tendría que card X = card Y, por lo tanto $X \prec Y$. Por la parte contraria, si $X \prec Y$ no puede pasar que card $Y \leq$ card X porque eso implicaría que Y sería equivalente a un subconjunto de X lo cuál no pasa, por lo tanto card $X <$ card Y.

Las propiedades de los números ordinales, se convierten en propiedades de los números cardinales. Este hecho simplifica la teoría, pero también tiene desventajas. En la aritmética de cardinales y ordinales, para dos cardinales a y b, $a+b$ tiene dos significados, y solo el contexto y la notación podrán salvarnos de caer en la ambigüedad.

Un símbolo especial para los números cardinales es la primera letra del alfabeto hebreo, $\aleph$ se leé *aleph* y es un número ordinal infinito que es un cardinal. El número ordinal transfinito más pequeño ω que además es un ordinal se denota por $\aleph_0$, y éste es contable (card $\omega = \aleph_0$). El número ordinal transfinito más pequeño e incontable se denota por Ω. Las propiedades análogas entre ellos son que ω es un conjunto infinito bien ordenado que consta de puros elementos finitos; para Ω la característica es que es un conjunto incontable infinito bien ordenado cuyos elementos o segmentos iniciales son contables. Ω satisface la definición de número cardinal (debido a que es el más pequeño) y su papel en cardinal se denota por $\aleph_1$. Otra definición equivalente de $\aleph_1$ es ser el sucesor inmediato de $\aleph_0$.

Otra definición de los números cardinales $\aleph_\alpha$ será dada. Para todo número cardinal infinito a, si denotamos por $c(a)$ a todos los números cardinales infinitos estrictamente menores que a, entonces se tiene que $c(a)$ es un conjunto bien ordenado, con un número ordinal α, de esta manera se tiene $a = \aleph_\alpha$, además, si $\alpha > 0$, $\aleph_\alpha$ es el número cardinal más pequeño que es estrictamente mayor a todos los $\aleph_\beta$´s con $\beta < \alpha$. Se cumple de esta manera que si $\operatorname{card} X = \aleph_\alpha$ y $\operatorname{card} Y = \aleph_\beta$, $\operatorname{card} X < \operatorname{card} Y$ ssi $\aleph_\alpha < \aleph_\beta$.

1.10. INDUCCIÓN TRANSFINITA

Como se vio en la página 26, existe un método para generalizar que una propiedad se cumpla para cualquier número natural, que es la inducción matemática. Pero hay un tipo de inducción que no solo aplica a los elementos de ω, sino también a los elementos de cada ordinal α, llamado principio de inducción transfinita, y es algo así como una generalización de la inducción matemática.

Éste principio dice que si un subconjunto S de α tiene la propiedad de que contiene a un ordinal $\beta \in \alpha$ siempre que contenga todos los ordinales menores que β, entonces se cumplirá que $S = \alpha$, esto claro, aplicado a ordinales $\alpha > \omega$. Generalmente asume esta forma una demostración por inducción transfinita:

1. demostrar que 0 tiene la propiedad pedida,

2. para ordinales β que están entre $0 < \beta < \alpha$ se descompone en dos casos: β es un ordinal sucesor y por eso se asumirá que el ordinal del que es sucesor tiene la propiedad; β es un ordinal límite y por eso se asumirá que todos los ordinales menores que β tienen la propiedad requerida,

3. en cualquiera de los dos casos del caso anterior, demostrar que β tiene dicha propiedad.

Una familia indexada por un número natural n o por el conjunto ω es llamada sucesión, pero cuando el índice es mayor que ω se dirá que se tiene una sucesión transfinita.

También se tienen definiciones por inducción transfinita. Se verán algunas en ésta tesis. Para definir por medio de inducción transfinita una sucesión $\{x_\beta\}$ indexada por un número ordinal α (i.e, que está indexada por el conjunto de todos los ordinales $\beta < \alpha$), se tiene que definir cada x_β con $\beta < \alpha$ en términos de los x_ξ con $\xi < \beta$. Muchas veces, la definición toma la siguiente forma:

1. definir x_0,

2. para definir x_β para ordinales β que están entre $0 < \beta < \alpha$ se descompone en dos casos: β es un ordinal sucesor $\beta = \xi + 1$ y se define a x_β en términos de x_ξ; β es un ordinal límite y por eso se define x_β en términos de los x_ξ con $\xi < \beta$,

Capítulo 2

PRELIMINARES HISTÓRICAS

The Axiom of Choice (together with the Continuum Hypothesis) is probably the most interesting and most discussed axiom in mathematics after Euclid´s Axiom of Parallels.

P. Bernays y A.A. Fraenkel [Her, p. 1].

Antes que nada empezaré introduciendo al lector en el contexto histórico del Axioma. En esta parte se verán desde el punto de vista histórico los acontecimientos precedentes y subsecuentes del Axioma, sin embargo, se introducirán matemáticamente conceptos y definiciones que se usarán en el texto. Se relata la historia del Axioma, escribiendo los acontecimientos de relevancia para entender el contexto histórico, dichos acontecimientos principalmente eran resultados, teoremas y proposiciones a los que llegaban los matemáticos. La sección 2.1 menciona las varias formas en que los matemáticos ocupaban al Axioma sin siquiera notarlo, y cómo en el mundo era más necesario, a veces sin saber que la demostración de algún teorema dependía del Axioma, relata las diferencias en los usos del Axioma (uso inevitable, uso evitable, uso de alguna forma más débil del Axioma). Hubo una época donde los matemáticos ponían especial interés a los métodos y las formas en que hacían sus demostraciones, debido a que no sabían definir con precisión las palabras conjunto, existencia, y de eso trata la secc. 2.2. Luego de tanta confusión, se crearon sistemas axiomáticos (2.3) con la intención de burlar dichos baches en la teoría, para no dejar lugar a paradojas ni construcciones no sustentadas. Con dichas Axiomatizaciones, se crearon nuevas posturas de pensamiento respecto a las matemáticas, aceptando o rechazando el Axioma o alguna forma de éste, debido principalmente, a las pruebas de independencia del Axioma con los sistemas axiomáticos de esas fechas, con ello, se difundió, profundizó y expandió el campo de aplicación del Axioma (2.4, 2.5, 2.6, 2.7).

Hoy en día la mayoría de los matemáticos acepta el Axioma debido a que sin él, como se verá adelante, las matemáticas modernas se verían tan afectadas que serían irreconocibles para muchos (incluso no matemáticos), debido a que conceptos tan fundamentales como el de conjunto finito (véase la subsección 3.7.1), se verían afectados debido a que existirían varias definiciones, y dependiendo cual sea la definición aceptada, sería el «finito» que se manejaría.

Lo que hoy se conoce como «teoría de conjuntos» es una ciencia que casi desarrolló solo Cantor. El «comienzo matemático» cuando Cantor intentó demostrar la unicidad de los desarrollos en series trigonométricas de funciones arbitrarias. Consiguió el resultado para series convergentes en $[0,2\pi]$, y aunque la convergencia no se cumplía en algunos puntos excepcionales, la unicidad se seguía cumpliendo. Luego, para estudiar sobre las excepciones definió lo que es un *conjunto derivado* de X escribiéndolo como $X´$. Además, calculó derivados sucesivos, $X^{(1)}$, $X^{(2)}$, $X^{(3)}$, $\ldots$, dijo que un conjunto X era de *primera especie* si $\exists n \in \mathbb{N}$ con $X^{(n)} = \emptyset$. Un conjunto es de *segunda especie* si no es de primera especie. Él trató de entender la diferencia entre estas dos nociones. Un conjunto de primera especie es numerable, Cantor lo intuyó, y con esto intentó saber si existía una función biyectiva entre $\mathbb{N}$ y $\mathbb{R}$. Se lo preguntó en noviembre de 1873 a su amigo Richard Dedekind y lo demostró al final del año. A principios del siguiente año, preguntó si se podían biyectar los puntos de una superficie (como un cuadrado) con los puntos de un segmento de recta. Tres años después, Cantor anunció en contra de lo que muchos creían, que tal cosa era posible. Con esto demostró la existencia de dos potencias diferentes, la de $\mathbb{R}^2$ y la de $\mathbb{N}$. De ahí surgió la duda de si existían conjuntos de diferente potencia que la de los números reales y de los números naturales. Se ayudó de los conjuntos de primera y segunda especie, encontrando que en conjunto de segunda especie, se forma una sucesión decreciente $X^{(1)} \supset X^{(2)} \supset X^{(3)} \supset \cdots$, y estudió a $X^{(\omega)} = \bigcap_{n=1}^{\infty} X^{(n)}$. Siguiendo su razonamiento, consideró el conjunto derivado de éste conjunto, y prosiguió con la formación de sus conjuntos derivados. Pero lo importante, es que enfatizó esos símbolos infinitos que había ocupado. A partir de este punto, le surge la necesidad de trabajar con los nuevos conceptos que definió como números ordinales y cardinales, abriendo con esto, las puertas a lo que sería un vasto y nuevo mundo matemático.

Por otra parte, la historia del axioma comienza cuando los filósofos del siglo XIX empezaron a revolucionar en el campo del pensamiento, mas específicamente, la concepción del *infinito*. En esa época, era frecuente que hubiera un pensamiento Aristotélico sobre el *horror infiniti* que perdió fuerza cuando Cantor lo defendió, y dijo que era necesario aceptarlo. Sin embargo, las actitudes filosóficas que los Alemanes tomaron sobre el infinito (que tendían más y más a aceptar el concepto), dieron lugar a cambios de formas de pensar, posturas, incluso a la percepción del mundo y la naturaleza. La filosofía de la naturaleza estaba repleta de implicaciones respecto al problema del infinito. Las discusiones se basaban principalmente en decidir si el espacio y el tiempo son infinitos o están acotados; o si el universo físico es o no hecho de elementos simples. Estos son los personificadores del antagonismo que discutía el influyente Kant en su libro «Kritik der reinen Vernunft [455-471]», refiriéndose a posturas sobre infinito. Pensadores filosóficos prominentes (algunos también eran matemáticos) influenciaron a muchos matemáticos de la época, como Riemann, Leibniz, Bolzano, Dedekind e incluso a Gauss y Cantor. Siendo éstos sumergidos en problemáticas sobre el infinito, tomaron posturas que en su mayoría eran las de sus mentores. De aquí que al final del siglo, el desarrollo de temas matemáticos como el cálculo de límites (Leibniz), fundamentos del cálculo (Leibniz), geometría no Euclidiana (Gauss), ecuaciones diferenciales parciales y geometría diferencial (Riemann), dieron lugar a un uso inevitable del infinito, pero casi sin haber tenido fundamentos matemáticos para hacerlo, debido en parte a que los artículos que se escribieron en ésa época eran muy atacados y escasos.

Un problema grande que ésto trajo consigo, fue el uso intuitivo de generalizaciones en casos infinitos, casi sin dudar de su veracidad debido al poco conocimiento científico que se tenía sobre el concepto. Fue hasta cuando salió a relucir el Axioma, que el matemático empezó a estudiar profundamente colecciones infinitas de conjuntos, al haberse dado cuenta que tales generalizaciones no eran del todo claras, y que era necesario reformular varios postulados debido a su «dudoso» sustento. De hecho, el Axioma concatena los cambios fundamentales (filosóficos, matemáticos y psicológicos) que dieron lugar en esa época, comenzando los pensadores a poner énfasis en los pasos «obvios» o «intuitivamente claros» de sus demostraciones, empezando con una revolución que culminó con el *Teorema de la incompletez de Gödel* y la *axiomatización de la Teoría de Conjuntos* entre otras cosas, dando lugar así, a las matemáticas modernas. En esa revolución hacia las matemáticas nuevas, los trabajos de Cantor mencionados antes fueron fundamentales, ellos lo llevaron a la consideración de clases arbitrarias y conjuntos infinitos, conllevando una formalización acerca de la noción de conjunto.

El Axioma de Elección (Axiom of Choice, Auswahlpostulat, Axiome du Choix) dice que para todo conjunto S, existe una función f que asocia a cada conjunto no vacío A de S, un único elemento $f(A) \in A$, a tal f se le llama función de elección para S. En palabras simples, podemos siempre «escoger» un elemento de cada subconjunto A de S (claro, el subconjunto debe ser no vacío). Pero si tenemos subconjuntos no vacíos de S, ¿porqué no podríamos escoger elementos de ahí, si al fin y al cabo son no vacíos?, el problema radica cuando tenemos conjuntos infinitos, se debe escoger un elemento de "todos" los conjuntos. Tales elecciones no se pueden justificar diciendo: escogemos un elemento de «cada uno» de sus subconjuntos, por que regresamos donde empezamos: ¿cómo asegurar que «siempre» podemos escoger tal elemento?. Una forma de asegurarlo es dando una regla que te diga exactamente, cómo escoger ese elemento de cada subconjunto, así, no necesitamos del Axioma debido a que siempre que busquemos tal elemento, aplicamos la regla y lo encontramos. Pero, ¿cuando no hay forma de encontrar tal regla?. Es ahí, cuando falta tal método explícito, cuando el infinito desafía nuestra intuición y no podemos enfrentar las dificultades que nuestra capacidad tiene, cuando requerimos del Axioma. Algunos, como los constructivistas (véase la p. 84 que es la sección 2.7), y críticos que incluyen prominencias como J.E. Littlewood y B. Russell, decían que tal regla (el Axioma) no era del todo válida si no se da la forma explícita de cómo aplicarla; Russell dijo: «*the apparent evidence of the axiom tends to dissipate upon the influence of reflection*» [Her, p. VIII].

2.1. ANTECEDENTES

Cerca de la época de la formulación por Zermelo de una versión del Axioma, en 1908, se piensa que hasta septiembre de 1904 muchos de los matemáticos no eran concientes de utilizar elecciones arbitrarias y formas débiles o equivalentes[1] del Axioma, incluso algunos de ellos sabían que no hacían algo fuera de lo normal, o algo dudoso, excepto por (hasta donde se sabe) 3 italianos, que evitaban hacer tales elecciones arbitrarias en sus demostraciones durante el periodo 1890-1902. Uno de ellos escribió esto haciendo referencia explícita de que encontraba un problema en su demostración:

> *...but as one cannot apply infinitely many times an «arbitrary» rule by wich one assigns to a class A an individual of this class, a «determinate» rule is stated here.*
>
> *Giuseppe Peano, 1890*

A veces, es posible evitar que el Axioma sea necesario en la demostración, hay ocasiones en las que alguien más modifica la demostración para que no lo necesite; otras, cuando es inevitable, es cuando la afirmación o proposición no puede ser probada en la *Teoría de Conjuntos de Zermelo-Fraenkel* (véase la sección 1.8.2), pero puede ser demostrada en la *Teoría de Conjuntos de Zermelo-Fraenkel con el Axioma de Elección*. Antes de 1904, el uso casi inconsciente del Axioma era principalmente en áreas como Análisis Real, Teoría de Números, Topología y Teoría de Conjuntos. Después de eso, varios buscaban saber si en un teorema era necesario o no el uso del Axioma, buscaban conocer de qué axiomas dependía el teorema.

La primera proposición a considerar (que ayudó a las investigaciones de Cantor en las épocas de 1870), que es muy utilizada y que depende de una forma débil del Axioma es:

Proposición 2.1.1. *La unión contable[2] de conjuntos contables es contable.*

---

[1] Una afirmación P es equivalente a otra Q, si una implica la otra y viceversa (i.e $P \Leftrightarrow Q$). Una afirmación P es más debil que otra Q, si ésta implica a la primera y no al revés (i.e $Q \Rightarrow P \wedge \neg(P \Rightarrow Q)$). Una afirmación es equivalente al Axioma, si tal equivalencia puede ser probada mediante los axiomas de **ZF**.

[2] Un conjunto se dice *contable* (*a lo más numerable*) cuando es *finito* o *denumerable*. Un conjunto se dice *denumerable* si existe una biyección entre él y $\mathbb{N}$. Se da otra demostración en la página 159.

A esta proposición le llamaremos *Teorema de la Unión Contable*. Para demostrarla, sean $A_1, A_2, \ldots$ conjuntos denumerables, entonces, cada uno de ellos, tiene elementos que llamaremos $a_{i1}, a_{i2}, \ldots$ para el $i-$ésimo conjunto. Luego, la unión A de todos los A_i tiene elementos a_{ij}, con i, j números naturales. Como Cantor demostró, el conjunto de los números racionales es denumerable, y esto implica que A, que tiene elementos a los que se puede poner en correspondencia con un subconjunto de un conjunto denumerable, es contable.

¿Dónde esta la dependencia de la proposición con el Axioma?, ésta radica en la numeración de los elementos de todos los A_i, ya que existen muchas biyecciones entre los números racionales y A. Hay infinitos A_i, para los cuales hay infinitas biyecciones entre éstos y $\mathbb{N}$. Por el Axioma, a cada A_i le asignamos una única biyección que llamaremos a_i (de hecho, la función asigna un único elemento de A_i para cada subconjunto de $\mathbb{N}$, y solo lo restringimos a los subconjuntos de un único elemento), entonces, a ésta biyección evaluada en cada número natural le corresponde un único elemento de A_i, es decir, $a_i(j)$, o lo que sería lo mismo, a_{ij}. Así, todo elemento de A estaría bien definido. En este caso la dependencia del Axioma como se dijo, es débil, ya que si éste aplicara solo a conjuntos denumerables, en vez de conjuntos arbitrarios, la demostración seguiría siendo válida. Cuando ese es el caso, se le llama el *Axioma de Elección Denumerable*, llamado en adelante el *Axioma Denumerable*. Además, sin el Axioma Denumerable, no se puede demostrar la proposición (2.1.1), ya que existen modelos[3] en *ZF* en que el Axioma Denumerable es falso, y también la proposición (2.1.1)[4].

La siguiente proposición es muy importante:

Proposición 2.1.2. *Todo conjunto infinito tiene un subconjunto denumerable*[5].

En la demostración de Russell en 1902, se nota cómo el Axioma está implicado: ya que A es infinito, existen subconjuntos $A_1, A_2, \ldots$ de A tales que, para cada n, A_n tiene exactamente n elementos, y además $A_n \subset A_{n+1}$. Luego, la unión B de todos los A_n es un conjunto denumerable, por la proposición 2.1.1.

Parece ser que Russell no se dio cuenta de las infinitas elecciones denumerables que hizo en la demostración: una vez obtenido A_n, para encontrar A_{n+1} uno debe de elegir un elemento de $A - A_n$, y esto se hace infinitas numerables ocasiones.

En realidad fue Cantor quien dio la primera demostración publicada de esta proposición, y en las demostraciones que dieron los matemáticos en las décadas posteriores, se ve la influencia de la demostración de Cantor. Éste dijo: Si uno ha quitado distintos elementos $t_1, t_2, \ldots, t_n$ de A por una regla, entonces es siempre posible quitar otro elemento t_{n+1}, y el conjunto $\{t_1, t_2, \ldots, t_{n+1}, \ldots\}$ es un subconjunto denumerable de A. Aquí Cantor burló las elecciones arbitrarias diciendo que con alguna regla se podían quitar tales elementos, el problema de su argumento es que aunque se seleccionen con una regla los primeros n para todo $n \in \mathbb{N}$, no se da una regla para el subconjunto denumerable «completo».

En 1901, Cassius Keyser no hizo las elecciones tan visibles, solo afirmó que si para todo n, el conjunto $\{1, 2, \ldots, n\}$ puede ser mandado mediante una función inyectiva a A, entonces claramente éste contenía un subconjunto denumerable.

De varias formas resolvieron o trataron de resolver el problema estos matemáticos, pero las elecciones arbitrarias o infinitas numerables están o muy obvias o muy escondidas, sin embargo, inevitables.

Es importante destacar que entre las proposiciones (2.1.1) y (2.1.2) no hay implicación alguna, esto en *ZF*.

Otro ejemplo es el llamado *Principio de Partición* :

---

[3]Véase la p. 23 para la referencia de qué es un modelo.

[4]Existe un modelo en *ZF* donde $\mathbb{R}$ aunque es un conjunto incontable, también es la unión contable de conjuntos contables.

[5]Un conjunto A es *finito* si es vacío, o si para algún $n \in \mathbb{N}$ existe una biyección entre A y $\{1,2,\ldots,n\}$. Si un conjunto no es *finito* se llama *infinito*.

Proposición 2.1.3. *Si un conjunto A es particionado*[6] *en una familia S de conjuntos disjuntos no vacíos, entonces S es* equipotente[7] *a un subconjunto de A, i.e.* $(S \sim B) \wedge (B \subset A)$.

En términos de una función f con dominio A, el principio nos dice que $(f(A) \sim B) \wedge (B \subset A)$. Cantor usó cerca de 1880 una forma de este principio para estudiar propiedades topológicas de la recta real. La prueba de (2.1.3) se basa en la elección de un elemento de cada conjunto en S para obtener una biyección entre S y un subconjunto de A. Ahí se ve implicado el Axioma.

La última proposición esta íntimamente ligada con la proposición de que todo conjunto tiene un buen-orden[8]:

Proposición 2.1.4. *Para todo cardinal* $\mathfrak{m}$ *y* $\mathfrak{n}$*, se cumple alguna de estas tres condiciones:* $\mathfrak{m}<\mathfrak{n}$ *o* $\mathfrak{m}=\mathfrak{n}$ *o* $\mathfrak{m}>\mathfrak{n}$[9].

Hablando de conjuntos, la proposición (2.1.4) dice que para cualesquier dos conjuntos A y B, éstos son comparables[10], i.e., uno es equipotente a un subconjunto del otro. Fue en 1895 cuando Cantor afirmó la proposición (2.1.4), aunque sin demostración. 4 años más tarde, aseguró que si todo conjunto tiene un buen-orden, entonces se sigue (2.1.4). No fue sino hasta 1915, cuando Friedrich Hartogs demostró que estas dos afirmaciones eran equivalentes, dando una ventaja al Axioma, debido a que cualquier teoría sobre Cardinales que se requiera amplia y fructífera, debería aceptarlo.

Estas 4 proposiciones, forman el principio de la historia del Axioma. La siguiente subsección tratará de introducir los orígenes del Axioma.

2.1.1. Eventos Preliminares

Algunos autores (véase [Moo, p. 11]) delimitan 4 etapas referentes al uso del Axioma durante los eventos preliminares a la formulación del mismo:

1. *Escoger un elemento no especificado de un solo conjunto.* Vestigios de esto se encuentran en el libro de Euclides «Elementos». Tales elecciones forman la base para la generalización de demostrar, escogiendo un objeto, y luego aplicar un argumento a tal. También se incluye aquí la elección de un elemento de cada uno de un numero finito de conjuntos; es importante señalar que en este caso no es necesario el Axioma, incluso si los conjuntos tuvieran infinitos elementos, debido a que para un solo conjunto, la elección arbitraria de un elemento, es posible por definición, y procediendo por inducción, se puede aplicar el procedimiento para un número finito de conjuntos.

2. *Hacer un número infinito de elecciones, dando una regla.* Los mejores candidatos para esta etapa son los matemáticos sumergidos en Teoría de Números o en Análisis del siglo XIX, por ejemplo unos escogiendo elementos de clases de equivalencia[11] y los otros de sucesiones. Ellos supusieron la existencia de familias infinitas de conjuntos.

3. *Hacer un número infinito de elecciones, dejando la regla sin especificar.* Cauchy, Cantor y otros aplicaron este procedimiento. Esto predispuso la última etapa debido a que por creerlo obvio, o por falta de cautela, muchos no vieron el problema de hacer tales elecciones.

[6] La p. 15 dice qué es una partición.

[7] Recuerde que dos conjuntos son *equipotentes* si existe una biyección entre ellos.

[8] Véase la p. 20 para la definición de buen orden.

[9] $\mathfrak{m}<\mathfrak{n}$ si para todo conjunto A y B de cardinalidad $\mathfrak{m}$ y $\mathfrak{n}$ respectivamente, A es equipotente a un subconjunto de B pero no a B. Para recordar lo que es cardinalidad véase la p. 35.

[10] Véase la p. 16.

[11] Véase la p. 4.

4. *Lo mismo que las dos anteriores, pero para este procedimiento no era posible dar tal regla.* En 1871, Cantor hizo elecciones arbitrarias de una secuencia infinita en la cual no hay regla para tales elecciones, y así, el Axioma denumerable fue necesario por primera vez. Sin embargo, él no supo que tal regla no era posible encontrar, ni vio la línea que cruzó después de tal procedimiento. Después de eso, varios matemáticos usaron tales elecciones de forma prominente sin aclarar que una regla sería necesaria, o que por lo menos, daría un poco más de formalidad. De esta etapa surgió la solución de Zermelo del problema del buen-orden, y la formulación explícita del Axioma también por Zermelo.

Las demostraciones en épocas pasadas eran del tipo constructivista en su mayoría, si uno quería demostrar que cierto objeto existía, uno construía o encontraba dicho objeto, tal vez usando los objetos que antes se habían demostrado existentes.

Pero ésto cambió con el paso del tiempo, debido a la especialización de los temas de estudio. Una tendencia visible fue en el trabajo más significativo de Teoría de Números de las fechas de 1800, hecho por Gauss y llamado *Disquisitiones Arithmeticae*. Cuando demostraba algo sobre formas cuadráticas binarias ($ax^2+2bxy+cy^2$), él demostró que para las que tienen discriminante $d=b^2-ac$, existe un único entero n, tal que las formas podían ser particionadas en n clases mediante cierta relación de equivalencia. Gauss se dio cuenta que había muchas formas de seleccionar a un representante de cada clase, así que dio una forma explícita de hallar cada uno de manera única; encontrándose así entre la primera y la segunda etapa.

Hasta 1821, cuando Cauchy escrutiñaba Análisis, la tercera etapa sale a relucir mientras demostraba una versión del *Teorema del Valor Intermedio* :

Teorema 2.1.1. *Cualquier función real f continua en un intervalo cerrado $[a,b]$ tiene una raíz ahí, si $f(a)$ y $f(b)$ tienen signos contrarios.*

Demostración. Para algún entero dado $m>1$, Cauchy encontró que la sucesión finita

$$f(a), f\left(a+\frac{b-a}{m}\right), f\left(a+\frac{2(b-a)}{m}\right), \ldots, f(b)$$

debe tener un par consecutivo con signos opuestos. Sean a_1 y b_1 tal par, con $a_1<b_1$ entonces $b_1-a_1=(b-a)/m$. Luego consideró los m puntos que dividen el intervalo $[a_1,b_1]$ en partes iguales, y de la misma manera se obtiene

$$f(a_1), f\left(a_1+\frac{b_1-a_1}{m}\right), f\left(a_1+\frac{2(b_1-a_1)}{m}\right), \ldots, f(b_1)$$

$$= f(a_1), f\left(a_1+\frac{b-a}{m^2}\right), f\left(a_1+\frac{2(b-a)}{m^2}\right), \ldots, f(b_1)$$

y se escogen a_2 y b_2 con propiedades: $a_2<b_2$, $b_2-a_2=(b-a)/m^2$. Procediendo así, se encuentran las sucesiones $a_1, a_2, \ldots,$ y $b_1, b_2, \ldots,$ que convergen al mismo punto p $\big($esto porque $\lim_{n\to\infty}(b_n - a_n) = \lim_{n\to\infty}\dfrac{b-a}{m^n}=0\big)$. Luego, como $f(a_n)$ y $f(b_n)$ tienen diferentes signos para todo n, y como f es continua, se sigue que $f(p)=0$. □

Cauchy usó en su demostración una forma especial del Axioma, el llamado *Principio de Elecciones Dependientes* :

(2.1.1.1) Si S es una relación en un conjunto A tal que para todo $x \in A$ existe algún $y \in A$ con xSy, entonces existe una secuencia $a_1,\ a_2,\ \ldots$ tal que para todo $n \in \mathbb{N}$, $a_n \in A$ y $a_n S a_{n+1}$ son verdaderas.

Pero en realidad ese postulado no es necesario, ya que se conoce una regla para escoger los elementos de la sucesión. Si se toma para cada n la pareja más cercana a $f(a)$ que tenga signos contrarios, y esa elección en cada paso siempre es posible ya que es una secuencia finita de donde escoger, entonces ya no se necesita el Principio de Elecciones Dependientes.

La cuarta etapa empezó en octubre de 1871, cuando escribió Eduard Heine un artículo sobre Análisis, postulando:

Teorema 2.1.2 (Heine)**.** *Una función real es continua en un punto p ssi es secuencialmente continua en el punto p*[12].

Éste teorema depende del Axioma Denumerable, pero pasaron 10 años hasta que la comunidad se diera cuenta de esto. Cantor lo ocupó y se le atribuye la dispersión del teorema a la comunidad, haciendo que se reproduzca su demostración u ocupándola como resultado. A lo que me refiero con esto, es que como solía suceder, algunos de los teoremas que se demostraban después de la aparición de un resultado importante, contenían la misma idea en su demostración que la idea del resultado, y por lo tanto el uso (no visible) del Axioma o una equivalencia de éste en algunos casos.

En esas épocas, específicamente en 1871, Dedekind utilizó la definición de módulo de A como un conjunto de números complejos, cerrado bajo la suma y la resta, disponiendo que $a \equiv b(mod\ A)$ fuera lo mismo que $a - b \in A$. Resultando en este teorema:

Teorema 2.1.3. *Si A y B son módulos, entonces existe un subconjunto B_1 de B tal que para todo $b \in B$ existe un único $b_1 \in B$ con $b \equiv b_1(mod\ A)$.*

Lo que hace B_1 es ser un *sistema completo de representantes* para el módulo B con respecto al módulo A. Para demostrarlo no hubo mucha diferencia que en su demostración para la obtención de representantes de formas binarias. Pero no hay una regla para seleccionar elementos en este teorema. Además, si $A = \mathbb{Q}$ y $B = \mathbb{R}$, se encuentra que B_1 sería un *conjunto no medible*[13], con esto, Dedekind fue el primero en usar una forma más fuerte del Axioma.

Con los avances de Cantor, se propasó una línea que no sabían era delimitada por el Axioma. Esta línea divide a los conceptos de Análisis definidos por vecindades[14] de los conceptos definidos por secuencias. Sin el axioma, se tendrían diferentes nociones de continuidad, punto límite, conjunto cerrado, conexidad, etc., pero con él, ellas son equivalentes.

En realidad, Cantor introdujo sus nociones topológicas en términos de vecindades, mientras que Jordan 10 años más tarde las introdujo respecto a sucesiones.

Un ejemplo de estas definiciones es:

Definición 2.1.1. *Un punto $p \in \mathbb{R}^n$ es un* punto límite *de A si toda vecindad de p tiene un punto de A aparte de p.*

Definición 2.1.2. *Un punto $p \in \mathbb{R}^n$ es un* punto límite secuencial *de A si existe una secuencia $a_1,\ a_2,\ \ldots$ de elementos en $A - \{p\}$ que convergen a p.*

En la ausencia del axioma, no todo punto límite es un punto límite secuencial, esto se verá más detalladamente en la p. 155 en la sección 3.7.3.

Incluso el teorema de Weierstrass cambia de noción por esta diferencia de definiciones:

[12] La demostración del teorema se dará en la p. 154.

[13] Ver la subsección 3.6.3.

[14] Ver la definición 3.6.3 en la p. 136 para la definición de vecindad.

Teorema 2.1.4. *Todo subconjunto infinito y* acotado[15] *de* $\mathbb{R}^n$ *tiene un punto límite.*

Ya que si cambiamos el teorema por punto límite secuencial, se necesitaría el Axioma para su demostración, mientras que la demostración de la formulación original puede hacerse sin el Axioma.

Cohen exhibió un modelo en el que existe un conjunto S de números reales que es infinito, y no tiene un subconjunto denumerable. Con esto, si $f : \mathbb{R} \mapsto (0,1)$ es una biyección, se tiene que $f(S)$ no cumple con el análogo del teorema de Weierstrass ya que $f(S) \subset f(R) = (0,1)$, obteniendo un conjunto acotado. Se tiene que $\forall y \in (0,1) \exists x \in S$ que cumple $f(x) = y$, entonces $f(S)$ es infinito; sin embargo, no tiene punto límite secuencial ya que si lo tuviera, encontraríamos una secuencia en $f(S)$, ésta sería un subconjunto denumerable de $f(S)$, y por lo tanto como f es biyectiva, podríamos encontrar una secuencia en S, contradiciendo que no tiene un subconjunto denumerable. Toda esta demostración recae en la existencia de S (no se demostrará que tal conjunto existe).

También se tienen como análogos, el conjunto secuencialmente compacto, el conjunto secuencialmente derivado de un conjunto, el conjunto secuencialmente cerrado, secuencialmente perfecto, secuencialmente isolado, todos ellos definidos mediante sucesiones, que resultan diferentes sin el Axioma a los que Cantor introdujo mediante vecindades[16].

Jordan fue el que usó ampliamente estas definiciones, considerando que son equivalentes a sus análogos, sin darse cuenta que todas dependían del Axioma Denumerable para ser definiciones iguales.

Fue al final del siglo, cuando muchos ocupaban elecciones denumerables que traían impresa la idea de Cauchy de 70 años antes, varios de ellos críticos fervientes del Axioma cuando se postuló. Cuando Cantor ofreció el resultado (y que la mayoría de los matemáticos creían inaceptable) de que todo conjunto puede ser bien ordenado, también hubo mucho debate entre la comunidad. Irónicamente, varios de ellos no se dieron cuenta que utilizaban un postulado que tachaban de implausible en sus propios trabajos.

Otro concepto fundamental en la historia del Axioma es el de *infinito*. Fue Bernard Bolzano quien trató de aclarar las propiedades que un conjunto infinito debía tener en su libro «*Paradoxes of the Infinite*» de 1851. Unas de las propiedades que debían cumplir los conjuntos infinitos son:

Definición 2.1.3. *Un conjunto A se llama* finito *si es vacío o si para algún $n \in \mathbb{N}$ se cumple $A \sim \{1,2,\ldots,n\}$, i.e., son equipotentes. De otra manera, se dice que A es* infinito.

La otra definición, se le atribuye a Dedekind[17],

Definición 2.1.4. *Un conjunto A es* Dedekind-infinito *(D-infinito) si un subconjunto propio de A es equipotente a A. Si esto no se cumple, se dice que A es* Dedekind-finito *(D-finito).*

Varios matemáticos del siglo XIX suponían que estas dos definiciones son equivalentes, es decir, un conjunto es finito ssi es D-finito. Para esta demostración[18], se utiliza el Axioma Denumerable, aunque ellos no lo sabían. Hay que aclarar, que bajo el Axioma Denumerable, estas definiciones son equivalentes, pero que sean equivalentes, no implica en ZF el Axioma Denumerable.

Fue Dedekind (ni Bolzano ni Cantor lo hicieron rigurosamente) quien en su libro *«Was sind und was sollen die Zahlen?»* demostró el caso donde se requiere un equivalente al Axioma, todo conjunto finito es D-finito. Lo hizo deduciéndolo de la proposición de que si para todo n el conjunto $Z_n = \{1,2,\ldots,n\}$ es equipotente a un subconjunto de un conjunto S, entonces se cumple que S es D-infinito (obsérvese

[15] Un conjunto acotado es el que está contenido en una bola centrada en el origen y de radio $r > 0$.

[16] Algunas definiciones en subsecc. 3.7.3, secc. 3.6, otras no se discutirán.

[17] Obsérvese que una diferencia importante respecto a la teoría de estas dos definiciones, es que la de Dedekind, no utiliza el concepto de número natural.

[18] Véase en la p. 153, la demostración de que todo conjunto infinito es D-infinito es la que depende del Axioma Denumerable.

que se ocupa $(p \Rightarrow q) \Leftrightarrow (\neg q \Rightarrow \neg p))$. El uso de elecciones se hizo cuando afirmó que para todo n, existe una función $a_n : Z_n \mapsto S$ que es inyectiva, seleccionando una función del conjunto de todas las funciones inyectivas posibles. Fue Rodolfo Bettazzi quien reconoció en 1896 que tal demostración no era «demostración» como se suponía fuera:

> *«But since there is more than one such correspondence between "any" Z_n and S, then one must take any of them "arbitrarily", ...,which does not seem rigorous; unless one wishes to accept as a postulate that such a choice can be carried out-something, however, wich seems ill-advised to us.»* [Moo, p. 26].

No solo reconoció Bettazzi ese paso dudoso en la demostración, poniendo en duda a la equivalencia de las definiciones, sino que además dio pauta para buscar un nuevo postulado que elija elementos de un conjunto y lo rechazó inmediatamente.

Se creé que el artículo de Bettazzi pudo haber generado debate en la teoría de conjuntos. Cesare Burali-Forti, un colega de Bettazzi, presentó un artículo sobre conjuntos finitos que convenció a Bettazzi la equivalencia de las definiciones 2.1.3 y 2.1.4. Él adoptó la definición de Dedekind, pero difirió de él en que no todas las propiedades de conjuntos finitos podrían ser probadas si no es considerando un postulado adicional: Si S es una familia de clases no vacías, S es equipotente a una subclase de la unión de S. Él añadió que este axioma debía ser considerado en la teoría de conjuntos que nadie había podido formalizar hasta ese momento. Esto presupone la necesidad de una teoría de conjuntos «adecuada» para las matemáticas de aquellos tiempos.

Después Russell demostró que si se le pide a la familia S que sea disjunta, entonces se tiene una equivalencia del principio de Partición. Burali-Forti no ocupó esta condición, haciendo falsas[19] varias demostraciones de su artículo (que son ciertas si se le pide a S ser una familia disjunta). Se dan 4 importantes que se demostrarán algunas en la sección 3.7:

Teorema 2.1.5. *Si se tiene un conjunto D-finito, su conjunto potencia es también D-finito.*

Teorema 2.1.6. *Si el conjunto A es D-finito y B es D-infinito, entonces se tiene que A es equipotente a un subconjunto de B.*

Teorema 2.1.7. *La unión de una familia D-finita (el conjunto de índices es D-finito) de conjuntos D-finitos es D-finita.*

Teorema 2.1.8. *Si una clase V contiene a toda clase de exactamente un elemento y si $(A \in V \ \vee \ b \notin A) \Rightarrow A \cup \{b\} \in V$, entonces V contiene a todo conjunto no vacío D-finito.*

Todas ellas dependen del Axioma Denumerable, y cualquiera implica que todo conjunto D-finito es finito. Ya para esas fechas, Dedekind había demostrado que todo conjunto D-infinito contiene a un subconjunto numerable.

Todos estos acontecimientos, nos dicen que el uso casi indiscriminado del Axioma se hizo por matemáticos de esa época (influenciados por Cantor o Dedekind) sin notarlo la mayoría. Hasta cuando Zermelo explicitó el Axioma de Elección, demostró la equivalencia de las definiciones 2.1.3 y 2.1.4. Sin embargo no solo se hicieron concientes del error provocado al creer su equivalencia, sino que se encontró que no solo dos definiciones, sino muchísimas podían ser dadas en la ausencia del Axioma[20].

---

[19]Existen conjuntos que son D-finitos y su conjunto potencia es D-infinito.

[20]Otras son: Un conjunto es finito ssi toda relación que ordena parcialmente al conjunto también lo bien ordena. Tarski lo definió como finito si toda familia de sus subconjuntos no vacíos contiene un elemento $\subset$-minimal. Otras más están en [Moo, p. 211].

2.1.2. El Problema del Buen Orden y la Hipótesis del Continuo

Estos dos problemas son de relevancia central en la historia del Axioma, se formularon antes que éste y no supieron el debate ni el alcance, ni siquiera la estrecha relación con el Axioma, que traerían cuando se dieron a luz por primera vez. Salieron a flote como una simple «consecuencia» del orden de los cardinales, y el otro como una «ley de pensamiento»:

Proposición 2.1.5 (Cantor)**.** El Principio del Buen Orden. *Todo conjunto tiene un buen orden.*

Proposición 2.1.6 (Cantor)**.** Tricotomía de los Cardinales. $\mathfrak{m}<\mathfrak{n}$ *o* $\mathfrak{m}=\mathfrak{n}$ *o* $\mathfrak{m}>\mathfrak{n}$.

En 1883 propuso el Principio del Buen orden como una ley lógica evidente por sí misma. En 1878 fue el primero en mencionar la Tricotomía, en un artículo que mostraba que $\mathbb{R}$ y $\mathbb{R}^n$ son equipotentes. Y se propuso a demostrar a éstas en la última década del siglo debido a que no creían en ellas la mayoría de los demás.

La idea que motivó a Cantor a formular el buen orden de los conjuntos, fue su deseo de extender la secuencia de los números naturales de forma «natural», llevándolos a números como ∞, $\infty+1$, $\infty+2$, $\ldots$,, ocupándolos como superíndices para sus conjuntos derivados $X^{(\omega)}$ vistos en la p. 38, pero fue en una carta a Dedekind cuando les dio significado formal, poniendo ω como el límite de la secuencia 1, 2, $\ldots$ en vez del signo de infinito. Con esto, definía el segundo principio de generación de ordinales, que permite la formación de ordinales límite. El tercer principio de Cantor dice que el número α está en la $(\beta+1)$-ésima clase-número si el conjunto de todos los predecesores ordinales de α tiene la potencia de la β-ésima clase-número. El primero permite sumar uno a cualquier ordinal (generar el ordinal sucesor de un ordinal). La clase-número unía los conceptos de ordinal y cardinal. La secuencia de los cardinales, que después llamó *alephs* (las potencias de conjuntos infinitos bien ordenados) salió a flote después de sus principios.

Como conclusión de su artículo que mostraba que $\mathbb{R}$ y $\mathbb{R}^n$ son equipotentes, puso esta conjetura que llamó «two-class theorem»:

(2.1.2.1) Todo subconjunto de $\mathbb{R}$ infinito es denumerable o tiene la potencia del continuo.

Fue a $\mathbb{R}$ a quien se le solía llamar *continuo*, y ésta fue la primera formulación de la *Hipótesis del Continuo* o HC, pero Cantor parece que nunca usó ese nombre. Luego aseveró una formulación más fuerte que la anterior,

(2.1.2.2) $\mathbb{R}$ tiene la potencia de la *segunda clase-número*[21].

En su notación con $\aleph$ de 1895[22], la HC (2.1.2.2) formula que

$$2^{\aleph_0}=\aleph_1.$$

La diferencia entre (2.1.2.1) y (2.1.2.2) es importante para el Axioma de Elección. La segunda traé como consecuencia que $\mathbb{R}$ podía ser bien ordenado, mientras que la primera no. En realidad, son equivalentes la primera junto con el hecho de que en $\mathbb{R}$ existe un buen orden y la segunda afirmación.

Después de que creyó que $\mathbb{R}$ puede ser bien ordenado, quiso generalizar y estudiar el hecho de que «todo» conjunto pueda ser bien ordenado. Escribió Moore sobre Cantor en [Moo, p. 42] al respecto:

---

[21] Segunda clase número significa el conjunto de todos los ordinales denumerables.

[22] Véase la p. 35.

«*The concept of well-ordered set turns out to be essential to the entire theory of point-sets. It is always possible to bring any "well-defined" set into the form of a well-ordered set. Since this law of thought appears to me to be fundamental, rich in consequences, and particularly marvelous for its general validity...*»

Parece que, excepto por Cantor, el Principio del Buen Orden no le interesó a los matemáticos sino hasta que Hilbert lo enfatizó en 1900.

Ni la HC, ni el principio de Buen Orden fueron justificados por él impresos con su total satisfacción. Luego de varias afirmaciones, una importante fue la de que

$$2^{\aleph_1} = \aleph_2,$$

encontrando así, que el conjunto de todas las funciones reales puede ser bien ordenado y tiene la siguiente potencia de $\aleph_1$. Después, se fijó que el Teorema de Equivalencia de 1882, estaba muy relacionado con la HC:

Teorema 2.1.9 (Teorema de Equivalencia)**.** *Si $A \subset B \subset C$ y tienen la misma potencia A y C, se cumple que tiene B la misma potencia que éstos.*

El teorema anterior se basa en la HC, como Cantor señaló, aunque no pudo ofrecer demostración alguna. De hecho, el *Beiträge* de Cantor, su más comprehensiva y última publicación en teoría de conjuntos, dejó sin resolver varios de los más fundamentales problemas: la Tricotomía de los Cardinales, el Teorema de Equivalencia, el Principio del Buen Orden y la Hipótesis del Continuo, todos ellos involucrando elecciones arbitrarias.

Burali-Forti leyó esa publicación. Con ello, hizo aportes como el siguiente:

(2.1.2.3) Para cualesquiera clases A y B, una es equipotente a una subclase de la otra.

Burali-Forti dedujo esta forma de la tricotomía de los cardinales, y se basó en dos postulados nuevos, su versión del Principio de Partición, y:

(2.1.2.4) Si A y B son clases incontables, existe una función $f : A \mapsto B$ que es inyectiva o suprayectiva.

En la actualidad, se sabe que éste postulado y el Axioma son equivalentes, incluso sin usar el Principio de Partición.

C. S. Pierce fue otro de los que enfrentaron el problema de la Tricotomía y logró un buen acercamiento, sin embargo la demostración rigurosa esperó hasta que Zermelo llegara. dio un argumento que pasó inadvertido en la comunidad, pero que ocupaba dos postulados equivalentes al Axioma, que para cuando lo notó, no supo decidir la validez de sus postulados:

(2.1.2.5) Contenida en toda relación sobre un conjunto existe una función con el mismo dominio.

(2.1.2.6) Toda función contiene una función con el mismo dominio que es inyectiva o suprayectiva.

Por otro lado, el Principio del Buen Orden, trajo aportes significativos, muchos de Cantor. Respecto a los avances que se tuvieron, en una prueba de que si existen conjuntos infinitos de potencia distinta a un aleph, Cantor distinguió entre dos tipos de clases o multitudes (*Vielheinten*). A un tipo de clase la llamó *absolutamente infinita* (*absolut unendliche*) o *inconsistente* (*inkonsistente*), esto debido a que si es considerada como una colección se llega a una contradicción. Al otro tipo de clase la llamó *multitud consistente* (*konsistente Vielheit*) o *conjunto* (*Menge*) y sus elementos pueden ser considerados una entidad, un solo objeto, sin llevar a contradicciones.

Lo que Cantor remarcaba era que ciertas colecciones son «demasiado grandes» como para ser consideradas conjuntos, pero no las dejó a un lado como otros matemáticos, incluso las utilizó en sus demostraciones, y trató no como paradojas ni como entidades que hay que «evitar», sino como herramientas para descubrir nuevas cosas.

Estas multitudes absolutamente infinitas le hicieron preguntarse sobre la consistencia de sus alephs, sin embargo, en una carta a Dedekind en 1899, dijo que uno no puede demostrar siquiera la consistencia de todo conjunto finito, tal consistencia era una simple «verdad indemostrable» y la llamó el *Axioma de la Aritmética*; de manera similar, llamó el *Axioma de la Aritmética Transfinita Extendida* a la consistencia de cada aleph. Esto último debido a que no sabía si existe un conjunto infinito cuyo cardinal no sea un aleph.

El único axioma en la Teoría de Conjuntos que introdujo fue el de la Aritmética Transfinita Extendida, todos los demás los concebía como principios de lógica, como leyes del pensamiento, por ejemplo, en su carta de 1899 metió como verdades más que como axiomas[23], afirmaciones que se parecen al Axioma de Unión, de Separación y de Reemplazo de ZF. En verdad, la visión Platonista de Cantor, le hacía buscar verdades en vez de asegurar las suposiciones mínimas necesarias para un sistema deductivo.

Esto presupuso la necesidad de una axiomatización de los trabajos que involucraban conjuntos, debido en gran parte a una serie de paradojas que se encontraron en las multitudes inconsistentes, se requería un acercamiento más riguroso a los trabajos hechos, e indudablemente, una teoría más firme.

Ya cuando comenzaba el siglo XX, Hilbert en su lectura famosa, hecha en Paris en el Segundo Congreso Internacional de Matemáticos, trató el Problema del Continuo junto con la existencia de un buen orden para $\mathbb{R}$ como el primero de sus 23 problemas centrales para el desarrollo de las matemáticas del siglo naciente.

Para su primer problema, en realidad, Hilbert buscaba un buen orden efectivo de los reales, que se mostrara explícitamente la relación que hiciera que cada subconjunto no vacío de $\mathbb{R}$ tuviera un elemento menor, quería que se diera una prueba directa de tal proposición.

Para el segundo, que clamaba por la consistencia de los números reales[24], Hilbert aseguró que también se podía dar la consistencia de cada aleph y clase-número. Para Hilbert, la consistencia de un conjunto de axiomas, significaba la existencia de objetos matemáticos que satisfacían dichos axiomas, por eso, decía que cada aleph existía, pero el conjunto de todos ellos, no podía ser dado por un sistema consistente, así que para él no existía dicha colección.

Al sacarlos a la comunidad matemática, y motivar a que se intentaran resolver esos dos problemas y la Tricotomía, apareció lo que ayudó a Zermelo a demostrar el Principio del Buen Orden: elecciones arbitrarias infinitas. Éstas ocurrieron en el reporte de Schoenflies de 1900, para deducir que una sucesión $X_1, X_2, \ldots$ de subconjuntos de $\mathbb{R}^n$ que cumplen $X_{n+1} \subset X_n$ para todo n, y que son cerrados y acotados[25], tiene una intersección no vacía. También al mostrar que el conjunto P de todas las particiones[26] de un conjunto X en subconjuntos de potencia $\mathfrak{n}$ tiene potencia $|X|^{\mathfrak{n}}$. También las utilizó en su aseveración de que el producto finitas o infinitas veces de un cardinal puede ser representado por exponenciación cardinal.

Los problemas de Hilbert se extendieron al extranjero, principalmente a Inglaterra. Ahí entra en esce-

[23]Enfatizó que la diferencia radica en ver una verdad como algo inherente del pensamiento, y a un axioma como algo inherente del sistema donde queremos trabajar, concibiendo diferentes modelos respecto a diferentes axiomas.

[24]Mostrar que no tiene subconjunto similar a W el conjunto de todos los ordinales.

[25]Véase la p. 147 para ver el teorema en espacios métricos.

[26]Véase la p. 15.

na Russell, que también se interesó en los escritos de Cantor a través de Arthur Hannequin, quién junto con Kant influenciaron a Russell, debido a que criticaba el uso de átomos o indivisibles tanto en las matemáticas como en la física, rechazaba el número ordinal de Cantor ω porque decía que la sucesión 1, 2, 3, ... no tiene límite. Así, Russell creyó que la teoría de conjuntos tenía contradicciones. En su ensayo de 1896 «On Some Difficulties of Continuous Quantity», intentó «*to show, what mathematicians are in danger of forgetting, that philosophical antinomies find their counterpart in mathematical fallacies. These fallacies seem ... to pervade the Calculus, and even the more elaborate machinery of Cantor´s collections (Mengen).*»[27]. 4 años más tarde, la influencia de Peano[28] con su lógica simbólica le ayudó a encontrar su famosa paradoja[29].

En su *The Principles of Mathematics* de 1903, Russell creó lo que ahora se conoce como la paradoja de Burali-Forti de la p. 25. Mientras que antes de esa fecha, pocos estaban en desacuerdo, después de ahí muchos empezaron a ver las cosas desde la perspectiva de Russell.

G. H. Hardy, amigo de Russell, en esas fechas intentó demostrar un corolario relacionado a la Hipótesis Generalizada del Continuo, que dice que $\aleph_{\alpha+1} \leq 2^{\aleph_\alpha}$ $\forall\alpha$ ordinal, corolario del teorema que demostró ocupando elecciones arbitrarias que afirma que todo cardinal infinito es un aleph o es mayor a todos los alephs. También se ocupó del Principio del Buen Orden. En las demostraciones de sus teoremas ocupa elecciones arbitrarias, y se ve claramente que para aquél entonces, la comunidad sabía que el uso de elecciones arbitrarias implicaban el Principio del Buen Orden.

Influenciado por Hardy, Philip Jourdain intentó demostrar su teorema, pero modificándolo para que implicara el Principio del Buen Orden. También le mandó la demostración a Cantor de que todo cardinal infinito es un aleph, afirmando que es equivalente a decir que todo conjunto consistente, puede ser puesto en una forma de un conjunto bien ordenado. Pero para su sorpresa, Cantor le respondió que esa demostración ya la tenía desde 7 años antes. Pero lo incitó a que la publicara por su cuenta, tal vez, debido a que Cantor creía que su propia demostración contenía muchas ambigüedades y puntos débiles.

El detalle principal, residía en las multitudes consistentes e inconsistentes, debido a que solo las primeras fueron definidas como conjuntos, lo que no se sabía era, cómo determinar si una multitud dada era consistente. En eso, Cantor dijo: «*A consistent multitude is one such that the assumption of collecting all its elements into one object does not lead to a contradiction...*»[30]. Las multitudes inconsistentes que se habían encontrado eran la multitud de todos los ordinales, de todos los alephs, y la de Dedekind de todo lo imaginable[31].

Por otro lado, ¿cómo saber si una multitud dada es inconsistente?. Jourdain dio un criterio que se aproximó al que conocemos ahora: una clase es inconsistente si contiene una subclase equipotente a la clase W de todos los ordinales. Para burlar la paradoja basada en W de Burali-Forti, Jourdain permitía a las clases consistentes tener un número ordinal o un cardinal. Lo que le faltó a la definición de Jourdain, fue decir la condición de que toda clase inconsistente fuera equipotente a la clase de todos los conjuntos (recuerde que es lo que los hace «muy grandes»), condición impuesta por von Neumann 20 años después encajada en la teoría de conjuntos de NBG[32]. Aún así, Jourdain fue el que más se acercó para tratar con cuidado a las clases inconsistentes, creyendo que no son dañinas para la teoría.

---

[27]Citado de [Moo, p. 58]. Una *paradoja* es también llamada *antinomía*.

[28]Russell y Whitehead acudieron al Primer congreso Internacional de Filosofía, en agosto de 1900, en París, ahí Peano los impresionó tanto que después de eso adoptaron su sistema de lógica simbólica en varios de sus trabajos.

[29]La paradoja se encuentra en la p. 24. La definición de paradoja está en 24.

[30]Citado de [Moo, p. 61].

[31]Véase la p. 12, ahí hay referencias sobre este concepto.

[32]Moore en su libro [Moo], hace una indicación que aquí reproduzco. No se sabe en que sistemas formales de la teoría de conjuntos, estas dos afirmaciones son equivalentes (la de von Neumann y la de Jourdain). Jourdain la utilizó para demostrar el Principio del Buen orden, y utilizó elecciones arbitrarias; pero von Neumann la utilizó para deducir que la clase de todos los conjuntos esta bien ordenada, sin utilizar siquiera el Axioma de Elección.

Fue después que Jourdain se dio cuenta del postulado que utilizó en varias de sus demostraciones y que comenta de esta forma[33]:

> *The validity of the process of making an infinite series of arbitrary selections was simply assumed by me in consequence of Hardy´s work (1903); but, in common with most other mathematicians, I was quite unconscious at that time of the fact that any unproved assumption was made by the admission of the principle of selection (Axiom of Choice).*

En cuanto a historia, solo faltaba que Zermelo se diera cuenta de que el uso de elecciones arbitrarias fuera postulado para abrir las puertas a un mundo matemático nuevo.

Después de que Zermelo demostrara el Principio del Buen Orden en 1904, varios matemáticos cambiaron su forma de pensar en cuanto al constructivismo matemático y a volverse intolerantes de los métodos no constructivos como el Axioma. Irónicamente, varios de esos matemáticos (principalmente los que utilizaban resultados de Cantor, o estudiaban análisis real) ocupaban el Axioma en sus trabajos.

René Baire, hizo grandes resultados en su tesis doctoral de 1899, respecto a la Teoría de Funciones Reales. Aceptó la idea de la definición general de Dirichlet de función real como cualquier correspondencia $x \mapsto f(x)$ entre números reales. En esta definición, Baire remarcó que no viene al caso preguntar cómo la correspondencia puede ser establecida efectivamente, o incluso si es posible establecerla. Su forma de pensar antes de los avances de Zermelo, dejan claro que era más tolerante respecto a las construcciones de objetos matemáticos.

La tesis de Baire, se centró en la clasificación de las funciones reales que después se vio es equivalente en esencia a los conjuntos de Borel de $\mathbb{R}$. Afirmó sin demostración una versión general del *Teorema de Weierstrass de Aproximación*, que dice que toda función real continua es el límite de una sucesión de polinomios[34]. También intuyó la existencia de conjuntos medibles que no sean de Borel, afirmación que depende del Axioma. Pero su mayor contribución fue en Topología General, escribió algo de nociones preliminares de primera y segunda categoría. Un subconjunto A de $\mathbb{R}$ es de *primera categoría* si es la unión contable de conjuntos *diseminados*[35] (en inglés, *nowhere dense*); es de *segunda categoría* si no es de primera categoría. Su famoso teorema dice:

Teorema 2.1.10 (Teorema de Baire)**.** $\mathbb{R}$ *es de segunda categoría.*

El siguiente teorema depende del Axioma:

Teorema 2.1.11 (Teorema de Baire)**.** *Trabajando en* $\mathbb{R}$*, la unión contable de conjuntos, cada uno de primera categoría, es también de primera categoría*[36].

Henri Lebesgue también fue uno de los que hizo las primeras contribuciones a la Teoría de Funciones Reales, respecto a medida de integración. Fue ahí donde planteó el *Problema de la Medida*, que preguntaba por una función m tal que para todo conjunto acotado $A \subset \mathbb{R}^n$ se cumpliera[37]:

1. $m(A)$ es un número real no negativo.
2. $\exists A$ que cumple $m(A) > 0$.

[33] Citado de [Moo, p. 64].

[34] Continuidad se ve en la subsección 3.7.2.

[35] Un conjunto es diseminado si su complemento contiene un abierto denso, o lo que es lo mismo, si el interior de su cerradura es vacío. Véase la subsección 3.6.1 para la aclaración de estos conceptos.

[36] La demostración está en la subsección 3.6.5.

[37] Al número $m(A)$ se le llama la *medida de A*.

3. Conjuntos congruentes[38] tienen la misma medida.

4. La medida de conjuntos disjuntos contables es la suma de las medidas de esos conjuntos.

La última propiedad es llamada *aditividad contable*. Para resolver dicho problema, introdujo los conjuntos *medibles*[39], extensiones de la familia de los conjuntos de Borel, que le permitían definir la integral de Lebesgue. No pasó mucho tiempo en que se descubrieron conjuntos acotados que no eran medibles[40], una de las consecuencias controversiales del Axioma. Además, ocupó en su demostración de que los conjuntos medibles de $\mathbb{R}$ son contablemente aditivos, el Axioma Denumerable[41].

Que la aditividad contable de la medida de Lebesgue dependiera del Axioma Denumerable (no reconocido al principio, demostrado en 1916 por Sierpinski), dejó en una encrucijada a Lebesgue cuando se dio cuenta, ya que tenía dos opciones: aceptar el Axioma Denumerable o restringir en gran medida su Teoría de Integración.

Whitehead también fue uno de los que contribuyeron notablemente, pero que no se dieron cuenta al principio que utilizaban el Axioma, o una forma equivalente de él. De hecho, Whitehead en un artículo que trataba sobre números cardinales, queriendo extender la multiplicación cardinal de Cantor a factores infinitos, introdujo la clase multiplicativa A^X de toda clase A de clases no vacías disjuntas. Definió a la clase A^X como la clase de todas las subclases M de $\bigcup A$ tal que para todo B en A, tuviera exactamente un elemento el conjunto $M \cap B$. Se acercó a establecer una forma equivalente[42] al Axioma, una formulación del llamado Axioma Multiplicativo[43]:

Axioma 2.1.1. *A^X es no vacío para toda familia A de clases no vacías.*

Con esta arma en mano, fue capaz de demostrar varios teoremas equivalentes al Axioma Multiplicativo:

Teorema 2.1.12. *Para toda familia disjunta de clases no vacías se cumple que $\bigcup(A^X) = \bigcup A$.*

Teorema 2.1.13. *Para todas las familias disjuntas A y B, se cumple que si $A^X = B^X$, entonces $A = B$.*

Teorema 2.1.14. *Si A es una familia de clases disjuntas no vacías, y $S \subset \bigcup A$ y si ningún miembro de A contiene más de un elemento de S, entonces $S \subset M$ para algún $M \in A^X$.*

El último teorema nos dice que una clase puede ser extendida a un miembro de A^X. Este teorema tiene mucha conexión con los *Principios Maximales*[44], que se verán algunos en la subsección 3.1.3, utilizables en Álgebra y Análisis principalmente.

---

[38] Dos espacios métricos (A,d) y (B,d') son congruentes (o isométricos) si existe una congruencia (o isometría o transformación congruente) entre ellos, que es una función biyectiva entre los conjuntos $f : A \mapsto B$ que preserva distancias, si $P,Q \in A \Rightarrow d(P,Q) = d'(f(P), f(Q))$. La definición de espacio métrico está en la p. 146.

[39] Si tenemos un conjunto $A \subset \mathbb{R}$ y un intervalo fijo I que contiene a A y de longitud k, la *medida externa* de A es la menor de las cotas superiores de la suma de los intervalos que contienen a A. La *medida interna* es $k - b$, con b la menor de las cotas superiores de la suma de las longitudes de los intervalos que contienen a $I - A$. A será *medible* si la medida interna es igual a la medida externa.

[40] Véase la subsección 142.

[41] Existe un modelo de ZF en el que $\mathbb{R}$ es la unión contable de conjuntos contables, en este modelo, la medida de Lebesgue no es contablemente aditiva, ya que $\mathbb{R}$ sería la unión contable de conjuntos de medida cero, teniendo medida cero. Que un subconjunto $X \subset \mathbb{R}^n$ tenga *medida cero* significa que $\forall \varepsilon > 0$, existe una familia $\{A_i\}_{i \in \mathbb{N}}$ tal que $X \subset \bigcup_{i \in \mathbb{N}} A_i$ y que cumpla $\sum_{i=0}^{\infty} m(A_i) < \varepsilon$.

[42] En el teorema 3.1.3 se ve la equivalencia.

[43] El Producto Cartesiano generalizado aún no era introducido.

[44] En realidad, deberían llamarse *principios extremos*, debido a que a muchos principios maximales les corresponde un principio minimal.

Pero en todos esos casos trató a las clases multiplicativas como no vacías. El uso de el Axioma estuvo presente en la medida de Lebesgue, la clasificación de Baire de las funciones reales, los conjuntos Borelianos y la aritmética cardinal de Whitehead, todos estos trabajos realizados antes de la demostración de Zermelo. Sin embargo, todos ellos rechazaron el Axioma que es la maquinaria en la que se basaba la demostración de Zermelo.

Las diferentes posturas no se hicieron esperar, [Moo, p. 75] dice que Russell y Whitehead explicitaron los teoremas que dependían del Axioma Multiplicativo en su libro "Principia Mathematica", aunque aún tenían dudas de su veracidad. Uno que a veces daba su voto al Axioma y a veces declinaba de él fue Borel (era inclinado a aceptar el Axioma Denumerable). Baire fue firme y sin equívoco no creyó en el Axioma. Lebesgue no creyó en él dudosamente.

2.2. EL PERIODO DE DEBATE

Las matemáticas después de la demostración de Zermelo de que todo conjunto puede ser bien ordenado quedaron casi devastadas.

La visión de Frege que vuelvo a reproducir, después de que salió la paradoja de Russell publicada en 1903, dice el dolor que sentía por ello: «*Difícilmente puede sobrevenirle a un escritor científico algo más infortunado que ver vacilar uno de los fundamentos de su edificio, después de que su trabajo está terminado.*». Quedaron muchas dudas en el aire. Las definiciones tambaleaban. Las formas de pensar cambiaban radicalmente, se formaban grupos de matemáticos que veían a su universo de diferente manera. No solo las bases matemáticas sucumbieron, las filosóficas y las psicológicas se vieron afectadas o por lo menos, puestas a prueba. Nunca antes una demostración había sido tan debatida públicamente. Y pocas veces se preocupaban los matemáticos de lo que se preocuparon después de dicha demostración.

En esa época, resultaba incierto qué era un conjunto. No se sabía si la clase W de todos los ordinales hacía no válida la demostración de Zermelo, ni qué tanto se relacionaba con otras paradojas. Y la pregunta más importante: ¿Acaso era verdad el Axioma de Elección?.

Más aún, había un problema del que no se establecía siquiera una línea de ataque debido a que era raro meterse en esos asuntos: ¿qué métodos se permiten en matemáticas?. No se sabía definir incluso qué representa una «construcción» ni cuando un objeto matemático existe.

Entre las cosas buenas, resulta que su demostración aclaró varias ideas inconclusas y aplicaciones erróneas. Una de ellas, era la de Julius König, un investigador prominente y brillante, que concluyó que $\mathbb{R}$ no podía ser bien ordenado, así como resultados erróneos sobre la HC. Escondido entre sus argumentos estaba el Principio de Tricotomía de los cardinales, resultados sobre cardinales de Cantor e incluso elecciones arbitrarias. Luego de la demostración de Zermelo, aclaró sus resultados y se ordenaron varios principios sobre cardinales que estaban vagos entre algunos miembros de la comunidad.

Como dije, la demostración de Zermelo tuvo relevancia trascendental, pero, ¿cómo empezó?. Ni siquiera se doctoró acerca de Teoría de Conjuntos en 1894, sino sobre Cálculo de Variaciones. Luego de eso, se cambió al área de Matemática Estadística, llegando a ser *privatdozent* en el Instituto de Física Teórica de Berlín, donde fue asistente de Max Planck. Zermelo escribió[45] años después:

> *Bericht an die Notgemeinschaft der Deutschen Wissenschaft über meine Forschungen betreffend die Grundlagen der Mathematik.*
>
> *«Thirty years ago, when I was a Privatdozent at Göttingen, I came under the influence of D. Hilbert, to whom I am surely the most indebted for my mathematical development. As a*

[45] Citado de [Moo, p. 89].

result I began to do research on the foundations of mathematics, especially on the fundamental problems of Cantorian set theory, whose true significance I learned to appreciate through the fruitfil collaboration of the mathematicians at Göttingen.»

Cerca de 1899 Zermelo descubrió en la Lógica Algebraica de Schröder lo que después fue conocido como la paradoja de Russell, y se lo hizo saber a Hilbert dos años antes que Russell. La diferencia con respecto a la paradoja de Russell, fue que Zermelo no concluyó que las matemáticas de esas épocas se veían afectadas, sino que solo consideró que cualquier conjunto que tenga a todos sus subconjuntos como elementos es en sí mismo contradictorio, pensó que el conjunto era el problema. En una publicación acerca de Teoría de Conjuntos, Zermelo se basó en gran medida en resultados de Cantor, y ocupó consciente e inconscientemente elecciones arbitrarias, sin remarcar esos hechos. De pronto, su línea de investigación cambió al ver el error en los argumentos de König, llevándolo ésto a relacionar sus investigaciones sobre la HC, y finalmente, se adentró en el Principio del Buen Orden. Era 24 de Septiembre de 1904 cuando su demostración del Principio del Buen Orden concluyó. Luego de ser enviada a Hilbert, apareció en la *Mathematische Annalen*.

La demostración del teorema se dará en la p. 96, sin embargo señalaré la propia anotación que Zermelo hizo cuando empezó su demostración escogiendo un conjunto no vacío arbitrario M y S el conjunto de todos sus subconjuntos no vacíos $M´$: «*To each subset $M´$ one associates any element m_1' wich occurs in $M´$ itself and wich may be named the "distinguised" element of $M´$.*»

Ésta fue la primera aparición explícita del principio[46] después llamado Axioma de Elección, ahí formuló la existencia de una *cubierta* de S denotada por γ (o como ahora es llamada *función de elección* o *función electora* de S o en S), donde $\gamma : S \mapsto M$ es una función con $\gamma(M') \in M'$, $\forall M' \in S$.

Zermelo concluyó, con varias proposiciones que se seguían de su teorema, como el hecho de que el cardinal de todo conjunto infinito es un aleph, la veracidad de ecuaciones importantes en Teoría de Cardinales y alephs como:

$$\mathfrak{m} = 2\mathfrak{m} = \aleph_0\mathfrak{m} = \mathfrak{m}^2,$$

y además, la Tricotomía de los Cardinales. En su conclusión, remarcó la existencia de las funciones γ junto con su equivalencia con el Axioma Multiplicativo. Pero fue esta existencia la causal de muchas disputas, gran parte del rechazo respecto al Axioma se dio por la pregunta: ¿cómo es esa función encontrada?, además la crítica se preguntaba por conjuntos paradójicos como el de Burali-Forti[47], y pronto hubo más personas que rechazaban su demostración[48], y pocos apoyantes, incluso, intentaron varios buscar, la razón primitiva del Axioma que hiciera que la comunidad se separara en tantas clases.

Entre los críticos estaba Borel, quien encargado por Hilbert, escribió un artículo sobre la demostración de Zermelo, diciendo que el argumento de éste no podía ser aceptado por algún matemático (una cita de ésto se encuentra en la página 84), afirmando que eso recae fuera de los confines de lo que la matemática trata.

El analista Jacques Hadamard estaba a favor de la demostración debido a que las elecciones eran independientes una de cada otra, y no veía diferencia alguna entre elecciones incontables o contables, por ser independientes, sin embargo, para el caso de las elecciones dependientes la diferencia sería significativa. Hadamard sospechó que tales elecciones no pudieran ser hechas efectivamente de manera que alguien las pueda realizar. Remarcó una importante diferencia entre demostrar que[49]:

[46] Zermelo no le dio un nombre explícito aquí, pero en 1908 usó el nombre de «postulate of choice».

[47] Dado en la p. 25.

[48] A la siguiente publicación de la revista, se editaron artículos de Borel, Schoenflies, König, y Jourdain donde daban argumentos por los que no aceptaban la demostración.

[49] Su posición era que un objeto matemático existe incluso si no podemos definirlo de manera única.

1. una función existe,

2. o que existe pudiendo ser especificada únicamente.

Varias cuestiones matemáticas tendrían diferentes soluciones y significados si una se reemplaza por la otra[50].

Baire fue un poco extremista. Dijo que todo en las matemáticas tenía que ser reducido a el plano finito. Señaló además cuando uno tiene un conjunto infinito: «*I consider it false to regard the subsets of this set as given.*»[51].

Lebesgue se puso en una posición más analítica, y preguntó: «*Can one prove the existence of a mathematical object without defining it?* »[52], y se preguntaba si es legítima una demostración de existencia si no especifica «únicamente» al objeto, poniéndose en el lado de que eso es mera convención, aunque eventualmente se supo que la noción de definibilidad se refiere más a especificar el sistema formal que algo que depende del lenguaje natural. Se pueden tomar dos posturas al querer demostrar la existencia de un objeto con cierta propiedad, remarcadas como dije por Hadamard, se asume válida (la demostración) si se especifica una clase de objetos con dicha propiedad (con clase me refiero a una colección que podría no ser conjunto, no a un tipo de objetos), o se asume válida si se muestra un objeto específico con la propiedad. Está última posición es constructivista. Lebesgue tenía un buen punto, decía que aunque parezca infactible, eso no le quita lo posible: «*Although I seriously doubt that a set will ever be named which is neiter finite nor infinite, it has not ben proved to my satisfaction that such set is impossible.*»[53]. Sin embargo estaba muy entremezclado con el Axioma, en una carta a Borel se ve que evitaba usar esas elecciones, aunque en su libro sobre integración en algún punto hacía mención explícita sobre el uso de las elecciones mientras que las usaba en otro lado sin notarlo.

Lebesgue también añadió que la definibilidad no debía tomarse del todo literal, ya que uno no puede «definir» algún número irracional, debido a que su representación decimal es infinita y ninguna operación puede ser completada con ellos; aunque se supone nadie desea eliminar a los irracionales de la teoría, ellos son muy importantes y como remarca Lebesgue, resultan de manera natural en muchas cuestiones matemáticas, sin embargo, respecto el Axioma, éste también resulta de manera natural en muchas cuestiones en materia de Análisis y Cardinales[54].

Lebesgue fue uno de los que reconocieron y escribieron acerca de las posturas en las que los matemáticos se dividieron después de la demostración del Principio del Buen Orden por Zermelo a principios del siglo XX. Según Moore en [Moo, p. 100] Lebesgue dividió en su libro a los matemáticos en *Idealistas* y *Empiristas*[55], los últimos solo aceptaban las funciones que fueran determinadas unívocamente, mientras que los primeros aceptarían de otros tipos. Los empiristas ven a demostraciones que solo dicen que una función existe sin dar características que la determinen, como «formas carentes de significado»; mientras que los idealistas, cuando buscan determinar una función, no buscan determinar la característica que la defina como única, para ellos, una función está únicamente determinada porque ellos afirman que lo es, y cuando utilizan una demostración, no hay ambigüedad en que otro esté pensando en otra función porque siempre se está pensando en la misma función escogida, y estas cosas para el idealista son inverificables. Lebesgue añadió que esta coyuntura no se resolvería por medios teóricos, sino mas bien por los resultados

[50]Borel utilizó funciones en teoremas de la convergencia de series complejas que no podían ser determinadas únicamente.

[51]Citado de [Moo, p. 95], esa y varias conversaciones entre Borel, ,Lebesgue, Baire, y Hadamard están en «*Bulletin de la Société Mathématique de France*», 1905, y están traducidas en el apéndice de [Moo].

[52]También ésta frase está en esas conversaciones.

[53]También se encuentra entre las conversaciones, o véase [Moo, p. 96].

[54]Un ejemplo de esto está en la p. 57.

[55]En filosofía moderna, el término correcto para los idealistas sería *realistas* y los Empiristas de Lebesgue serían llamados Idealistas.

que logren dar los Idealistas, los alcances útiles de las funciones de éstos determinarían que tan bueno es ponerse de un lado con menor restricciones.

Claro que los alcances idealistas no fueron bien vistos del todo por Lebesgue, debido a que en 1905 Giuseppe Vitali demostró con la ayuda del Axioma la existencia de un conjunto de $\mathbb{R}$ que no era medible[56]. De la misma manera, demostró que el Problema de la Medida de Lebesgue[57] no tenía solución, que no existía una medida real positiva, invariante bajo traslaciones y aditiva contable para todos los subconjuntos acotados de $\mathbb{R}$.

En 1907 publicó la solución a un problema que propuso Carrado Serge sobre si existían otras funciones complejas a parte de la nula, la identidad, y la conjugada[58] que cumplieran para todo número complejo z y w que

$$f(z+w) = f(z) + f(w),$$
$$f(zw) = f(z)f(w).$$

Lebesgue demostró que no hay funciones medibles[59] que cumplan con esas condiciones, por otro lado, si se acepta que $\mathbb{R}$ puede ser bien ordenado, existen infinitas funciones que cumplen eso[60].

Una observación importante hizo Lebesgue acerca de las funciones de elección en la familia de todos los subconjuntos denumerables de $\mathbb{R}$, demostrando que cualquier tal función debe ser no medible e indicando qué tan obscuras e indefinibles han de ser esas funciones de elección que ocupaba Zermelo, claro, si uno siquiera pudiera hablar de ellas debido a que no estaba clara su existencia.

Otro de los que hicieron avances y dieron las bases para la Teoría de Análisis Funcional fue Maurice Fréchet. En su tesis de 1906 ocupó límites de secuencias para hacer resultados topológicos, generalizó un problema de Cantor sobre intersecciones denumerables ocupando en la demostración el Axioma Denumerable[61]:

Si se tienen subconjuntos $C_1, C_2, \ldots$ de un conjunto compacto C no vacíos, secuencialmente cerrados y que cumplen $C_{i+1} \subset C_i$ para $i \in \mathbb{N}$, entonces la intersección de todos los C_i es no vacía.

Definió el concepto de *conjunto extremo*, que es lo que conocemos nosotros como *secuencialmente compacto*, diciendo que es un subconjunto A compacto y secuencialmente cerrado de uno dado B; también definió una función $f : B \mapsto \mathbb{R}$ *secuencialmente continua* en A que cumplía con que $\forall a \in A$ y para toda sucesión $\{a_1, a_2, \ldots\} \subset A$ con límite a se tiene que $\lim_{n\to\infty} f(a_n) = f(a)$, con esto generalizó un teorema importante:

Si se tiene $f : B \mapsto \mathbb{R}$ una función secuencialmente continua en un conjunto extremo B, implica que f está acotada en B y que alcanza un máximo y un mínimo valores en ese conjunto[62].

Georg Hamel aceptó el Axioma e investigó que alcances trae éste. Obtuvo *funciones discontinuas* reales que satisfacían la ecuación

$$f(x+y) = f(x) + f(y)$$

para todos los valores reales, mediante el uso del Buen Orden de $\mathbb{R}$. Investigó acerca del Análisis Vectorial y obtuvo una base de $\mathbb{R}$ visto como un espacio vectorial sobre el campo de los números racionales[63], después nombrada *base de Hamel*. La existencia de bases sobre espacios vectoriales de dimensión arbitraria es equivalente al Axioma, ya que hay modelos donde dos bases tienen cardinalidad

[56] Véase la p. 143 para la demostración de dicha existencia.

[57] Véase la p. 50.

[58] Si $z = a+ib$, la función conjugada es $f(z) = \bar{z}$, donde $\bar{z} = a - ib$.

[59] Véase la subsección 3.7.4.

[60] El teorema 3.7.9 da una demostración de un hecho similar.

[61] Las definiciones de *compacto* están en la p. 137 y de secuencialmente cerrados en la p. 155.

[62] Una función $f : A \mapsto \mathbb{R}$ alcanza un máximo (mínimo) en A si existe $a \in A$ que hace $f(x) \leq f(a), \forall x \in A$ $(f(x) \geq f(a), \forall x \in A)$.

[63] En la subsección 3.4.1 se dan estas definiciones y la demostración general.

distinta, o incluso espacios vectoriales sin base. Respecto a esto, Lebesgue demostró con el Axioma que las bases de Hamel son medibles.

Alguien que no daba su posición al respecto, pero sacaba conclusiones de suponer el Axioma cierto o falso era Gerhard Hessenberg. Él estableció afirmaciones que tuvieron consecuencias importantes en el futuro. Dijo que uno puede hacer elecciones arbitrarias y sucesivas, dependientes, también, estableció lo que se conoció después como el Axioma Multiplicativo:

Axioma 2.2.1. *Si C es cualquier familia de conjuntos no vacíos disjuntos, existe un conjunto B tal que para todo $A \in C$, el conjunto $A \cap B$ contiene exactamente un elemento.*

Comentó Hessenberg que respecto a la demostración de Zermelo, no había alguien que haya hecho tales elecciones para todos los subconjuntos de $\mathbb{R}$, y que la demostración no acercaba a alguien a resolver el problema de Cantor sobre la HC. En [Moo, p. 114] se dice que como Hessenberg afirma, uno nunca podría especificar un conjunto no vacío en el que sea imposible seleccionar un elemento. Piense lector en esta afirmación, ¿qué tan lógica o qué tan matemática le resulta?, exhibir un conjunto donde no se pueda seleccionar un elemento suena extraño, pero recuerde las paradojas que resultan de considerar a «todos los conjuntos».

Felix Hausdorff cambió de posición, primero negando la veracidad del Axioma, y luego lo aceptó y lo utilizó creando resultados como el que veremos en la sección 3.5. En 1907 aplicó un buen orden de $\mathbb{R}$ para demostrar la existencia de un *pantachie*, que se basa en sucesiones de números reales con el orden parcial dado por:

Para todas las sucesiones f y g, $f < g$ significará que g es más grande que f después de cierto número, i.e., $\exists n \in \mathbb{N}$ tal que $f_m < g_m, \forall m > n$.

El pantachie sería un conjunto maximal X de esas sucesiones, ordenadas por $<$. Que sea *maximal* en este caso significa que no pueden añadirse sucesiones y que se siga cumpliendo el orden de $<$. Esta fue una base para lo que después sería el concepto de *Principio Maximal de Hausdorff*, equivalente al Axioma[64].

Otro resultado importante de Hausdorff sobre conjuntos ordenados fue:

(2.2.1) Todo conjunto denso X puede ser particionado en dos conjuntos, los cuales también son densos en X.

La definición de un conjunto parcialmente ordenado X que es *denso* es que siempre hay un elemento de X en medio de cualesquier dos elementos distintos de X, y $Y \subset X$ será *denso en X* si para cualesquier dos elementos distintos de X existe uno en medio de ellos que está en el subconjunto Y.

Rechazó aclarar cuándo había usado el Teorema del Buen Orden luego de haberlo aceptado como verdadero, con esto hecho, definió conjuntos que su existencia dependía de lo que formuló como:

Axioma 2.2.2 (Hipótesis Generalizada del Continuo)**.** *La suma de todos los $\aleph_\alpha^{\mathfrak{m}}$ con $\mathfrak{m}$ corriendo sobre todos los cardinales menores a $\aleph_\alpha$ es igual a $\aleph_\alpha$.*

El corolario de esta afirmación es:

Corolario 2.2.1. *Para todos los ordinales α se cumple que $2^{\aleph_\alpha} = \aleph_{\alpha+1}$.*

En 1908 en un artículo enfatizó la necesidad de una formalización del concepto de conjuntos, dijo que incluso se hiciera ésto axiomáticamente. El problema que veían muchos en esa época de hacer esto posible, era que, no se tenía una forma de comparar intuitivamente, como lo era la geometría por ejemplo,

[64] Véase la página 93, donde viene la demostración de dicha equivalencia en el teorema 3.1.4.

donde los axiomas podían compararse con lo que visualmente entendemos para darse una «idea» de adonde ir. Otro problema, era la importancia de esa Teoría de Conjuntos, ya que los fundamentos de ramas como la Lógica y la aritmética estaban mezclados en la Teoría de Conjuntos.

Del otro lado, la escuela Inglesa no aceptó al principio con buenas miras a la demostración de Zermelo. Un personaje que añadió mucho a la historia fue Russell.

Antes de que Zermelo formulara el Axioma, Whitehead y Russell crearon el Axioma Multiplicativo[65]. En realidad lo tomaron como teorema, el cuál ninguno de los dos pudo demostrar. Se empezaron a dar cuenta que iba a ser muy difícil hacer esto, diciendo Russell que: «...but gradually I saw that, if there is a proof, it must be very recondite.»[66].

Una cuestión importante a remarcar, es ver cómo se convierte en axioma ese «teorema» de Russell y Whitehead. El Axioma de Elección resultó de los trabajos de Zermelo sobre el Buen Orden, para dar un Buen Orden en $\mathbb{R}$, el Axioma Multiplicativo resultó de considerar el producto infinito de conjuntos disjuntos, de su clase multiplicativa, para definir después el producto de una infinitud de cardinales, diciendo que A^X es no vacío si A es una familia de clases disjuntas y no vacías. Luego de que intentaron resolverlo por bastante tiempo, empezaron a convertirse escépticos de su veracidad, a dudar qué tanto su afirmación tiene una verdad implícita y si es así, que si quiera pudieran encontrarla. Russell después fue quien, en realidad tomó a esta afirmación como una *«suposición fundamental no demostrada»*, que es diferente a lo que axioma se entiende por *verdad por sí misma evidente*. Pero, ¿cómo es que tiene fundamento la idea de suponer algo cierto sin saber si sea verdadero?, recaen ahí dos cuestiones, la intuitiva es por tratarse de algo que sale de manera «natural» en la investigación, y en la formal en demostrar que esa suposición es independiente de los demás axiomas supuestos, para así estar seguros que no es contradictorio en nuestra teoría.

Hardy dio 4 argumentos del hecho de suponer a la clase multiplicativa vacía:

1. parece paradójico, «no satisface el intelecto» como dirían los idealistas,
2. no parece lógicamente necesario,
3. no parece resolver dificultades insolubles de otra forma y
4. parece dejar pasar muchas cosas matemáticas interesantes.

Hay que notar lo intuitivo de sus razones, basándose en la necesidad y no en la veracidad del Axioma Multiplicativo.

Cuando Russell leyó la demostración de Zermelo, notó que su Axioma implicaba el Axioma Multiplicativo, y escribió:

> *«a theorem without wich much of cardinal arithmetic would be impossible ... I do not know if it is true or false, and I find it too complicated for a Pp (postulate) ... one could deduce everything necessary from the axiom:*
>
> $\exists f:\ f(u) \in u$ *for all non-empty set u.*
>
> *But I do not know if this axiom is true. ... since one should hardly admit an axiom that is to complicated and so dubious».*[67]

[65] Se habló de él en la p. 51.
[66] Lo dijo en una carta el 15 de marzo de 1906, está en Grattan-Guinnes 1977, p. 80.
[67] Citado de [Moo, p. 124], fue escrito en una carta para Couturat el 26 de abril de 1905.

Lo que afirmó Russell fue lo que conocemos hoy por el *Axioma de Elección Global*, que implica al Axioma de Elección en ZF. Aunque hubo muchas dudas sobre ésto de parte de Russell, quien era regañado por Couturat, diciéndole éste que lo sorprendía el hecho de que Russell dude de la veracidad del Axioma Multiplicativo, enfatizando que un sistema lógico debe darnos todo lo que es dado por el sentido común. A esto, Russell respondió que el sentido común a dado a relucir muchas falacias y por eso no podemos confiarnos de él.

Así, Jourdain y Russell se dieron a la tarea de determinar qué teoremas necesitaban del Axioma y cuáles no. Pero hasta que no hubiera una Teoría de Conjuntos Axiomatizada, sería muy difícil lograrlo. Incluso cuando Zermelo dio una axiomatización en 1908, sólo se podía decir cuando una proposición era equivalente al Axioma y cuando no, ya que no había forma de demostrar que por ejemplo, el Axioma era más fuerte que el teorema.

Un americano que demostró un hecho importante se llamó Edward Van Vleck, quien estudió sobre los conjuntos que no son Lebesgue medibles, remarcando cómo ellos dependen de elecciones arbitrarias, esto en 1908. Dio un ejemplo de un conjunto que no es medible según Lebesgue independiente de los resultados de Lebesgue sobre este tipo de conjuntos. Lo describió como el conjunto de pares

$$P_x = \left\{ \left\{ \left(\frac{x}{2^{\pm p}}\right) \pm \left(\frac{m}{2^n}\right) : m,\ n,\ p \in \mathbb{N} \right\}, \left\{ 1 - \left(\frac{x}{2^{\pm p}}\right) \pm \left(\frac{m}{2^n}\right) : m,\ n,\ p \in \mathbb{N} \right\} \right\}$$

para todo número irracional $x \in (0,1)$. De esta forma, al elegir un conjunto de cada par contenido en P_x, y poniéndolo en un conjunto A, y al otro restante metiéndolo en un conjunto B, se particionan a los irracionales en dos conjuntos A y B, éstos no medibles y que no tienen la *propiedad de Baire*[68].

El resultado importante que obtuvo, fue que su argumento, al utilizar el Axioma aplicado a familias que tienen como elementos pares, estas familias con potencia $2^{\aleph_0}$, resulta en un conjunto que no es medible y tampoco de Baire.

A pesar de todos los esfuerzos y las críticas dadas, ninguno de los mencionados aquí encontraron cómo formular el término «definible»; solo cuando Hadamard intercambió cartas con Baire, Borel y Lebesgue fue que preguntaron si la existencia de un objeto con una propiedad puede ser tomada válida si la clase de objetos que tiene dicha propiedad es no vacía, o se tiene que definir un objeto único de dicha clase (es como preguntar si uno se conforma con tener un conjunto no vacío o se requiere especificar elementos de ahí). Pronto tuvo que venir el periodo de formalización de las ideas que se tenían vagas en el mundo matemático.

2.3. EL PERIODO DE AXIOMATIZACIÓN

Desde antes de la salida de las paradojas de Russell y Burali-Forti, algunos intentaron dar una axiomatización de la Teoría de Conjuntos. El que lo intentó mediante términos matemáticos y no términos lógicos fue Zermelo, como forma de defender su demostración del Teorema del Buen Orden, buscando las suposiciones que utilizó, que darían una base firme a dicha demostración y además serían la base de todas las matemáticas. Después de que su axiomatización salió a la luz, siguió criticada por más de 10 años, no ayudando al convencimiento de la comunidad sobre la validez de su prueba (en realidad, la nueva demostración dependía del Axioma tanto como la anterior). La primera axiomatización de la Teoría de Conjuntos la publicó él en 1908, según lo que escribió Moore en [Moo, p. 143], eran dos artículos en los que intentaba defenderse de la comunidad que no aceptaba su prueba, intentando ser más específico para

[68] Si tenemos un espacio topológico X, se dice que $A \subset X$ tiene la propiedad de Baire si existe un abierto U de X tal que $A - U$ y $U - A$ son de primera categoría. Los conceptos topológicos están en la sección 3.6, el de primera categoría en la p. 50.

aclarar dudas, interpretaciones y subjetividades que conciernen a los conceptos. Definió el concepto de conjunto bien ordenado, el Axioma de Separación y el del Conjunto Potencia, como los postulados que ocupó en la demostración nueva que dio sobre el Principio del Buen Orden[69], definió la generalización de lo que ahora conocemos como familia ordenada por inclusión[70], que era su llamada θ-*cadena*:

Sea un conjunto X y una función de elección f en la familia de todos los subconjuntos no vacíos de X, T será llamada θ-cadena si

1. T es un conjunto de subconjuntos de M,
2. $X \in T$,
3. si $A \in T \Rightarrow A - \{f(A)\} \in T$
4. y si $B \subset T \Rightarrow \bigcap B \in T$.

Como tuvo que hacerlo, la formulación matemática del Axioma que tomó como evidente por su carácter objetivo, la dio luego de su demostración[71]:

Axioma 2.3.1. *Si un conjunto S es particionado en una familia disjunta A de conjuntos no vacíos, entonces existe al menos un subconjunto T de S que tiene exactamente un elemento en común con cada miembro de A.*

De este axioma, derivó el que había postulado en 1904, que lo llamó aquí *Principio General de Elección* (Allemeines Auswahlprinzip)

(2.3.1) Para toda familia T de conjuntos no vacíos, existe una función f que cumple $f(S) \in S$ para todo $S \in T$ no vacío.

Zermelo también dio resultados importantes que creía dependían del Axioma como argumentos para la necesidad de tal:

1. Todo conjunto D-finito es finito.
2. Las uniones respectivas de 2 familias de conjuntos disjuntos y equipotentes son equipotentes.
3. La clase multiplicativa de una familia disjunta de conjuntos es vacía solo si uno de sus miembros es vacío.
4. Existe una base de Hamel.
5. Existen funciones reales discontinuas f tales que $f(x+y) = f(x) + f(y)\ \forall x, y \in \mathbb{R}$.
6. Teorema de la Unión Contable.
7. Principio de Partición.

La última aplicación que le dio Zermelo al Axioma en ese artículo, fue:

---

[69] En su segundo artículo definió 7 axiomas parecidos a los que dimos en ZF, tres de los cuales son éstos.

[70] Visto en la p. 18.

[71] No se refirió al Axioma Multiplicativo de Russell de 2 años antes, esta nueva formulación del Axioma es lo mismo, pero se sabe lo formuló independientemente en ese mismo año que Russell.

(2.3.2) Si T y S son familias de conjuntos disjuntos, y que existe una biyección entre ellas $f : T \mapsto S$, y para cada $M \in T$ se tiene que $M \prec f(M)$, entonces $\bigcup T \prec P$, con P el producto (clase multiplicativa) de S.

Este resultado es atribuido a König, y se le llama *Teorema de König*; en términos de cardinales dice que:

(2.3.3) si para todo $t \in T$ se cumple que $\mathfrak{m}_t < \mathfrak{n}_t$, entonces

$$\sum_{t \in T} \mathfrak{m}_t < \prod_{t \in T} \mathfrak{n}_t.$$

Zermelo se defendió contra sus atacantes ocupando sus mismos argumentos contra ellos, y para evitar la paradoja de Burali-Forti, dijo que no se debía considerar a W (la clase de todos los ordinales) como conjunto; para defenderse de la crítica sobre los principios de Cantor para generar ordinales[72] (decían que no todo ordinal tiene sucesor, como el ordinal de W, ya que sería un ordinal estrictamente mayor a el ordinal de todos los ordinales) se necesitaba un axioma que garantizara la existencia de un conjunto del tipo de ω, predisponiendo con ésto, el Axioma de Infinitud; y por último, para defenderse sobre la crítica de la veracidad del Axioma, objetó que era un principio indemostrable.

Zermelo creía que el criterio que se tome para decidir si un objeto cae en la definición de un concepto, debe ser objetivo e independiente del concepto definido. Cuando una relación especial existe entre el objeto y el concepto, dice que el objeto no fue creado, sino solamente descrito. Ésta es la forma de ver de un *realista*.

Para tener una demostración segura y que permitiera que su demostración fuera «válida», Zermelo debió incluir en su formalización al Axioma. Con estos contraataques, fue como las críticas disminuyeron sobre el Principio del Buen Orden, aunque los matemáticos no aceptaron el Axioma al principio por toda la controversia que lo rodeaba, y casi nadie ocupó su sistema axiomático en años subsecuentes a su aparición, tampoco se supo sobre su consistencia ni sobre la independencia de dichos Axiomas[73]. Pero cuando ganó adeptos, los matemáticos comenzaban a utilizar y a explotar al Axioma, a ser precavidos y notar en sus demostraciones, cuándo utilizaban el Axioma.

La idea de una axiomatización, es obtener las demostraciones a partir de los axiomas dados, sin referirse en momento alguno al significado físico de los conceptos. La mejor influencia de Zermelo fue Hilbert con su libro «*Grundlagen der Geometrie*» que era un libro axiomático sobre Geometría, que involucró un dominio de objetos (puntos, líneas y planos) con una relación primitiva, las suposiciones usadas escritas de manera explícita como axiomas y la investigación de la independencia y consistencia de los axiomas.

Respecto a la Teoría de Conjuntos, el que primero propuso postulados fue Burali-Forti, en 1986 propuso el Principio de Partición sin condición de que sean disjuntos, y que para todas las clases incontables S y T existía una función inyectiva o suprayectiva $f : S \mapsto T$. Esto dice que la necesidad de la axiomatización en realidad fue independiente de las paradojas encontradas.

Dos años antes de la axiomatización de Zermelo de 1908, indicó a la Academia de Ciencias de Göttingen, que él tomaba como base de su axiomatización, un dominio de distintos «objetos», algunos que son conjuntos y otros que son objetos indivisibles. Con esto se propusieron por vez primera los *urelements* (o *átomos*) como objetos que no son conjuntos, no tienen elementos y pueden pertenecer a conjuntos[74].

[72] Véase la p. 46.

[73] Zermelo mismo afirmó que era incapaz de probar tal independencia.

[74] En la p. 30 se dan anotaciones sobre los átomos.

No solo los 7 axiomas que planteó Zermelo fueron vistos como escogidos arbitrariamente, sino que el Axioma de Separación causó conflicto entre la comunidad. Éste tenía encajada la noción básica de «propiedad definida», un concepto vago y recurrente en los debates de esa época, debido a que los trabajos de Zermelo principalmente, buscaban una seguridad en sus bases mediante la definición precisa de estos conceptos, la aclaración de «definitez» desde un punto de vista matemático que salvaguarde a la teoría, evite las paradojas y permita la construcción (o la existencialidad si no pueden ser construidos) de los objetos matemáticos importantes de la época.

La palabra que usó Zermelo era *definit*, refiriéndose en un sentido técnico a decidible en la base de la relación de pertenencia. Los Axiomas de los conjuntos definidos se encuentran en la página 156 de [Moo].

Esa idea de propiedad definida estaba en medio de la lógica subyacente, de la matemática o del pensamiento en los círculos matemáticos, que no tenían bases sólidas para establecer de qué dependía dicha concepción; y fue hasta que en 1922 Thoralf Skolem metió esta noción de propiedad definida en la lógica de primer orden.

Entre 1904 y 1908 fueron concebidos muchos principios sobre la Teoría de Conjuntos, sin embargo, era difícil saber cuáles adoptar en un sistema, que no generara contradicciones.

2.4. ¿QUÉ RESULTÓ DEL AXIOMA Y LA AXIOMATIZACIÓN?

El Axioma Multiplicativo y el Principio del Buen Orden se demostraron equivalentes en 1908, para cerca de 1920 se agregó a la lista la Tricotomía de los Cardinales[75], pero los resultados que se agregaron a la lista y expandieron las aplicaciones del Axioma fueron los Principios Maximales, que bajo condiciones dadas, establecen la existencia de elementos R-maximales[76].

Uno de los que obtuvieron algo importante fue Felix Hausdorff en sus investigaciones de 1907[77]. Dos años después demostró de otra manera el resultado, para el cual ocupó como hipótesis que S era una familia de conjuntos tales que, para todo ordinal límite α se cumplía que si: $\{B_\beta | \beta < \alpha\} \subset S$, y si $B_\gamma \subset B_\delta$ siempre que $\gamma < \delta < \alpha$, entonces existirá un conjunto $B_\alpha \in S$ tal que $B_\beta \subset B_\alpha \; \forall \beta < \alpha$.

A tal S le llamó *extendible por límites* (*extendable by limits*), y lo ocupó para su teorema cuya referencia está en [Moo, p. 168]:

Teorema 2.4.1. *Sea M un conjunto infinito y sea S una familia no vacía de subconjuntos de M. Si S es extendible por límites, S contendrá un subconjunto que no es un subconjunto propio de algún miembro de S.*

Este teorema es lo hoy llamamos *Lema de Zorn*. Otros dos que demostró según Moore [Moo, p. 168] y que se deducen del Axioma, son:

Teorema 2.4.2. *Todo conjunto parcialmente ordenado X tiene un subconjunto parcialmente ordenado A, que es el mayor $\subset$-maximal entre los subconjuntos ordenados parcialmente de X.*

Teorema 2.4.3. *Todo conjunto parcialmente ordenado X tiene un subconjunto A parcialmente ordenado, que es el mayor respecto a la relación $\subset$ entre todos los subconjuntos parcialmente ordenados y también A incluye a B, un subconjunto dado parcialmente ordenado de X.*

[75] Friedrich Hartogs en 1915 demostró la implicación que faltaba para ser equivalencia entre ellos dos, demostró que en el sistema de Zermelo sin el Axioma, la Tricotomía de los Cardinales implica que todo conjunto puede ser bien ordenado.

[76] Si se tiene el conjunto parcialmente ordenado (X,R), un elemento de X es R-maximal si no hay otro elemento estrictamente más grande que él respecto a la relación R.

[77] Ya se habló de su concepto de pantachie en la p. 56, un conjunto maximal de sucesiones reales.

No fueron muy conocidos en la comunidad, ni los propuso como axiomas ni como postulados, pero en el teorema 3.1.4 se demuestra que son equivalentes al Axioma.

Russell, interesado por la veracidad o falsedad del Axioma, en el 1er volumen del libro *Principia Mathematica* de 1910 (véase [Moo, p. 169]), incluyó que el Axioma era equivalente a estas afirmaciones ya vistas antes:

(2.4.1) Para toda clase K y toda relación S con $K \subset domS$, existe una función f con $f \subset S$ y con $domf = K$,

y a:

(2.4.2) Para toda clase K y toda función f que cumple con que $K \subset ranf$, existe una función inyectiva g que cumple con que $g \subset f$ y $rang = K$.

Respecto a los resultados algebraicos, éstos se tuvieron primordialmente en Teoría de Campos, Dominios Integrales y Álgebra Abstracta. Ernst Steinitz ocupó el Axioma en *campos algebráicamente cerrados*[78], diciendo que

(2.4.3) Existe una única cerradura algebráica para el campo de los números racionales, salvo isomorfismos[79].

Steinitz acertó que para un campo arbitrario F, el Axioma no solo se necesita para demostrar la unicidad, sino también para demostrar la existencia. Aunque se ha demostrado que para el campo de los racionales, el Axioma no es necesario. Su trabajo lo constituyó de manera que solo ocupe el Axioma cuando no se conozcan métodos para evitarlo, en los teoremas que propuso hizo mención explícita de cuándo ocupó el Axioma, dos de ellos son:

(2.4.4) Cualesquiera dos *campos absolútamente algebráicos* que tienen un *subcampo primo* finito y los mismos subcampos propios finitos son isomorfos[80].

(2.4.5) Si un campo F_2 es una extensión de un campo F, se sigue que F_2 puede ser obtenido de F por medio de una *extensión puramente trascendental*[81] F_1 seguida de una *extensión algebraica* con respecto a F_1.

Sus demostraciones dependieron del Axioma por que primero afirmó la existencia de un buen orden en los campos, para luego aplicar *Inducción Transfinita*[82], un método que puede decirse, es la generalización del Principio de Inducción Matemática.

Otra forma en la que ocupó el Axioma fue en el uso de esto:

---

[78] Un campo algebráicamente cerrado es un campo en el que todo polinomio de una indeterminada puede ser descompuesto en factores lineales. La *cerradura algebraica* de un campo es el campo algebráicamente cerrado más pequeño que incluye al campo dado. La definición de un *campo* está en la p. 114.

[79] Cuando se dice que algo pasa salvo isomorfismo, uno se refiere a que el resultado sucede aunque no es tan único, pero los demás objetos que lo cumplan serán isomorfos (y por lo tanto muy «parecidos») al objeto «único» obtenido.

[80] Un campo absolútamente algebráico es un campo que tiene a todos sus elementos algebráicos relativos a su subcampo primo. Un campo primo es un campo que no tiene subcampo propio.

[81] Un campo F_1 es una extensión puramente trascendental de un campo F si F_1 es isomorfo a un campo de cocientes obtenido al adjuntar algún número de indeterminadas a F. Se dice que F_2 es una extensión algebraica de F_1 si todo elemento de F_2 es algebraico con respecto a F_1.

[82] Se ve en la p. 36.

(2.4.6) Si se tiene un conjunto infinito, la familia de todos sus subconjuntos finitos tiene la misma cardinalidad que el conjunto[83],

que le ayudó a demostrar que una extensión algebraica en un campo no tiene mayor cardinalidad que el campo, debido a que la cardinalidad de polinomios de una indeterminada en un campo infinito es de la misma cardinalidad que el campo.

Respecto a los trabajos que se hicieron, hubo quienes intentaron desarrollar la teoría evitando el Axioma, y así buscando conclusiones importantes sin él. Entre ellas, unos creían que sin el Axioma se puede demostrar la existencia de un conjunto que no sea Lebesgue medible, simplemente ocupando este teorema:

Teorema 2.4.4. *La familia de todos los subconjuntos de* $\mathbb{R}$ *denumerables tiene la potencia del continuo.*

El analista Michele Cipola, en 1913, tuvo una idea que funcionó, juntó en un conjunto todos los números reales que satisfacían una condición, formando una sucesión denumerable de conjuntos en vez de elegir un real de cada subconjunto. Extendió el concepto de punto límite a familias de conjuntos, estableciendo que si se tiene una sucesión de conjuntos no vacíos A_1, $A_2, \ldots$ de los reales, ésta converge a un real p si para todo $\varepsilon > 0$ existe un entero positivo m que para todo $n > m$ se cumple que A_n es un subconjunto de $(p-\varepsilon, p+\varepsilon)$. Con esto en mano, demostró sin el Axioma que[84]:

Teorema 2.4.5. *Si p es un punto límite de un conjunto X, uno puede construir una* secuencia no constante[85] *de conjuntos contenidos en X que convergen a p.*

También Nikolai Luzin y Mikhail Suslin obtuvieron resultados sin el Axioma. Utilizaron una nueva definición para demostrar que la definición de Conjuntos Borelianos en $\mathbb{R}$ puede no utilizar el Axioma. Llamaron a una función $f : A \mapsto B$, donde A consta de sucesiones finitas de números naturales y B de intervalos cerrados[86] un *sistema definidor* (defining system). Se dice entonces que un número real p está asociado con f si existe una sucesión a_1, a_2, ... de números naturales tales que p está en $f(a_1)$, $f(a_1, a_2)$, ... Llamaron a un subconjunto $A \subset \mathbb{R}$ *analítico* si existe un sistema definidor f que hace que A contenga todos y solo esos números reales asociados con f. Concluyendo con todo esto que:

Teorema 2.4.6. *Todo conjunto de* $\mathbb{R}$ *es de Borel*[87] *ssi son analíticos él y su complemento.*

Al cabo de varios esfuerzos, intentando ver la consistencia o no consistencia de los axiomas de Zermelo, Hausdorff encontró una paradoja en el sistema que movió los suelos donde varios pisaban. Fue encontrada por él en 1914, y establece que la mitad de una esfera es congruente con un tercio de la misma esfera.

Él lo que quería era resolver el problema de Lebesgue de la Medida, visto ya en la p. 50, replicando que la definición constructiva de la función m no le asocia un valor a todos los conjuntos acotados de $\mathbb{R}$, y esto hace que no se tenga solución al problema. Hausdorff pudo demostrar que una solución para $\mathbb{R}^{n+1}$ daría una solución para $\mathbb{R}^n$, y al principio se concentró en una respuesta para la recta real, encontrando que ni el círculo unitario (i.e. el intervalo $[-1,1]$) la tenía.

En realidad Hausdorff ocupó estas 4 condiciones parecidas a las que pedía Lebesgue:

[83] Si el conjunto es denumerable o tiene la potencia del continuo, no se necesita del Axioma.

[84] En el caso de convergencia de sucesiones de números reales, éste teorema requiere del Axioma.

[85] Refiriéndonos a que no se cumple que $A_n = A$ para infinitos $n \in \mathbb{N}$.

[86] Subconjuntos de $\mathbb{R}$ tales que su complemento es abierto, i.e., que pertenece a la topología generada por los intervalos abiertos. Los conceptos se darán en la sección 3.6.

[87] Existen modelos como los de Feferman y Levy donde todo subconjunto de los reales es Boreliano, este resultado de Suslin no se sabe si pude ser deducido en ZF.

1. El cubo unitario n-dimensional tiene medida 1.

2. Conjuntos congruentes tienen la misma medida.

3. Si se tienen dos conjuntos disjuntos y acotados A y B, $m(A \cup B) = m(A) + m(B)$.

4. Si los conjuntos A_i con $i \in \mathbb{N}$ son disjuntos y su unión está acotada, entonces $m(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} m(A_i)$.

La solución se puede encontrar para el plano y la recta real si se desecha la cuarta condición, sin embargo, ni desechando esa solución se puede encontrar una respuesta afirmativa al problema para el caso de dimensiones mayores.

Al encontrar soluciones para los problemas de Lebesgue, sabiendo la necesidad del Axioma para varios, Hasudorff bajó de grado las condiciones para formular un nuevo problema de la Medida de Lebesgue, requiriendo las mismas condiciones que requería el anterior (las que acabo de enunciar menos la última), pero pidiendo solo aditividad finita, debido a que la contable dependía del Axioma. Con esto, encontró la *paradoja de Hausdorff* al encontrar que no existía solución para cuando la dimensión era mayor a 2; lo hizo particionando la esfera en cuatro conjuntos A, B, C y D, con los tres primeros y la unión de B y C congruentes entre ellos, y D un conjunto contable.

Dos preguntas quedaban por resolver para las dimensiones pequeñas, y eso hizo que saliera a flote una paradoja más especial: la *paradoja de Banach-Tarski*. Ésta fue una razón que a varios, como Borel, les reafirmaba el hecho de negar al Axioma, debido a que el resultado dependía de él; por lo mientras, Banach y Tarski, no veían a su demostración como algo fuera de lo intuitivo ni paradójico del Axioma, sino simplemente como un hecho o propiedad.

El intrigante resultado lo obtuvieron al establecer que cualesquier dos subconjuntos acotados de $\mathbb{R}^n$ ($n \leq 3$), que tuvieran elementos en su *interior*[88], son *equivalentes mediante una descomposición finita* (o equivalent by finite decomposition). La definición de equivalentes mediante descomposición finita es lo que cumplen dos conjuntos A y B, para los cuales existe una $n \in \mathbb{N}$, tal que para A, el i-ésimo elemento de una partición $\{A_1, A_2, \ldots, A_n\}$ es congruente con el i-ésimo elemento de una partición $\{B_1, B_2, \ldots, B_n\}$ de B, con $i \in \{1, 2, \ldots, n\}$. El resultado fue publicado en 1924, el libro (hecho por los dos, aunque descubrieron la paradoja independientemente) lo escribieron bajo el nombre de «*Sur la décomposition des ensembles de points en parties respectivement congruentes.*» [Moo, p. 284], y de él se concluye que cualesquiera dos esferas de radio cualquiera, son equivalentes mediante una descomposición finita, o como otra conclusión, que una esfera de radio arbitrario puede ser particionada en un número finito de piezas y reensamblada en dos esferas del mismo radio.

Banach generalizó el problema para convertirlo en el *Problema de la Medida Generalizado*, donde preguntaba por la existencia de un conjunto X y una medida m (función) la cual, asigne un número real no negativo a cada subconjunto de X, y además que satisfaga:

1. $\exists A$ que satisface $m(A) \neq 0$,

2. $m(A) = 0$ si A tiene exactamente un elemento,

3. m es contablemente aditiva.

[88] La definición de interior de un conjunto está en la p. 136.

Al cardinal del conjunto que satisfaga dicho problema se le llama *cardinal medible real valuado*[89]. Los resultados a los que se arribó: si a dicha generalización se le restringe menos pidiéndole sólo aditividad finita, del Axioma de deduce que todo conjunto infinito cuenta con dicha medida; Stanislaw Ulam demostró ocupando el Axioma, que un conjunto que satisfaga el problema generalizado tiene una potencia mayor o igual que el primer cardinal fuertemente inaccesible, siempre y cuando no haya cardinal débilmente inaccesible menor o igual a $2^{\aleph_0}$.

Borel en su libro de 1914 (véase [Moo, p. 188]), anotó que era una paradoja a lo que llegó Hausdorff (ya que Hausdorff no lo había visto así), y culpó al Axioma, que es necesario para la demostración. Dijo:

> «*If, then, we designate by a, b, c the probability that a point in S belongs to A, B or C respectively and if we grant that the probability of a point belonging to a set E is not changed by a rotation around a diameter (this is what Lebesgue expresses by saying that two congruent sets have the same measure), one obtains the contradictory equalities:* $a+b+c=1$, $a=b$, $a=c$, $a=b+c$. *...The set A is homogeneous on the spere; but it is at the same time a half and a third of it...*»

Otra anotación de Borel, fue decir que dicho conjunto A no estaba definido, en el sentido lógico de la palabra definido, debido a que se llegaba a una contradicción asumiendo su existencia, sin embargo, Kurt Gödel demostró que los conjuntos no medibles de Hausdorff A, B y C son definibles por medio de ordinales en la teoría ZF[90].

2.5. UNA ETAPA QUE APACIGUA LA CONTROVERSIA

El Axioma fue ganando adeptos al paso del tiempo, o por lo menos, matemáticos que investigaban sobre sus consecuencias. Algunos de los resultados obtenidos fueron en Campos Reales (Artin y Schreier), Álgebras Booleanas (Stone), el Lema de Zorn, la postulación del Axioma en la lógica llamado el ε-axioma (Hilbert)[91], Teoría de Modelos (Löwenheim y Skolem), Topología (Tychonoff), y demás. Se llegó a su independencia en el caso denumerable (Fraenkel), su consistencia relativa mediante la definición de conjuntos construibles (Gödel).

El cambio radical respecto a los avances del Axioma, empezó cuando Waclaw Sierpinski ayudó a fundar en 1918 una escuela de matemáticas en Varsovia. Incitando a sus alumnos a que se investiguen las consecuencias del Axioma, esa escuela tuvo grandes resultados en el tema de manera casi inmediata. De los primeros fue el de Tarski y Banach, quienes con la idea de Hausdorff, crearon un teorema conocido ahora con el nombre de la *paradoja de Banach-Tarski* que descompone una esfera en un número finito de piezas y las vuelve a reconstruir para formar dos esferas cada una con el mismo radio que la esfera original.

En los primeros años de vida de dicha escuela destacan las investigaciones de Sierpinski, donde demostró dónde se requería del Axioma y dónde se podía evitar, lo hizo en varias partes de las matemáticas. Junto con Luzin demostró que se puede evitar el Axioma en estos dos teoremas:

Teorema 2.5.1. *Sea una partición de* $\mathbb{R}$ *en dos conjuntos, entonces al menos uno de ellos tiene la potencia del continuo.*

[89] Cuando se restringe a la función m a solo tomar los valores 0 y 1, se dice que el cardinal de dicho conjunto es un *cardinal medible*.

[90] Construibles de orden algún ordinal, véase la página 79.

[91] Es un análogo a la existencia de funciones de elección sobre conjuntos.

Teorema 2.5.2. *La* proyección[92] *de un conjunto $A \subset \mathbb{R}^2$ en los reales, tiene menor potencia que el conjunto A.*

Este último teorema (que depende del Axioma) es una particularización de $\forall A|f(A)| < |A|$, con A conjunto, que es equivalente al Principio de Partición.

También se puede evitar el Axioma en la demostración de que la familia de todas las sucesiones infinitas tiene la potencia del continuo, así como la familia de todos los conjuntos finitos de $\mathbb{R}$. Sobre cardinales, si $\mathfrak{m} \leq \mathfrak{m}_0$ y $\mathfrak{n} \leq \mathfrak{n}_0$, se cumple que $\mathfrak{m} + \mathfrak{m}_0 \leq \mathfrak{n} + \mathfrak{n}_0$ y también $\mathfrak{m}\mathfrak{m}_0 \leq \mathfrak{n}\mathfrak{n}_0$ sin necesidad del Axioma[93].

En cuanto a topología, demostró varios hechos que dependían del Axioma Denumerable al reemplazar la definición de punto límite, por punto límite secuencial, que si se toman los teoremas con la noción de punto límite, no se necesita el Axioma en su demostración[94]:

Teorema 2.5.3. *Todo subconjunto secuencialmente isolado de $\mathbb{R}^n$ es contable.*

Teorema 2.5.4. *Todo subconjunto secuencialmente cerrado de $\mathbb{R}$, que está acotado por arriba, contiene su cota superior menor.*

Teorema 2.5.5. *El conjunto secuencialmente derivado de X, denotado por X', cumple que $X'' \subset X'$.*

Teorema 2.5.6. *Todo conjunto secuencialmente perfecto[95] tiene la potencia del continuo.*

Teorema 2.5.7 (Lindelöf Covering Theorem)**.** *Toda familia de intervalos abiertos de los reales, que cubren[96] a un subconjunto X de $\mathbb{R}$, contiene una subfamilia contable que aún cubre a X.*

Resultados sobre la Teoría de la Medida y Análisis Real dependen mucho del Axioma:

Teorema 2.5.8 (Borel)**.** *Si una* función medible[97] *f, que está acotada* casi en todos lados[98]. *y se tienen dos números positivos y pequeños ε y ε_1 cualesquiera, entonces existe un polinomio $P(x)$ tal que $m\big(\{x : |f(x) - P(x)| > \varepsilon\}\big) < \varepsilon_1$.*

Teoremas sorprendentes por la conclusión de ellos, como algunas cosas antiintuitivas, son:

Teorema 2.5.9 (Mazurkiewicz)**.** *Existe un subconjunto X del plano, tal que toda línea tiene exactamente dos puntos en común con X.*

Teorema 2.5.10 (Luzin)**.** *Existe un subconjunto incontable en los reales, que es de primera categoría en todo conjunto perfecto.*

Teorema 2.5.11 (Luzin y Sierpinski)**.** *Existe un subconjunto incontable X de un intervalo dado I que cumple con que toda biyección continua de I transforma a X en un conjunto de medida cero.*

Teorema 2.5.12 (Sierpinski)**.** *Existe una familia F* casi disjunta[99] *de subconjuntos no medibles del intervalo $[0,1]$ tal que $|F| > 2^{\aleph_0}$.*

---

[92] Si $f : \mathbb{R}^2 \mapsto \mathbb{R}$, $f(A)$ es la proyección del conjunto A en $\mathbb{R}$.

[93] Reemplazando por desigualdad estricta todas las relaciones, la demostración si requiere del Axioma.

[94] También los de Bolzano-Weierstrass, Cantor-Bendixson y Heine-Borel entre otros.

[95] Un conjunto perfecto es un conjunto cerrado y el cuál todos sus puntos son puntos límite.

[96] Una familia $\{A_\alpha\}_{\alpha\in\mathscr{A}}$ cubre a un conjunto X si $X \subset \bigcup_{\alpha\in\mathscr{A}} A_\alpha$.

[97] Una función es medible si existe una sucesión de funciones escalonadas (continuas excepto un número finito de puntos) tales que la función y el límite de la sucesión es igual excepto en un conjunto de medida cero.

[98] Que no esté acotada en un conjunto de medida cero, i.e., en un conjunto de muy «poco volumen».

[99] $\{A_\alpha\}_{\alpha\in\mathscr{A}}$ es una familia casi disjunta si $|\bigcap_{\alpha\in\mathscr{A}} A_\alpha| < |A_\beta|$ para todo $\beta \in \mathscr{A}$.

Tarski fue el que dio gran avance a la teoría de Cardinales según Moore [Moo, p. 213], con su publicación en 1924 dio 7 equivalencias al Axioma:

(2.5.1) Para todos los cardinales infinitos $\mathfrak{m}\mathfrak{n} = \mathfrak{m} + \mathfrak{n}$.

(2.5.2) Para todo cardinal infinito $\mathfrak{m}^2 = \mathfrak{m}$.

(2.5.3) Si se cumple que $\mathfrak{m}^2 = \mathfrak{n}^2$, entonces debe de ser que $\mathfrak{m} = \mathfrak{n}$.

(2.5.4) Si $\mathfrak{m} < \mathfrak{n}$ y $\mathfrak{p} < \mathfrak{q}$, se cumple $\mathfrak{m} + \mathfrak{p} < \mathfrak{n} + \mathfrak{q}$.

(2.5.5) Si $\mathfrak{m} < \mathfrak{n}$ y $\mathfrak{p} < \mathfrak{q}$, se tiene entonces que $\mathfrak{m}\mathfrak{p} < \mathfrak{n}\mathfrak{q}$.

(2.5.6) Si se cumple $\mathfrak{m} + \mathfrak{p} < \mathfrak{n} + \mathfrak{p}$, se sigue que $\mathfrak{m} < \mathfrak{n}$.

(2.5.7) Cuando $\mathfrak{m}\mathfrak{p} < \mathfrak{n}\mathfrak{p}$, se puede cancelar para hacer $\mathfrak{m} < \mathfrak{n}$.

Dice Moore en [Moo, p. 216], que después de eso, Tarski junto con Lindenbaum hizo un artículo en 1924 llamado «*Communication sur les recherches de la théorie des ensembles*». Ahí añadieron una proposición equivalente al Axioma, definiendo primero que $|A| \leq^* |B|$ significa que $A = \emptyset$ y si no es vacío, significa la existencia de una función suprayectiva de B sobre A, la proposición es:

(2.5.8) Las relaciones $\leq^*$ y $\leq$ en cardinales, son equivalentes[100].

También la Tricotomía de los Cardinales con esta nueva definición es equivalente al Axioma.

Además, refirieron sin demostración una propiedad más fuerte que el Axioma, que es la *Hipótesis Generalizada del Continuo* y a una propiedad más débil que el Axioma de Partición respectivamente:

(2.5.9) Para todo ordinal α, se cumple $2^{\aleph_\alpha} = \aleph_{\alpha+1}$.

(2.5.10) Si se cumple $\mathfrak{m} \leq^* \mathfrak{n}$ entonces no se cumplirá que $\mathfrak{n} < \mathfrak{m}$.

Entre resultados que abrieron ramas que aún siguen creciendo, Tarski en 1938 (según Moore en [Moo, p. 218] investigó sobre los *cardinales inaccesibles*. Un cardinal se dirá *débilmente inaccesible* si es igual a $\aleph_\alpha$ para algún ordinal *regular*[101] límite α.

La cuestión era la existencia de un cardinal débilmente inaccesible. Junto con Sierpinski ya había introducido la noción de *cardinal fuertemente inaccesible*, que es el cardinal de un conjunto X que no es la unión de menos de $|X|$ conjuntos de potencia menor que $|X|$ y que X no sea equipotente a la potencia de algún conjunto. Con el Axioma, todo cardinal fuertemente inaccesible es débilmente inaccesible.

Además, añadió el *Axioma de Conjuntos Inaccesibles*, que aseguraba la existencia de un conjunto más grande para todo conjunto, cuyo cardinal es fuertemente inaccesible:

Axioma 2.5.1 (Axioma de Conjuntos Inaccesibles)**.** *Para todo conjunto A, existe otro B que cumple las 4 condiciones:*

[100] Recuerde que $|A| \leq |B|$ significa que existe una función inyectiva de A a B.

[101] Se dice que un ordinal es *regular* si no es cofinal con otro ordinal menor. La *cofinalidad* de β, denotada $cf\,\beta$, es el menor ordinal α, tal que existe una aplicación cofinal $f : \alpha \mapsto \beta$. Una aplicación $f : \alpha \mapsto \beta$ entre ordinales es *cofinal* si $f(\alpha)$ no está acotado estrictamente en β, es decir, $\forall \gamma < \beta$, $\exists \lambda < \alpha$ que cumple con que $\gamma \leq f(\lambda)$.

1. *$A \in B$,*

2. *si $X \in B$ y también $Y \subset X$, se sigue que $Y \in B$,*

3. *si $X \in B$, entonces $\mathscr{P}(X) \in B$ y*

4. *si $Y \subset B$ y cumple que $|Y| < |B|$, entonces $Y \in B$.*

De su axioma, Tarski demostró el Axioma de Elección. Ya cerca de 1939, Tarski llevaba en su lista de equivalencias al Axioma, 16 proposiciones sobre cardinales.

Respecto a los principios maximales, Hausdorff fue el que introdujo a la comunidad el concepto de *conjunto maximal* (según Moore en [Moo, p. 220]), hoy conocido por elemento $\subset$-maximal, un conjunto A de una familia que no está contenido propiamente en otro conjunto de la familia ($\neg(\exists B)B \neq A \wedge A \subset B$). Como aplicación de los ordinales transfinitos, escribió que un conjunto X es *extendible superiormente por límites* (o *extendable above by limits*) si para cada una de sus subfamilias bien ordenadas por $\subset$, existe una cota superior perteneciente a X con respecto a $\subset$, si se toma la inclusión contraria se habla de un conjunto *extendible inferiormente por límites* (o *extendable below by limits*), con esto, demostró[102] que:

(2.5.11) Si se tiene una familia no vacía X de conjuntos que es extendible superiormente por límites, se tendrá que X contiene un elemento $\subset$-maximal.

Este teorema anterior tiene su análogo si se cambia la inclusión y se llega a un elemento $\subset$-minimal. Kazimierz Kuratowski dio un principio minimal 5 años antes que Hausdorff[103]:

(2.5.12) Si S es una familia de los subconjuntos de un conjunto X que tienen una propiedad dada P, y si $X \in S$ y también para todo subconjunto no vacío A de S bien ordenado por $\supset$ se cumple que $\bigcap A \in S$, entonces resultará que S tiene un conjunto *irreducible*.

Era 1935 cuando salió el Lema de Zorn, que fue muy conocido por la comunidad, pero al principio, no fue tratado como teorema por Max Zorn sino como axioma:

Axioma 2.5.2 (Lema de Zorn)**.** *Si se tiene una familia de conjuntos A tales que la unión de toda familia ordenada por inclusión $B \subset A$ está en A, entonces existe un elemento $C \in A$ que no es un subconjunto propio de otro conjunto de A.*

El resultado del Lema de Zorn es un elemento C que es $\subset$-maximal, y dio tanto a algebristas como a topólogos que utilizaban el Principio del Buen Orden y la Inducción Transfinita, otra herramienta para trabajar con la cuál ellos se sentían menos incómodos. Pero fue hasta 1940 cuando Tukey demostró la equivalencia de éste con el Axioma, no dando ventaja al Lema de Zorn en contra del Axioma.

Nicolas Bourbaki, como se hacían llamar unos matemáticos franceses, publicó(aron) teoremas relacionados, Moore en [Moo, p. 224] lo llama Teorema de Zorn:

Teorema 2.5.13 (Teorema de Zorn)**.** *Si se tiene un conjunto parcialmente ordenado X y si toda cadena de X tiene una cota superior en X, entonces X tiene un elemento maximal.*

Otro que dio en ese libro fue:

---

[102] Se sabe ahora que es equivalente al Axioma.

[103] Kuratowski llamaba irreducible a lo que llamamos conjunto $\subset$-minimal.

Teorema 2.5.14. *Si se tiene una familia no vacía de* carácter finito[104] *de subconjuntos de un conjunto X, entonces la familia contendrá un elemento máximo.*

El algebrista Oswald Teichmüller también demostró este teorema, junto con otro sobre cálculo proposicional. Demostró varias equivalencias con el Axioma que Zorn no pudo, y además mostró parte del alcance que tienen en Álgebra y Análisis los Principios Maximales.

En realidad, la mayoría de los Algebristas que usaban métodos abstractos, requerían de conjuntos maximales, y por eso, aceptaban el Axioma o el Lema de Zorn sin dudarlo, aunque hubo un periodo en el que se quiso hacer al Álgebra constructiva.

En el año de 1937, Erich Kamke fue el que descubrió la existencia de funciones complejas discontinuas que satisfacían las ecuaciones $f(z+w) = f(z) + f(w)$, y $f(zw) = f(z)f(w)$ vistas ya en la p. 55; sin embargo, utilizó una *base trascendente*[105] de $\mathbb{C}$ sobre $\mathbb{Q}$.

Ya para el año de 1932, Hausdorff demostró en su libro (véase [Moo, p. 228])que también contenía conceptos de espacios métricos uno de los resultados más importantes del Álgebra, debido a su generalidad (suponiendo válido el Principio del Buen Orden):

(2.5.13) Todo espacio vectorial[106] tiene una base.

Después se demostró que ésta afirmación es equivalente al Axioma. Teichmüller demostró otro resultado parecido:

(2.5.14) Todo *espacio de Hilbert*[107] tiene una base ortonormal.

La *Teoría de Campos Reales*[108] (rama del Álgebra Abstracta) y la Teoría de Anillos creció gracias al Axioma. Emil Artin y Otto Schreier en 1927 (según Moore en [Moo, p. 228]), introdujeron los conceptos de *campo real* y *campo real cerrado*, para demostrar por medio del Principio del Buen Orden que:

(2.5.15) Todo campo real tiene al menos una extensión algebraica que también es un campo real cerrado.

También demostraron por el Axioma que:

(2.5.16) Todo campo real es un *campo ordenado*[109].

Emmy Noether y Wolfgang Krull aportaron a la *Teoría de Anillos*[110], encontrando métodos para la clasificación de anillos y demostrando en su libro sobre Teoría de Anillos que (véase [Moo, p. 229]):

(2.5.17) Todo *anillo conmutativo*[111] con unidad puede ser extendido a un anillo algebráicamente cerrado con unidad.

---

[104] Una familia $\mathscr{A}$ se dice que tiene *carácter finito* si $X \in \mathscr{A}$ ssi todo subconjunto finito de X pertenece a $\mathscr{A}$.

[105] Una *base trascendente* de un campo F sobre un subcampo F_1 es un subconjunto maximal de F algebráicamente independiente con respecto a F_1.

[106] Las definiciones de *Espacio Vectorial* y de *base* se verán en la introducción de la sec. 3.4.

[107] Espacios normados completos, i.e., en donde toda sucesión de Cauchy converge.

[108] Un campo es un *Campo Real* si -1 no es una suma de cuadrados. Un *Campo Real Cerrado* es un Campo Real en el cuál ninguna de sus extensiones algebraicas es un campo real.

[109] Un *campo ordenado* es un campo en el que hay un orden $<$ tal que cumple que si $0 < a$ y también $0 < b$, se tendrá que $0 < a+b$ y $0 < ab$.

[110] Lo que es un anillo está en la p. 114.

[111] Véanse las definiciones en la introducción de la sección 3.4.

Krull quiso generalizar investigaciones sobre la *Teoría de Ideales*, y obtuvo que:

(2.5.18) En todo anillo conmutativo, todo ideal propio puede ser extendido a un *ideal maximal primo*.

Era 1936 cuando Marshall Stone encontró representaciones para los *Anillos Booleanos*, que son anillos, con al menos dos elementos, en los cuales $a^2 = a$ para todo elemento. Se define una *Álgebra Booleana* como un anillo Booleano con unidad; y un *ideal*[112] como un subconjunto I de un anillo conmutativo X, tal que $\forall a,\ b \in I$ y también $\forall c \in X$, se tiene $a+b, ac \in I$. A Marshall se le debe este importante teorema más débil que el Axioma:

Teorema 2.5.15 (Boolean Prime Ideal Theorem)**.** *En todo anillo Booleano existe al menos un ideal primo.*

Su referencia se encuentra en [Moo, p. 230]. De este teorema, dedujo otro también famoso:

Teorema 2.5.16 (Stone Representation Theorem)**.** *Si se tiene un anillo Booleano, éste será isomorfo a un anillo de conjuntos*[113].

En otras áreas (Álgebra combinado con Teoría de Conjuntos) se aplicaron resultados del Axioma, por ejemplo, Dénes König ocupó en *Teoría de Grafos* y a la *Teoría de Juegos*, el siguiente Lema de Infinitud:

(2.5.19) Si una *gráfica* infinita A tiene subgráficas disjuntas $A_1, A_2, \ldots$, que cumplen que cada punto de A_{n+1} está unido a algún punto de A_n, entonces existirá una *rama*[114] infinita $a_1, a_2, \ldots$ con $a_n \in A_n$, para todo natural.

Respecto a los avances en Análisis Funcional, se tiene el de Stefan Banach de 1929, ahora llamado:

Teorema 2.5.17 (Teorema de Hahn-Banach)**.** *Si se tiene una funcional real*[115] *p sobre un espacio vectorial normado y completo V, que cumple con que $p(x+y) \leq p(x)+p(y)$ y $p(ax) = ap(x)$, para todos los $x,\ y \in V$ y $a \in \mathbb{R}^+$, entonces se tendrá la existencia de una* funcional aditiva[116] *f tal que para todo $x \in V$, pasa la desigualdad $-p(-x) \leq f(x) \leq p(x)$.*

Respecto a la Topología General, no se tenían nociones claras de espacios que no fueran $\mathbb{R}^n$ hasta que Maurice Fréchet generalizó y definió unos espacios en su tesis de 1906 (según Moore en [Moo, p. 235]). Definió los *L-espacios* y *los espacios métricos*[117], luego de tener más herramientas sobre vecindades, generalizó éstos y definió los *espacios topológicos*. Hausdorff con ello demostró el teorema que depende del Axioma Denumerable:

Teorema 2.5.18. *Todo subespacio de un espacio métrico* separable[118]*, será separable.*

---

[112] Término definido en la página 114.

[113] Un *anillo de conjuntos* es una familia de conjuntos con las operaciones de diferencia simétrica representada por unión y la de intersección por multiplicación.

[114] No se darán las definiciones de Rama y Gráfica, ni nos meteremos tanto en esos temas.

[115] Las definiciones están en la subsec. 3.6.2.

[116] Una *funcional aditiva* es una funcional que cumple para todos $a,\ b \in \mathbb{R}$ y para todos $x,\ y \in V$ que $p(ax+by) = ap(x) + bp(y)$.

[117] Su definición en la p. 146.

[118] *Separable* significa que contenga un subconjunto denso contable. Un subconjunto D de un espacio topológico es *denso* si todo punto del espacio está en D o es un punto límite de D.

Sierpinski, Hausdorff y Fréchet fueron innovadores en este campo, usando elecciones sin siquiera explicitarlo. Mediante el Axioma, Sierpinski demostró que si un espacio topológico X tiene una base contable, entonces toda cubierta abierta de X tendrá una subcubierta abierta contable[119].

Un espacio se dice *segundo contable* si tiene una *base* contable. Con esto, Sierpinski demostró que:

Teorema 2.5.19. *En un espacio de Hausdorff que también es segundo contable, un punto x será punto límite de un conjunto X ssi x es un punto límite secuencial.*

Teorema 2.5.20. *En un espacio de Hausdorff segundo contable X, una función $f : X \mapsto X$ será* continua *en un punto x ssi f es* secuencialmente continua *en el punto x.*

Kuratowski fue el que demostró que todo espacio topológico segundo contable es separable.

Dos topólogos rusos crearon un concepto fundamental en esta materia. Paul Alexandroff y Paul Urysohn definieron en 1923 el término *bicompactness*, que ahora se llama compacidad, definida como: un espacio topológico X es compacto si toda cubierta abierta de él contiene una subcubierta abierta finita. Mediante su definición de *punto de acumulación completo* de un conjunto X, que es un punto p tal que para todo abierto V con $p \in V$, se tiene que $V \cap A \sim A$, generalizaron el teorema de Bolzano-Weierstrass (utilizando el Axioma):

Teorema 2.5.21. *Todo subconjunto infinito de un espacio topológico tiene un punto de acumulación completo.*

Alexandroff publicó los resultados de Urysohn después de que éste murió, entre ellos esta el Lema de Urysohn (véase [Moo, p. 237]):

(2.5.20) Si tenemos dos subconjuntos $A,\ B \subset X$ disjuntos y cerrados, y X es un espacio de Hausdorff *normal*, implica que existe una función real f continua en X con $f(x) = 0$ si $x \in A$ y $f(x) = 1$ si $x \in B$.

Donde se define a un espacio topológico *normal* el que para todos A y B disjuntos y cerrados, existen abiertos U y V con $A \subset U$ y $B \subset V$.

Gracias a los avances de Alexandroff y Urysohn, de otro topólogo salió un teorema que demostraremos en la p. 138, llamado por su creador Andrei Tychonoff, aunque solo demostró una versión débil del teorema[120], gracias a su método se generaliza:

Teorema 2.5.22. *El producto de cualquier familia de espacios compactos es compacto.*

Tychonoff dependió del Axioma debido a que utilizó en su demostración el Lema de Urysohn; Eduard Čech lo aplicó a espacios topológicos más generales en 1937 mediante el Principio del Buen Orden, y por eso a veces se le llama Teorema de Čech-Tychonoff, se demostró en 1950 que es equivalente al Axioma. Luego, Čech definió el concepto de *compactificación de Stone-Čech* de un espacio de Tychonoff[121] S (o con su notación original $\beta(S)$). Un problema fundamental respecto a su compactificación, fue que Čech pensó que ni siquiera se podría construir un elemento del conjunto $\beta(\mathbb{N}) - \mathbb{N}$. Un elemento de este

[119] Véase el principio de la sec. 3.6.1 para las definiciones.

[120] En [Moo, p. 238] esta la referencia del libro donde probó que el producto de cualquier número de copias de $[0,1]$ es compacto, donde una copia es el mismo conjunto.

[121] Un espacio de Hausdorff completamente regular se dice que es de Tychonoff. Un espacio topológico X es *completamente regular* cuando para todo cerrado F y todo punto $p \notin F$, existe una función real f en X continua que cumple con $\forall x \in F,\ f(x) = 1$ y $f(p) = 0$.

conjunto puede ser visto como un *ultrafiltro no principal*[122] en $\mathbb{N}$, y de hecho, su existencia depende del Axioma. Un ultrafiltro F en X se dirá *principal* si consiste de todos los subconjuntos de X que contienen un punto dado $x \in X$, los no principales no cumplen con esta condición. En la actualidad, se sabe que la existencia de la compactificación de Stone-Čech para espacios Tychonoff es equivalente al Teorema del Ideal Booleano Primo.

Como era de esperarse, se generalizaron a espacios menos restrictivos los conceptos de convergencia de sucesiones. Un aproximamiento lo hizo Birkhoff, mediante la definición de *conjuntos directos*, $\{x_\alpha\}_{\alpha\in\mathscr{A}}$ es un conjunto directo si es una función inyectiva con una relación transitiva definida en $\mathscr{A}$, y con la existencia de un $c \in \mathscr{A}$ que cumple $a \leq c$ y $b \leq c$ para todos los $a,\ b \in \mathscr{A}$. La convergencia del conjunto directo $\{x_\alpha\}_{\alpha\in\mathscr{A}}$ a x ($x_\alpha \mapsto x$) la definió si se cumple que para todo abierto V con $x \in V$, existe un $b \in A$ tal que $b \leq c$ implicara que $x_c \in V$.

El otro concepto generalizado de convergencia, lo dio Henri Cartan en 1937, por medio de *filtros* sobre un conjunto X, un filtro en X es una familia F no vacía de subconjuntos de X que cumple:

1. $\emptyset \notin F$,
2. si $A, B \in F \Rightarrow A \cap B \in F$ y
3. si se tiene que $A \subset B \subset X$ y también $A \in F \Rightarrow B \in F$.

La referencia al libro donde se publicó esta definición está en [Moo, p. 240], en su texto publicado el mismo año, refiriéndose a la emergente Teoría de Filtros y Ultrafiltros escribió[123]:

Teorema 2.5.23 (Ultrafilter Theorem)**.** *Todo filtro puede ser extendido a un ultrafiltro.*

La convergencia de un filtro F fue definida mediante un punto límite p de F, que no era otra cosa más que un punto en el que todo abierto que lo contiene está contenido en F (i.e. $p \in V \subset F$, $\forall V$ abierto tal que $p \in V$).

Cartan demostró en ese mismo texto que un espacio de Hausdorff es compacto ssi todo ultrafiltro en el espacio tiene un punto límite.

Actualmente, se sabe que un filtro en un espacio es el dual de un ideal en una Álgebra Booleana, lo que da la equivalencia del Teorema de Ultrafiltros con el Teorema del Ideal Booleano Primo.

2.5.1. Sobre Lógica

En las épocas de 1920 no se sabía sobre la *consistencia* del Axioma, que significa que no se sabía si añadir al sistema ZF el Axioma de Elección llevaría a contradicciones o paradojas, ni había métodos que ayudaran a resolver el problema de una manera segura. Algunos dudaban de la «absolutez» de una Teoría de Conjuntos, dudaban de no poder encontrar una teoría universal, pensando en que podría haber cardinales inaccesibles en un universo teórico, mientras que en otro ni siquiera existían. La alternativa de los matemáticos que rechazaban el Axioma era buscar otras teorías[124], pero ellos no solo tenían el mismo problema de demostrar la consistencia de su postulado alternativo, sino que también lidiaban con encontrar utilidad (matemática) a su teoría nueva, como la que encontraron años antes con el Axioma los matemáticos mencionados en las secciones pasadas; pero hubo varios intentos que trajeron resultados mencionables, dos de ellos son el Axioma de Constructibilidad y el Axioma de Determinancia que se verán en las páginas 79 y 83 respectivamente.

[122] La definición de ultrafiltro está en la p. 114.
[123] Un *ultrafiltro* es un filtro maximal.
[124] Véase la sec. 3.2.

Fue en los años siguientes de 1920 cuando empezó a investigarse a profundidad la rama de la Lógica Matemática, y con ello, el aspecto lógico del Axioma (¿era un postulado de la lógica?). Los primeros niveles de la Lógica, que se dieron los primeros dos en la sección 1.7, se dividen conforme las clases: cálculo proposicional con conectores Booleanos ($\wedge$, $\vee$, $\neg$), cálculo predicativo de 1er orden (los cuantificadores solo se permiten en variables objeto), cálculo predicativo de segundo orden (los cuantificadores también se permiten en relatores y funciones), y así. Además, se dividían en semántica y sintaxis, elaboradas en trabajos de Álgebra de la Lógica por George Boole, Pierce y Schröder y en trabajos de Frege, Russell y Hilbert, respectivamente; esa división tomó una forma más homogénea cuando Tarski y Gödel entraron en escena. Ellos dos respondían cuestiones de sintaxis y de semántica como ¿Se puede llegar a una contradicción del sistema axiomático escogido?, o como ¿tiene el sistema axiomático un modelo?.

Mientras, en Berlin, Leopold Löwenheim fue el que utilizó a la lógica de 1er orden como contenida en la lógica, permitió que las sentencias tengan cadenas denumerables de cuantificadores y relaciones, para demostrar el:

Teorema 2.5.24. *Si una sentencia de primer orden se* satisface[125] *en un conjunto M, entonces se satisface en un subconjunto contable.*

La generalización de este teorema, lleva a uno importante:

Teorema 2.5.25 (Teorema de Löwenheim-Skolem)**.** *Si un conjunto de fórmulas sobre un lenguaje contable tiene un modelo, entonces tendrá un modelo contable.*

Que un sistema axiomático sea *categórico* significa que cualesquiera dos de sus modelos son isomorfos. Éste teorema de Löwenheim-Skolem traé como consecuencia de fondo el hecho de que ningún sistema axiomático para la Teoría de Conjuntos que recaiga en la lógica de primer orden, puede ser categórico.

Tarski hizo una proposición que hasta después se demostró equivalente al Axioma (ahora conocida por el Teorema de Löwenheim-Skolem-Tarski), sin embargo, no publicó su demostración:

Proposición 2.5.1. *Si un conjunto contable de sentencias tiene un modelo infinito, tendrá un modelo infinito de cualquier cardinalidad.*

Hilbert también se propuso a la tarea de acercar a la matemática con la lógica por medios formales. Buscaba principalmente la consistencia de los axiomas de $\mathbb{R}$, demostrar que puede ser resuelto todo problema matemático *bien formulado* (o well-formulated). En la década de 1920 se creó una lógica basada en lenguajes de primer orden con un número contable de símbolos, aceptando de esta manera, lenguajes con un número de conjuntos de sentencias denumerables.

Para tratar a las matemáticas como cadenas de símbolos sin sentido, se formuló la *Teoría de Demostraciones Matemáticas* (o *Mathematical Proof Theory*), que servía como fundamento de los razonamientos matemáticos, encontrando así, a las demostraciones, como una serie de deducciones formadas por reglas mecánicas, entendiendo su significado solo en el nivel *metamatemático*[126].

Cerca de 1928, los matemáticos Hilbert y Ackermann sacaron el libro «*Grundzüge der theoretischen Logik*», donde formulaban dos problemas de relevancia profunda (véase [Moo, p. 256] donde viene la referencia del libro): el Problema de Decisión (todo problema matemático bien formulado es resolvible) y el mostrar que la lógica de primer orden es completa (Teorema de Completitud).

[125] En inglés, *satisfiable*.

[126] Es decir, que no importa que signifiquen las demostraciones, pero el nivel metamatemático se preocupa por dichos significados y hasta posiblemente representaciones del mundo real.

La mente prominente y capaz de Kurt Gödel logró resolver éste último, en su tesis doctoral de 1930 (referencia en [Moo, p. 256]), primero discerniendo entre la consistencia y el hecho de la «satisfabilidad», estableciendo que en general, estos dos conceptos no son equivalentes. El hecho de que no sean equivalentes lo demostró gracias a su *Teorema de la Incompletitud* (referencia en [Moo, p. 257]), que habla acerca de que si se tiene una teoría axiomática de primer orden que fuera tan útil como para expresar la aritmética de los números naturales, se tendría la existencia de una sentencia verdadera en la teoría que no pudiese ser deducida por algún axioma de la teoría; el resultado de su *Segundo Teorema de la Incompletitud* establece que ninguna teoría puede demostrar su propia consistencia.

Además, Gödel demostró una forma para caracterizar la satisfabilidad general según Moore escribe en [Moo, p. 257]:

Teorema 2.5.26 (Teorema de Compacidad)**.** *Un conjunto denumerable de sentencias de primer orden es satisfable ssi todo su subconjunto finito es satisfable.*

Luego de eso, demostró que todo conjunto contable de sentencias de primer orden tiene un modelo.

Lo limitado de estos resultados lo eliminó Anatolii Malcev, ya que los demostró para conjuntos incontables[127], generalizó el Teorema de Löwenheim-Skolem, el Teorema de Compacidad, y relacionó la lógica de primer orden con la proposicional:

(2.5.1.1) Para todo conjunto S de sentencias de primer orden existe un conjunto T de sentencias de lógica proposicional tales que S es satisfable en lógica de primer orden ssi T es satisfable en lógica proposicional.

La Lógica Matemática emergente no solo se creó como una rama que crecería por su propio camino, también se introdujo en la axiomatización de la Teoría de Conjuntos para lograr el sueño de muchos: crear una base para las matemáticas. Ahí es donde destacan Fraenkel, Skolem y Gödel, que empiezan sugiriendo cambios en la estructura, la adición de axiomas nuevos, y dando aportaciones a la validez lógica del sistema de Zermelo. De los primeros aportes, ellos tres dotaron a la Teoría de Conjuntos de lógica de primer orden, en parte por la difusa idea de propiedad «definida» en la que recaía el Axioma de Separación de Zermelo, en parte por la consistencia misma de éste sistema.

Durante el Quinto Congreso de Matemáticos Escandinavos de 1922, Skolem basó sus trabajos en lecturas de ahí que le darían la idea para definir lo que sería *propiedad definida*. Al siguiente año (véase [Moo, p. 261]), escribió que una propiedad definida era una proposición que era de la forma $x \in y$, de la forma $x = y$ o era obtenida de éstas mediante el uso finito de la negación, conjunción, disjunción, cuantificación existencial sobre individuos, o la substitución de una constante por una variable. Como hemos visto, dicha cuantificación hizo al Axioma de Separación un esquema axiomático de lógica de primer orden.

Otro planteamiento dudoso del sistema de Zermelo, y uno que Dimitry Mirimanoff en 1917 consideró el más importante de la Teoría de Conjuntos, era: ¿existen condiciones necesarias y suficientes para garantizar la existencia de un conjunto?. Éste problema lo llevó a considerar sucesiones infinitas de conjuntos de la forma

$$\cdots \in A_3 \in A_2 \in A_1 \in A_0,$$

Al primer elemento A_0 de una sucesión así, le dio el nombre de *conjunto extraordinario*, y a los demás de *conjuntos ordinarios*. Lo que ésto trajo consigo, fue la definición de *rango*, una idea parecida a la de Zermelo sobre jerarquía cumulativa de clase[128] (o cumulative type hierarchy): el rango de los

[127] Fue el uso de conjuntos incontables que relacionó a la lógica con el Axioma, inevitablemente, ya que varias proposiciones son equivalentes a éste.

[128] Véase la p. 77.

átomos y el vacío era cero, y el de un conjunto ordinario A lo definió como el menor ordinal mayor que el rango de todos los elementos de A, con esto, todo conjunto ordinario estaba dotado de un rango, y la clase de todos los conjuntos ordinarios de rango dado, sería un conjunto, respondiendo en parte el problema dado: una colección de conjuntos ordinarios es un conjunto ssi existe un ordinal mayor que el rango de cada conjunto de la colección, esto debido a que tal existencia diría que la colección estaría en otro conjunto y suponiendo que la clase de todos los átomos es un conjunto. Como conclusión afirmó que: si una colección de conjuntos ordinarios es equipotente a un conjunto, la colección será un conjunto ordinario.

La siguiente proposición a considerar, es parecida al Axioma de Reemplazo, aunque éste en verdad fue explicitado por Fraenkel y Skolem (independientemente), y añadido como postulado al sistema de Zermelo. Otra razón de la carencia del sistema de Zermelo, era que el conjunto que contiene a los naturales $\mathbb{N}$ como elemento, y a su conjunto potencia, y así *ad infinitum*, no puede ser establecido con los axiomas de Zermelo.

Moore escribe en [Moo, p. 263] el axioma de Skolem:

Axioma 2.5.3 (Skolem)**.** *Supón que $A(x,z)$ es una propiedad definida y que para todo x, existe a lo más un z haciendo $A(x,z)$ verdadera, entonces para todo conjunto M existe un conjunto M_A tal que $z \in M_A \Leftrightarrow \exists x \in M$ para el cuál $A(x,z)$ es verdadera,*

mientras que Frankel lo postuló como:

Axioma 2.5.4 (Frankel)**.** *Si M es un conjunto y M' es obtenido al reemplazar cada elemento de M por algún objeto del dominio, M' seguirá siendo un conjunto.*

El matemático John von Newmann quería hacer unas revisiones al axioma de Fraenkel, y establecer algunos criterios que la Teoría Axiomática debería de cumplir y tener:

1. Los conjuntos serían obtenidos de las nociones primitivas de función y argumento.

2. El Axioma de Reemplazo de Fraenkel sería mejorado.

3. Las colecciones «muy grandes» serían permitidas, pero no estarían como elementos de algún conjunto.

Como lo requirió, von Neumann creó su propia axiomatización de la Teoría de Conjuntos, publicándola con el nombre de «*Eine Axiomatisierung der Mengenlehre*» en 1925 (referencia en [Moo, p. 264]), dando también la condición necesaria y suficiente para que una colección sea muy grande (que sea una clase propia):

(2.5.1.2) Una colección es una clase propia ssi es equipotente a la clase de todos los conjuntos.

Una nota importante es el hecho de que ésta condición implica el Axioma de Reemplazo y el Axioma de Separación, como lo demostró von Neumann, al igual que el Principio del Buen Orden y el Axioma de Elección Global, concluyendo que:

(2.5.1.3) La clase de todos los conjuntos puede ser bien ordenada.

Pero el problema ya no era la limitación de trabajar con modelos contables, sino, que recaía en trabajar con lógica de primer orden, y a pesar de que, cualquier teoría lógica de primer orden en la cual se pueda deducir la existencia de conjuntos incontables no podría ser categórica, fue Zermelo quien supo distinguir las diferencias del modelo al trabajar ahora con lógica de segundo orden, y durante el periodo de 1929 a 1935, aportó al tema varias cosas. Pero había algo más dentro de su propósito, y era el hecho de que no solo por el Axioma de Elección su sistema no era aceptado, sino que como dijimos, el Axioma de Separación no establecía precisamente la noción de propiedad «definida», y como lo hizo con su axiomatización para rescatar su demostración del Principio del Buen Orden, no se quedó atras e intentó salvar también a su noción de propiedad definida.

Para axiomatizar tal noción, según Moore en [Moo, p. 268] Zermelo estableció que un sistema axiomático tuviera sus propias relaciones primitivas, y si se consideran a P y Q proposiciones definidas, se considerarían *definidas*:

1. las relaciones primitivas para cualquier valor de las variables,
2. $\neg P$,
3. $P \wedge Q$,
4. $P \vee Q$,
5. $\forall P(x)$ siempre que $P(x)$ estuviera definida para todo valor de la variable libre individual x,
6. $\exists P(x)$ siempre que $P(x)$ estuviera definida para todo valor de la variable libre individual x,
7. $\forall f P(f)$ siempre que $P(f)$ estuviera definida para toda función proposicional f que tiene solamente variables individuales,
8. $\exists f P(f)$ siempre que $P(f)$ estuviera definida para toda función proposicional f que tiene solamente variables individuales.

De todas esas reglas, se denominaría la clase de las *proposiciones definidas* como la intersección de las clases de proposiciones cerradas bajo estas las operaciones lógicas[129].

Para 1930 Zermelo creó la teoría axiomática que llamó Teoría de Conjuntos de Zermelo-Fraenkel, parecida a la dada en la sec. 1.8, donde postulaba 7 axiomas y permitía la existencia de átomos: el Axioma de Extensionalidad, el Axioma de Potencia, el Axioma de Unión (estos eran de su primera axiomatización), el Axioma de Par, el Axioma de Separación (estos también pero estaban modificados), el Axioma de Reemplazo y el Axioma de Fundación.

El Axioma de Separación lo formuló: si se tiene un conjunto X y una función proposicional de primer o segundo orden $P(x)$, será un conjunto $\{x \in X | P(x)\}$; el de Reemplazo como: si $y = P(x)$ es una función proposicional de primer o segundo orden, con cada $P(x)$ conjunto o átomo, entonces será un conjunto $\{P(x) | x \in X\}$; el de Fundación lo estableció de la forma: no hay una $\in$-sucesión infinita y decreciente.

Aunque estableció de esa forma el Axioma de Fundación, y dijo que era equivalente a éste otro: todo conjunto no vacío tiene un elemento $\in$-minimal (éste es el Axioma de Regularidad de la p. 27), uno que ningún miembro de él esté en el conjunto, se demostró después que se requiere el Axioma para tal equivalencia.

[129]Que un espacio sea cerrado bajo alguna operación, es que al aplicarle a cualesquiera elementos del espacio la operación, el elemento resultante esté en el espacio.

El Axioma de Elección y de Infinitud no los agregó, el último porque creía no pertenecía a la Teoría de Conjuntos general, y el segundo porque no serviría para delimitar modelos en Teoría de Conjuntos, aunque lo tomó como un principio de lógica general.

Estableciendo la caracterización salvo isomorfismos de cualquier modelo *estándar transitivo*[130] o *dominio normal* $\mathfrak{M}$ de ZF mediante dos cardinales: el del conjunto de todos los átomos (la *base* del modelo) y el menor ordinal mayor que todos los ordinales en $\mathfrak{M}$ (la característica de $\mathfrak{M}$), Zermelo definió la *jerarquía cumulativa de clase*. Se define recursivamente para un modelo $\mathfrak{M}$ en ZF con característica κ:

$$P_1 = Q_0$$

con Q_0 el conjunto de todos los átomos,

$$P_{\alpha+1} = Q_\alpha \cup Q_\alpha,$$
$$P_\beta = \bigcup_{\alpha<\beta} P_\alpha,$$

donde β es todo el ordinal límite menor que κ, y habiendo probado antes que $\mathfrak{M}$ puede ser descompuesto en κ niveles disjuntos llamados Q_α, consistiendo éstos (excepto por Q_0) en todos los conjuntos X que no están en niveles anteriores pero que sus elementos si. A los P_α se les llama *niveles cumulativos*.

A pesar del intento de Zermelo por salvar su sistema, pocos usaron la lógica de segundo orden y emplearon la de primero. Con esto en mente, Bernays modificó la Teoría Axiomática de von Neumann en un libro de 3 que se publicarían (referencia en [Moo, p. 272]), utilizó lógica de primer orden e hizo que se pareciera más al sistema de Zermelo (dando más importancia al concepto de conjunto que al de función).

Los que intentaron axiomatizar la Teoría de Conjuntos, tenían la disyuntiva sobre la validez de la *independencia* del Axioma, que debían resolver si querían ser escuchados por la comunidad. La independencia de una proposición se refiere al hecho de que en un sistema axiomático, la proposición no pueda ser deducida, que no sea un teorema en ese sistema, lo que es equivalente a que el sistema concatenado con la negación de la proposición, sea consistente.

Pasaba el año de 1922 cuando se publicó el artículo de Fraenkel en una revista matemática (referencia en [Moo, p. 273]), que trataba sobre la independencia del Axioma de Separación, de Elección y de Extensionalidad. La idea es, utilizando los axiomas restantes, encontrar un *dominio* en el que no se satisfaga el axioma que se desea mostrar independiente, es decir, que se encuentre un sistema junto con la negación del axioma. El argumento que dio para demostrar la independencia del Axioma de Elección, es tomar el conjunto vacío y el conjunto ω, con ellos, se genera un dominio con las uniones, y los conjuntos potencia y sacando parejas no ordenadas (utilizando el Axioma de Unión, de Potencia y de Par) entre ellos un número cualquiera finito de veces. Con ello, consideró una sucesión de conjuntos no vacíos X_1, X_2, X_3, $\ldots$ del dominio, que cumpliera ser disjunta y que toda propiedad definida verdadera para al menos algún elemento de la sucesión (algún X_i), fuera verdadera para todos los elementos de otro conjunto de la sucesión (los $x \in X_j$). Se sigue que al tener la familia no vacía de conjuntos no vacíos X_1, X_2, X_3, $\ldots$, esa familia no tendrá función de elección, debido a que si la tuviera, sería una propiedad que tomaría un elemento de un conjunto, pero debería cumplirse para todos los demás conjuntos, lo cuál no pasa porque son disjuntos los conjuntos. La única cuestión que faltaba por resolver, era la existencia de dicha familia, y para ello ocupó un conjunto denumerable de átomos, como los elementos de la familia.

[130]Modelo estándar de la Teoría de Conjuntos es el que la relación de pertenencia es $\in$. Modelo transitivo $\mathfrak{M}$ es el que tiene la característica de conjunto transitivo, i.e., si $x \in y \wedge y \in \mathfrak{M} \Rightarrow x \in \mathfrak{M}$.

El mismo año publicó un libro cuya referencia esta en [Moo, p. 274], donde mejora su argumento y prueba la independencia del Axioma Denumerable, mediante el uso de denumerables átomos, los cuales serían diferentes y vendrían en pares: $X = \{\{a_1, a_1^1\}, \{a_2, a_2^2\}, \ldots\}$, con X la familia requerida. Con esto, demostró la independencia del *Axioma Denumerable restringido a familias de conjuntos de pares no ordenados*.

Este Axioma se enuncia así:

Axioma 2.5.5. *Para toda familia denumerable de conjuntos que son pares no ordenados, $X = \{\{a,b\}, \{a_1,b_1\}, \ldots\}$, existe una función de elección.*

Dieron un aporte significativo a pruebas de independencia, Adolf Lindenbaum y Andrzej Mostowski[131], mejorando técnicas de Fraenkel, debido a que argüían que la construcción de objetos llevada por Fraenkel, era imprecisa por la vaguedad con que utilizaba conceptos en el lenguaje matemático y en el metamatemático. Sacaron un escrito en 1938 con 9 proposiciones que eran demostradas independientes en ZF (referencia del escrito en [Moo, p. 276]): 5 sobre equivalencias del concepto de finito[132], una sobre cardinales, una sobre orden, el Axioma Denumerable restringido a pares no ordenados y el Principio de Partición. Ya para 1948 fueron los que demostraron que el Axioma de Elecciones Dependientes no implica al Axioma de Elección.

También demostraron la *consistencia relativa* de la Teoría de Conjuntos con átomos en esa publicación (añadir átomos a la teoría no genera contradicciones, siempre y cuando la misma teoría no sea contradictoria). Vieron la diferencia entre el sistema axiomático U_0 y el sistema con el que se trabaja éste (metamatemáticamente), llamado U, aceptaron a U_0 como el sistema de Bernays con átomos, y a U como el de von Neumann; al suponer que éste ultimo es consistente, produjeron una *interpretación* del sistema U_0 con él, cambiando la interpretación (debido al sistema metamatemático) de los átomos como conjuntos en U, estableciendo así la consistencia relativa con átomos.

En esas fechas, ya había demostrado Gödel que la consistencia del sistema de von Neuman sin el Axioma de Elección como axioma, implicaba la consistencia del sistema con el Axioma, es decir, ya había probado su consistencia relativa. Algunos libros se refieren a esto de forma equivalente y tal vez más educativa: supón que con la teoría de ZF (con el Axioma) se encuentra una contradicción, entonces también se encontrará una contradicción en la teoría ZF sin el Axioma.

Pero para las pruebas de independencia, se asumía la consistencia del sistema axiomático base, sin aún poder demostrar la consistencia del sistema de Zermelo o de von Neumann, hasta que Gödel en 1931 tuvo éxito en demostrar que los axiomas de los números naturales no podrían demostrar su propia consistencia con la lógica de primer orden, esto está en [Moo, p. 279]. Se vuelve un círculo sin fin tal situación: como el sistema de los números naturales está encajado en la teoría de ZF, y no solo en ella, sino en cualquier teoría que sea lo suficientemente útil para contener a los naturales[133], éstas tampoco podrán demostrar su consistencia (ya que una parte de éstas no lo es), y para demostrar la consistencia de ZF o las otras teorías se requeriría de un sistema axiomático más fuerte, en el cuál pasaría los mismo[134].

Un ejemplo que se da en [Moo, p. 279] es el siguiente: sabiendo que Zermelo había demostrado que la jerarquía cumulativa de clases[135] hasta el primer ordinal inaccesible fuertemente es un modelo de ZF, supongamos que la proposición IN denota la existencia de un ordinal fuertemente inaccesible, entonces

[131] Éste viajó a Zurich para convertirse en Actuario, pero se aburrió con esas matemáticas tan triviales y regresó a Varsovia.

[132] Como que no todo conjunto D-finito es finito en ZF.

[133] Para teorías axiomáticas, casi es una condición ésta, debido a que de los naturales se puede deducir mucha de la matemática que se ocupa.

[134] Véase la p. 74

[135] Véase la p. 77.

pasa a ser teorema la consistencia de ZF en el sistema nuevo y más fuerte ZF+IN, sin embargo, no se ha probado la existencia de dicho ordinal (en esa teoría, no hablo metamatemáticamente de: en general).

Como consuelo, se tiene solo la *consistencia relativa* para demostrar la consistencia de un axioma. La idea es construir un modelo para la Teoría de Conjuntos, y la proposición dentro de un modelo dado de la Teoría de Conjuntos (construir un *modelo interno*). Una aplicación de ésto la dio Gödel en el año 1938, en su libro prominente y decisivo para el Axioma «*The Consistency of the Axiom of Choice and of the Generalized Continuum Hypothesis*» (véase [Moo, p. 280]), demostrando la consistencia relativa del Axioma de Elección y Hipótesis del Continuo Generalizada mediante la herramienta conocida como «conjuntos construibles» del sistema axiomático mejorado por él; también demostró que la negación del Axioma y la negación de la HC son ambas, independientes de su mismo sistema. La consistencia relativa del Axioma la dedujo, al construir un modelo en el sistema de von Neumann con lógica de primer orden, mostrando que la clase de conjuntos construibles era un modelo interno del sistema que también satisfacía el Axioma de Elección, este sistema era muy parecido a la jerarquía cumulativa de conjuntos ya vista antes.

Pero la característica a remarcar de éste modelo es que satisfacía el:

Axioma 2.5.6 (Axioma de Constructibilidad)**.** *Todo conjunto es construible.*

Recalcó la importancia de su proposición, refiriéndose a la determinación que le da a los conceptos no precisos como los de conjuntos infinitos, ya que los vuelve construibles (véase [Moo, p. 280]). Los resultados los trasladó al sistema axiomático ZF al siguiente año (referencia del libro en [Moo, p. 281]), para ello, definió al conjunto X' como el de todos los subconjuntos de X definibles por una función proposicional de primer orden, permitiendo solo el símbolo de relación $\in$ en dicha función y además las jerarquías de conjuntos construibles:

$$\begin{gathered} X_0 = 0, \\ X_{\alpha+1} = (X_\alpha)', \text{ y} \\ X_\beta = \bigcup_{\alpha<\beta} X_\alpha \end{gathered}$$

cuando β es un ordinal límite. Éstos son los conjuntos construibles a los que nos referimos de Gödel, llamando a un conjunto Y *construible de orden* α si $Y \in X_{\alpha+1} - X_\alpha$.

Gödel propuso una nueva axiomatización de la teoría de Bernays, y también dio una equivalencia a sus conjuntos construibles que los creaba mediante 8 operaciones elementales sobre conjuntos, y con recursión transfinita (referencia del libro en [Moo, p. 282]). Creando con esto una teoría más sólida, aunque como él demostró, no podría demostrarse su consistencia en un sistema menos fuerte que ese mismo.

Con los grandes avances que hizo Gödel, una etapa de dudas y búsquedas que no tendrían fin, terminó, sus teoremas de Incompletez hicieron entender a la comunidad, que los sistemas con los que estudiaban, serían incompletos o inconsistentes.

La Teoría de Conjuntos Descriptiva es una de las ramas de la Teoría de conjuntos que se basa principalmente en la definibilidad y clasificación de conjuntos de números reales y también de espacios Polacos[136] (Polish Spaces). Dado un conjunto de números reales, se intenta ver si se puede dar una definición explícita en un lenguaje formal, y que tan «complicado» es hacer esto, entendiéndose por cómo se construye el conjunto a partir de operaciones de la teoría de conjuntos y de unos conjuntos más simples[137].

[136] Espacios separables, completamente regulares y metrizables.

[137] Shoenfiel definió a la *definibilidad* de un conjunto X como la complejidad de los cuantificadores que se necesitan para definir a X.

Por ejemplo, los conjuntos Borelianos, se obtienen a partir de una σ-álgebra generada por un conjunto, y se pueden especificar esos Borelianos viendo cuántas operaciones de complementos y uniones fueron necesarias para crearlos; con esto, podríamos dotar de niveles a la familia de conjuntos Borelianos, y asignarles una jerarquía que dependa de las operaciones para crearlos de los conjuntos abiertos simples. También se puede ver la definición más simple de un conjunto dependiendo de cuántos cuantificadores fueron necesarios para crearlo.

2.5.2. Sobre definibilidad

En su tesis para doctor en Filosofía, John Daniel Clemens [Cle, p. 2], habla sobre las consideraciones pertinentes a la Teoría de Conjuntos Descriptiva, entre esas están la de *definibilidad* de conjuntos, dice que por lo menos se requiere que el conjunto sea *bien portado* (well-behaved), ya que de los conjuntos que no son definibles explícitamente, como los dados por el Axioma (conjuntos no Lebesgue medibles, conjuntos no Borelianos, etc.), no se pueden dar consideraciones generales. Habla acerca de que los conjuntos mejor definibles explícitamente y bien portados son los Borelianos, también son bien definibles los *conjuntos proyectivos*[138], los conjuntos de Baire, los conjuntos medibles.

Claro que dependiendo el modelo que trabajemos resulta la definibilidad de los conjuntos. Considerando hipótesis más fuertes que las de ZFC, se obtienen clases más grandes, una de éstas es la forma natural de definibilidad de reales que nos da el modelo $L(\mathbb{R})$, que son esos conjuntos construibles a partir de parámetros reales.

Parece que para Russell, no afectaron sus pensamientos filosóficos para aceptar o no el Axioma, como escribe Moore citándolo:

> *«All philosophical differences ... ought not to affect the detail of mathematics, but only the interpretation. Mathematics would be in a bad way if it could not procees until the dispute between idealism and realism has been settled. When a new entity is introduced, Dr. Hobson regards the entity as* created *by the activity of the mind, while I regard it as merely* discerned*; but this difference of interpretation can hardly affect the question whether the introduction of an entity is legitimate or not, wich is the only question with which mathematics, as opposed to philosofy, is concerned.»* [Moo, p. 132]

Zermelo creía que si un objeto A caía o no bajo una definición de un concepto, debía ser decidido por un criterio objetivo, independientemente del concepto definido.

Pensemos ahora en el conjunto A de todos los conjuntos cuyos elementos son números naturales. No podemos asignar un contenido metamatemático a esta definición. Nos encontramos con este fenómeno: sabemos lo que es el conjunto de los números pares, el de los números primos, el de las potencias de dos, e infinitos más, pero no tenemos alguna definición precisa de lo que es un conjunto de números naturales en abstracto, ni tenemos, en particular, representación alguna de la totalidad de tales conjuntos. Todas las contradicciones de la teoría de conjuntos surgen de la pretensión de hablar de colecciones de objetos en sentido abstracto como si supiéramos de qué estamos hablando.

[138] La definición de conjunto proyectivo está en la p. 83.

2.6. LA ÚLTIMA ETAPA

Varias de las aseveraciones[139] que dieron matemáticos durante las primeras décadas de vida del Axioma, fueron demostradas en ésta etapa, y también se intensificó la búsqueda de la equivalencia del Axioma con resultados ya obtenidos, para encontrar la necesidad matemática del mismo.

En 1953, concerniente a Álgebra, Jürgen Schmidt demostró equivalencias al Axioma. Primero terminología: se dice que un álgebra $\mathfrak{U} = (A, \mathfrak{F})$ es un par ordenado, con $\mathfrak{F}$ un familia de funciones $f : A^n \mapsto A$, para valores especificados de n, y A un conjunto. Una subálgebra de $\mathfrak{U}$ será un par ordenado $(B, \mathfrak{F}')$ con $B \subset A$ y $\mathfrak{F}'$ la familia de todas las funciones en $\mathfrak{F}$ restringidas a B, i.e., las funciones $f|_B$. Se le llamará *base* de $\mathfrak{U}$ a $B \subset A$ si la única subálgebra que incluye a B es $\mathfrak{U}$ y además si cualquier subconjunto propio de B está incluido en una subálgebra propia de $\mathfrak{U}$.

Con ello, Schmidt demostró:

(2.6.1) Si $\mathfrak{U}$ es un Álgebra con una base finita y $\mathfrak{T}$ una subálgebra propia de $\mathfrak{U}$, existirá una subálgebra propia de $\mathfrak{U}$ que incluye a $\mathfrak{T}$ y además es $\subset$-maximal.

Una demostración relacionada con la HGC, fue la de Sierpinski: si no existe un cardinal infinito $\mathfrak{m}$ que cumpla con que $\mathfrak{n} < \mathfrak{m} < 2^{\mathfrak{n}}$, entonces se cumple el Axioma de Elección.

Después demostró que la afirmación vista en la página 67, es decir, el hecho de que la desigualdad equivalente de cardinales $\leq^*$ implique que no pase la contraria en la desigualdad normal $>$, implica que existe un subconjunto no medible de $\mathbb{R}$ y además que $\aleph_1 \leq 2^{\aleph_0}$.

Para crear un trabajo que tenga resultados parecidos a los del Axioma, pero que sea más intuitivo y débil, crearon W. Kinna y K. Wagner en 1955 (referencia del libro en [Moo, p. 298]), esta proposición:

Proposición 2.6.1 (Principio de Kinna-Wagner)**.** *Para todo conjunto X existe una función f con $f(A)$ un subconjunto propio y no vacío de A, para todo $A \subset X$ con más de un elemento.*

Con esta proposición (que necesitaba la técnica de forcing para mostrar que no implica al Axioma), demostraron esta otra:

Proposición 2.6.2. *Para todo conjunto X existe un ordinal α tal que su conjunto potencia tiene un subconjunto equipotente a X.*

En esa época, otro intento de hacer lo mismo apareció, y escribo intento ya que se demostró un año después, que es equivalente al Axioma; nos referimos a esta serie de proposiciones donde cada n da una diferente:

Proposición 2.6.3. *Para toda familia X de conjuntos no vacíos existe una función f que hace a $f(A)$ un subconjunto no vacío de A de a lo más n elementos, para todo $A \in X$.*

Tal equivalencia con el Axioma y esta proposición, se cumple para todos los naturales (cada natural, asigna una equivalencia), demostrado por A. Levy (referencia del libro en [Moo, p. 299]).

Así como las varias definiciones de conjunto finito de Dedekind, Cantor y Tarski, la Teoría de Cardinales tiene su análogo. La teoría cuenta con tres definiciones, que se refieren cada una a que $\mathfrak{m}$ tiene un sucesor inmediato $\mathfrak{n}$ respecto a su perspectiva:

[139] Formalmente se les llama *conjeturas*, que son afirmaciones las que no se sabe si son verdaderas o falsas, debido a que no se han demostrado.

Definición 2.6.1. *Dicho* $\mathfrak{n}$ *cumple con que:*
$\mathfrak{m} < \mathfrak{n}$,
y no existe $\mathfrak{k}$ *que cumpla* $\mathfrak{m} < \mathfrak{k} < \mathfrak{n}$.

Definición 2.6.2. *Dicho* $\mathfrak{n}$ *cumple con que:*
$\mathfrak{m} < \mathfrak{n}$,
y para todo $\mathfrak{k}$ *que cumpla* $\mathfrak{m} < \mathfrak{k}$*, se tendrá* $\mathfrak{n} \leq \mathfrak{k}$.

Definición 2.6.3. *Dicho* $\mathfrak{n}$ *cumple con que:*
$\mathfrak{m} < \mathfrak{n}$,
y para todo $\mathfrak{k} < \mathfrak{n}$ *se tendrá* $\mathfrak{k} \leq \mathfrak{m}$.

La primera se puede deducir en ZF, la segunda es equivalente al Axioma, y para la tercera demostró Cohen que no se puede deducir de ZF, lo hizo utilizando su método llamado forcing.

Ernst Specker también intentó dar proposiciones menos fuertes que el Axioma, propuso como alternativas:

Proposición 2.6.4. $\mathbb{R}$ *es una unión contable de conjuntos contables*[140].

Dando además como generalización:

Proposición 2.6.5. *El conjunto potencia de cualquier conjunto de cardinalidad* $\mathfrak{n}$ *es la unión de un número infinito contable de conjuntos con cardinalidad* $\mathfrak{n}$.

Con esta proposición aceptada, él logró demostrar que la consistencia de ésta implica la existencia de un cardinal inaccesible, concluyendo que la existencia de éstos, no es meramente materia del Axioma, esto ya que en otros universos donde la negación del Axioma o algunas proposiciones más débiles son ciertas, dichos cardinales existen, y por lo tanto, abriendo las puertas para la *Teoría de Cardinales Largos*.

Con la posibilidad del estudio de cardinales inaccesibles, se trabaja un axioma alternativo al Axioma de Elección, llamado *Axioma de Determinancia*.

Dentro de los métodos de independencia conocidos antes de 1963, varios ocupaban átomos y no lograban demostrar la independencia de varias proposiciones para $\mathbb{R}$, y los que lograron evitar los átomos, aún tenían deficiencias en el dominio de su aplicación. El que se denomina *forcing* es de los más conocidos, y traé respuesta a varias limitaciones sobre pruebas de independencia, siendo amplio su campo de aplicación. Paul Cohen fue el que lo desarrolló públicamente en 1963, y no necesitaba aplicar en el método átomos, así que pudo demostrar la independencia de la Hipótesis del Continuo y además del Axioma del sistema ZF, cambiando la perspectiva del Axioma para lados más metamatemáticos (respecto a su investigación, no tanto a su aplicación) e iniciando un periodo de investigación sobre pruebas de independencia.

Gödel dijo que la técnica de forcing es un método para hacer afirmaciones verdaderas acerca de una cosa sobre la que no conocemos algo.

Los diferentes puntos de vista sobre el método de Forcing son[Avi, p. 2]: Para el que trabaja en teoría de conjuntos, el método Forcing provee maneras de extender un modelo de teoría de conjuntos al añadir un objeto "genérico", de tal manera que la verdad en la extensión es determinada por aproximaciones del objeto genérico que vive en el modelo original. Para el que trabaja en lógica modal, el método provee maneras para que podamos explicar la noción de verdad necesaria, o verdad en todos los posibles mundos. Forcing también provee las semánticas para la lógica intuisionista basada en una noción de información

[140]Como se ve en la subsec. 3.7.5, para demostrar la negación de ésto se requiere del Axioma, así que es una alternativa a una negación del Axioma la proposición de Specker.

parcial, o estados de conocimiento en el tiempo. Para el que trabaja en teoría de la recursión, provee una manera conveniente de describir contrucciones en las cuáles una secuencia de condiciones se satisface cada una a la vez. Para el que trabaja en teoría de modelos, es una construcción que provee un conveniente modelo genérico de cualquier teoría inductiva $(\forall\exists)$. Del punto de vista de "Sheaf Theory", el método provee una manera de describir la lógica interna de un "topos". Para el que trabaja en teoría de conjuntos descriptiva, provee una manera de decir que significa para una propiedad ser "genéricamente verdadera" en un espacio Polaco.

Aunque no fue lógico matemático ni estudió Teoría de Conjuntos, sino que fue Analista, hizo aportaciones a estas dos ramas que trascendieron en la historia. En su artículo (véase [Moo, p. 302]), él demostró la existencia de un modelo mínimo de la Teoría de Conjuntos, allí, no hay una fórmula de primer orden de ZF que pueda definir un modelo en ZF para ZF y la negación del Axioma de Constructibilidad. Debido a que este axioma implica la HGC y al Axioma, no hay un modelo en ZF que no cumpliera con alguno de esos tres axiomas (Constructibilidad, HGC y Axioma de Elección). En su siguiente artículo dedujo que existen modelos de ZF en los que: el Axioma y la HGC se cumple y algún subconjunto de $\mathbb{N}$ no es construible; no puede ser bien ordenado $\mathbb{R}$; el Axioma se cumple, pero la HGC no; el Axioma Denumerable no se cumple para alguna familia de pares de funciones reales. Con ellos, concluyó que sus demostraciones pueden ser llevadas a la Teoría de Números de primer orden, para demostrar que la consistencia del sistema ZF implicaba la consistencia de ZF y alguna de estas 4 proposiciones anteriores. Las demostraciones que hizo, dependían fuertemente de su método Forcing, que en cierto sentido, fuerza a un modelo a tener las propiedades que uno requiere.

Mediante la técnica forcing, se establecieron independencias como: con asumir la existencia de un modelo de ZF con un ordinal inaccesible, se puede obtener un modelo de ZF, el Axioma de Elecciones Dependientes y la proposición de que todo subconjunto de los reales es medible según Lebesgue, tiene la propiedad de Baire, y si es incontable, tendrá un subconjunto perfecto, ésto demostrado por Solovay, además, bajo las mismas hipótesis, Levy concluyó que existe un modelo de ZFC, en el que la HGC es válida, junto con la proposición de que todo subconjunto de $\mathbb{R}$ que es *proyectivo*[141], será Lebesgue medible, tendrá la propiedad de Baire y además, si es incontable, tendrá un subconjunto perfecto.

Dentro de la categoría de las alternativas sobre el Axioma, y bastante productiva, entra la que dieron Jan Mycielski y Hugo Steinhaus por 1962 (referencia en [Moo, p. 306]), llamado el *Axioma de Determinancia* (o *Axiom of Deteminacy* o *Axiom of Determinateness*), que asegura que un *juego*[142] es *determinado*, cuando se tiene un conjunto de sucesiones denumerables de 0´s y 1´s y se tienen dos *jugadores* que escogen términos de 0´s y 1´s alternativamente formando una sucesión infinita. La materia es *juegos posicionales infinitos*, y el Axioma de Determinancia concluye que existe una estrategia que siempre asegurará que la secuencia de los jugadores estará o no en el conjunto de sucesiones denumerables.

Este axioma, tiene como consecuencias el hecho de que todo subconjunto de $\mathbb{R}$ es medible según Lebesgue, así como la aditividad contable de la medida de Lebesgue, que la continuidad secuencial y la continuidad de funciones reales es equivalente, el Teorema de la Unión Contable de la p. 39 para subconjuntos de $\mathbb{R}^n$, el Axioma Denumerable para subconjuntos de $\mathbb{R}$, la regularidad[143] de $\aleph_1$, y otros más.

---

[141] Un conjunto proyectivo X es uno tal que para toda función $f : X \mapsto Y$ y toda función suprayectiva $g : Z \mapsto Y$ existe una función $k : X \mapsto Z$ tal que $f = g \circ k$.

[142] Las definiciones sobre qué es un juego, y estos conceptos están en [Her, p. 137].

[143] Término visto en la p. 67.

2.7. DIFERENTES PUNTOS DE VISTA

Any argument where one supposes an arbitrary choice to be made a non-denumerably infinite number of times... is outside the domain of mathematics.
Emile Borel, tomado de [Moo, p. 85]

Para cuando Cantor publicó en 1874 una nueva demostración del teorema de Liouville de que existen infinitos números trascendentes, hubo una transformación respecto a demostrar que un objeto matemático «existe». En tal demostración Cantor no definió o construyó tales números, sino que, al contrario, asumiendo que no existen, uno es llevado a una contradicción. Por el lado de las matemáticas que se remontan desde los tiempos de Euclides, hasta épocas de Cantor, la existencia de una objeto matemático con una propiedad dada se establecía construyendo un ejemplo particular de él; y aunque Euclides también utilizó la técnica que utilizó Cantor, una *demostración indirecta*, para mostrar que un objeto existe, la diferencia, que no está en la demostración del teorema de Liouville pero si en las otras demostraciones que ocupan el Axioma, es que la demostración indirecta de Cantor «no» puede ser reemplazada por una demostración constructiva, cambiando así el Axioma el significado de la existencia matemática.

B. Russell escribió sobre el Axioma: «*Numerous mathematicians, like Zermelo himself, assert that this axiom is as obvious as the other axioms and that it can be accepted without hesitation. Others say that there is no reason to believe that the axiom is true. Peano, having proved the independence of the axiom, devotes to the consideration of its truth only the following remark: "Are we to believe now that the proposition is true or that it is false? Our attitude is neutral". Peano maintains in the same paper that the question of obviousness is a psychlogical problem, not concerning logic.*» (sacado de [Moo, p. 176].

De entre los empirístas y los intuisionistas, aclaro unas ideas importantes:

Los empirístas veían a usar al Axioma dentro de una demostración como un método heurístico para resolver problemas.

El principio más importante entre el intuisionismo es la identificación de la existencia matemática con la constructibilidad.

Los Platonistas se permiten trabajar con cualquier tipo de conjuntos, incluyendo conjuntos infinitos.

Es el filosófico *problema de existencia* el que desencadena controversias acerca del Axioma. Podemos mantener la posición de que en las matemáticas, el mostrar que un objeto existe y que tiene cierta propiedad es lo mismo que negar la proposición de que no existe un objeto con la propiedad dada. Los intuisionistas no comparten esta forma de ver la existencia matemática. Es su forma no constructiva y existencial la que hace que los intuisionistas de un tipo radical y conservativo vean al Axioma no importante e incluso inadmisible, debido a que la identificación de la existencia matemática con la construcción es el principio fundamental del intuisionismo, según Sierpinski en [Sie, p. 95].

Una objeción que von Neumann dio contra el intuisionismo de Brouwer, fue el hecho de la categoricidad de la teoría de conjuntos, vista su definición en la página 73, su réplica fue el hecho de que una axiomatización de la noción de conjunto finito es ambigua, debido a que un conjunto infinito en un modelo puede ser finito en otro.

Otra objeción en contra del Axioma era el hecho de la incapacidad del procedimiento humano para realizar las elecciones que se hacen de los subconjuntos, o el bien orden de algún conjunto. Como idea intuitiva de esto último, daba Hausdorff el argumento de escoger un elemento a_0, luego escoger otro a_1, y así sucesivamente hasta agotar todos los elementos del conjunto; a cualquier subconjunto que le saquemos solo buscamos el elemento de menor índice y así tenemos el menor de todos ellos, obteniendo un buen orden del conjunto. La objeción en contra del argumento es la elección de «todos» los elementos del conjunto, ¿cómo saber que serán contados todos?, porque la clase de todos los ordinales no podría ser vaciada y aún así tener elementos sin contar del conjunto dado. Pero, si pasa aunque sea una cantidad muy pequeña de tiempo al escoger cada elemento, ¿se llega a contar siquiera a_ω?, seguro no, debido a que el tiempo de conteo terminaría en infinito, pero Hausdorff aseguraba que se deben abandonar las consideraciones practicas y psicológicas, debido a que el Axioma asegura la elección simultanea de los elementos del conjunto.

Herrlich en [Her, p. 7] escribe lo que comentó M.D. Potter:

> «*The formal consistency of making this assumption (validity of the Axioma of Strong Choice) can hardly be doubted, but it ascribes to us abilities wich I for one am not aware of possessing.*»

Las objeciones seguían siendo claras: no da un método para escoger los elementos que dice que escoge, y tampoco define una función de elección únicamente.

> «*The axiom of choice has many elegant consequences, but that is an argument for its mathematical interest, not for its truth.*»

Las razones por las que Russell no aceptaba el axioma, eran primordialmente por sus concepciones lógicas. En [Moo, p. 176] se escribe sobre lo que Russell dice respecto al Axioma:

> *Peano says... that the question of self-evidence is a psychological question wich does not concern logic. But logic depends on the axioms of logic, and these axioms are accepted because they are self-evident; at least, the self-evidence either of the axioms or of their consequences is the sole reason wich makes us accept the axioms. Thus we may ask if the multiplicative axiom is true... But in my opinion, this is not at all self-evident. Thus I find myself led to the conclusion that the axiom ceases to be self-evident as soon as we grasp what it means.*

Russell y Whitehead, en su libro *Principia Mathematica* (véase [Moo, p. 176]), trataron al Axioma como un postulado de lógica, escribiéndolo de forma hipotética como «si el Axioma Multiplicativo es cierto, entonces P es cierto»[144].

Moore escribe que Poincaré (véase [Moo, p. 177]) habló de dos escuelas de matemáticos, los *Pragmáticos* y los *Cantorianos*, que eran parecidas a las clases de matemáticos que hablamos antes hechas por Lebesgue en la p. 54. Decía Poincaré que los pragmáticos casi no creían en la demostración del Principio del Buen Orden de Zermelo, debido a que no daba una regla expresable en un número finito de palabras, y los Cantorianos si creían en ella.

En contra de los Pragmáticos que cuestionaban el hecho de encajar palabras como *indefinido* e *infinitamente* en definiciones, Borel estableció que no importaba el carácter psicológico, sino el matemático, importaba para Borel la determinación matemática del objeto en cuestión.

Respecto a la objeción contra los Pragmáticos, Borel dio como ejemplo (sacado de [Moo, p. 178]):

> *Of course it is possible to reason about a class of mathematical objets, for example all the real numbers...this class is defined by means of a finite number of words, although not all the members can be defined in that way. Thus one obtains the general properties of that class...*

Hadamard añadió a Borel[145]:

---

[144] Donde P es una proposición.

[145] Esto fue en una carta de Hadamard a Borel que está en el apéndice de [Moo].

> *For indeed, all the members must exist in some way in order to form the class... Zermelo would thus have demonstrated... at least that there exists a (non-empty) class of such orderings (of the continuum)...This means, in sum, that (if Zermelo´s argument is not ultimately perfected) we shall only be able to reason about the properties common to all these orderings... There are so many other things that we shall never know.*

Desde ese punto de vista, Borel decía que se llegaban a conclusiones verbales al no especificar miembros definidos de las clases de las que hablan, estas conclusiones no relacionadas con la realidad.

Incluso Sierpinski creía que el nombre de Axioma de Elección estaba mal empleado, ya que no permitía a alguien escoger efectivamente un elemento, sino esa elección solo era teórica. Por ejemplo, ni siquiera si a un conjunto no vacío A le aplicamos el Axioma, podemos especificar características de un elemento del conjunto, ya que en este caso, el Axioma lo «único» que dice es que existe un elemento $f(A)$, tal que $f(A) \in A$. Un conjunto del que no podemos especificar sus elementos es el de todos los subconjuntos no medibles de $\mathbb{R}$, no es vacío debido al Axioma, pero no se conoce cómo determinar un único elemento, una propiedad que lo separe de los demás.

Pero debido a no poder exhibir elementos de dichos conjuntos, ¿qué significa que una propiedad sea demostrada gracias al Axioma?, para Sierpinski y Luzin, solo demuestra que la propiedad no es falsa, más no demuestra que sea verdadera.

Dice [Moo, p. 242] sobre la abstracción de la topología y el uso del Axioma sin pensar su validez: «*Perhaps topologists sensed that it was inappropiate, when engaged in so abstract an enterprise, to insist on constructive methods. Perhaps the enchanting landscape of abstraction and generalization freed them from any lingering constructivist scrupules.*»

Con la controversia del Axioma, y en gran parte para resolverlo, se dejó que un enfoque lógico entrara en el área matemática. Después de que Gödel afirmó que si los axiomas de NBG sin el Axioma de Elección eran consistentes (aunque se sabe que ésto no puede demostrarse), también lo sería la teoría NBG más la hipótesis del continuo.

Con esto Gödel dio una solución a un problema abatidor. En la teoría de conjuntos había sentencias sobre números naturales indecidibles, y si algo no tan complejo como la aritmética resultaba incompleto, con mayor razón la teoría de conjuntos. Paul Cohen en 1963 terminó de sellar el hueco del problema con su técnica forcing que dejó demostrar la independencia de la HC y el Axioma con la teoría de conjuntos.

El Axioma de Elección y la HC son afirmaciones que no pueden decidirse a partir de los axiomas postulados por NBG sin el Axioma y ZF, en realidad son dos de tantas. Esto resulta «inconveniente» al parecer para muchos matemáticos. Sucede como el shock del descubrimiento de nuevas geometrías, las llamadas «geometrías no euclidianas»; primero eran inconcebibles, luego posibles, sin embargo absurdas, y al final, se cree que una buena postura es que todas las geometrías tienen el mismo «valor matemático», esto debido a que estudiar una y creerla más natural, hoy se sabe es relativo. De la misma forma, existen axiomas que tienen como consecuencia la negación del Axioma, o de la HC, con resultados igual de interesantes y resultados igual de antiintuitivos. Unas matemáticas con el Axioma son un universo de alguien que toma ese modelo por un momento como base, otras matemáticas son las que no tienen al Axioma o tienen su negación, y éstas no son mas o menos interesantes, simplemente son varias matemáticas.

A veces, se suele creer que la noción natural y bien definida de conjunto existe, y esto hace creer que diferentes teorías son erróneas porque la noción debe llevarnos por el camino de la correcta. Pero no es así, se carece de una noción intuitiva de conjunto que corresponda a los axiomas de ZFC, significando que dados dos modelos de ZFC no podríamos decidir cuál está dentro de nuestra noción de conjunto y cuál no. Los resultados que no nos parezcan correctos con una teoría no deben tomarse como tales ni tampoco considerar malas las propiedades deducidas de conjuntos de esta manera, sino como una propiedad de los conjuntos o axiomas que estamos considerando.

Una diferencia de pensamiento cambia la forma de estudiar el mundo que se estudia. Arquímedes llamaba a la circunferencia «la curva plana que describe el extremo de un segmento cuando gira sobre su otro extremo», y esta es una forma constructiva de verla. Por otro lado, Euclides la describe como «el conjunto de puntos en el plano que equidistan de otro», y él lo que hacía era describir. La diferencia no solo es la evidente de entendimiento, sino al construir afirmas la existencia, algo de lo que se puede obtener algo debido a que en tu mundo esta ahí, mientras que cuando describes, solo «evades» en ciertas ocasiones la abstracción y realización del objeto.

En realidad, podría pensarse que la matemática estudia objetos que nosotros nombramos como conjuntos, pero lo que en verdad hacemos es estudiar estructuras llamadas modelos de ZFC o NBG. Cuando el matemático estudia la medibilidad de conjuntos, dice «sea X un conjunto» y si en alguna parte requiere que sea no medible, entonces dice «supongamos que X es no medible, entonces ...», y esto lo hace (o debería hacer) sin plantearse la duda lógico-filosófica de si puede o no ser no medible su conjunto, pero debe tener en cuenta en qué universo esta él trabajando; él solo llega a conclusiones en un modelo, entra a un mundo donde así se dan las cosas, donde la no medibilidad es posible, y de donde puede salir y llegar al polo opuesto y ocupar cuando lo requiera o sea de su agrado el hecho de que no existen conjuntos no medibles.

Pero lejos de las abstracciones, ¿qué hay acerca de los modelos y su aplicación en el mundo real?. Aquí hay que hacer varias anotaciones y aclaraciones. En el mundo real no se encuentran directamente las abstracciones, no hay números imaginarios, ni naturales. Usamos nuestras imágenes mentales, y esos objetos los jalamos de nuestra mente a nuestra realidad, para así ocuparlos como nos plazca. Pero para que esos conceptos tan abstractos como el conjunto de los números naturales tengan sentido y no sean ambiguos en una mente razonablemente entrenada, se da una propiedad que cumplan esos objetos y de ahí creamos una imagen «única» de tal, intentando quitar toda intuición de nuestra definición. Como objeción, alguien podría preguntar ¿qué es una propiedad?, y no importa de qué manera definamos los conceptos en términos de más definiciones, el proceso no puede nunca acabar, y llega un punto en que la abstracción debe tocar la intuición y lo que creemos como razonable, una base que comúnmente se acepta como lógica.

Enderton H. B. aclara en [End, p. 125] que hay dos formas de introducir objetos matemáticos:

1. Objetos nuevos a partir de otros conocidos de antes.

2. Objetos nuevos introducidos como nociones primitivas o como axiomas.

El autor aclara que la segunda no es una forma de responder cuestiones fundamentales acerca de los objetos sino una forma de evadirlas. Para que tengan sentido nuestros conceptos, deben reflejar con precisión lo que queremos que reflejen, aunque esto demanda un enlace entre un concepto matemático y una idea informal, no hay forma de evitar esto. Esto puede causar errores en los modelos, al definir sobre nuestra intuición debemos de tener precisión en lo que queremos hacer. Las líneas rectas de la vida real podrían no parecerse a intervalos de $\mathbb{R}$, hoy sabemos que la mecánica clásica es errónea debido a que en las fórmulas faltan las ecuaciones que Einstein demostró y que ocupan peculiaridades del espacio, como la atracción que hace la masa a la luz, doblándola. Esto requeriría un cambio radical en nuestros conceptos de línea recta y geometría real, sin embargo, puede uno ocupar geometrías no euclídeas u otros modelos para aproximar la teoría al mundo real[146]. En mediciones pequeñas (las que ocupamos en el planeta que habitamos), la mecánica clásica y la cuántica se parecen tanto que se ocupa la más cómoda, pero para mediciones en años-luz u otras astronómicas, la cuántica juega el papel de la más aproximada. En esos

[146] Podrían considerarse geodésicas.

casos, aunque las propiedades que tienen las rectas, las esferas, los planos, siguen siendo verdaderas en la matemática que conocemos, dejan de ser válidas algunas de esas propiedades en esos sistemas, debido a que dejan de ser aplicables, pero señalo, dejan de ser aplicables en «esos» sistemas, y eso no las hace menos interesantes, o falsas, sino falsas en «esos» sistemas.

Los conceptos matemáticos son buenos para resolver problemas de la vida real, si ellos miden aproximadamente las características de esos sistemas. Para esto, se consta de tres partes, con el problema de la vida real en mano y esperando a ser resuelto, hacemos un modelo matemático del problema, encontramos una solución matemática al modelo hecho y al final interpretamos esa solución en términos del problema real. Cuando esa interpretación no se corresponde o ajusta al problema es cuando hay que cambiar el modelo. Y el problema radica que en la amplia e infinita caja de modelos con la que contamos, la que mejores características para el problema tenga, sea la apropiada.

Bajo tal situación, hablar sobre lo que debería o no debería ser un conjunto no tiene objetivo, si no está ya definida la estructura sobre la que ocuparemos nuestra definición o nuestra idea de lo que conjunto debería de ser.

Un formalista radical solo acepta razonamientos en los que se expliciten los conceptos y axiomas que se usarán. Una definición en broma sobre un formalista (radical) es la de alguien incapaz de entender algo a menos de que carezca de significado.

El formalismo (fundado por Hilbert) fue creado en parte para formular teorías matemáticas libres de contradicciones, pero también como la forma «correcta» de «hacer» matemáticas. Para esto, Hilbert esperaba salvar a las matemáticas de lo que la escuela de pensamiento intuisionista de Brouwer había destruido. Para un formalista el significado es irrelevante, si no tiene que meterse con conjuntos infinitos esta perfecto.

Un formalista aplica lógica clásica, mientras que para varios intuisionistas, por ejemplo la doble negación no es removible. Diferente a un Platonista, un pensador formalista de la escuela de Hilbert no cree en conjuntos infinitos, mientras que el intuisionista si, aunque los dos son finitistas, permiten la obtención de cualquier número que se nos antoje, siempre que ellos sean finitamente generados por cadenas de cualquier tamaño «n», i.e., cualquier cantidad en cualquier pedido, siempre que se pidan finitos cada vez.

Para los constructivistas, la existencia de algún objeto matemático significa «hallarlo». Los constructivistas, en particular los intuisionistas, niegan la existencia de todas las cosas que no pueden ser «construidas»:

> *«A formal system in wich $\exists x G(x)$ is provable, but wich provides no method for finding the x in question, is one in wich the existencial quantifier fails to fulfill its intended function.»*
>
> R.L Goostein (1968) (tomado de [Her, p. 6])

Capítulo 3

PRINCIPAL

En este capítulo se demuestran equivalencias y consecuencias del Axioma, se enuncian axiomas alternativos, e incluso axiomas que implican al Axioma. Se abarcan diferentes áreas matemáticas como Cardinales y Ordinales, demostrando que todos los cardinales se pueden comparar con una relación “menor igual que” (subsecc. 3.1.5); Teoría de Conjuntos demostrando que todo conjunto tiene un buen orden (3.1.4), el Lema de Zorn (3.1.3), el producto no vacío de conjuntos no vacíos es no vacío (3.1.1); la hipótesis generalizada del continuo implica al Axioma (p. 103); en Álgebra se demuestra que todo espacio vectorial tiene una base (p. 116); en Topología se demuestra la Paradoja de Banach-Tarski, que tiene como aplicación descomponer una esfera en pedazos finitos y reensamblar los pedazos para formar dos esferas iguales a la original (p. 126). En la secc. 3.7 se culmina la tesis con temas puramente de la carrera y con casos interesantes del Axioma, sobre Cálculo se enuncian definiciones distintas de conjunto finito, cerrado, función continua, que dependiendo del esquema axiomático son equivalentes o no equivalentes; en la subsecc. 3.7.4 se muestra un tipo de función que cumple propiedades no intuitivas: es densa en los reales, no medible y existen funciones de ese tipo en mayor cantidad a las soluciones continuas de la ecuación de Cauchy; las últimas 3 secciones demuestran aplicaciones del Axioma a acertijos (subsecc. 3.7.6), una clasificación de conjuntos que no puede existir (subsecc. 3.7.7), y por último, una aplicación mía del Axioma de Elección (subsecc. 3.7.8), donde demuestro que todos los conjuntos son interesantes.

Mucha de la matemática moderna que estudiamos se basa en la forma Aristotélica de ver la ciencia matemática; pensar en los números, figuras, objetos matemáticos y teoremas son solo pensamientos sobre cosas existentes en la mente del matemático. La existencia del número «3» es puramente formal, existe en el universo mental de las matemáticas. A pesar de esto, varios universos matemáticos son posibles, explicaré esto: cuando queremos estudiar algo debemos hacer suposiciones básicas (*axiomas*) para empezar a sacar conclusiones sobre lo que estudiamos (*teoremas*); ¿porqué esto debe ser así?, alguien más puede llegar a conclusiones contrarias a las que llegamos nosotros y no sabríamos quién tiene la razón, ¿y cómo saber quién la tiene?, pues dependiendo de los axiomas del investigador, en el modelo que uno aplica llega a conclusiones para «ése» modelo.

Al crear una teoría, en ocasiones se requieren objetos nuevos para el pensamiento, como el concepto de “el conjunto de los números naturales”. Después de introducir objetos matemáticos nuevos, a veces quedan implícitas suposiciones adyacentes de los objetos, o pasan inadvertidas (¿existe el conjunto de los números natuales?, ¿se requiere de alguna suposición para probar su existencia?); cuando alguien se da cuenta que tales objetos dependen de dicha suposición, es cuando la dependencia entre el objeto y la razón que lo sustenta sale a la luz, y es descrita explícitamente, mejorada, corregida o hasta cuestionada; incluso, hay suposiciones hechas premeditadamente para demostrar teoremas. Pero ¿cómo saber cuáles suposiciones son «buenas» o «necesarias»?, ¿cuándo es necesario cuestionarse acerca de su «obviedad»?. En este capítulo se intentará responder a esas preguntas.

En el libro que contiene muchísimas equivalencias del Axioma [RubRub] (383 formas del Axioma), Rubin y Rubin separan las ciencias matemáticas que tienen presente el Axioma en: Principios Maximales, Formas Algebraicas, Formas de Números Cardinales, y en conjunto las Formas de Topología, Análisis y Lógica.

Los Principios Maximales se generaron principalmente después de la publicación del Lema de Zorn en 1935, aunque el primero de ellos lo dijo Hausdorff desde 1909. Los Algebraicos tuvieron prominencia después de 1950. Los que se refieren a Cardinales, tuvieron su origen gracias al aporte de Hartogs en 1915 sobre la Tricotomía; las contribuciones más significativas sobre Cardinales las dio Tarski entre 1924 y 1926[1].

3.1. EQUIVALENCIAS

Un nido (en inglés, nest) es una clase que está linealmente ordenada[2] por la relación de inclusión $\subset$.

Recordemos que una propiedad se refiere a una fórmula en un lenguaje (ZF o NBG) que tiene al menos una variable libre, y decir que $P(x)$ pasa, significa que es verdadera al tomar la variable x. Se dirá que $P(X)$ es una propiedad de carácter finito si $P(\emptyset)$ y para toda clase X:

$$P(X) \Leftrightarrow \text{para todo subconjunto finito } x \subset X, P(x).$$

Algunos libros, dan una forma equivalente de enunciar esto, si se pone:

$$Q = \{x : \; x \; \text{ es finito y } P(x)\},$$

se dirá que $P(X)$ es una propiedad de carácter finito ssi

$$P(X) \Leftrightarrow (\forall x)(x \subset X \text{ y } x \text{ es finito } \Rightarrow x \in Q).$$

Demostraremos que para el caso finito del conjunto de índices, el Axioma no se requiere para encontrar un elemento de cada subconjunto.

Teorema 3.1.1. *Siempre existe una función de elección para una familia de conjuntos no vacíos $\{X_i\}_{i\in\{1,2,\ldots,n\}}$ para todo $n \in \mathbb{N}$.*

Demostración. Se demostrará por inducción, cuando $n = 1$, se tiene que como X_1 es no vacío, se puede encontrar un elemento x_1. Supongamos que se puede encontrar una función de elección f para la familia de conjuntos no vacíos indexada por $\{1,2,\ldots,n\}$, i.e., $\forall i \in \{1,2,\ldots,n\}$ tal que $f(i) \in X_i$, como el conjunto X_{n+1} es no vacío, elegimos un elemento de ese conjunto, digamos x_{n+1}, la función extendida

$$f(i) = \begin{cases} x_i & \text{si } i \in \{1,2,\ldots,n\} \\ x_{n+1} & \text{si } i = n+1 \end{cases}$$

es una función de elección para $\{X_i\}_{i\in\{1,2,\ldots,n+1\}}$, demostrando que para todo natural se cumple la afirmación. $\square$

Hay que notar el hecho de que para $n = 1$ el caso se reduce a la definición de conjunto no vacío, incluso si X_1 es un conjunto no numerable, se puede «siempre» escoger un elemento de ahí, hay que notar que, el hecho de que se cumpla el teorema anterior para todo natural, no significa que se cumpla para el conjunto de todos los naturales, refiriéndonos a que no significa que se cumpla cuando el conjunto de índices es ω, debido a que ésta afirmación requiere del Axioma Denumerable.

[1] Véase la página 67.

[2] Véa la p. 18, ahí viene la definición de orden lineal.

3.1.1. Producto Cartesiano Generalizado

Lo primero que necesitamos es definir el producto cartesiano generalizado. Si tenemos dos conjuntos no vacíos X_1 y X_2, podemos formar su producto cartesiano, definido como el conjunto de todos los pares ordenados, con la primera componente elemento de X_1 y la segunda elemento de X_2. Es decir, $X_1 \times X_2$ es no vacío si los dos son no vacíos, ¿esto por qué?, porque la definición dice que si $x \in X_1$ y $y \in X_2$, se tendrá que $z = (x,y) \in X_1 \times X_2$. Lo mismo se puede aplicar para cualquier colección de conjuntos de tamaño $n \in \mathbb{N}$, debido a que si se tiene una colección de conjuntos no vacíos $\{X_i\}_{i\in A}$ y $A = \{1,2,\ldots,n\}$, y además $x_i \in X_i$, por la definición de producto cartesiano, $(x_1,x_2,\ldots,x_n) \in X_1 \times X_2 \times \cdots \times X_n = \prod_{i=1}^{n} X_i$, por lo tanto, éste último conjunto será no vacío si cada elemento del producto es no vacío. ¿Qué pasa cuando pensamos en un conjunto de índices infinito?, de la misma manera, supongamos que tenemos una colección de conjuntos no vacíos $\{X_i\}_{i\in\mathscr{A}}$ y $\mathscr{A}$ un conjunto de índices arbitrario, y además $\forall i \in \mathscr{A}$ se tiene que $x_i \in X_i$, ¿se tendrá que $\exists x$ tal que $x \in \prod_{i\in\mathscr{A}} X_i$?.

Definición 3.1.1. *Sea $\{X_i | i \in \mathscr{A}\}$ una familia de conjuntos. El producto cartesiano generalizado, denotado por $\prod_{i\in\mathscr{A}} X_i$, es el conjunto de todas las funciones $c : \mathscr{A} \mapsto \bigcup_i X_i$ que cumplan con la propiedad de que $\forall i \in \mathscr{A}$ se tiene que $c(i) \in X_i$.*

A una función $c \in \prod_{i\in\mathscr{A}} X_i$, también se le representa por $\{x_i\}$, para expresar que $c(i) = x_i$, para toda i; a $x_i \in X_i$ se le llamará la i-ésima coordenada de $\{x_i\}$, y al conjunto X_i se le llamará el i-ésimo factor de $\prod_{i\in\mathscr{A}} X_i$.

El siguiente teorema de [Dug, p. 23], utiliza una técnica que se empleará otras veces en demostraciones. La idea es simple, si se pide encontrar la suma de 2 y de 3, generalmente, un niño utiliza elementos que tenga al alcance, como su mano, de una mano ocupa 3 dedos y de la otra 2 y los suma; aunque parezca «natural», pero es sorprendente, casi nadie ocupa para sumar conjuntos no disjuntos, y por eso, las sumas de ese tipo funcionan, cuando se eligiera un dedo el cual está en los dos conjuntos a sumar, no se sabe si contarlo una vez, o dos veces. Una forma para evadirlo, es identificar los elementos de cada conjunto, ¿y como se identifican?, marcándolos para distinguirlos si están en otro conjunto, y la forma más fácil de marcarlos es enumerarlos y poniéndoles el número correspondiente, si se tiene una familia $\{A_i | i \in \mathscr{A}\}$, para hacerla disjunta se toma otra familia $\{A_i' | i \in \mathscr{A}\}$, con $A_i' = \{i\} \times A_i$, si $j \neq i \Rightarrow A_i' \cap A_j' = (\{i\} \cap \{j\}) \times (A_i \cap A_j) = \emptyset$, ya que el primer componente es vacío. Y lo bueno de ésta técnica, es que se tiene una biyección, entre la familia de conjuntos A_i' y A_i, con ésto, casi todas las propiedades que tiene la familia no disjunta, se trasladan a la nueva familia. La técnica también se aplicará para sumar cardinales.

Teorema 3.1.2. *Son equivalentes:*

i) *Sea una familia no vacía de conjuntos no vacíos $\{A_i | i \in \mathscr{A}\}$, se cumple que $\prod_{i\in\mathscr{A}} X_i \neq \emptyset$.*

ii) *El Axioma de Elección.*

iii) *Sea una familia no vacía de conjuntos no vacíos $\{X_i | i \in \mathscr{A}\}$, existe una función de elección para dicha familia, i.e., $c : \mathscr{A} \mapsto \bigcup_i X_i$, cumpliendo que $\forall i \in \mathscr{A}$ se tiene que $(i) \in X_i$.*

Demostración. $i) \Rightarrow ii)$
Demostraremos que $i)$ implica el punto $iii)$ del siguiente teorema (3.1.3), que dice que si se tiene una familia no vacía de conjuntos disjuntos no vacíos, habrá un conjunto que contiene exactamente un elemento de cada miembro de la familia. Además, en dicho teorema demostramos que el punto $iii)$ es equivalente al Axioma (que es equivalente a $iv)$), concluyendo lo pedido.

Si para cada familia $\{X_i'\}_{i\in\mathscr{A}}$ con $\mathscr{A}\neq\emptyset$ y también $A_i\neq\emptyset$, se tiene que $\prod_{i\in\mathscr{A}} X_i'\neq\emptyset$, en particular para una familia de conjuntos disjuntos $\{X_i\}_{i\in\mathscr{A}}$ se cumplirá, pudiendo encontrar un $c\in\prod_{i\in\mathscr{A}} X_i$, que cumpla con que $\forall i\in\mathscr{A}$ se tiene que $c(i)\in X_i$, por lo tanto, $c(\mathscr{A})=\{c(i):\ i\in\mathscr{A}\}$ contiene exactamente un elemento de cada miembro de la familia $\{X_i\}_{i\in\mathscr{A}}$.

$ii)\Rightarrow iii)$

Definamos $\forall i\in\mathscr{A}$ a $X_i'=\{i\}\times X_i$, con ello, $\{X_i'\}_{i\in\mathscr{A}}$ es una familia no vacía de conjuntos disjuntos no vacíos, y por la equivalencia del Axioma de Elección, existe un conjunto S con exactamente un elemento de cada X_i', por eso, $\forall i\in\mathscr{A}$ se tiene que $((i,x_i)\in S)\wedge(x_i\in X_i)$, y dicha S será una función de elección $S:\mathscr{A}\mapsto\bigcup_i X_i$ porque, ya que $\{i\}\subset\mathscr{A}$ y que $\bigcup\mathscr{A}=\mathscr{A}$ se cumple:

$$S\subset\bigcup_i(\{i\}\times X_i)\subset\bigcup_i(\mathscr{A}\times X_i)=\mathscr{A}\times\bigcup X_i.$$

$iii)\Rightarrow i)$

Si tenemos dicha función de elección $c:\mathscr{A}\mapsto\bigcup_i X_i$, se tendrá por definición de producto cartesiano generalizado que $c\in\prod_i X_i$, diciendo que es no vacío. □

3.1.2. Formas del Axioma de Elección

Teorema 3.1.3. *Son equivalentes:*

i) *Para toda relación R, existe una función f que cumple $f\subset R$ y $dom\ f=dom\ R$.*

ii) *El producto cartesiano de una familia no vacía de conjuntos no vacíos es no vacío.*

iii) *Si $\mathscr{A}$ es un conjunto que cumple con que cada uno de sus elementos son no vacíos y cualesquier dos elementos no tienen elementos en común (una familia disjunta de conjuntos no vacíos), entonces, existirá un conjunto C que contiene exactamente un elemento de cada miembro de $\mathscr{A}$, i.e., para todo $A\in\mathscr{A}$, se cumple para algún $x\in A$ que $C\cap A=\{x\}$[3].*

iv) *Para cualquier conjunto A existe una función de elección f cumpliendo que su dominio son los subconjuntos no vacíos de A y además que $\forall B\subset A$ se tiene que $f(B)\in B$.*

Demostración. $i)\Rightarrow ii)$

Sea $\{X_i\}_{i\in\mathscr{I}}$ una familia de conjuntos no vacíos con conjunto de índices $\mathscr{I}$ y sea R la relación definida por

$$R=\{(i,x)|i\in\mathscr{I}\wedge x\in X_i\}$$

cuyos elementos siempre existen porque $\mathscr{I}$ y X_i son distintos del vacío. Por hipótesis, tendremos que existe una función $f\subset R$ y que además $domf=domR=\mathscr{I}$, por lo tanto, $(i,f(i))\in f\subset R$, entonces existirá un x tal que $f(i)=x$, lo que nos dice que $\forall i\in\mathscr{I}$ se tiene que $f(i)\in X_i$. Por lo tanto, hemos encontrado una función que cumple $f:\mathscr{I}\mapsto\bigcup_{i\in\mathscr{I}} X_i$ y que asigna a todo i, $f(i)\in X_i$, por eso $\prod_{i\in\mathscr{I}} X_i\neq\emptyset$.

$ii)\Rightarrow iii)$

Sea un conjunto $\mathscr{A}$ que tenga elementos no vacíos y disjuntos, y sea $1_{\mathscr{A}}$ la función identidad en $\mathscr{A}$. Se cumple que $\forall B\in\mathscr{A}$ se tiene que $1_{\mathscr{A}}(B)=B\neq\emptyset$, y por hipótesis existe una función f con $domf=\mathscr{A}$ y $\forall B\in\mathscr{A}$ se cumple que $f(B)\in 1_{\mathscr{A}}(B)=B$. Abreviando $ranf=C$, para todo $B\in\mathscr{A}$ se cumple que solo

[3] Otra forma de denotar esto, es decir que toda partición, tiene un *conjunto de representantes*, en este caso C sería dicho conjunto, un subconjunto de $\mathscr{A}$ que tiene un único elemento con la intersección de cualquier elemento de $\mathscr{A}$.

tiene un elemento $B \cap C = \{f(B)\}$. Para ver que $\{f(B)\} \subset B \cap C$, por la propiedad de C éste tiene un elemento en común con B y además por la propiedad de la función de elección se tiene que $f(B) \in B$. Para la contención contraria, si hay un elemento distinto de $f(B)$ llamado x que cumple que $B \cap C = \{f(B), x\}$ se tiene que existe A que satisface $x = f(A)$ con $A \in \mathscr{A}$, debido a que $x \in ranf$, y por el lado izquierdo de la igualdad, $f(A) \in B$, pero por la propiedad de la función $f(A) \in 1_{\mathscr{A}}(A) = A$, por lo tanto, se concluye que se tiene $f(A) \in B$ y $f(A) \in A$, contradiciendo que la familia $\mathscr{A}$ es disjunta. De aquí, se tiene que el conjunto C tiene únicamente un elemento en común con todo elemento de $\mathscr{A}$.

iii) $\Rightarrow$ *iv*)

Si tenemos un conjunto A, definamos

$$\mathscr{A} = \{\{B\} \times B | \emptyset \neq B \subset A\},$$

Todo elemento de $\mathscr{A}$ será no vacío, además, serán disjuntos dos elementos diferentes, porque si $(x,y) \in (\{B\} \times B) \cap (\{D\} \times D)$, se tendrá que $x = B = D$ por el primer elemento de los pares ordenados. Si llamamos a C el conjunto que tiene un solo elemento cuando lo intersectamos con cada elemento de $\mathscr{A}$, cuya existencia depende de la hipótesis, se tiene

$$C \cap (\{B\} \times B) = \{(B,x)\},$$

para algún $x \in B$, debido a que $(B,x) \in \{B\} \times B$. Un punto a destacar, es que la hipótesis solo menciona que la intersección de C con los elementos de $\mathscr{A}$ será un solo elemento, pero el conjunto C podría tener elementos que no estuvieran en algún miembro de $\mathscr{A}$, por lo tanto, tomamos otro conjunto con solo elementos de los miembros de $\mathscr{A}$, obtenido así: $f = C \cap \bigcup \mathscr{A}$. Tenemos así una función de elección f para A, ya que como $C \cap (\{B\} \times B) = \{(B,x)\}$, los elementos de f serán de la forma $(B,x) = (B, f(B))$ con $x \in B$ (es decir, se tiene $(B,x) \in f$), y para todo subconjunto no vacío $B \subset A$, hay un único x tal que $(B,x) \in f$ porque si hubiera otro distinto, llamado z, se tendría que $((B,z),(B,x) \in f) \wedge ((B,z),(B,x) \in (\{B\}, B))$, contradiciendo que solo hay un elemento en la intersección. Encontramos así, que f es una función de elección en A.

iv) $\Rightarrow$ *i*)

Si tenemos cualquier relación R, la hipótesis nos asegura la existencia de una función de elección g sobre *ranR*, que cumple que $g(B) \in B$ para todo subconjunto no vacío de *ranR*. Definamos la función f mediante

$$f(x) = g(\{y | xRy\}),$$

y con dominio igual al dominio de R. Se cumple entonces que

$$f(x) = g(\{y | xRy\}) \in \{y | xRy\},$$

por lo tanto, $(x, f(x)) \in R$ debido a que $f(x) \in ranR$, se concluye que $f \subset R$. □

3.1.3. Formas de principios maximales

Para demostrar equivalencias en cuanto a principios maximales, hay que notar que cuando hablamos de un nido (como subconjunto) de un conjunto y que dicho nido es $\subset$-maximal, o una cota superior o su unión, éstos términos coinciden, debido a que si es una cota superior contiene a todos sus subconjuntos, si es $\subset$-maximal, no habrá un conjunto que lo contenga propiamente, por eso, contendrá a todos los conjuntos, y si hablamos de la unión del nido, no habrá un conjunto más grande que ésta, también gracias a que se tiene una cadena en el orden lineal, con ello podemos demostrar el:

Teorema 3.1.4. *Son equivalentes:*

M1. Lema de Zorn. Si R es un orden parcial en X $\neq \emptyset$ *, y si* $\forall A \subset X$ *que sea linealmente ordenado por R, A tiene una R-cota superior, se tendrá que X tiene un elemento R-maximal.*

M2. Si R es un orden parcial en X no vacío, y $\forall A \subset X$ *que sea bien ordenado por R, A tiene una R-cota superior, se tendrá que X tiene un elemento R-maximal.*

M3. Si todo subconjunto no vacío que sea un nido, de un conjunto X no vacío tiene su unión en X, entonces X tendrá un elemento $\subset$*-maximal.*

M4. Si todo subconjunto no vacío y bien ordenado que sea un nido, de un conjunto X no vacío tiene su unión en X, entonces X tendrá un elemento $\subset$*-maximal.*

M5. Condición de Cadena Máxima de Hausdorff. Si R es una relación que hace que (X,R) *sea un conjunto parcialmente ordenado, entonces existe un subconjunto* $\subset$*-maximal de X que está linealmente ordenado por R*[4].

M6. Para todo conjunto X, existe un subconjunto de X $\subset$*-maximal que es un nido.*

M7. Teichmüller-Tukey Lemma. Para todo conjunto X y toda propiedad de carácter finito P, existe un subconjunto de X $\subset$*-maximal que tiene la propiedad P.*

Demostración. $M2 \Rightarrow M1)$ En la p. 20 se ve que todo conjunto bien ordenado, será un conjunto linealmente ordenado, de ahí se desprende que $M1$ es un caso particular de la hipótesis.

$M1 \Rightarrow M3$ Un nido es un orden lineal con la relación $\subset$, por lo tanto $M3$ es un caso particular de $M1$, ya que la unión está en X, éste será su cota $\subset$-superior del subconjunto, y $M1$ afirma que tendrá X un elemento $\subset$-maximal, que en este caso será $\subset$-maximal.

$M3 \Rightarrow M4$ M4 es una caso particular de M3, ya que la diferencia en los enunciados, es que en M4 se pide que el conjunto también sea bien ordenado.

$M7 \Rightarrow M5$ Una propiedad de carácter finito es que X sea linealmente ordenado por R, debido a que si X es linealmente ordenado por R, todo subconjunto finito de X tendrá un orden lineal, ya que como es una cadena X, toda pareja de elementos $\{x,y\}$ de X se relaciona: xRy o yRx, y también esa pareja es un orden parcial, por lo tanto, para todo conjunto finito se cumple lo mismo (por inducción); luego, si se tiene un orden lineal en toda pareja de elementos $\{x,y\}$ de X, para demostrar que hay un orden lineal para X, como cualesquiera dos elementos se relacionan por ésta hipótesis, será una cadena, y para demostrar que es un orden parcial, hay que notar que se demuestra para cualesquiera 1,2,3 elementos de X para ver que la relación es reflexiva, antisimétrica y transitiva respectivamente, por ello, es una propiedad de carácter finito.

Si se toma esa propiedad, el conjunto (X,R) en particular será un conjunto parcialmente ordenado, y la hipótesis afirma la existencia de un subconjunto $\subset$-maximal que también será linealmente ordenado.

$M5 \Rightarrow M6$ Si se tiene un conjunto X, la relación $\subset$ hace que (X,R) sea un conjunto parcialmente ordenado, porque es reflexiva, antisimétrica y transitiva, por hipótesis se tiene un conjunto $\subset$-maximal que también será linealmente ordenado por $\subset$, i.e., un nido.

$M3 \Rightarrow M7$ Se toma un conjunto X y a una propiedad de carácter finito llamada P, denotando por $Y = \{t : t \subset$ y también $P(t)\}$, se va a demostrar que si se tiene un nido $n \subset Y$ que sea no vacío, se cumple que $\bigcup n \in Y$. Como $n \subset Y \subset \mathscr{P}(X)$, se tiene que $\bigcup n \subset \bigcup Y \subset \bigcup \mathscr{P}(X) = X$. Al tomar un subconjunto $u \subset \bigcup n$ finito, del hecho de que n es un nido, dicho u será un subconjunto de algún elemento de n, i.e. $u \subset t \in n$, ya que todos los elementos de n están relacionados por la relación $\subset$, y debido a que

[4]También se enuncia así: Todo conjunto parcialmente ordenado contiene una cadena máxima.

todo elemento de n tiene la propiedad P porque $n \subset Y$, i.e. $P(t)$, se cumple que $P(u)$, porque u es un subconjunto finito de t. Al haber encontrado que para todo subconjunto finito de $\bigcup n$ se cumple P, se tiene que $P(\bigcup n)$, además, se cumple $P(\emptyset)$ debido a que P es de carácter finito, por lo tanto, Y cumple con la hipótesis de $M3$, y será un elemento maximal el elemento encontrado y que cumple con P.

$M4 \Rightarrow M7$ Se acaba de demostrar que M3 implica M7, y al cumplir M4 las hipótesis de M3 junto con una adicional, la demostración sería igual a la anterior.

$M6 \Rightarrow M5$ Sea R una relación de orden parcial en X, sea

$$Y = \{A \subset X : \ (A,R) \text{ es un conjunto linealmente ordenado}\},$$

Y no es igual al vacío porque $\{x\}$ con $x \in X$ es linealmente ordenado por R ya que éste es reflexivo. Por la hipótesis de $M6$, existe un n nido que también es $\subset$-maximal en Y. Se demuestra que $\bigcup n$ es un elemento $\subset$-maximal de Y, y que también está linealmente ordenado por R. Se tiene que $\bigcup n \subset X$ por la misma razón que la demostración anterior. Tomando dos elementos distintos de $\bigcup n$, llamados u, v, entonces, $\exists w$ tal que u, $v \in w$ y $w \in n$, porque n es un nido y por eso, contendrá elementos que se relacionan con dichos elementos por medio de $\subset$, demostrando que es una cadena; para demostrar que es un orden lineal $(\bigcup n, R)$, se debe notar que (w,R) lo es, porque $w \in n \subset Y$, por eso se cumple que uRv o que vRu, que sea un orden parcial se sigue directo de que $\bigcup n \subset X$. Solo falta demostrar que $\bigcup n$ es $\subset$-maximal. Si no lo fuera, existiría un $z \in Y$ con $\bigcup n \subset z$, pero es un hecho que $n \cup \{z\}$ es un nido, porque si $A \in n$ y si $x \in A$, se tiene que $x \in \bigcup n \subset z$, se concluye que $A \subset \bigcup n \subset z$, es decir, todo elemento de n es subconjunto de z, lo que dice que se relaciona con z, además, de la $\subset$-maximidad de n se debe cumplir que $\{z\} \subset n$ porque z es un conjunto linealmente ordenado al estar en Y, o lo que es lo mismo, $z \in n$, implica que $z \subset \bigcup n$ (por definición de unión), pero ésto contradice la maximidad de z, por lo tanto, $\bigcup n$ es un elemento $\subset$-maximal de X que es linealmente ordenado por R.

$M5 \Rightarrow M2$ Se toma un conjunto parcialmente ordenado (X,R), y a

$$Y = \{A \subset X : \ A \ \text{ es bien ordenado por } R\},$$

se denota a:

$$S = \{(t,u) : \ t\ ,\ u \in Y,\ t \ \text{ es un } R\text{-segmento inicial de } u\},$$

recordar que un conjunto es un R-segmento inicial de otro fue visto en la p. 18. Se demuestra que S es una relación de orden parcial en Y: $A \in Y \Rightarrow ASA$ porque $A \subset A$ y $\forall a \in A$, como R es reflexiva, aRa, y esto implica que $a \in A$; Si A, $B \in Y$ y ASB y BSA, por definición de S, A es un R-segmento inicial de B y viceversa, por lo tanto, por definición de R-segmento inicial, $A = B$; si A, B, $C \in Y$, con ASB y BSC, se cumple que $A \subset B \subset C$, por eso, $A \subset C$, y como $(\forall a \in A)(\forall b \in B)(\forall c \in C)(bRa \wedge cRb \Rightarrow b \in A \wedge c \in B)$, en particular, como R es transitiva cRa, y además, para el $c \in B$ encontrado, y aplicándose a la condición de nuevo, $c \in A$, por lo tanto, A es un R-segmento inicial de C.

Por $M5$, existe en Y un subconjunto n que es $\subset$-maximal, y además, está linealmente ordenado por S, se demuestra que el elemento buscado es $\bigcup n$, éste tiene una cota R-superior que es R-maximal en X.

Primero se ve que $\bigcup n \in Y$, como las demostraciones pasadas, $\bigcup n \subset X$, si se toma $\emptyset \neq z \subset \bigcup n$, cuando $x \in z \Rightarrow x \in A$ con $A \in n$ por la definición de unión, por lo tanto $z \cap A \neq \emptyset$. Si u es el R-primer elemento de $z \cap A$, que existe porque $z \cap A \subset A \in Y$ que contiene conjuntos bien ordenados por R, se tiene que demostrar que u es el R-primero en z, para concluir que está bien ordenado por R. Al considerar $v \in z$ con $v \neq u$, se dividen dos casos: cuando $v \in A$, por la propiedad de u, uRv y $\neg(vRu)$ porque son diferentes; cuando $v \notin A$, entonces ($\exists t$ tal que $v \in t$) para algún $t \in n$, ésto por la misma razón que al principio de este párrafo cuando se tomó $z \subset \bigcup n$, luego, se tiene que A, $t \in n$, y ya que n es $\subset$-maximal y está linealmente ordenado por S, se tendrá que ASt o que tSA, pero $v \notin A$, por lo que ASt (por definición de S), lo que implica que uRv y que $\neg(vRu)$, la primera porque t es un buen orden y en particular un orden lineal (demostrado en la p. 20) y la segunda porque si lo hiciera, estaría $v \in A$, contradiciendo lo hipótesis. Con todo esto, se demostró que para todo elemento v de todo elemento de z se tiene que uRv, lo que dice que es el único R-primer elemento de z, por eso $(\bigcup n, R)$ es un buen orden, y se tiene que $\bigcup n \in Y$.

Ahora se demostrará que $\bigcup n$ es maximal en Y, de igual forma, si no lo fuera, existiría un $z \in Y$, con $\bigcup n \subset z$, para llegar a la contradicción, se tiene que demostrar que $n \cup \{z\}$ está linealmente ordenado por S, contradiciendo la maximidad de n. Como n es un orden lineal por S debido a su construcción, y todo elemento de n es subconjunto de z por lo mismo que la demostración anterior, solo falta ver que $n \cup \{z\}$ sea una cadena: todos los elementos de n se relacionan por ser éste orden lineal, si $A \in n \Rightarrow A \subset \bigcup n \subset z$, por ello $A \subset z$, solo falta ver que A es un R-segmento inicial de z para demostrar que ASz, y concluir que es un orden lineal, pero del siguiente hecho resulta de inmediato: como $z \in Y$ y como n es maximal, se tiene que $\{z\} \subset n \Rightarrow z \in n$ por eso y como n es cadena, se tiene que ASz, además $z \in n \Rightarrow z \subset \bigcup n$, lo cual contradice la hipótesis sobre z.

Por último, $\bigcup n$ no tiene una cota R-superior estricta, si b lo fuera, el conjunto $\bigcup n \cup \{b\} \in Y$, debido a que solo se agrega un elemento al conjunto bien ordenado $\bigcup n$, y que es mayor que todos los demás, no quitando el elemento menor de los subconjuntos, y con lo que se acaba de demostrar, se contradice la maximidad de m, por ser $\bigcup n \cup \{b\} \in Y$ más grande. La hipótesis de $M2$ nos dice que como $\bigcup n \in Y$, éste tiene una cota R-superior llamada b, si se cumple que $z \in X$ y que bRz, se debe tener que zRb concluyendo que $b = z$, si no pasara, como R es transitiva, $\forall a \in \bigcup n$ se tendría que $(aRb) \wedge (bRz) \Rightarrow aRz$, demostrando que z es una cota R-superior estricta, que se vio que no puede ocurrir. Esto que se acaba de afirmar, dice que b es un elemento R-maximal de X, que es lo que se requería. □

3.1.4. El principio del Buen Orden

Ahora introduciremos un concepto, que nos ayudará a demostrar un teorema elemental, dentro de las equivalencias del Axioma.

Definición 3.1.2. *Para todo conjunto infinito X, el menor $\aleph$ que cumpla $\aleph \not\leq |X|$ se llamará el número de Hartog de X denotado por $h(X)$.*

El número de Hartog de un conjunto siempre existe, y dicha existencia no depende del Axioma. Recuerde lector que un aleph es un número ordinal inicial, lo mismo es decir que es la cardinalidad de un conjunto infinito bien ordenable. La demostración de que dicho aleph siempre existe y que dicha existencia no depende del Axioma de Elección está en [Sie, p. 410]. La equivalencia que demostraremos es la del Axioma de Elección y el Teorema del Buen Orden.

Teorema 3.1.5. *Son equivalentes*[5]*:*

i) El Axioma de Elección. Todo conjunto tiene una función de elección.

ii) El principio del Buen Orden. Dado un conjunto X, existe una relación $R \subset X \times X$ tal que (X,R) es un buen orden.

Demostración. $ii) \Rightarrow i)$
Si se tiene un conjunto X, sea $Z = \{X_i\}_{i \in \mathscr{I}}$ la familia que tiene como elementos a todos los subconjuntos de X no vacíos y además, los consideraremos disjuntos (la familia puede ser considerada como todos sus subconjuntos indexados con ellos mismos, visto en la p. 14.). Sea $S = \bigcup_{i \in \mathscr{I}} X_i$, S es un conjunto

[5]La demostración es para conjuntos infinitos, ya que para el caso finito, las equivalencias aún se cumplen.

porque es un subconjunto del conjunto potencia de X, por hipótesis, S tendrá un buen orden, i.e., existirá una relación R que asigne a cada subconjunto de S un primer elemento. Todos los elementos de Z, son subconjuntos de S, porque $X_i \subset \bigcup_{i\in\mathscr{I}} X_i, \forall i \in \mathscr{I}$, así que podemos escoger un primer elemento de todo X_i denotado por $f(X_i)$, éste será un único elemento ya que el primer elemento es único, por lo tanto, a todo subconjunto no vacío de X le encontramos un elemento que pertenece a ese subconjunto, y eso nos dice que el Axioma se cumple.

$i) \Rightarrow ii)$

La idea es enumerar a todos los elementos del conjunto X infinito. Escogemos un elemento de X, que será el primero x_1, luego escogemos un elemento de $X - \{x_1\}$, que será el segundo llamado x_2, y seguimos así hasta acabar con todos los elementos del conjunto X. Cuando tengamos un subconjunto de X, nos fijamos en el elemento de menor índice y ese será el menor en el subconjunto, así encontramos que X tiene un buen orden.

La demostración la haremos definiendo vía recursión transfinita, por el Axioma, existe una función de elección f en X, tomamos un elemento $a \notin X$ (el cual existe debido a que X no es la clase de todos los conjuntos) y a $g : h(X) \mapsto X \cup a$ (con $h(X)$ el número de Hartogs de X) como

$$g(\alpha) = \begin{cases} f(X - \{g(\beta)|\beta < \alpha\}) & \text{si } X \neq \{g(\beta)|\beta < \alpha\} \\ a & \text{si } X = \{g(\beta)|\beta < \alpha\} \end{cases}$$

Hay que demostrar que en algún número ordinal $\lambda < h(X)$, terminamos de contar todos los elementos de X, i.e., existe algún λ tal que $g(\lambda) = a$, y de allí tomaremos al menor que lo cumple. Para $\beta < \alpha$ y supongamos que $g(\alpha) \neq a$, tenemos que por definición de g, $g(\alpha) \in X - \{g(\beta)|\beta < \alpha\}$, que es el elemento que escoge f de este conjunto, y además $g(\beta) \in \{g(\beta)|\beta < \alpha\}$ ya que $\beta < \alpha$, juntando estas dos pertenencias llegamos a que $g(\beta) \neq g(\alpha)$, concluyendo que g es inyectiva. ¿Cómo la hacemos suprayectiva?, tenemos que detener el conteo hasta el primer ordinal λ tal que pase $g(\lambda) = a$. Supongamos que no existe tal ordinal, i.e., $\forall \alpha < h(X)$ se tiene que $g(\alpha) \neq a$, tendríamos que $g : h(X) \mapsto X$ sería una función inyectiva, por lo que se cumpliría que $h(X) \leq |X|$, contradiciendo la definición de $h(X)$ el menor ordinal que no puede ser mapeado a X por una función inyectiva.

Sea $\lambda = \text{mín}\{\alpha < h(X)|g(\alpha) = a\}$, tendremos que $g|_\lambda : \lambda \mapsto X$ será una biyección. Esto ya que $ran(g|_\lambda) \subset X$ por la definición de g, para la contención contraria, si $X - ran(g|_\lambda) \neq \emptyset$, tendriamos por la definición de g que $g(\lambda) \neq a$, contradiciendo la definición de λ.

Concluimos que existe una biyección entre un conjunto bien ordenado (el ordinal λ) y X, por lo tanto, X tiene un buen orden. □

Teorema 3.1.6. *Son equivalentes:*

i) Todo cardinal que no es finito es un aleph.

ii) El Principio del Buen orden.

Demostración. $i) \Rightarrow ii)$ Si todo cardinal infinito es un aleph, sea X un conjunto infinito, por lo tanto, su cardinalidad será un aleph, y por definición de aleph, existirá una relación que le de un buen orden a X, lo que demuestra que todo conjunto infinito es bien ordenado. Para el caso de conjuntos finitos no es necesaria la hipótesis.

$ii) \Rightarrow i)$ Si el número cardinal $\mathfrak{n}$ no es finito, y si X es un conjunto de cardinalidad $\mathfrak{n}$, se tendrá que el conjunto X es infinito, ya que si fuera finito, sería una biyección entre un conjunto finito y uno infinito (porque card $X \sim X$). Por hipótesis, existe una relación que bien ordena a X, y por definición, por ser $\mathfrak{n}$ la cardinalidad de un conjunto infinito bien ordenado, éste será un aleph. □

Aquí esta la equivalencia del Lema de Zorn mediante el principio del buen orden, el cuál se acabó de demostrar equivalente al Axioma:

Teorema 3.1.7. *Son equivalentes:*

i) El principio del buen orden.

ii) El Lema de Zorn. Si R es un orden parcial en $X \neq \emptyset$, y si $\forall A \subset X$ que sea linealmente ordenado por R, A tiene una R-cota superior, se tendrá que X tiene un elemento R-maximal.

Demostración. $i) \Rightarrow ii)$ Sea $(P, <)$ un conjunto parcialmente ordenado y supongamos que toda cadena en P tiene una cota superior, debemos encontrar mediante la hipótesis un elemento maximal de P. Hacemos ésto, al bien ordenar al conjunto P, enumerando todos sus elementos:

$$P = \{p_0,\ p_1,\ \ldots,\ p_\xi,\ \ldots\} \qquad \xi < \alpha$$

para algún número ordinal κ (tal como el número de Hartogs de P, cuya existencia no depende del Axioma). Definamos por recursión transfinita:

$$c_0 = p_0,$$

$$c_\xi = p_\gamma$$

si γ es el menor ordinal tal que p_γ es una cota superior de la cadena $C = \{c_\eta :\ \eta < \xi\}$ y que también cumple con que $p_\gamma \notin C$. Lo que estamos haciendo es nombrar a un elemento de la sucesión, por ejemplo c_ξ, y a todos los anteriores (c_η con $\eta < \xi$) ponerlos en un conjunto donde c_ξ es una cota y no pertenece a él.

Demostraremos que para cualquier $\xi < \kappa$, el conjunto $C = \{c_\eta :\ \eta < \xi\}$ es una cadena, que es lo mismo a demostrar que si tenemos dos elementos c_α y c_β en C, con $\alpha,\ \beta < \xi$, tendremos que $c_\alpha < c_\beta$ o que $c_\beta < c_\alpha$. Como los números ordinales están bien ordenados, sin pérdida de generalidad supongamos que $\alpha < \beta$, entonces, para el elemento de la sucesión $c_\beta = p_{\beta_0}$ se cumple que si $\eta < \beta$ se tiene que $c_\eta < p_{\beta_0} = c_\beta$ por ser cota superior; aplicando esto a c_α lo cuál es posible porque $\alpha < \beta$, se cumple que $c_\alpha < c_\beta$, mientras que si suponemos $\beta < \alpha$ llegaremos a la relación inversa, por lo tanto, es una cadena el conjunto C.

El número p_γ de la definición de la sucesión casi siempre existe, porque la hipótesis dice que cualquier cadena tiene una cota superior, solo no existe cuando $c_{\xi-1}$ es un elemento maximal de P, porque el siguiente elemento de la sucesión ya no podría cumplir que $c_\xi = p_\gamma \notin C = \{c_\eta :\ \eta < \xi\}$. Pero dicha construcción de la sucesión, en algún momento, tiene que detenerse, porque solo enumeramos hasta algún ordinal κ, al llegar al tope, encontramos un elemento maximal de P.

$ii) \Rightarrow i)$ Demostraremos que cualquier familia de conjuntos no vacíos S tiene una función de elección. Sea F la familia de todas las funciones f que tienen $dom\ f \subset S$ y además $f(X) \in X$ para cualquier $X \in S$, es decir, funciones de elección sobre S. Ordenamos al conjunto F por inclusión $\subset$. Si tenemos a F_0 un subconjunto linealmente ordenado de $(F, \subset)$, entonces el conjunto $f_0 = \bigcup F_0$ es una función por el teorema 1.5.3, ya que al ser también los dominios de las funciones de F_0 linealmente ordenados, las funciones tendrán el mismo valor en la intersección de sus dominios, teniendo así a un sistema de funciones compatibles F_0. El dominio es $dom\ f_0 = \bigcup\{dom\ f | f \in F_0\} \subset \bigcup S = S$, y además si $X \in S$, entonces $f_0(X) = f_i(X) \in S$ para alguna $f_i \in F_0 \subset F$. Eso nos lleva a concluir que $f_0 \in F$ y claramente, como f_0 es la unión, será también una cota superior de F_0 en $(F, \subset)$.

Al haber demostrado que las hipótesis del lema de Zorn se satisfacen, concluimos que $(F, \subset)$ tiene un elemento maximal, el cuál llamaremos g. Sólo basta demostrar que $dom\ g = S$ para haber encontrado una función de elección sobre S. Supongamos que no se cumple, sean $X \in S - dom\ g$ y $x \in X$, se cumple que $g_0 = g \cup \{(X,x)\} \in F$, porque $domg_0 = dom\ g \cup X \subset S \cup X = S$ y además $g_0(X) = x \in X$ y por construcción $X \in S$. La función g_0 cumple que $g \subset g_0$ y contradice la maximidad de g, por lo tanto, se cumple el Axioma de Elección. □

3.1.5. Tricotomía de los Cardinales

Las siguientes definiciones se obtienen de [Sie].

Como el concepto de equivalencia o equipotencia entre conjuntos nos trae el concepto de número cardinal, el concepto de similaridad entre conjuntos ordenados[6] lo hace con el de tipos ordenados (en inglés, order types). De la misma forma para construir la definición de números cardinales, dividimos a los conjuntos ordenados en clases, asignando dos a la misma clase ssi son similares. Dicha relación es reflexiva, simétrica y transitiva, creando una relación de equivalencia entre los subconjuntos ordenados de un conjunto dado. Dichas clases disjuntas son llamadas tipos ordenados. Por ello, todo conjunto ordenado A determina un cierto tipo ordenado, que es el tipo determinado por la clase de todos los conjuntos ordenados similares al conjunto A. El tipo ordenado de un conjunto se denota por $\overline{A}$, es como la propiedad del conjunto que queda cuando no nos fijamos en los elementos pero si en su orden, y cuando se quita además el orden, según Cantor, uno se queda con el concepto de potencia del conjunto, notación vista en la p. 16. Otra forma más clara de definir un tipo ordenado, la hace [HrbJec, p. 113], si W es un conjunto bien ordenado, el tipo ordenado de W es el único número ordinal isomorfo a W.

Resumiendo, dos conjuntos ordenados son similares ssi $\overline{A} = \overline{B}$, y además, eso implica que $\overline{\overline{A}} = \overline{\overline{B}}$, pero la afirmación contraria no siempre pasa.

Se han dividido a los números ordinales infinitos en clases, estando dos de ellos en la misma clase ssi tienen la misma potencia, se puede encontrar además, un elemento menor en cada clase, que será un ordinal inicial (visto en la p. 35).

Siendo Z una clase dada de números ordinales mayores o iguales que ω, ψ su número ordinal inicial, y P_ψ el conjunto de todos los números ordinales transfinitos e iniciales menores que ψ, el conjunto P_ψ será bien ordenado por las magnitudes de los números en él, por ello, su tipo será un cierto número ordinal α, si renombramos al número $\psi = \omega_\alpha$, concluimos que todo número ordinal inicial tiene un índice α, que es su número ordinal asociado. Si escribimos $\overline{\omega_\alpha} = \aleph_\alpha$, éste se interpreta como la potencia de todo número ordinal que está en la clase con ordinal inicial ω_α. Dicha clase es denotada por $Z(\aleph_\alpha)$, y está formada por números ordinales con potencia $\aleph_\alpha$.

Los números $\aleph_\alpha$ son llamados alephs, la potencia de todo conjunto bien ordenado es un aleph y la implicación contraria también se mantiene.

Cuando a es un número ordinal tal que $a \in Z(\aleph_\alpha)$, entonces $\overline{a} = \overline{\omega_\alpha} = \aleph_\alpha$, con $\overline{\omega_\alpha}$ el número ordinal inicial de $Z(\aleph_\alpha)$.

Del hecho de que todo subconjunto de un conjunto bien ordenado está bien ordenado, se tiene que todo número cardinal menor que un aleph, o es finito o es un aleph, porque si es infinito y está bien ordenado, será aleph, esto lo ocuparemos varias veces en las páginas siguientes.

Si se tiene un ordinal transfinito α, se dirá que es un ordinal inicial ssi $\forall \beta < \alpha$ con β un ordinal, se tiene que $\beta \prec \alpha$. Se suelen denotar los ordinales iniciales por ω_α, con α un ordinal.

Resultados importantes sobre ordinales son:

[6]Sierpinski, como mencionamos en la página 17, utiliza el concepto de conjunto ordenado para lo que definimos como orden simple estricto, utilizaremos su notación, para que el lector interesado no encuentre problema al leer directo el libro, que es una buena referencia para el tema.

$$a \in \mathrm{On} \Rightarrow a = \overline{A} \Leftrightarrow A \cong a,$$

$$a \in \mathrm{On} \Rightarrow a = \overline{P_a} = \overline{\{\omega_b : \ \omega b < \omega a\}}$$

donde On denota la clase de todos los números ordinales.

Ahora supongamos que $\alpha < \beta$ y que existen $\aleph_\alpha$ y $\aleph_\beta$, como α es el tipo del conjunto P_α de todos los números iniciales menores que ω_α, y lo mismo para β, se tiene que $\omega_\alpha < \omega_\beta$, si fuera cierto que $\omega_\alpha \geq \omega_\beta$, se tendría que $P_\beta \subset P_\alpha = \{\omega_\gamma : \ \omega_\gamma < \omega_\alpha\}$, y por lo tanto $\alpha \cong P_\alpha \subset P_\beta \cong \beta$, y eso dice que α será similar a un subconjunto de β, i.e., $\beta \leq \alpha$, contradiciendo nuestra hipótesis de dichos números. Eso implica que se cumple $\overline{\omega_\alpha} \leq \overline{\omega_\beta}$, porque ω_α es similar a un subconjunto de ω_β, por ésto, será equipotente a dicho subconjunto; pero no se cumple $\overline{\omega_\alpha} = \overline{\omega_\beta}$, porque pertenecen a diferentes clases ya que $\alpha < \beta$ y por ello, no son iguales, lo que dice que $\overline{\omega_\alpha} < \overline{\omega_\beta}$, que por nuestra definición, $\aleph_\alpha < \aleph_\beta$.

De hecho se tiene que $\alpha < \beta \Leftrightarrow \aleph_\alpha < \aleph_\beta$.

De la definición de ordinales, se tiene que ellos están ordenados por la relación <, por eso, se acaba de demostrar que[7]:

Teorema 3.1.8. *La tricotomía para alephs se cumple.*

Teorema 3.1.9. *Son equivalentes:*

i) *El Axioma de Elección.*

ii) *La Ley de la Tricotomía de los Cardinales.*

Demostración. $i) \Rightarrow ii)$ El Axioma es equivalente al Principio del Buen Orden, que implica que todo cardinal que no es finito es un aleph, afirmación que se acaba de demostrar, por lo tanto, por el teorema anterior, la tricotomía se aplica para todo cardinal infinito, y como el caso de los cardinales finitos se reduce a la relación de comparación de números naturales, se cumple la Tricotomía de los Cardinales.

$ii) \Rightarrow i)$ Se demostrará que la Tricotomía implica el Principio del buen orden, equivalente al Axioma. Si se tiene un conjunto M infinito[8], con su respectiva cardinalidad $\mathfrak{m}$, existe un aleph $\aleph(\mathfrak{m}) = h(X)$ que no cumple $\aleph(\mathfrak{m}) \leq \mathfrak{m}$, por lo tanto, la Tricotomía implica que $\mathfrak{m} \leq \aleph(\mathfrak{m})$, por eso, la definición de la desigualdad implica que $M \sim X$ con X un subconjunto de un conjunto bien ordenado de potencia $\aleph(\mathfrak{m})$, y eso dice que existe una relación que bien ordena a M. □

3.2. ALTERNATIVAS

Como se vio antes, el problema de aceptar elecciones infinitas no contables parece para algunos algo implausible. Las alternativas son varias, y algunas ofrecen caminos para los matemáticos que no aceptan el Axioma de Elección en su forma más completa, los cuáles también cuentan con sus contradicciones, o caminos donde faltan condiciones y axiomas para llegar a resultados importantes o resultados a los que se busca llegar sin aceptar completamente el Axioma, y esta sección menciona algunos ejemplos, aunque no daré demostraciones, ya que algunos resultados son casos muy particulares, muy generales, o fuera del alcance de la tesis.

Hay que aclarar algo importante, si se quiere tener una demostración bien fundamentada, hay que explicitar si se utilizó el Axioma. Cuando una demostración depende de él, se puede (o más bien, debe):

[7] En la página 392 de [Sie], se demuestra que para todo ordinal, dicho aleph indexado por el ordinal, existe.

[8] Para el caso de conjuntos finitos, se reduce la demostración al buen orden de los naturales.

1. Aclarar que la demostración depende de un caso particular del Axioma, indicando en que punto se utiliza (incluso se podría aclarar que todas las demostraciones conocidas del teorema utilizan el Axioma).

2. Determinar el caso particular del Axioma que es suficiente para la demostración y el caso particular del Axioma que es necesario para la demostración (decir que un caso particular es necesario es aclarar que el teorema implica la veracidad de ese caso particular).

3. Determinar el caso particular de Axioma que es necesario y suficiente para la demostración del teorema.

3.2.1. Debilitando las hipótesis

Se pueden dar más restricciones a las condiciones del Axioma como:

1. Permitir un conjunto de índices menos general.

2. Restringir a los conjuntos de la familia.

3. Reemplazar la afirmación de que se puede elegir para cada X_i un elemento distinguido x_i por alguna afirmación más débil.

El Axioma de Elecciones Contables resulta de permitir las elecciones sobre un conjunto de índices más restringido: los naturales. Es decir:

Definición 3.2.1. *El Axioma de Elecciones Contables*[9]*: para toda sucesión de conjuntos no vacíos* $\{X_n\}_{n\in\mathbb{N}}$*, el producto cartesiano es no vacío* $\prod_{n\in\mathbb{N}} X_n$.

Cuando se restringen a los conjuntos X_i queda:

Definición 3.2.2. *AC(fin): para cada familia* $\{X_i\}_{i\in I}$ *de conjuntos no vacíos y finitos, el producto* $\prod_{i\in I} X_i$ *es no vacío.*

Definición 3.2.3. *AC(n): para cada familia* $\{X_i\}_{i\in I}$ *de conjuntos no vacíos y de cardinalidad* $n \in \mathbb{N}$*, el producto* $\prod_{i\in I} X_i$ *es no vacío.*

Otra alternativa que trae como consecuencia varios principios:

Definición 3.2.4. *El Axioma de Elecciones Múltiples*[10]*: para cada familia* $\{X_i\}_{i\in I}$ *de conjuntos no vacíos, existe una familia* $\{F_i\}_{i\in I}$ *tal que* $F_i \subset X_i$ *y* F_i *es un conjunto finito no vacío.*

Definición 3.2.5. *El Principio de Selección de Kinna-Wagner: para cada familia* $\{X_i\}_{i\in I}$ *donde cada* X_i *tiene al menos dos elementos, existe una familia* $\{Y_i\}_{i\in I}$ *de conjuntos no vacíos y propios* $Y_i \subset X_i$.

Al aceptar el primer y el segundo método quedan:

Definición 3.2.6. *CC(*$\mathbb{R}$*): para cada sucesión de subconjuntos no vacíos de* $\mathbb{R}$*, su producto es no vacío.*

Definición 3.2.7. *CC(fin) y CC(n): que para cada sucesión de subconjuntos no vacíos y finitos (de n elementos), su producto es no vacío (respectivamente)*[11].

[9] Otra definición es que existe una función de elección sobre este tipo de familias.
[10] También llamado AMC.
[11] La diferencia con AC(n) radica en que se utiliza una sucesión de conjuntos $\{X_n\}_{n\in\mathbb{N}}$, no una familia arbitraria $\{X_i\}_{i\in I}$.

Al utilizar el segundo y tercer método queda:

Definición 3.2.8. *El Axioma de Elecciones Múltiples Contables: para cada sucesión $\{X_n\}_{n\in\mathbb{N}}$ de conjuntos no vacíos, existe una sucesión $\{F_n\}_{n\in\mathbb{N}}$ tal que $F_n \subset X_n$ y F_n es un conjunto finito no vacío.*

Otras formas de flanquear el problema son:

Definición 3.2.9. *El Principio de Elecciones Dependientes: para toda pareja (X,R) donde X es no vacío y R es una relación en X tal que $\forall x \in X$ se tiene que $\exists y \in X$ tal que xRy,existe una sucesión $\{x_n\}_{n\in\mathbb{N}}$ en X tal que x_nRx_{n+1} para cada $n \in \mathbb{N}$.*

Definición 3.2.10. *El Axioma de Elecciones Contables Parcial: para cada sucesión $\{X_n\}_{n\in\mathbb{N}}$ de conjuntos no vacíos, existe un subconjunto $M \subset \mathbb{N}$ infinito, tal que $\prod_{m\in\mathbb{M}} X_m \neq \emptyset$.*

Teorema 3.2.1. *El Principio de Elecciones Dependientes implica el Axioma de Elecciones Contables.*

Teorema 3.2.2. *El Axioma de Elecciones Contables implica que todo conjunto infinito contiene un subconjunto denumerable, o lo que es lo mismo, que todo número cardinal puede ser comparado con $\aleph_0$.*

Ninguna de las implicaciones contrarias de los dos teoremas anteriores se cumple. Ahora introduciremos algunas generalizaciones de dichos teoremas. Sea κ un aleph, y llamemos:

Definición 3.2.11 (DC$_\kappa$)**.** *Sea S un conjunto arbitrario y R una relación binaria en él tal que para todo $\alpha < \kappa$ y toda α-sucesión $s = \{x_\xi : \xi < \alpha\}$ de elementos de S existe un $y \in S$ tal que sRy. Entonces existe una función $f : \kappa \mapsto S$ tal que para toda $\alpha < \kappa$, se tiene que $(f|_\alpha)Rf(\alpha)$.*

La anterior es una generalización del Principio de Elecciones Dependientes, fue creada por Levy en 1964, llamada también *Principio de Elecciones $\aleph_\alpha$-Dependientes*.

Definición 3.2.12 (AC$_\kappa$)**.** *Para toda familia $\mathscr{X}$ de conjuntos no vacíos tal que $\mathscr{X} = \kappa$ se tiene una función de elección.*

Definición 3.2.13 (W$_\kappa$)**.** *Para todo conjunto X se cumple que $|X| \leq \kappa$ o se cumple que $|X| \geq \kappa$.*

La afirmación AC$_{\aleph_0}$ es el Axioma de Elecciones Contables, W$_{\aleph_0}$ dice que todo conjunto infinito tiene un subconjunto denumerable y DC$_{\aleph_0}$ es lo mismo que el Principio de Elecciones Dependientes.

El siguiente teorema resume varios resultados que se pueden utilizar para generalizar el Axioma, son implicaciones entre las 3 definiciones anteriores, y nos demuestra cuántos otros modelos matemáticos se pueden crear al cambiar los axiomas utilizados:

Teorema 3.2.3. *Se cumplen:*

i) *si $\kappa < \lambda$ entonces DC_λ implica DC_κ, AC_λ implica AC_κ y W_λ implica W_κ.*

ii) *DC_κ implica a las dos afirmaciones AC_κ y W_κ.*

iii) *$(\forall\kappa)DC_\kappa$ es equivalente al Axioma de Elección, así como también $(\forall\kappa)W_\kappa$ es equivalente al Axioma.*

iv) *$(\forall\kappa)AC_\kappa$ implica el Principio de Elecciones Dependientes.*

3.2.2. Buscando otros Axiomas

En esta sección daremos pocos resultados que dan otro mundo matemático, otro sistema axiomático que se parece al conocido pero con hipótesis diferentes al Axioma.

Una condición más fuerte que el Axioma es:

(3.2.1) La Hipótesis Generalizada del Continuo afirma que para cardinales infinitos n y m, las desigualdades $m \leq n < 2^m$ implican que $n = m$.

Otra condición muy ligada:

(3.2.2) La Hipótesis del Aleph, dice que $2^{\aleph_a} = \aleph_{a+1}$ para todo ordinal a.

Teorema 3.2.4. *Son equivalentes:*

i) La Hipótesis Generalizada del Continuo.

ii) La hipótesis del Aleph.

Teorema 3.2.5. *La Hipótesis Generalizada del Continuo implica el Axioma de Elección.*

Un axioma que implica al varios axiomas, entre ellos el Axioma de Elección y el Axioma del Conjunto Potencia, es el que creó Tarski para conjuntos inaccesibles.

Dicho axioma surgió debido a la duda sobre la existencia de cardinales u ordinales inaccesibles. Para cualquier conjunto de cardinales A, se cumple que existe un cardinal más pequeño que es mayor a todos los miembros de A, llamado $\sup A$. Los ejemplos son:

$$\aleph_0 = \sup\{1,\ 2\ ,\ \ldots,\ n,\ \ldots\},$$

$$\aleph_1 = \sup\{\aleph_0\}.$$

Decimos que:

Definición 3.2.14. *Un cardinal m distinto de 0 es inaccesible si:*

1. *Para todo conjunto de cardinales A tal que $|A| < m$ y que $n < m$ para cualquier $n \in A$ se cumple que $\sup A < m$.*

2. *Si $n < m$ y si $p < m$ se cumple que $n^p < m$.*

Mediante esta definición $\aleph_0$ es inaccesible, pero la pregunta era si es posible demostrar la existencia de otros números cardinales inaccesibles bajo el sistema de los axiomas de ZFC. Se demostró ya que el postulado de que no hay otros tales cardinales inaccesibles es consistente con los axiomas de NBG, y se mantiene la consistencia con los de ZF.

Para establecer la existencia de los cardinales inaccesibles Tarski propuso el:

Axioma 3.2.1 (Axioma para Conjuntos Inaccesibles)**.** *Para todo conjunto N existe un conjunto M cumpliendo:*

1. *N es equipotente a un subconjunto de M.*

2. *$\{A :\ (A \subset M) \vee (|A| < M)\}$ es equipotente a M.*

3. *No existe un subconjunto R tal que $\mathscr{P}(R)$ es equipotente a M.*

Tarski demostró que el número cardinal de un conjunto M es infinito e inaccesible ssi M satisface las condiciones 2 y 3 de su axioma.

Otra forma de encontrar alternativas es negar alguna implicación no reversible del Axioma. En un modelo como el de Feferman y Levy donde el Axioma Denumerable es falso, $\mathbb{R}$ sería la unión contable de conjuntos contables. Ahí la aditividad contable de Lebesgue fallaría, todo conjunto medible sería uno boreliano, habría conjuntos borelianos que fueran no medibles, que les faltaría la propiedad de Baire y aunque fueran incontables, no tendrían subconjunto alguno perfecto.

3.3. CARDINALES Y ORDINALES

Para esta sección ocuparemos un teorema muy importante, la demostración está en [Dug, p. 47], y nos ayudará a establecer muchos resultados:

Teorema 3.3.1 (Teorema de Bernstein-Schröder)**.** *Si existe una inyección $X \mapsto Y$ y una inyección $Y \mapsto X$, entonces existirá una biyección $X \mapsto Y$.*

Un ordinal transfinito α, se dirá que es un ordinal inicial ssi $\forall \beta < \alpha$ se tiene que $\beta \prec \alpha$. Se suelen denotar los ordinales iniciales por ω_α, con α un ordinal.

Para un número ordinal α, todos los conjuntos de tipo α, i.e. que son los similares a α, son equivalentes. El número cardinal que les corresponde es denotado por $\overline{\alpha}$, y es la potencia de α; también, $\overline{\alpha}$ es la potencia del conjunto de todos los números ordinales menores que α. Se tiene que $a \in$On, el conjunto de todos los ordinales, y un conjunto A son similares, escrito por $A \cong a$, ssi a es el tipo ordenado de A, i.e., $a = \overline{A}$. Por esta razón, se escribe a veces doble raya arriba de un conjunto para denotar su cardinalidad, ya que si $a \in$On, $\overline{a} = \overline{\overline{A}} = |A|$.

Comenzaremos a definir las operaciones aritméticas para los ordinales y los cardinales.

Dos conjuntos E y F, se pueden hacer disjuntos si se denotan por $E' = \{(x,0) : x \in E\}$ y $F' = \{(z,1) : z \in F\}$. Las funciones biyectivas $x \mapsto (x,0)$ y $z \mapsto (z,1)$ pueden transportar estructuras que tengan los conjuntos del dominio, al contradominio. Así, asumiremos que dados dos conjuntos cualesquiera, ellos son disjuntos. La generalización, se vió ya en la página 91.

Ahora, para dos conjuntos disjuntos bien ordenados, E y F, definiremos el orden de $E \cup F$ de tal forma que las parejas de elementos en E y parejas en F, retengan su orden original cada una, y que todo elemento de E anteceda a todo elemento de F. Para verlo más matemáticamente: si R y S son las respectivas relaciones de E y F, $E \cup F$ estará ordenado por $R \cup S \cup (E \times F)$, los primeros dos elementos de la unión mantienen el orden original entre elementos del mismo conjunto, y el producto antepone a todo elemento de E para los de F. Al ser bien ordenados los dos conjuntos, se tiene que $E \cup F$ está bien ordenado. A éste conjunto se le llama suma ordinal de los conjuntos bien ordenados E y F.

Generalizando: para una familia de conjuntos disjuntos bien ordenados $\{A_i : i \in I\}$, indexada por el conjunto bien ordenado I, la suma ordinal de la familia será la unión $\bigcup_{i \in I} A_i$, y para dos elementos a y b de la unión, pertenecientes a A_i y A_j respectivamente, $a < b$ significará que $i < j$ o que $i = j$ y que a antecede a b en el orden de A_i, éste será el orden de la unión.

Ahora se puede definir la suma de dos números ordinales: para todo conjunto bien ordenado X, sea ord X el único número ordinal similar a X, si α y β son números ordinales, se pueden encontrar A y B disjuntos y además bien ordenados, cumpliendo que ord $A = \alpha$ y que ord $B = \beta$, y si C es la suma ordinal de los conjuntos A y B, la suma $\alpha + \beta$, será el número ordinal de C por definición, porque ord A+ord B=ord C. Una propiedad que se cumple, es que dicha definición no depende de los representantes elegidos A y B como conjuntos con números ordinales α y β, cualesquiera otros con los mismos ordinales funcionarían.

Generalizando: sea una familia bien ordenada $\{\alpha_i : i \in I\}$ de números ordinales indexados por un conjunto bien ordenado I, sea la familia $\{A_i : i \in I\}$ tal que $\forall i \in I$ se cumple que $\operatorname{ord} A_i = \alpha_i$, y sea A la suma ordinal de $\{A_i : i \in I\}$, con ello, $\sum_{i\in I} \alpha_i = \operatorname{ord} A$. Tampoco importa en este caso la elección representante de la familia $\{a_i : i \in I\}$.

La ley conmutativa falla en esta suma definida.

El producto ordinal de dos conjuntos A y B se define como el producto cartesiano de A y B con el orden lexicográfico revertido[12]. Si $(a,b),\ (c,d) \in A \times B$, $(a,b) < (c,d)$ significa que $b < d$ o que $b = d$ y $a < c$.

Sean α y β números ordinales, A y B tales que $\operatorname{ord} A = \alpha$ y $\operatorname{ord} B = \beta$, y C el producto ordinal de A y B, el producto $\alpha\beta$ es el número ordinal de C, es decir, $(\operatorname{ord} A)(\operatorname{ord} B) = \operatorname{ord} C$.

La multiplicación puede ser usada para definir los exponentes ordinales. Pero para definir α^β mediante recursión transfinita sobre β, se escribe $\alpha^0 = 1$ y $\alpha^{\beta+1} = \alpha^\beta\alpha$, para cuando β es un número ordinal límite, $\alpha^\beta = \sup\{\alpha^\lambda : \lambda < \beta\}$.

Hay que notar que el par no ordenado 2^ω significa todas las funciones de ω a 2, pero el conjunto bien ordenado 2^ω significa la menor cota superior de la sucesión de números ordinales $2,\ 2\cdot 2,\ 2\cdot 2\cdot 2,\ \ldots$, y no son lo mismo, así que hay que tener cuidado sobre a qué se refiere nuestra notación.

Para obtener conjuntos que tengan un ordinal dado, una forma de bordear el hecho de escoger conjuntos arbitrarios con dicha propiedad, es escoger el mismo ordinal, cuyo número ordinal es él mismo.

La aritmética cardinal tiene menos dificultades.

Si a y b son números cardinales, y A y B conjuntos disjuntos con $|A| = a$ y $|B| = b$, se define $a + b = |A \cup B|$, el hecho de que si C y D son disjuntos y además $|C| = a$ y $|B| = b$, se cumple que $A \sim C$, $B \sim D$ y esto implica que $A \cup B \sim C \cup D$, se encuentra en la página 131, ahí se ocupa la G-equidescomponibilidad de conjuntos, propiedad parecida a la equipotencia de éstos. Con esta definición, la adición cardinal, es conmutativa. La adición para términos infinitos es parecida: si $\{a_i : i \in I\}$ es una familia de números cardinales y si $\{A_i : i \in I\}$ una familia disjunta tal que para todo i en I se tiene que $|A_i| = a_i$, se escribe $\sum_{i\in I} a_i = |\bigcup_{i\in I} A_i|$. La demostración de que los representantes de la familia no importa, está en [HrbJec, p. 156].

Resta de números cardinales: si a y b son números cardinales, decimos que la diferencia $a - b$ existe si existe un y solo un número cardinal p que cumple que $a = b + p$, se denota la resta por $p = a - b$.

Para la definición del producto de los números cardinales a y b, denotamos a conjuntos tales que $|A| = a$ y $|B| = b$, se escribe $ab = |A \times B|$. También se puede definir ab como la adición de a un número b de veces, es decir $\sum_{i\in I} a_i$, con $|I| = b$ y $\forall i \in I$ se cumple que $a_i = a$. La multiplicación para factores infinitos es: si $\{a_i : i \in I\}$ es una familia de números cardinales, y si $\{A_i : i \in I\}$ su correspondiente familia de conjuntos tales que $|A_i| = a_i$, se escribirá $\prod_{i\in I} a_i = |\prod_{i\in I} A_i|$. La demostración de que los representantes de la familia no importan, está en [HrbJec, p. 157].

Para la exponenciación de números cardinales a^b, primero se tienen A y B conjuntos tales que $|A| = a$ y $|B| = b$, y se define $a^b = |A^B|$, la cardinalidad del conjunto de las funciones de B a A. Otra forma de hacerlo es multiplicar a por él mismo b veces, i.e., si $|I| = b$ y $\forall i \in I$ se escribe $a_i = a$, entonces $\prod_{i\in I} a_i = a^b$.

Propiedades de la suma de cardinales son la asociatividad, la conmutatividad, y ([HrbJec, p. 94]):

1. $a \leq a + b$,

[12]Éste es su nombre porque hace resemblancia a cómo se ponen los nombres en una lista, por orden alfabético, y si dos letras se repiten, se toma la siguiente letra para comparar, el revertido significa que se hace de adelante para atrás.

2. $a_1 \leq a_2,\ b_1 \leq b_2 \Rightarrow a_1 + b_1 \leq a_2 + b_2$.

Propiedades del producto de cardinales son la asociatividad, la conmutatividad, y ([HrbJec, p. 94]):

1. $a(b+c) = ab + ac$,
2. si $b > 0$, $a \leq ab$,
3. $a_1 \leq a_2,\ b_1 \leq b_2 \Rightarrow a_1 b_1 \leq a_2 b_2$,
4. $a + a = 2a$,
5. si $a > 1$, $a + a \leq aa$.

Propiedades de la exponenciación cardinal son ([HrbJec, p. 95]):

1. si $b > 0$, $a \leq a^b$,
2. si $a > 1$, $b \leq a^b$,
3. $a_1 \leq a_2,\ b_1 \leq b_2 \Rightarrow a_1^{b_1} \leq a_2^{b_2}$,
4. $aa = a^2$,
5. $a^{b+c} = a^b a^c$,
6. $(a^b)^c = a^{ab}$,
7. $(ab)^c = a^c b^c$.

Se utilizarán algunos hechos sobre alephs para las equivalencias, algunas demostraciones que salgan del alcance de la tesis se omitirán, mientras que las relevantes se pondrán, sin hacer mención cuando se aplique el Teorema de Schröder-Bernstein.

En [Sie, p. 392] se demuestra que:

Teorema 3.3.2. *Para todo número ordinal α existe $\aleph_\alpha$.*

El siguiente teorema, nos simplificará enormemente los cálculos y varios resultados, dice que:

Teorema 3.3.3. *Para todo número ordinal α, $\aleph_\alpha \aleph_\alpha = \aleph_\alpha$.*

Demostración. La demostración será por inducción transfinita. Para todo a ordinal, se construirá un buen orden denotado por $\prec$ en el conjunto $\omega_a \times \omega_a$, se usará como hipótesis de inducción que $\aleph_b \aleph_b \leq \aleph_b$ para todo $b < a$, y que el tipo ordenado del conjunto bien ordenado $(\omega_a \times \omega_a, \prec)$ es a lo más ω_a, es decir, $\aleph_a \aleph_a \leq \aleph_a$ ($\overline{\omega_a \times \omega_a} = \overline{\omega_a \omega_a} = \aleph_a \aleph_a$), luego de eso, por la propiedad de que $a < b \Rightarrow \aleph_a < \aleph_b$ demostrada en la p. 100, y por la propiedad 2 del producto de cardinales, $\aleph_a \leq \aleph_a \aleph_a$, concluyendo que $\aleph_a \aleph_a = \aleph_a$.

El buen orden $\prec$ de $\omega_a \times \omega_a$ se define para todo ω_a como

$$(a_1, a_2) \prec (b_1, b_2) \Leftrightarrow \begin{cases} \max\{a_1, a_2\} < \max\{b_1, b_2\} \\ \max\{a_1, a_2\} = \max\{b_1, b_2\} & ,\ a_1 < b_1 \\ \max\{a_1, a_2\} = \max\{b_1, b_2\} & ,\ a_1 = b_1, a_2 < b_2. \end{cases}$$

Demostraremos que es un buen orden.

La transitividad, con $(a_1,a_2) \prec (b_1,b_2)$ y $(b_1,b_2) \prec (c_1,c_2)$, por definición $\text{máx}\{a_1,a_2\} \leq \text{máx}\{b_1,b_2\} \leq \text{máx}\{c_1,c_2\}$, por lo tanto $\text{máx}\{a_1,a_2\} \leq \text{máx}\{c_1,c_2\}$. Si $\text{máx}\{a_1,a_2\} < \text{máx}\{c_1,c_2\}$, por definición $(a_1,a_2) \prec (c_1,c_2)$; si $\text{máx}\{a_1,a_2\} = \text{máx}\{b_1,b_2\} = \text{máx}\{c_1,c_2\}$, entonces $a_1 \leq b_1 \leq c_1$, por la segunda definición de $\prec$. Hay dos casos otra vez, si $a_1 < c_1$, se cumple $(a_1,a_2) \prec (c_1,c_2)$ por la segunda definición de $\prec$; el otro caso es $a_1 = b_1 = c_1$, aquí, los máximos son iguales y las primeras componentes también, por la tercera definición de $\prec$ es cierto que $a_2 < b_2 < c_2$, i.e., $a_2 < c_2$, por eso $(a_1,a_2) \prec (c_1,c_2)$.

Para la tricotomía y su propiedad de ser mutuamente excluyente: $(a_1,a_2) \prec (b_1,b_2)$ o $(b_1,b_2) \prec (a_1,a_2)$ o $(a_1,a_2) = (b_1,b_2)$, comparamos dados (a_1,a_2) y (b_1,b_2), los máximos de las componentes, luego se comparan los ordinales que están en las primeras componentes, y luego en las segundas, y así se cae en algún caso de los tres.

Para demostrar que $(\omega_a \times \omega_a, \prec)$ es un buen orden, solo falta demostrar que cualquier conjunto no vacío X de parejas de ordinales, tiene un elemento $\prec$-menor en X. Utilizaremos el hecho mencionado en la p. 33, de que todo conjunto de ordinales está bien ordenado, y la idea será ir construyendo elementos menores conforme la definición de $\prec$. Considerando el conjunto de ordinales $A = \{\text{máx}\{a,b\} | (a,b) \in X\}$, sea c su primer elemento respecto al orden natural de los ordinales. Sea

$$Y = \{(a,b) \in X | \text{máx}\{a,b\} = c\},$$

que es no vacío porque $c \in A$ y $Y \subset X$. Para todos los $(a,b) \in Y$, por definición de Y $\text{máx}\{a,b\} = c$, además para $(a_1,b_1) \in X - Y$, se cumple $c < \text{máx}\, a_1, b_1$ por la definición de c. Por lo tanto, se cumple que para $(a,b) \in Y$ y $(a_1,b_1) \in X - Y$, $(a,b) \prec (a_1,b_1)$ por la primera condición de $\prec$, lo que significa que si existe el primer elemento de Y, también lo será de X.

Sea d el primer ordinal del conjunto de ordinales bien ordenado $\{a | (a,b) \in Y \text{ para algún } b\}$, y sea

$$Z = \{(a,b) \in Y | a = d\},$$

que es no vacío porque $(d,b) \in Z$ si $(d,b) \in Y$,además, $Z \subset Y$. Cuando $(a,b) \in Z$ y $(a_1,b_1) \in Y - Z$, es válido que $(a,b) \prec (a_1,b_1)$, debido a que si sus máximos son iguales, su primer elemento es distinto por la definición de d.

Por último y análogamente, e será el primer elemento del conjunto $\{b | (d,b) \in Z\}$, el elemento buscado es (d,e), que es el $\prec$-menor en Z y al juntar las conclusiones anteriores también en X .

Ahora probaremos por inducción transfinita sobre a, que $|\omega_a \times \omega_a| \leq \aleph_a$, o lo que es lo mismo, $\aleph_a \aleph_a \leq \aleph_a$.

$\aleph_0 \aleph_0 \leq \aleph_0$ es verdadero porque los números racionales son numerables. Sea $a > 0$, y supongamos que para todo $b < a$, $\aleph_b \aleph_b \leq \aleph_b$. Se tiene que demostrar que la propiedad también se cumple para a. De la propiedad que si a es ordinal, el tipo ordenado de a es igual a la potencia del conjunto de todos los ordinales menores que a, se sigue que si el tipo ordenado del conjunto $\omega_a \times \omega_a$ fuera mayor que del de ω_a, existiría un $(a_1,a_2) \in \omega_a \times \omega_a$, tal que el conjunto X de todos los ordinales que se relacionan con (a_1,a_2) por medio de $\prec$, tendría cardinalidad mayor que $\aleph_a$. Por eso, mediante la negación de la conclusión y la hipótesis, basta probar que para todo $(a_1,a_2) \in \omega_a \times \omega_a$, se cumple $|X| < \aleph_a$, con $X = \{(b_1,b_2) \in \omega_a \times \omega_a | (b_1,b_2) \prec (a_1,a_2)\}$.

Sea $b = \text{máx}\{a_1,a_2\} + 1$, $b \in \omega_a$, y para todos los $(b_1,b_2) \in X$, se cumple por la definición de X que $\text{máx}\{b_1,b_2\} \leq \text{máx}\{a_1,a_2\} < b$, esto último por ser el sucesor, concluyendo que $b_1 \in b$ y también $b_2 \in b$, por ser menores. Como fueron arbitrarios, $X \subset b \times b$. Agarremos cualquier $g < a$ que cumpla $|b| \leq \aleph_g$, este existe porque como $b \in \omega_a$, $|b| \leq \aleph_a$, y solo elegimos uno en medio. Operando, $|X| \leq |b \times b| = |b||b| \leq \aleph_g \aleph_g \leq \aleph_g$, la primera desigualdad es cierta porque la cardinalidad es monótona, la segunda por definición de multiplicación, la tercera por la propiedad 3 del producto de cardinales y la propiedad de g, y la última por la hipótesis de inducción, por lo tanto, $|X| \leq \aleph_g < \aleph_a$ que se demostró para la tricotomía de alephs en la página 100, demostrando lo que se requería.

La conclusión de la inducción nos dice que $|\omega_a \times \omega_a| \leq \aleph_a$, por lo tanto, $\forall a \in \mathrm{On}$ se tiene que $\aleph_a \aleph_a \leq \aleph_a$, concluimos que $\aleph_a \aleph_a = \aleph_a$. ☐

Corolario 3.3.1. *$\aleph_a \aleph_b = \aleph_b$, con $a \leq b$.*

Demostración. La desigualdad $a \leq b$ implica $\aleph_a \leq \aleph_b$, por la propiedad 3 del producto de cardinales $\aleph_a \aleph_b \leq \aleph_b \aleph_b = \aleph_b$, y por la misma propiedad $\aleph_b \leq \aleph_b \aleph_a$, demostrando lo pedido. ☐

Corolario 3.3.2. *$\aleph_b + \aleph_a = \aleph_b$ si se tiene que $a \leq b$.*

Demostración. La desigualdad $a \leq b$ implica $\aleph_a \leq \aleph_b$, la propiedad 2 de la suma de cardinales nos dice que

$$\aleph_a + \aleph_b \leq \aleph_b + \aleph_b = 2\aleph_b \leq \aleph_b \aleph_b = \aleph_b,$$

primera igualdad por la propiedad 4 del producto, la siguiente desigualdad por la propiedad 4 del producto, y la última por el teorema anterior. Del otro lado, la propiedad 1 de la suma de cardinales, nos da como resultado que $\aleph_b \leq \aleph_b + \aleph_a$, concluyendo lo pedido. ☐

Corolario 3.3.3. *Para todo número ordinal α y $n \in \omega$, $n + \aleph_\alpha = \aleph_\alpha$.*

Demostración. Se tiene que $\aleph_\alpha \leq n + \aleph_\alpha \leq \aleph_\alpha + \aleph_\alpha \leq \aleph_\alpha \aleph_\alpha = \aleph_\alpha^2 = \aleph_\alpha$, donde se utilizaron las propiedades 5 del producto de cardinales, la 4 de la exponenciación cardinal para la primera y la segunda igualdades. ☐

Corolario 3.3.4. *Para todo número ordinal $\alpha \leq \beta$, $\aleph_\alpha^{\aleph_\beta} = 2^{\aleph_\beta}$.*

Demostración. Como $\alpha \leq \beta$, se cumple siempre que $2 < \aleph_\alpha < 2^{\aleph_\alpha} \leq 2^{\aleph_\beta}$, por el teorema de Cantor de la p. 34, y la propiedad 3 de la exponenciación cardinal, y luego, $2^{\aleph_\beta} \leq \aleph_\alpha^{\aleph_\beta} \leq (2^{\aleph_\beta})^{\aleph_\beta} = 2^{\aleph_\beta^2} = 2^{\aleph_\beta}$, por la propiedad 3 de exponenciación, el primer resultado, la propiedad 6 de exponenciación, la 4, y el teorema anterior respectivamente. ☐

Corolario 3.3.5. *Para todo número ordinal α y $n \in \omega$, $\aleph_\alpha^n = \aleph_\alpha$.*

Demostración. Se demuestra por inducción, cuando $n = 2$, el teorema anterior afirma que es cierto, supongámoslo cierto para $n \in \omega$, i.e., se cumple que $\aleph_\alpha^n = \aleph_\alpha$, usamos la propiedad 5 de la exponenciación de cardinales $\aleph_\alpha^{n+1} = \aleph_\alpha^n \aleph_\alpha = \aleph_\alpha \aleph_\alpha = \aleph_\alpha$, otra vez por el teorema anterior. Se demuestra así lo que se quería. ☐

Corolario 3.3.6. *Para todo número ordinal α y $n \in \omega$, $n\aleph_\alpha = \aleph_\alpha$.*

Demostración. Por las propiedades 2 y 3 de la multiplicación de cardinales, y por el teorema anterior:

$$\aleph_\alpha \leq n\aleph_\alpha \leq \aleph_\alpha \aleph_\alpha = \aleph_\alpha.$$

☐

Teorema 3.3.4. *Para todo número ordinal $\beta < \alpha$, $\aleph_\alpha - \aleph_\beta = \aleph_\alpha$.*

Demostración. Sea a un ordinal dado, y n un número cardinal con $n < \aleph_a$, que existe porque $\aleph_a$ es un cardinal infinito por definición. Supongamos por el momento[13] que existe un número cardinal p tal que $\aleph_a = n + p$. Si n es finito, p no es finito porque el aleph no es finito, y como $p \leq \aleph_a$, por lo escrito en la p. 99, p es un aleph, digamos $p = \aleph_b$. Por el corolario 3.3.3, si n es finito, $n + \aleph_b = \aleph_b = \aleph_a$, y $p = \aleph_a$. Si n no es finito, como $n < \aleph_a$, $n = \aleph_c$, y $\aleph_c < \aleph_a \Rightarrow c < a$, esto nos afirma que el número cardinal p no puede ser finito, porque pasaría que $\aleph_a = \aleph_c + p = \aleph_c$, contradiciendo que $\aleph_c < \aleph_a$; por eso $p = \aleph_b$. Cuando $b \leq c$, se cumple que $\aleph_a = \aleph_b + \aleph_c = \aleph_c$ por el corolario 3.3.2, contradiciendo la hipótesis de nuevo, por eso $b > c$, y vemos que $\aleph_a = \aleph_c + \aleph_b = \aleph_b$, por eso $p = \aleph_a$, en cualquier caso, $p = \aleph_a$.

Para $n < \aleph_a$, siempre existe un número cardinal p tal que $\aleph_a = n + p$, y precisamente es $p = \aleph_a$.

Como conclusiones, de la definición de resta de cardinales para todo cardinal $n < \aleph_a$, $\aleph_a - n = \aleph_a$ existe, y que si $b < a \Rightarrow \aleph_a - \aleph_b = \aleph_a$. □

Para los siguientes teoremas, se asume que $\mathfrak{n}$ no es finito.

Teorema 3.3.5. *Son equivalentes:*

i) *El Axioma de Elección.*

ii) *Si* $0 < \mathfrak{m} \leq \mathfrak{n}$, *entonces* $\mathfrak{m} + \mathfrak{n} = \mathfrak{m}\mathfrak{n} = \mathfrak{n}$

Demostración. $i) \Rightarrow ii)$ Sabemos que el Axioma es equivalente al hecho de que todo cardinal infinito es un aleph, y por el corolario 3.3.1 y el corolario 3.3.2, se cumplen las dos igualdades.

$ii) \Rightarrow i)$ Si esa igualdad se cumple (que sabemos no se cumple para cardinales finitos), y que ocurre $\neg(m + n > n)$ y que $\neg(m + n > m)$, la propiedad 1 de la suma de cardinales nos dice que $m + n \geq m$ y que $m + n \geq n$, por lo tanto, $m + n = n$ implica que $m \leq n$ y $m + n = m$ que $n \leq m$; por eso, cualesquiera dos números cardinales infinitos se pueden comparar, i.e., se cumple la Tricotomía de los Cardinales, demostrada ya equivalente al Axioma en el teorema 3.1.9. □

Teorema 3.3.6. *Son equivalentes:*

i) *El Axioma de Elección.*

ii) *Cuando* $\mathfrak{m} > \mathfrak{n}$, *se tiene que la diferencia* $\mathfrak{m} - \mathfrak{n}$ *siempre existe, y además* $\mathfrak{m} - \mathfrak{n} = \mathfrak{m}$.

Demostración. $i) \Rightarrow ii)$ Supongamos que m es finito, entonces la diferencia existe. Sea m un número cardinal infinito. Para demostrar la existencia de la diferencia es necesario y suficiente que exista un y solo un número cardinal p con $m = n + p$, debido a la definición de diferencia.

Como $n \leq m$, de la primera de estas equivalencias que demostramos, $m = n + m$, solo falta demostrar que si un número cardinal p cumple que $m = n + p$ con $n < m$, se tendrá que $p = m$. Como $m = n + p$, sabemos que $p \leq m$, supongamos que $p < m$, el número m será la suma de dos números menores que él, contradiciendo que la suma de dos cardinales no finitos es el mayor de ellos, por lo tanto $p = m$.

$ii) \Rightarrow i)$ Si se cumple la proposición que la diferencia de $m - n$ existe para todos los números cardinales n, m con $n < m$, demostraremos que todo cardinal infinito es un aleph, y por lo tanto, el Axioma de Elección.

Sea m un número cardinal infinito. De la definición del número de Hartog de m, que no ocupa el Axioma, existe un aleph $h(m)$, que denotaremos por conveniencia por $\aleph(m)$, para entender algunas propiedades de éste.

[13] Se demuestra después que si existe.

Dicho $\aleph(m)$ cumple que no es cierto $\aleph(m) \leq m$ por definición, y se cumple por la propiedad 1 de la suma de cardinales, que $\aleph(m) \leq m + \aleph(m)$. Si fuera ésta una desigualdad estricta, nuestra hipótesis asegura que existe el número $(m + \aleph(m)) - \aleph(m)$, y un único número cardinal p con $m + \aleph(m) = p + \aleph(m)$, esta formula es cierta para $p = m$ pero también para $p = m + \aleph(m)$, porque sustituyendo queda $\aleph(m) + \aleph(m) = \aleph(m)$, demostrado en la p. 108 como el segundo corolario. Ya que el número p es único, $m = m + \aleph(m)$, por eso $\aleph(m) \leq m$, contradiciendo la propiedad del número de Hartogs $\aleph(m)$. Por lo tanto, no se cumple que $m + \aleph(m) > \aleph(m)$, y como se da la igualdad de éstos, implica que $m \leq \aleph(m)$, que por lo ya dicho en la p. 99, un subconjunto de un aleph es un aleph.

El suponer la hipótesis cierta, nos lleva a que todo cardinal infinito es un aleph, y por el teorema 3.1.6 se cumple el Principio del Buen Orden, y como sabemos, el Axioma. □

Teorema 3.3.7. *Son equivalentes:*

i) El Axioma de Elección.

ii) Si $\mathfrak{m} < \mathfrak{n}$ *y* $\mathfrak{m}_1 < \mathfrak{n}_1$*, se cumple* $\mathfrak{m} + \mathfrak{m}_1 < \mathfrak{n} + \mathfrak{n}_1$.

Demostración. $i) \Rightarrow ii)$ Supongamos que el Axioma y la hipótesis de $ii)$ se cumplen: si los números m y m_1 son finitos, el número $m + m_1$ también, siendo menor que $n + n_1$ si es finito o si al menos alguno de los dos es infinito.

Supongamos que m o m_1 son infinitos, es posible que lo sean los dos. Como el Axioma implica la Tricotomía de Cardinales, $m \leq m_1$ o $m_1 \leq m$, sin pérdida de generalidad, si se cumple la primera, se tiene que m_1 es no finito, porque si lo fuera, también lo sería m. Por la primera de las equivalencias que demostramos $m + m_1 = m_1$, y por la hipótesis más la primera propiedad de sumas de cardinales $m + m_1 = m_1 < n_1 < n + n_1$, que es lo que se quería demostrar.

$ii) \Rightarrow i)$ Sea m un número cardinal que cumple no ser finito, y sea $n = m\aleph_0$ infinito.

Se cumple que

$$2n = 2m\aleph_0 = m(2\aleph_0) = m\aleph_0 = n,$$

porque la multiplicación cardinal es conmutativa y por el último corolario 3.3.6. Para dicho n existe su número de Hartogs $h(n) = \aleph(n)$, cuya existencia no depende del Axioma y que no cumple que $\aleph(n) \leq n$. Se tiene $\aleph(n) \leq n + \aleph(n)$ por la propiedad 1 de la suma de cardinales.

Supongamos que la desigualdad es estricta, que se cumple $\aleph(n) < n + \aleph(n)$, el objetivo es llegar a una contradicción. Como $n \leq n + \aleph(n)$, si $n = n + \aleph(n)$ se tendría que $\aleph(n) \leq n$, lo cuál no puede ser por la definición de $\aleph(n)$, por lo tanto $n < n + \aleph(n)$; por la hipótesis, sumamos éstas dos desigualdades obteniendo

$$n + \aleph(n) < (n + \aleph(n)) + (n + \aleph(n)) = 2n + 2\aleph(n),$$

porque la suma es conmutativa y asociativa. Del corolario 3.3.6, y de la propiedad demostrada para n

$$n + \aleph(n) < n + \aleph(n),$$

lo cuál nunca pasa porque significa que $a \in a$ con a un número ordinal, por lo que el Axioma de Regularidad no se cumpliría. Por ésto, no es cierto que $\aleph(n) < n + \aleph(n)$, y se obtiene $\aleph(n) = n + \aleph(n)$, implicando $n \leq \aleph(n)$ o lo que es lo mismo $m\aleph_0 \leq \aleph(n)\aleph(n) = \aleph(n)$ por el teorema 3.3.3. Llegamos a que el número cardinal m que supusimos no finito es un aleph porque es un subconjunto de uno.

Demostramos así, que todo cardinal infinito es un aleph, que por el teorema 3.1.6, es equivalente al Axioma. □

Teorema 3.3.8. *Son equivalentes:*

i) *El Axioma de Elección.*

ii) *Si* $\mathfrak{m} < \mathfrak{n}$ *y* $\mathfrak{m}_1 < \mathfrak{n}_1$*, se tiene entonces que* $\mathfrak{m}\mathfrak{m}_1 < \mathfrak{n}\mathfrak{n}_1$.

Demostración. $i) \Rightarrow ii)$ La demostración es casi análoga que la pasada, reproduzco lo que necesitamos:

Si el Axioma y la hipótesis de $ii)$ se cumplen: si los números m y m_1 son finitos, el número $m+m_1$ también, y será menor que $n+n_1$ si es finito, o si al menos alguno de los dos es infinito.

Supongamos que m o m_1 son infinitos, pudiendo serlo los dos. Como el Axioma implica la Tricotomía de Cardinales, $m \leq m_1$ o $m_1 \leq m$, sin pérdida de generalidad, si se cumple la primera, se tiene que m_1 es no finito, porque si lo fuera, también lo sería m. Por el primer corolario 3.3.1, la hipótesis de $ii)$ y la segunda propiedad de multiplicación cardinal, se cumple $mm_1 = m_1 < n_1 \leq nn_1$, y por transitividad $mm_1 < nn_1$, que es lo requerido.

$ii) \Rightarrow i)$ Sea el número cardinal m infinito y formemos al número cardinal $n = m^{\aleph_0}$. Por la propiedad 6 de exponenciación y el último corolario 3.3.6, se cumple $n^2 = m^{2\aleph_0} = m^{\aleph_0} = n$. La propiedad 2 del producto de cardinales, dice que $n \leq n\aleph(n)$ con $\aleph(n)$ el número de Hartogs de n. Supondremos que la desigualdad es estricta, y arrivaremos a una contradicción.

Si $\aleph(n) = n\aleph(n)$, siendo el producto mayor que ese número, se debe cumplir $\aleph(n) \leq n$, contradiciendo la definición de $\aleph(n)$, por lo que $\aleph(n) < n\aleph(n)$, con ésta y suponiendo la desigualdad $n < n\aleph(n)$, mas la hipótesis nos dan:

$$n\aleph(n) < n\aleph(n)n\aleph(n) = n^2(\aleph(n))^2,$$

porque la multiplicación es conmutativa, mientras que las propiedades de n y de $\aleph(n)$ (la última por el corolario 3.3.5) darían que $n\aleph(n) < n\aleph(n)$, lo cuál nunca sucede.

Para no llegar a contradicción, se debe tener $n\aleph(n) = n$, lo cuál significa que $n \leq n\aleph(n)$ y que $m \leq m^{\aleph_0} = n = n\aleph(n)$, ésto por la primera propiedad de exponenciación cardinal, la definición de n y la segunda del producto respectivamente; diciéndonos que m como número ordinal infinito es un aleph, que es equivalente al Axioma. □

Teorema 3.3.9. *Son equivalentes:*

i) *El Axioma de Elección.*

ii) $\mathfrak{m}^2 = \mathfrak{m}$.

Demostración. $i) \Rightarrow ii)$ El Axioma es equivalente a la primera equivalencia demostrada, poniendo $\mathfrak{m} = \mathfrak{n}$, se tiene lo pedido.

$ii) \Rightarrow i)$ Supongamos que se cumple para todo cardinal infinito, y tomemos uno denotado por m, sea $n = h(m)$ su número de Hartog, que también es infinito debido a que m lo es. Se cumplen las ecuaciones:

$$m^2 = m,$$
$$n^2 = n,$$
$$(m+n)^2 = m+n,$$
$$m+n = m+2mn+n,$$

esta última porque la multiplicación y suma con conmutativas y asociativas. De la propiedad 2 del producto de cardinales, la propiedad 1 de la suma de éstos, y la transitividad, se tiene que

$$mn \leq 2mn \leq m+2mn+n = m+n \Rightarrow mn \leq m+n.$$

Como la hipótesis se cumple para todos los cardinales infinitos, también es cierta para cuando m y n son sucesores de algunos números cardinales m_1 y n_1 respectivamente; con ellos, se convierte $mn = (m_1+1)(n_1+1) = m_1n_1 + m_1 + n_1 + 1 \geq 1 + m_1 + n_1 + 1 = m+n$ debido a la propiedad 3 de la multiplicación cardinal y la definición de m_1 y n_1. Por lo tanto se cumple $m+n = mn$.

Tomemos dos conjuntos M y N que cumplan tener cardinalidad m y n respectivamente, se puede suponer que son disjuntos, y como $n = \aleph(m)$, el conjunto N está bien ordenado porque su potencia es un aleph, este hecho lo ocuparemos para tomar elementos menores.

Sea $P = M \times N$, por definición de multiplicación:

$$|P| = mn,$$

y ya que $m+n = mn$:

$$|P| = m+n,\ P = M_1 \cup N_1,$$

con $|M_1| = m$ y $|N_1| = n$ y dichos conjuntos disjuntos, por ello, $M \times N = M_1 \cup N_1$.

Tomamos un elemento del lado izquierdo y fijamos la primera componente, cuando se varía la componente de la derecha del elemento del producto, también cae en el producto por la definición de éste, y al ser iguales, estará toda esa colección en solo uno de M_1 o N_1 por ser disjuntos. De forma simétrica, si tomamos un elemento del producto $M \times N$, y fijamos la componente de la derecha y variamos el de la izquierda, esos elementos están en el producto, y deberán estar en solo uno de los M_1 o N_1. Por esto, resultan dos casos: hay un elemento $m_1 \in M$ tal que para todo $n \in N$, se cumple que $(m_1, n) \in M_1$; o que para todo $m \in M$, exista al menos un $n \in N$ con $(m,n) \in N_1$.

Cuando ocurre el primer caso, si denotamos a $N_2 = \{(m_1,n) | n \in N\}$, se tiene que $|N_2| = |N| = n$ porque $1 \cdot a = a$, además $N_2 \subset M_1$ por estar en el primer caso, por lo tanto $n = |N| \leq |M_1| = m$, pero no puede ocurrir ésto porque la definición de n es $n = \aleph(m)$, y no se cumple que $\aleph(m) \leq m$.

En el segundo caso, si $m \in M$, sea $f(m)$ el primer elemento del conjunto bien ordenado N, tal que $(m,n) \in N_1$ (un m fijo, da al menos un n que cumpla que $(m,n) \in N_1$ porque estamos en el segundo caso). Sea $M_2 = \{(m, f(m)) | m \in M\}$, eso dice que $M_2 \subset N_1$, y que $|M_2| = |M|$, por las casi mismas razones que lo anterior, por eso, $m = |M| = |M_2| \leq |N_1| = n$, concluyendo que $m \leq n$.

Se cumple así que $m \leq \aleph(m)$, y significa que el cardinal arbitrario infinito m es un aleph, por lo tanto, por el teorema 3.1.6, se tiene una equivalencia del Axioma. □

Teorema 3.3.10. *Son equivalentes:*

i) El Axioma de Elección.

ii) Si se cumple que $\mathfrak{m}^2 = \mathfrak{n}^2$*, entonces debe de ser que* $\mathfrak{m} = \mathfrak{n}$.

Demostración. $i) \Rightarrow ii)$ Se sigue del corolario 3.3.5 de esta sección (ya que al elevar al cuadrado un aleph, éste es igual al aleph), y del hecho de que todo cardinal infinito es un aleph.

$ii) \Rightarrow i)$ Tomemos un número cardinal infinito m, y formemos estos otros 3:

$$n = m^{\aleph_0},\ p = n + \aleph(n) \text{ y } q = n\aleph(n).$$

Calcularemos los cuadrados de éstos:

$$n^2 = (m^{\aleph_0})^2 = m^{2\aleph_0} = m^{\aleph_0} = n,$$

que se cumple ocupando la propiedad 6 de exponenciación, el corolario 3.3.6, y la definición de n. Luego:

$$p^2 = (n + \aleph(n))^2 = n^2 + 2n\aleph(n) + (\aleph(n))^2$$
$$= n + n2\aleph(n) + \aleph(n) = (n + \aleph(n)) + n\aleph(n),$$

por las propiedades de conmutatividad, la anterior propiedad de n, el corolario 3.3.5, el corolario 3.3.6 y la asociatividad.

Por la propiedad 5 del producto de cardinales $n + \aleph(n) \leq n\aleph(n)$, ayudando a demostrar que:

$$n\aleph(n) \leq (n + \aleph(n)) + n\aleph(n) \leq n\aleph(n) + n\aleph(n) = n(2\aleph(n)) = n\aleph(n),$$

concluyendo que $n\aleph(n) = (n + \aleph(n)) + n\aleph(n) = p^2$.

Por último:

$$q^2 = (n\aleph(n))^2 = n^2(\aleph(n))^2 = n\aleph(n) = p^2,$$

por la propiedad 7 de la exponenciación de cardinales, y el corolario 3.3.5. Por la hipótesis $p = q$, es decir:

$$n + \aleph(n) = n\aleph(n),$$

pero en la demostración anterior, se probó que ésta igualdad (tomando a $n = \aleph(m)$), conduce a que todo número cardinal infinito es un aleph, equivalente al Axioma. □

3.4. ÁLGEBRA

Las siguientes definiciones se obtuvieron de [RubRub, p. 93-103], y se requieren para demostrar algunos teoremas de la sección:

$\mathscr{A} = (A, \{O_k : k \in K\})$ es llamada un álgebra (abstracta) o un sistema algebraico si A es un conjunto y para cada $k \in K = \mathbb{N}$, se cumple que O_k es una operación n-aria definida en A.

Una 0-aria operación es una constante y para $n > 0$, una operación n-aria es una función $O_k : A^n \mapsto A$.

A $\mathscr{B} = (B, \{O_k|_B : k \in K\})$ se le llamará subálgebra de A si $B \subset A$ y si además B es cerrado con respecto a las operaciones O_k para $k \in K$, i.e., si O_k es una operación n-aria, entonces para todos los $b_1, b_2, \ldots, b_n \in B$, se cumple que $O_k(b_1, b_2, \ldots, b_n) \in B$. Cuando $\mathscr{B} \neq \mathscr{A}$ se llama a $\mathscr{B}$ subálgebra propia de $\mathscr{A}$.

Sea L un conjunto, con 2 operaciones binarias sobre L: $\wedge$ y $\vee$, una *lattice* es $\mathscr{A} = (L, \{\vee, \wedge\})$ si se satisfacen para todos los $a, b, c \in L$:

1. L es un conjunto y $\vee$ y $\wedge$ son operaciones binarias,

2. $a \vee b \in L$ y también $a \wedge b \in L$,

3. $a \vee a = a$ y también $a \wedge a = a$,

4. se cumple que $a \vee b = b \vee a$ y $a \wedge b = b \wedge a$,

5. $a \vee (b \vee c) = (a \vee b) \vee c$ y $a \wedge (b \wedge c) = (a \wedge b) \wedge c$,

6. $a \vee (a \wedge b) = a$ y $a \wedge (a \vee b) = a$,

$1 \in L$ será el elemento identidad si $\forall a \in L$, $a \wedge 1 = a$

$0 \in L$ será el elemento nulo si $\forall a \in L$, $a \vee 0 = a$

Se llamará ideal a $E \subset L$ si siempre que $a,\ b \in E$, se cumple que $a \vee b \in E$, y siempre que $a \in L$ y $b \in E$ se tiene $a \wedge b \in E$. Si $E \neq L$ se llamará ideal propio, y cumplirá que $1 \notin E$.

Se llamará a $F \subset L$ un filtro si siempre que $a,\ b \in F$, se cumpla que $a \wedge b \in F$, y siempre que $a \in L$ y $b \in F$ se cumpla que $a \vee b \in F$. Si $F \neq L$ se llamará filtro propio, y se llamará ultrafiltro si es un filtro propio $\subset$-maximal.

Se dice que $E \subset L$ es un ideal primo si E es un ideal y si $\forall a,\ b \in L$ que cumplen $a \wedge b \in E$, se tendrá que $a \in E$ o que $b \in E$.

Un tipo especial de lattice es un álgebra booleana:

$\mathscr{B} = (B, \{\vee, \wedge\})$ es un álgebra booleana si se satisface:

1. $\mathscr{B}$ es una lattice,

2. $a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$, de manera similar $a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$, para todos los $a,\ b,\ c \in B$.

3. 0 y 1 están en B.

4. Para todo $a \in B$ existe $a^c \in B$ tal que $a \vee a^c = 1$ y también $a \wedge a^c = 0$.

Si se tiene $\mathscr{A} = (A, \{O_k :\ k \in K\})$ un sistema algebraico:

$B \subset A$ se dice que genera a A si para cada $a \in A$ se cumple que $a \in B$ o que existe un natural $n > 0$, $k \in K$ y $b_1,\ b_2,\ \ldots,\ b_n \in A$ tales que B genera $\{b_1,\ b_2,\ \ldots,\ b_n\}$ y que $a = O_k\{b_1,\ b_2,\ \ldots,\ b_n\}$.

$B \subset A$ es algebráicamente dependiente si existen un número natural $n > 0$, $k \in K$, $b \in B$ y $a_1,\ a_2,\ \ldots,\ a_n \in A$ que cumplen con que $\{b\}^c$ genera $\{a_1,\ a_2,\ \ldots,\ a_n\}$ y $b = O_k(a_1,\ a_2,\ \ldots,\ a_n)$. Si no se cumple, B es llamado algebráicamente independiente.

$B \subset A$ será una base para A si B es algebráicamente independiente y genera a A. Una base es un subconjunto de A que es algebráicamente independiente y $\subset$-maximal.

Un grupoide es un álgebra con una operación binaria $(A, \{o\})$.

Un grupoide cancelativo $(A, \{o\})$ es un grupoide que satisface las leyes de cancelación: $\forall a,\ b,\ c \in A$, se cumple que $aob = aoc \Rightarrow b = c$ y $aoc = boc \Rightarrow a = b$.

Un grupoide abeliano (o conmutativo) es $(A, \{o\})$ que satisface la ley asociativa: $\forall a,\ b,\ c \in A$ se tiene que $ao(boc) = (aob)oc$.

Un cuasigrupo $(A, \{o\})$ es un grupoide cancelativo que cumple con: $(\forall a,\ b \in A)(\exists x,\ z \in A)$ tales que $aox = b$ y $zoa = b$.

Un loop es un álgebra $(A, \{o, e\})$ con $(A, \{o\})$ un cuasigrupo y e el elemento identidad: $\forall a \in A$ se tiene que $aoe = eoa = a$

Un grupo es un álgebra $(A, \{o, e\})$ con $(A, \{o\})$ un cuasigrupo, o es asociativo y e es la identidad.

Un anillo es un álgebra $(A, \{+, *, 0\})$ con $(A, \{+, 0\})$ un grupo abeliano, $*$ asociativa y las leyes distributivas se mantienen: $\forall a,\ b,\ c \in A$ se tiene $a * (b + c) = a * b + a * c$ y $(a + b) * c = a * c + b * c$.

Un anillo será conmutativo si $*$ es conmutativo.

Un dominio entero $(A, \{+, *, 0\})$ es un anillo conmutativo el cual cumple: $\forall a,\ b \in A$ con $a,\ b \neq 0 \Rightarrow a * b \neq 0$.

Un campo es un álgebra $(A, \{+, *, 0, 1\})$ con $(A, \{+, *, 0\})$ un anillo conmutativo y $(A - \{0\}, \{*, 1\})$ un grupo abeliano.

Se dirá que C es un operador de cerradura finita (en inglés, finitary closure operator) si es una operación 1-aria monótona[14] que cumple que para toda clase X:

$$C(X) = \bigcup\{C(z) : \ z \subset X \text{ y } z \text{ es finito}\}.$$

Cuando se cumple que $C(X) \subset X$ se dirá que X es C-cerrado.

3.4.1. Espacios Vectoriales

En esta subsección daremos las definiciones necesarias primero, y después demostraciones. Algunos teoremas de esta subsección, se demuestra son equivalentes al Axioma, pero debido a que varios autores suponen cierto el Axioma, los llaman teoremas, por ello, aquí no se demuestran como tales, sino que se demuestran dichas equivalencias.

Si tenemos un campo $\big(F, \{+, *, 0, 1\}\big)$, el álgebra $\big(V, \{+', 0'\} \cup \{O_a : \ a \in F\}\big)$ será llamada un espacio vectorial sobre F ssi se cumplen:

1. $\big(V, \{+', 0'\}\big)$ es un grupo abeliano,
2. $(\forall a, b \in F)(\forall v \in V) O_{a+b}(v) = O_a(v) +' O_b(v)$,
3. $(\forall a \in F)(\forall u, v \in V) O_a(u +' v) = O_a(u) +' O_a(v)$,
4. $(\forall a, b \in F)(\forall v \in V) O_{a*b}(v) = O_a(O_b(v))$,
5. $(\forall v \in V) O_1(v) = v$.

Se usa en los textos actuales la notación av en vez de $O_a(v)$, indicando la multiplicación del escalar a (elemento del campo) por un vector v (elemento del espacio vectorial). Un subconjunto W de un espacio vectorial V se dice que es un subespacio vectorial de V (o más brevemente, subespacio de V) si es una subálgebra de V. Para demostrar que $W \subset V$ es subespacio vectorial, basta demostrar que para todos los $u, v \in W$ y todos los $a \in F$ se tiene que $av + u \in W$.

Un subconjunto A de un espacio vectorial V se dice linealmente dependiente si $(\exists n \in \omega)(\exists v_1, \ldots, v_n \in A)(\exists a_1, \ldots, a_n \in F) 0 = a_1 v_1 + \cdots + a_n v_n$ y para algún i se tiene que $a_i \neq 0$. De la otra manera, si $(\forall n \in \omega)(\forall v_1, \ldots, v_n \in A) 0 = a_1 v_1 + \cdots + a_n v_n \Rightarrow a_i = 0$ para toda $i \in \{1, 2, \ldots, n\}$ se dirá que A es linealmente independiente.

Todo subconjunto B de un conjunto linealmente independiente a es linealmente independiente (si el conjunto es él mismo, trivialmente se cumple), ya que, utilizando la notación del párrafo anterior, si $0 = a_1 v_1 + \cdots + a_n v_n$ entonces $a_i = 0$ para toda $i \in \{1, 2, \ldots, n\}$, siendo $B \subset A$, y $k < n$, al poner $a_i = 0$ para toda $i \in \{k, k+1, \ldots, n\}$, las igualdades $0 = a_1 v_1 + \cdots + a_k v_k = a_1 v_1 + \cdots + a_k v_k + 0 \cdot v_{k+1} + \cdots + 0 \cdot v_n$, dicen que se debe cumplir que $a_i = 0$ para toda $i \in \{1, 2, \ldots, k\}$, por lo que el subconjunto B es linealmente independiente.

Se le llama a $A \subset V$ generador o que genera a V si $(\forall v \in V)(\exists n \in \omega)(\exists v_1, \ldots, v_n \in A)(\exists a_1, \ldots, a_n \in F) v = a_1 v_1 + \cdots + a_n v_n$, y a dicha v se le llama combinación lineal de los v_i´s.

Un subconjunto $B \subset V$ es una base de V si es linealmente independiente y genera a V.

Si tenemos subespacios A, $B \subset V$ de un espacio vectorial, la suma de los subespacios es:

$$A + B = \{\alpha = \beta + \gamma | \beta \in A, \ \gamma \in B\},$$

[14]Véase el índice alfabético para la definición de función monótona.

y cuando se tiene que $A \cap B = \emptyset$, se dirá que $A \oplus B$ es la suma directa de los subespacios A y B.

Sea $A \subset V$ subespacio del espacio vectorial V, el espacio complementario de A es B si $A \oplus B = V$. Esta definición se ocupará aquí, aunque a veces se definen los subespacios complementarios de diferente manera.

Ocuparemos el resultado que un campo sobre un subcampo es un espacio vectorial sobre dicho subcampo.

También necesitamos un poco de terminología de polinomios, sacada del gran libro «Linear Algebra» [HofKun, p. 117-119]:

Sea F un campo. Se dice que un álgebra lineal sobre el campo F es un espacio vectorial L sobre F con la operación adicional de multiplicación de vectores, que se simboliza por uv para todos los $u,\ v \in L$ y que cumple:

1. la multiplicación es asociativa, i.e. $t(uv) = (tu)v$,
2. la multiplicación es distributiva respecto a la adición, i.e. $t(u+v) = tu + tv$ y también $(u+v)t = ut + vt$,
3. para todo escalar $c \in F$ se tiene que $c(tu) = (ct)u = t(cu)$.

Si tenemos a F un campo y a el conjunto de todos lo enteros no negativos $\mathbb{N}$, y consideramos al conjunto de todas las funciones de $\mathbb{N}$ a F, éste será un espacio vectorial, denotado por F^∞, que contendrá elementos de la forma $(f_0,\ f_1,\ f_2, \ldots)$. A el vector $(0,0,\ldots,0,1,0,\ldots)$ que tiene un uno en la n-ésima entrada (la n-ésima componente del vector) y todas las demás cero, se le denotará por x^n, con $x^0 = 1$. Tenemos que con esto, un elemento $f = (f_0,\ f_1,\ f_2,\ldots)$ de F^∞ se escribe de la forma $f = \sum_{n=0}^{\infty} f_n x^n$. Si $a,b \in F$ y $f,g \in F^\infty$ se operará a $af+bg$ y a fg respectivamente:

$$af + bg = (af_0 + bg_0, af_1 + bg_1, af_2 + bg_2, \ldots)$$

$$(fg)_n = \sum_{i=0}^{n} f_i g_{n-i} \text{ con } n = 0,1,2,\ldots$$

donde $(f)_n$ es la n-ésima entrada de f.

Cuando se tiene $F[x]$ el subespacio de F^∞ que es generado por los vectores $1,\ x,\ x^2,\ \ldots$, a los elementos de $F[x]$ se le llaman polinomios sobre F, también se dice que son todas las combinaciones lineales finitas de x y sus potencias. Si se tiene $f \in F[x]$ no nulo, éste será de la forma $f = f_0x^0 + f_1x^1 + f_2x^2 + \cdots + f_nx^n$, con $f_n \neq 0$, a los escalares f_i se les llamará coeficientes de f, y se dirá monomio y no polinomio cuando $f = f_nx^n$, con $f_n \neq 0$. Si n es el último natural para el cual $f_n \neq 0$ y para los siguientes naturales los coeficientes son cero, se dirá que n es el grado del polinomio.

El conjunto de todos los polinomios sobre un campo dado F junto con las operaciones de adición y multiplicación dadas aquí, será un álgebra lineal conmutativa con identidad sobre F. Y cuando se tienen divisiones de polinomios, es decir, $p = \frac{f}{g}$ con $f,g \in F[x]$, se dirá que son polinomios racionales, con la división definida por

$$p(x) = \left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)}.$$

Teorema 3.4.1. *Son equivalentes:*

1. *A1: Todo espacio vectorial tiene una base.*
2. *Axioma de Elección.*

Demostración. $i) \Rightarrow ii)$ Demostraremos que $i)$ implica AMC (definición 3.2.4). Sea $(X_i)_{i\in I}$ una familia de conjuntos disjuntos y no vacíos, y tomemos a k un campo arbitrario, y a $k[X]$ el campo de las funciones racionales en las variables $x \in X = \bigcup_{i\in I} X_i$ sobre el campo k (es decir, cada $f \in F$ es un polinomio racional con «variables» en X). Definiremos los grados de éstos polinomios. Para monomios, es decir, elementos p de $k[X]$ de la forma $p = \alpha x_1^{n_1} x_2^{n_2} \cdots x_m^{n_m}$, definimos para cada $i \in I$ el i-ésimo grado de p como $d_i(p) = \sum_{x_k \in X_i} n_k$ (i.e., será la suma de los exponentes de los elementos del monomio que aparezcan en X_i). Para los elementos $p \in k[X]$, que serán de la forma $p = \dfrac{p_1 + \cdots + p_n}{q_1 + \cdots + q_m}$ con cada p_k y q_k monomios, se dirá que son i-homogéneos de grado d si todos los q_k tienen el mismo i-grado, digamos d_1, y todos los p_k tienen el mismo i-grado que será $d_2 = d + d_1$, sin éstos ser restringidos a los números naturales.

Formemos el conjunto

$$K = \{a \in k[X] : \ a \ \text{ es } i\text{-homogéneo de grado 0 para cada } i \in I\},$$

éste estará compuesto de divisiones de polinomios del mismo grado, y será un subcampo de $k[X]$, por lo que $k[X]$ será un espacio vectorial sobre K.

La hipótesis nos afirma que $k[X]$ tiene una base B, implicando que todo elemento $x \in X$ (en realidad x es un monomio) puede ser expresado de manera única de la forma

$$x = \sum_{b \in B(x)} a_b(x) b$$

donde $B(x) \subset B$ es un conjunto finito, y $\forall b \in B(x)$ se tiene que $a_b(x) \in K - \{0\}$, que son los que fungen como escalares del monomio. Ésta representación es única debido a que tenemos una base, la combinación lineal es finita por la definición de base, y además los escalares son no nulos debido a que si lo fueran, no se necesitaría de ese vector de la base que lo multiplica para generar a x, pudiéndolo descartar y tomando el subconjunto de $B(x)$ que utiliza escalares no nulos en su representación. Para todos los $x,\ y \in X_i$ se tendrá que

$$y = \frac{y}{x} x = \sum_{b \in B(x)} \frac{y}{x} a_b(x) b = \sum_{b \in B(y)} a_b(y) b.$$

Demostraremos que para todos $x,\ y \in X_i$ se cumple que $B(x) = B(y)$ y ocupando ésta igualdad $\forall b \in B(x)$ se cumplirá que $\frac{a_b(x)}{x} = \frac{a_b(y)}{y}$. Primero se tiene que $\frac{y}{x} \in K$ debido a que es 1-homogéneo de grado cero por tener el mismo 1-grado en denominador y numerador; restando de la ecuación de arriba:

$$\sum_{b \in B(x)} \frac{y}{x} a_b(x) b - \sum_{b \in B(y)} a_b(y) b = 0,$$

sin pérdida de generalidad, suponiendo que existen $m \geq 1$ vectores $b_1,\ b_2, \ldots, b_m \in B(y)$ tales que no son iguales a algún vector de $B(x)$, í.e., $\forall b \in B(x)$ se tiene que $b_1,\ b_2, \ldots, b_m \neq b$, y:

$$\sum_{b \in B(x)} \left(\frac{y}{x} a_b(x) - a_b(y) \right) b - \sum_{b \in \{b_1,\ \ldots, b_m\}} a_b(y) b = 0,$$

pero sabiendo que $B(x), B(y) \subset B$ son subconjuntos de una base, serán linealmente independientes los vectores b de la combinación lineal anterior ya que quitamos a los repetidos y por la definición de independencia lineal, todos los escalares son cero, es decir:

$$\forall b \in B(x) : \frac{y}{x} a_b(x) - a_b(y) = 0 \text{ y también } \forall b \in \{b_1, \ldots, b_m\} : a_b(y) = 0,$$

la segunda ecuación nos lleva a una contradicción, debido a que $\forall b \in B(x)$ se tiene que $a_b(x) \in K - \{0\}$, por lo tanto, no se cumple lo que supusimos falso, y se cumple que $B(x) = B(y)$. Con esto, la primera ecuación nos lleva a que $\forall b \in B(x)$ se cumple que $\frac{a_b(x)}{x} = \frac{a_b(y)}{y}$. Debido a que x y y fueron arbitrarios, podemos renombrar $B(x) = B_i$ y $\frac{a_b(x)}{x} = \alpha(b,i)$. El hecho de que $a_b(x) \in K - \{0\}$ nos dice que éste es i-homogéneo de grado 0, por eso, $\alpha(b,i)$ por tener denominador un monomio de grado uno, es i-homogéneo de grado -1 (menos uno).

Al escribir a $\alpha(b,i)$ como cociente de polinomios en su forma reducida (recuerde que aunque $a_b(x)$ sea de grado 0, puede ser el cociente de polinomios de grados iguales y no necesariamente escalares), algún $x \in X_i$ ocurrirá en el denominador, por lo tanto, si llamamos a F_i el conjunto de todos los $x \in X_i$ que aparecen en el denominador de $\alpha(b,i)$ para algún $b \in B_i$, es no vacío y está contenido en X_i, concluyendo que $\forall i \in I$ se tiene que $\emptyset \neq F_i \subset X_i$.

$ii) \Rightarrow i)$ La demostración es casi igual a la última de ésta sección, que nos dice que M3 implica A4, debido a que cualquier conjunto con un solo elemento es un conjunto linealmente independiente, visto en la demostración de que A4 implica A2, por lo tanto encontramos una base para V. $\square$

Teorema 3.4.2. *Son equivalentes:*

1. *A2: Todo espacio vectorial V sobre k un campo, tiene la propiedad de que si $S \subset V$ es un subespacio de V, existirá un subespacio $S' \subset V$ cumpliendo que $S \cap S' = \{0\}$ y que $S \cup S'$ genera a V.*

2. *Axioma de Elección.*

Demostración. $i) \Rightarrow ii)$ Demostraremos que $i)$ implica AMC. Sea $(X_i)_{i\in I}$ una familia de conjuntos disjuntos y no vacíos, y tomemos para toda $i \in I$, $V_i = k^{(X_i)}$ la suma directa de X_i copias de k, utilizando la base canónica $(e_x)_{x\in X_i}$ (el vector e_x es 1 en su x-ésima coordenada, y en las demás cero). Si S_i es el subespacio de V_i consistiendo de todos los $v = \sum\limits_{x\in X_i} \alpha_x e_x$, con $\alpha_x \in k$, que cumplan con que $\sum\limits_{x\in X_i} \alpha_x = 0$, se cumplirá que es un subespacio de V_i, ya que si $r \in k$, y $v,\ u \in S_i$, se tendrá:

$$ru + v = r(\sum_{x\in X_i} \alpha_x e_x) + \sum_{x\in X_i} \beta_x e_x = \sum_{x\in X_i} (r\alpha_x + \beta_x) e_x,$$

y para demostrar que cumple la condición de S_i, sumamos sus escalares

$$\sum_{x\in X_i} (r\alpha_x + \beta_x) = r\sum_{x\in X_i} \alpha_x + \sum_{x\in X_i} \beta_x = r0 + 0 = 0,$$

ya que $v,\ u \in S_i$, cumpliendo lo pedido. Un ejemplo de un elemento en S_i será $v = \sum_{x\in X_i} \alpha_x e_x = e_x - e_y$, porque $\alpha_x + \alpha_y = 1 - 1 = 0$.

Tomemos la suma directa $S = \oplus S_i$ con $i \in I$ de subespacios como subespacio del espacio $V = \oplus V_i$ con $i \in I$, la hipótesis nos confirma la existencia de S', que cumple $S \cap S' = \{0\}$ y $S + S' = V$. Para cualquier $x \in X_i$, el elemento de la base canónica de V_i considerado como elemento de V será ocupado, ésto para la existencia de dos elementos únicos $s(x) \in S$ y $s'(x) \in S'$ que cumplen $e_x = s(x) + s'(x)$. Tomemos ahora dos elementos arbitrarios $x,\ y \in X_i$, la resta $s'(x) - s'(y)$ está en S' ya que es una combinación lineal de elementos de S', además

$$s'(x) - s'(y) = (e_x - s(x)) - (e_y - s(y)) = (e_x - e_y) + (s(y) - s(x)) \in S$$

debido a que $e_x - e_y \in S_i$ para algún i, por el ejemplo proporcionado, y por lo tanto $e_x - e_y \in S$, por eso, para los $s(y) - s(x) = 0$ implicaría que $s'(x) - s'(y)$ estaría en S. Encontramos así un elemento $s'(x) - s'(y)$ que está en la intersección de los subespacios S y S', por lo tanto la hipótesis dice que para cualesquiera $x,\ y \in X_i$ se tiene $s'(x) = s'(y) = s'_i$ (renombrando).

Tomemos esos elementos $s'_i = \sum_{j\in I} v_j$ del espacio, con cada $v_j \in V_j$, porque están en el espacio se cumple $v_i = \sum_{x\in X_i} \alpha_x e_x$. Así, si definimos a $F_i = \{x \in X_i | \alpha_x \neq 0\}$, probaremos que es finito y no vacío para toda $i \in I$.

Para cada $x \in X_i$, vimos que

$$s(x) = e_x - s'_i = e_x - \sum_{j\in I} v_j = (e_x - v_i) - \sum_{j\neq i, j\in I} v_j,$$

por construcción, $s(x) \in S$ y debido a que en la última ecuación hay una combinación lineal de elementos que no están en V_j (el segundo sumando) y los dos sumandos están en $S = \oplus_{i\in I} S_i$, se tendrá que $e_x - v_i \in S_i$. Por la condición de los elementos de S_i,

$$e_x - v_i = e_x - \sum_{y\in X_i} \alpha_y e_y = \sum_{y\in X_i} \beta_y e_y,$$

con $\beta_x = 1 - \alpha_x$ y $\forall y \neq x$ se tiene que $\beta_y = -\alpha_y$; como $\sum\limits_{x\in X_i} \beta_x = 0$ se cumple que:

$$0 = 1 - \alpha_x - \sum_{y\neq x, y\in X_i} \alpha_y = 1 - \sum_{y\in X_i} \alpha_y,$$

con eso demostramos que existen $\alpha_x \neq 0$. Ellos son finitos porque las combinaciones lineales de los vectores con elementos de la base son finitas. Por lo tanto $\emptyset \neq F_i \subset X_i$.

$ii) \Rightarrow i)$ La demostración es esencialmente la misma que A4 implica A2 de aquí abajo, tenemos un subespacio de V, encontramos una base de él, ya que todo espacio vectorial tiene base (ya demostrado), la extendemos a una base de V. El espacio complementario sería el generado por los vectores que están en la base de V y que no están en la base de S. □

Teorema 3.4.3. *A3: Para todo campo F y todo espacio vectorial V sobre F, todo subconjunto de V que genera a éste, contiene una base de V.*

Teorema 3.4.4. *A3 implica A1.*

Demostración. Si tenemos un espacio vectorial V sobre un campo F, y el mismo conjunto lo tomamos como subespacio $V \subset V$ que genera a V (es subespacio porque es cerrado bajo suma de vectores y multiplicación por escalares por definición de espacio, genera a V porque todo elemento de V se puede expresar con elementos de V, a saber, él mismo), por hipótesis, V tiene una base, y lo que demuestra éste argumento general es que todo espacio vectorial tiene base. □

Teorema 3.4.5. *A4: Todo espacio vectorial V sobre un campo F, tiene la propiedad de que si $S \subset V$ es un subespacio de V linealmente independiente, S podrá ser extendido a una base de V.*

Teorema 3.4.6. *A4 implica A2.*

Demostración. Supongamos que tenemos un espacio vectorial V sobre F un campo, y que $S \subset V$ es un subespacio (dicho existe, por ejemplo el conjunto con el elemento nulo es uno). Si el subespacio es nulo, su complemento será todo el espacio y se cumplen las propiedades de la conclusión. Si no, tomemos cualquier elemento no nulo $b \in S$, el conjunto $\{b\}$ es linealmente independiente debido a que si $a \in F$ y $ab = 0$, necesariamente $a = 0$.

La hipótesis nos dice que existe una base B_1 en S, debido a que extendemos el conjunto linealmente independiente encontrado antes. De esta manera, encontramos un subconjunto linealmente independiente de V que es B_1, por ello, aplicamos otra vez la hipótesis, y concluimos que encontramos una base B de V, Si llamamos a S' el subespacio generado por los elementos de $B - B_1$, tenemos que $S \cap S' = \{0\}$ debido a que uno es generado por B_1 y el otro por $B - B_1$; además, $S \cup S'$ genera a V porque B genera a V. $\square$

Para la siguiente demostración, nos valemos de una definición equivalente de base, tomada de [Fri, p. 58], dice que un subconjunto $B \subset S$ con $S \subset V$ un subconjunto de V un espacio vectorial es un subconjunto máximo linealmente independiente de S si:

(3.4.1) B es linealmente independiente,

(3.4.2) cualquier subconjunto de S que contenga propiamente a B es linealmente dependiente.

Denotando a $L(S)$ el espacio generado por el conjunto S, otro resultado que necesitamos es (p. 42):

(3.4.3) Sea S un subconjunto linealmente independiente de un espacio vectorial V, con $x \in V$ y $x \notin S$. $S \cup \{x\}$ es linealmente dependiente ssi $x \in L(S)$.

Y el último será (p. 36):

(3.4.4) Si $S_1 \subset S_2$ son subconjuntos de un espacio vectorial V, se tiene que $L(S_1) \subset L(S_2)$. En particular, cuando $L(S_1) = V \Rightarrow L(S_2) = V$.

Teorema 3.4.7. *A4 implica A3.*

Demostración. Tenemos un espacio vectorial V sobre un campo F, y un subespacio $S \subset V$ que genera a V, de la misma manera que la demostración anterior, aplicamos la hipótesis y encontramos un subconjunto B de S que es base para S y éste será un subconjunto máximo linealmente independiente de S, debido a que las definiciones son análogas. Demostraremos que B es una base para V.

Como es linealmente independiente, falta demostrar que genera a V. Si S contiene propiamente al espacio que genera B, sea $x \in S$ y $x \notin L(B)$, por el resultado de arriba $B \cup \{x\}$ es linealmente independiente, contradiciendo la maximidad de B, porque $B \cup \{x\}$ es un conjunto que lo contiene propiamente. Por eso $S \subset L(B)$, y ya que la hipótesis nos dice que $L(S) = V$. Se concluye que $L(B) = V$ por el último resultado de [Fri] visto (el inmediato anterior a este teorema). $\square$

Para probar que son equivalencias el Axioma, A1, A2, A3 y A4, falta unir éstos con M3, que se demostró antes (teorema 3.1.4 p. 93) ser equivalente al Axioma:

Teorema 3.4.8. *M3 implica A4.*

Demostración. Si tenemos un espacio vectorial V sobre un campo F, y un subespacio $S \subset V$ que es linealmente independiente, ocuparemos a T la familia de todos los subconjuntos linealmente independientes de V que contienen a S. Para aplicar la hipótesis, debemos demostrar que si tenemos un nido N en T, la unión $\bigcup N \in T$, o lo que es lo mismo, que $\bigcup N$ es un subconjunto linealmente independiente y que contiene a S.

Si $a \in N \Rightarrow S \subset a$ (por la definición de los elementos de T), por lo tanto $S \subset \bigcup N \subset V$. Si tenemos vectores $u_1, \ldots, u_n \in \bigcup N$ y escalares $a_1, \ldots, a_n \in F$ y se cumple $a_1 u_1 + \cdots + a_n u_n = 0$, debido a que $u_i \in \bigcup N$, se cumple $\exists A_i \in N$ con $u_i \in A_i$. Al ser una cadena N, habrá un A_j que contenga a todos los demás subconjuntos A_i, por lo tanto, $u_1, \ldots, u_n \in A_j$, éste es un conjunto linealmente independiente porque está en T, por eso los escalares a_i son cero, diciendo que $\bigcup N \in T$.

Aplicando M3 se concluye con un subconjunto maximal de V que además es linealmente independiente (porque está contenido en T), es la extensión del subconjunto S, y es una base de V. $\square$

3.4.2. Ideales y Filtros

Teorema 3.4.9. *Son equivalentes:*

i) *AL1: Teorema Maximal Ideal para Lattices. Toda lattice con unidad y al menos otro elemento diferente de la unidad, tiene un ideal $\subset$-maximal propio.*

ii) *AL-1: Todo ideal propio de una lattice con un elemento identidad puede ser extendido a un $\subset$-maximal ideal propio.*

iii) *AL-2: Si A es un sistema algebraico, B una subálgebra, $a \in A$ y además $a \notin B$, existirá una subálgebra $\subset$-maximal que tiene a B como subconjunto pero no tiene a a como elemento.*

iv) *AL-3: Si x es un conjunto arbitrario, C un operador de cerradura finita, P una propiedad de carácter finito, y z un subconjunto C-cerrado de x que tiene la propiedad P, entonces, existirá un subconjunto de x que será C-cerrado y $\subset$-maximal, tendrá la propiedad P y contendrá a z como subconjunto.*

v) *Axioma de Elección*

Por comodidad y para seguir la notación de [RubRub], el primer punto se denota AL1, el segundo AL-1, el tercero AL-2 y el cuarto AL-3. Se demostrará M3⇒AL-3⇒AL-2⇒AL-1⇒AL1⇒M7, mientras que en la subsección 3.1.3 se demostró la equivalencia de M3 y M7 con M1, y por lo tanto, con el Axioma, cerrando el círculo.

Demostración. $M3) \Rightarrow AL-3)$ Sea X un conjunto, P una propiedad de carácter finito, C una operación de cerradura finita, y por último, Y un subconjunto de X que es C-cerrado y cumple $P(Y)$. Tenemos que demostrar que $\exists w$ tal que $Y \subset w \subset X$ que es $\subset$-maximal en X, C-cerrado y además $P(w)$.

Llamemos a $k = \{t : \ t \subset X, \ t \text{ es } C\text{-cerrado} \ \text{ y } P(t)\}$. Ahora, escogiendo un subconjunto no vacío $n \subset k$ que sea un nido, hay que hacer que $\bigcup n \in k$. Como k solo consta de subconjuntos de X, se tiene $n \subset k \subset \mathscr{P}(X)$, ocupamos el hecho de que $\bigcup \mathscr{P}(X) = X$ para cualquier conjunto (debido a que la unión tiene como elementos a los elementos de los subconjuntos de X) y que la operación unión es una operación monótona, para concluír que $\bigcup n \subset X$. Al demostrar que $P(\bigcup n)$, que $\bigcup n$ es C-cerrado y ocupando la hipótesis, habremos encontrado el subconjunto w que es $\subset$-maximal.

Debemos demostrar que $P(\bigcup n)$. Sea $u \subset \bigcup n$ finito, el hecho de que n esté ordenado por inclusión, demuestra que hay un elemento de n al que u es subconjunto (i.e., $u \subset h \in n$), pero todo elemento de n cumple P porque $n \subset y$, entonces $P(u)$; acabamos de demostrar, que todo subconjunto finito de $\bigcup n$ cumple P, por la propiedad de carácter finito tenemos que $P(\bigcup n)$, y $P(\emptyset)$ también es cierta porque P es de carácter finito.

Ahora, hay que demostrar que $\bigcup n$ es C-cerrado. Necesitamos saber que si $a \in n \Rightarrow C(a) \subset a$ debido a que $n \subset y$, y como en el párrafo anterior, que si $t \subset \bigcup n$ y t finito, $t \subset h \in n$. Con ello, si $b \in C(\bigcup n) = \bigcup\{C(t) : \ t \subset \bigcup n \text{ y } t \text{ finito}\}$ por definición de C, tenemos que $b \in C(t)$ con t finito y $t \subset \bigcup n \Rightarrow t \subset h \in n \Rightarrow C(t) \subset t$, concluyendo que si $b \in C(\bigcup n) \Rightarrow \exists t \subset \bigcup n$ con $b \in C(t) \subset t \subset \bigcup n$, y eso dice que $\bigcup n$ es C-cerrado.

Así, demostramos que $\bigcup n \in k$ y como se satifacen las hipótesis de M3, se cumple que k tiene un elemento $\subset$-maximal m, y ya que $m \subset k$, entonces C-cerrado, $P(m)$. Por último $Y \subset m$ ya que $Y \in k$.

$AL-3) \Rightarrow AL-2)$ Supongamos que tenemos un sistema algebraico $\mathscr{A} = (A, \{O_k : \ k \in K\})$. Si se tiene $x \subset A$, definimos a $C(x)$ como: $a \in C(x)$ ssi existen $a_1, \ \ldots, \ a_n \in x$ y para algún $k \in K$ se tiene que $a = O_k(a_1, \ \ldots, \ a_n)$.

Demostraremos que O_k es una operación finitaria $\forall k \in K$. $O_k(\emptyset)$ es cierta ya que es una constante, además, si tenemos un subconjunto finito $x \subset A$, se cumplirá que $O_k(x)$ es cierta porque $O_k : A^n \mapsto A$.

Demostraremos que C es un operador de cerradura finita. Hay que probar que $C(X) = \bigcup\{C(z) : z \subset X \text{ y } z \text{ es finito}\}$ para todo $X \subset A$. Si $X \subset A$ arbitrario y $a \in C(X)$, existen $a_1, \ldots, a_n \in X$ tales que $a = O_k(a_1, \ldots, a_n)$, entonces $a_1, \ldots, a_n \in \{a_1, \ldots, a_n\}$, por ello y que $a = O_k(a_1, \ldots, a_n)$, se cumple $a \in C(\{a_1, \ldots, a_n\})$, por lo tanto $a \in \bigcup\{C(z) : z \subset X \text{ y } z \text{ es finito}\}$ $\big($i.e., $C(X) \subset \bigcup\{C(z) : z \subset X \text{ y } z \text{ es finito}\}\big)$ porque está en al menos uno. Para la contención contraria, si a está en la unión, estará en al menos un conjunto $C(z)$, con z finito y $z \subset X$, entonces existirán $a_1, \ldots, a_n \in z \subset X$ que cumplan $a = O_k(a_1, \ldots, a_n)$, juntando ésta igualdad con la pertenencia de cada a_i en X, se tiene que $a \in C(X)$.

Demostraremos que B es subálgebra de A ssi B es C-cerrado. Para demostrar que B es C-cerrado, como B es subálgebra de A, se cumple que $B \subset A$, y que $\forall a_1, \ldots, a_n \in B$ se cumple $O_k(a_1, \ldots, a_n) \in B$. Hay que demostrar que $C(B) \subset B$, si $a \in C(B) \Rightarrow$ existen $a_1, \ldots, a_n \in B$ que cumplen $a = O_k(a_1, \ldots, a_n)$, y como O_k es cerrada $a \in B$, demostrando una implicación. Para la otra implicación, como $C(B) \subset B$, sean $a_1, \ldots, a_n \in B$ tales que $a = O_k(a_1, \ldots, a_n)$, por la definición de $C(B)$ tenemos que $a \in C(B) \subset B$, por lo tanto, $a \in B$, lo que hemos demostrado, es que para cualesquier $a_1, \ldots, a_n \in B$, se tiene que $O_k(a_1, \ldots, a_n) \in B$, por ello B es una subálgebra.

La propiedad de excluir un elemento de un conjunto es una propiedad de carácter finito, ya que a todo subconjunto finito del conjunto dado, se le puede aplicar la propiedad, incluso si es vacío.

Concluimos que AL-2 es un caso particular de AL-3, con la propiedad P como no tener un elemento, y de ahí la implicación.

$AL-2) \Rightarrow AL-1)$ La conclusión es un caso especial de AL-2. AL-2 pide que haya un $a \in A$ y un $a \notin B$, diciendo que B es una subálgebra propia, y si se tiene un ideal propio E, éste será una subálgebra propia debido a que es cerrado bajo $\vee$ y $\wedge$, además, la propiedad de tener un ideal propio $\subset$-maximal que sea una extensión es un caso particular de tener una subálgebra propia $\subset$-maximal conteniendo a la subálgebra dada.

$AL-1) \Rightarrow AL1)$ Si tenemos una lattice con un elemento identidad y otro diferente a éste, en el caso de no tener ideal propio no vacío, el vacío será el ideal maximal propio buscado, si tiene al menos un ideal propio, se llega a la sentencia AL-1, por lo que se puede extender a un ideal maximal propio.

$AL1) \Rightarrow M7)$ Supongamos que tenemos un conjunto x y una propiedad de carácter finito P. Tenemos que demostrar que existe un subconjunto $\subset$-maximal de x que tiene la propiedad P. Asumiremos que x no cumple P y que al menos algún subconjunto no vacío de x la cumple, si no la cumpliera, en el primer caso x sería el subconjunto maximal que buscamos, en el segundo, el conjunto vacío sería dicho maximal. Definamos

$$y = \{t : t \subset x \text{ y } P(t)\} \cup \{x\},$$

si $s, t \in y$, se definirá

$$s \wedge t = s \cap t$$

$$s \vee t = \begin{cases} s \cup t & \text{si se cumple que } P(s \cup t) \\ x & \text{en cualquier otro caso.} \end{cases}$$

De esta manera obtenemos una lattice $(y, \{\vee, \wedge\})$, con elemento identidad x. Lo anterior es cierto debido a que se satisfacen las condiciones de la p. 113, ya que se cumplen para los conjuntos. Por hipótesis, $(y, \{\vee, \wedge\})$ tiene un ideal propio maximal, que llamaremos z, demostraremos a continuación que $\bigcup z$ es un subconjunto maximal de x que cumple con la propiedad P (es lo que se requiere para M7). y tiene solo como elementos subconjuntos de x y aparte a x, también tenemos $z \subset y$, por lo tanto, $z \subset y \subset \mathscr{P}(x)$, por lo que $\bigcup z \subset \bigcup \mathscr{P}(x) = x$. Escogamos cualquier subconjunto finito $t \subset \bigcup z$, existirán

n elementos $a_1, \ldots, a_n \in z$ que cumplen $t \subset \bigcup_{i=1}^{n}\{a_i\}$ debido a que t es finito. Como z es un ideal propio de y, $x \notin z$ (si no, $z = \{x\} \cup x \supset y$), por lo tanto, $\bigcup_{i=1}^{n}\{a_i\} \in z$ (un ideal es cerrado bajo $\vee$), pero de aquí se concluye por la definición de $\vee$ y por que $z \subset y$ que $P(\bigcup_{i=1}^{n}\{a_i\})$. Debido a ésto, se cumple $P(t)$, por ser un subconjunto finito de $\bigcup_{i=1}^{n}\{a_i\}$, pero acabamos de demostrar que $P(\bigcup z)$ es cierto porque todo subconjunto finito de dicha unión cumple P, por lo tanto, también $\bigcup z$ por ser P de carácter finito. Del hecho de que z es maximal, se sigue que $\bigcup z$ es maximal, ya que si w cumple P y también $z \subset w \subset y \subset \mathscr{P}(x)$ tenemos $w = z$, por eso si $\bigcup z \subset \bigcup w \subset \bigcup y \subset x$, $\bigcup z = \bigcup w$, demostrando lo pedido.

□

Un hecho importante para nuestros propósitos es saber que en un Álgebra Booleana un ideal será un ideal primo ssi es maximal, la demostración de ello está en la página 172 de [GivHal]. Con ello, tendremos que los teoremas demostrados para ideales maximales aplican a ideales primos, siempre y cuando se trabajen en álgebras Booleanas.

Hay que notar que por la dualidad de los ideales y los filtros, si se reemplaza en AL1 y AL-1 las palabras identidad por nulo e ideal propio $\subset$-maximal por ultrafiltro propio, se tendrán algunas equivalencias al Axioma. El siguiente teorema no es equivalente al Axioma:

Teorema 3.4.10. *Teorema Booleano del Ideal Primo*[15]*. Toda Álgebra Booleana tiene un ideal primo propio.*

Demostración. Acabamos de ver que un ideal es ideal primo ssi es maximal (en un Álgebra Booleana), además, por definición, un Álgebra Booleana es un tipo especial de Lattice, por lo tanto, el Teorema Maximal Ideal para Lattices, se aplica aquí, modificando «Álgebra Booleana» con «Lattice con elemento unidad y al menos otro elemento», por lo que existe un ideal maximal propio en el Álgebra Booleana, i.e., un ideal primo propio. □

El siguiente teorema es una particularización del segundo punto del primer teorema de ésta sección, dice que todo ideal propio de una lattice con identidad puede ser extendido a un ideal $\subset$-maximal propio.

Teorema 3.4.11 (Teorema Ideal Maximal)**.** *Todo ideal propio de una Álgebra Booleana puede ser extendido a un ideal primo propio.*

Mediante las dualidades, se tienen estas dos consecuencias:

Teorema 3.4.12. *Toda Álgebra Booleana tiene un ultrafiltro.*

Teorema 3.4.13. *Todo filtro de una Álgebra Booleana puede ser extendido a un ultrafiltro.*

En los libros se definen los conceptos de Álgebras Booleanas y Lattices dependiendo del enfoque que se quiera tomar, el enfoque puede ser de Teoría del Orden o de Teoría de Conjuntos, y algunos autores dicen que se utiliza el más cómodo para obtener los resultados requeridos. Presentaremos el enfoque de Teoría del Orden.

En un Álgebra Booleana B, cuando se requiere trabajar con un orden parcial, "$\leq$" se introduce en lugar de "+" por medio de

$$u \leq v \Leftrightarrow u + v = v.$$

Un ideal I en B, será un subconjunto no vacío y propio de B que satisfaga

1. $u \in I$ y $v \leq u$, entonces $v \in I$,

[15]Se abrevia BPI, por sus siglas en inglés, Boolean Prime Ideal Theorem.

2. $u \in I$ y $v \in I$, entonces $u + v \in I$.

Un ideal I será propio ssi $1 \notin I$.
Se llamará un ideal primo a un ideal I si

1. $\forall u \in B$ se cumplirá que $u \in I$ o que $-u \in I$,

Los conceptos duales de dichos ideales e ideales primos serán los de filtros y ultrafiltros, respectivamente:

1. $u \in F$ y $v \geq u$, entonces $v \in F$,

2. $u \in F$ y $v \in F$, entonces $u * v \in F$,

para un filtro F, mientras que para un ultrafiltro F se tendrá que

1. $\forall u \in B$ se cumplirá que $u \in F$ o que $-u \in F$,

Una definición alterna de filtro y ultrafiltro para conjuntos es que un filtro F sobre un conjunto S será un subconjunto propio no vacío de la familia de todos los subconjuntos de S tal que:

1. $X \in F$ y $X \subset Y$, entonces $Y \in F$,

2. $X \in F$ y $Y \in F$, entonces $X \cap Y \in F$, y un ultrafiltro será un filtro tal que:

3. para cada $X \subset S$, se cumplirá que $X \in F$ o que $S - X \in F$.

Con ello, al considerar al álgebra como el conjunto $\mathscr{P}(S)$ con las operaciones de unión, intersección y complemento, las dos definiciones de filtros y ultrafiltros coinciden, y como ya se vio para Álgebras Booleanas, también se cumplirá este

Teorema 3.4.14. *Teorema del Ultrafiltro. Todo filtro sobre un conjunto S puede ser extendido a un ultrafiltro.*

La afirmación de que para todo conjunto no vacío S existe un ultrafiltro sobre S es cierta sin necesidad del Axioma, ya que siempre será un ultrafiltro sobre S si $p \in S$

$$\{X \subset S : \; p \in X\},$$

a éste se le llama ultrafiltro principal sobre S y se denota por (p), el siguiente teorema requiere del Axioma, debido a que depende del Teorema Ideal Principal, pero primero necesitamos unos resultados, cada uno indicado de la página donde fue tomado, del libro [GivHal]:

El supremo de un subconjunto de un Álgebra Booleana, está definido igual que el supremo de un orden parcial, por que se pueden ver las dos definiciones de Álgebra Booleana mediante conjuntos y ordenes. Al supremo de un subconjunto $E \subset B$ de un Álgebra Booleana se le denota por $\bigvee E$ (p. 40).

Un átomo de una Álgebra Booleana es un elemento que no tiene subelementos no triviales propios. Si $q \neq 0$ es un átomo, solo hay dos elementos p que cumplen $p \leq q$ que son 0 y q. Para caracterizar a un átomo $q \in B$ un Álgebra Booleana se utiliza

(3.4.1) son equivalentes (p. 117):

1. q es un átomo,

2. para todo p se cumple solo una de las dos: $q \leq p, \ q \leq p^c$.

De esas equivalencias se siguen los siguientes lemas:

(3.4.2) Si un elemento $p \in B$ es el supremo de un conjunto arbitrario E de átomos, E es el conjunto de todos los átomos a_i tales que $a_i \leq p$.

Este lema nos dice que para un supremo de un conjunto de átomos, necesariamente ese conjunto tiene a todos los átomos menores que el supremo, no solo a los átomos «dados» (p. 118).

Cuando todo elemento distinto de 0 domina al menos un átomo (i.e. $\forall p \in B \exists a_i$ tal que $a_i \leq p$) se dirá que el Álgebra Booleana es atómica.

(3.4.3) Son equivalentes (p. 118):

1. el Álgebra Booleana es atómica,
2. la unidad 1 es el supremo del conjunto de todos los átomos.

A un isomorfismo entre un Álgebra Booleana B y una subalgebra de $\mathscr{P}(X)$ para algún conjunto X se le llama representación de B. Una representación que preserva todos los supremos como uniones se llama completa, con esto, se tiene este teorema (p. 119):

(3.4.4) Si se tiene un Álgebra Booleana atómica B, y X es el conjunto de todos sus átomos, existirá una representación completa de B sobre X.

Con este teorema, nos ayudaremos para decir que si tenemos un Álgebra Booleana atómica con un conjunto finito de átomos, el Álgebra Booleana será finita por existir un isomorfismo entre ella y el conjunto de átomos (que es finito).

El ideal generado por un subconjunto $E \subset B$ es la intersección de todos los ideales para los que E está contenido. Se tiene que (p. 155):

(3.4.5) Un elemento $p \in B$ está en el álgebra generada por un conjunto E ssi existe un subconjunto finito $F \subset E$ cumpliendo $p \leq \bigvee F$.

Con esto, caracterizamos a M el ideal generado por todos los átomos de una Álgebra Booleana, ya que todos sus elementos son de la forma $p = \bigvee_{i=1}^{n} Fi$ con F_i conjuntos de átomos.

Como propiedad de un ideal maximal se tiene (p. 171):

(3.4.6) Un ideal M de un Álgebra Booleana B es maximal ssi para todo $p \in B$ se tiene que $p \in M$ o que $p^c \in M$ pero no ambas.

Por último necesitamos el siguiente resultado,

(3.4.7) Un ideal principal (p) es maximal ssi p^c es un átomo.

Con todo esto, tenemos lo suficiente para demostrar que:

Teorema 3.4.15. *Para cada Álgebra Booleana infinita B, existe un ideal maximal no principal sobre B.*

Demostración. No se puede cumplir que $1 = \bigvee_{i=1}^{n}\{a_i\}$ (es decir, que 1 fuera el supremo de un conjunto finito) para a_i átomos y algún $n \in \mathbb{N}$, si se cumpliera, por (3.4.2) se tendría que $E = \{a_i\}_{i=1}^{n}$ sería el conjunto de todos lo átomos de B, eso implicaría por (3.4.3) que B es atómica, por lo que existiría una representación completa de B sobre E por (3.4.4), e implicaría que B sería finito. Sea M el ideal generado por el conjunto de todos los átomos de B, se tendrá que M consiste de todos los elementos $p \in B$ tales que $p = \bigvee_{i=1}^{n} A_i$ conjuntos finitos de átomos. Como 1 no se puede escribir de esa forma por lo visto anteriormente, $1 \notin M$, y recordando que un ideal es propio ssi 1 no pertenece al ideal, M es un ideal propio. Aplicando el Teorema Ideal Maximal a M, se extiende M a un ideal maximal propio N, como $M \subset N$, N tiene a todos los átomos de B, y por (3.4.6) N no tendrá a algún a_i^c para ningún i, concluyendo por (3.4.7) que N no es principal, ya que el ideal maximal principal es (p^c). □

Se cumple que son equivalentes el Teorema del Ideal Primo y el Teorema del Ultrafiltro, sin embargo, no daré la demostración.

3.5. LA PARADOJA DE BANACH-TARSKI

Las definiciones, teoremas y demás son tomados de [DumFoo] y de [Wag].

Esta definición es obtenida de [DumFoo, p. 41]:

Definición 3.5.1. *Una acción de grupo de un grupo G en un conjunto X, es una función de $G \times X$ a X, escrito por $g \cdot a$ para todos $g \in G$ y $a \in X$, que satisface:*

1. $g_1 \cdot (g_2 \cdot a) = (g_1 g_2) \cdot a$ *para todos los* $g_1,\ g_2 \in G$ *y* $a \in X$, *y*
2. $1 \cdot a = a$, *para toda* $a \in X$.

Se dice que G actúa sobre X en vez de decir que se tiene una acción de grupo sobre un conjunto. Además, $g \cdot a$ generalmente se escribe como ga, teniendo cuidado de no confundir esta operación con la del grupo.

Para todo $g \in G$ fijo, se obtiene la función σ_g definida por:

$$\sigma_g(a) = g \cdot a,$$

y a dicha σ_g se le llama permutación de X.

Cuando se tiene que $X = G$ y G es un grupo, una función de $G \times X$ a X definida por $g \cdot a = ga$, hace que G actúe sobre sí mismo, y que cada $g \in G$ permute los elementos de G por multiplicación por la izquierda:

$$g : a \mapsto ga \qquad \forall a \in G,$$

mientras que si G es escrito aditivo, la traslación será por la izquierda:

$$g : a \mapsto g + a \qquad \forall a \in G,$$

La idea de duplicar un conjunto usando ciertas funciones es simplificada si se ocupan biyecciones de un conjunto, y para hacerlo, se utilizan las acciones de grupos.

Definición 3.5.2. *Sea un grupo G que actúa sobre un conjunto X, sea $E \subset X$, se dirá que E es G-paradójico (o paradójico con respecto a G) si para algunos naturales m, n hay subconjuntos de E disjuntos a pares $A_1, A_2, \ldots, A_n, B_1, B_2, \ldots, B_m$ y también elementos $g_1, g_2, \ldots, g_n, h_1, h_2, \ldots, h_m \in G$, tales que se cumpla que $\bigcup g_i(A_i) = E = \bigcup h_j(B_j)$.*

Cuando dicho conjunto E es G-paradójico, habrá dos conjuntos disjuntos, $\bigcup A_i$ y $\bigcup B_j$, los cuales pueden ser tomados aparte y transformados mediante G para cubrir a E. Además, se pueden tomar dichos subconjuntos que hacen a E G-paradójico para que formen una partición de E, obteniendo las 3 particiones $\{A_i\} \cup \{B_j\}$, $\{g_i(A_i)\}$ y $\{h_j(B_j)\}$. Que la primera pueda ser hecha partición se verá en el corolario (3.5.1), para que las otras dos lo sean, se utiliza el método de la página 14, donde se sacan subconjuntos disjuntos con la misma unión que los conjuntos $\{A_i\}$, si definimos

$$C_1 = A_1 \text{ y}$$

$$C_i = A_i - \bigcup_{k=1}^{i-1} A_k \text{ para } 2 \leq i \leq n,$$

se tendrá que $\bigcup g_i(C_i) = E$, ya que, como $C_i = A_i - \bigcup_{k=1}^{i-1} A_k \subset A_i$, $\bigcup g_i(C_i) \subset \bigcup g_i(A_i) = E$, para la otra contención, si $y \in E$, por ser E G-paradójico, existe un $1 \leq k \leq n$, tal que $y \in g_k(A_k)$, luego, existe un $x \in A_k$ con $g(x) = y$, y de la definición de C_i, $x \in C_k$, por lo tanto $x \in \bigcup g_i(C_i)$. De la misma manera se prueba que la familia $\{h_j(B_j)\}$ puede ser puesta como una partición de E.

De [DumFoo, p. 26] obtenemos: Sea un grupo G y un subconjunto $S \subset G$ con la propiedad de que todo elemento de G puede ser escrito como un producto finito de elementos de S junto con sus inversas, se dirá que S es un conjunto de generadores de G. Cuando ésto pasa, se escribe $G = \langle S \rangle$, y se dice que G es generado por S o que S genera a G.

La forma de entender los elementos de $\langle S \rangle$, es mediante ([DumFoo, p. 63]):

Proposición 3.5.1. *$\langle S \rangle$ es el conjunto de todos los productos finitos de elementos de S e inversas de los elementos de S, i.e.:*

$$\langle S \rangle = \{a_1^{e_1} a_2^{e_2} \cdots a_n^{e_n} | n \geq 0, \forall i \geq 0 : (a_i \in S, e_i = \pm 1)\}.$$

Lo que se quiere demostrar, lo que se concluye de la paradoja de Banach-Tarski, es que cualquier bola en $\mathbb{R}^3$ es paradójica con respecto al grupo de las isometrías de ese espacio, llamado G_3. Las isometrías de $\mathbb{R}^n$ son funciones biyectivas de $\mathbb{R}^n$ a $\mathbb{R}^n$ que preservan la distancia[16]. En el caso de que G sea el grupo de las isometrías de un conjunto X, simplemente se dirá que X es paradójico. La idea central sobre los conjuntos paradójicos son los grupos libres sobre dos generadores. Un grupo actúa naturalmente en él mismo mediante la traslación por la izquierda, pero se buscan grupos que sean paradójicos respecto a ésta acción, o respecto a la de multiplicación.

La idea básica es sacada de [DumFoo, p. 215-217]: se define un grupo libre $F(S)$ generado por un conjunto S. Si se tiene por ejemplo $S = \{a, b\}$, los elementos del grupo libre sobre los dos generadores a y b serán de la forma aa, a, ab, $abab$, $baba$, etc., junto con las inversas de dichos elementos. A dichos elementos se les llama palabras (words en inglés), y pueden ser agrupados cuando varios se repiten, para ponerlos con notación exponencial, por ejemplo $aaa = a^3$, $a^{-1}a^{-1} = a^{-2}$. Los elementos de un grupo libre son multiplicados concatenando las palabras, el producto de ab con $a^{-1}b$ sería $aba^{-1}b$.

Definición 3.5.3. *El grupo libre $F(S)$ con conjunto generador S, es el grupo de todas las palabras finitas usando los elementos $\{a, a^{-1} : a \in S\}$, diciendo que dos palabras son equivalentes, si una puede ser transformada a la otra, mediante la reducción o adición adyacente de parejas finitas de la forma aa^{-1} o de la forma $a^{-1}a$. Una palabra sin dichos pares adyacentes será llamada reducida.*

Se supondrá de aquí en adelante, que todas las palabras de $F(S)$ están en su forma reducida.

[16] El concepto de distancia está en la p. 146, ejemplos de estas isometrías son las translaciones y rotaciones.

Definición 3.5.4. *El grupo $F(S)$ es llamado el grupo libre del conjunto S. Un grupo F es un grupo libre si existe un conjunto S que cumpla con que $F = F(S)$, y se llamará a S un conjunto de generadores libre o una base libre de F. A la cardinalidad de dicha S se le llama el rango del grupo libre.*

El conjunto $aW = \{ab : \; b \in W\}$ será «como» multiplicar un elemento por un conjunto.

Un teorema sencillo que solo utiliza 4 piezas:

Teorema 3.5.1. *Cualquier grupo libre F de rango 2 es F-paradójico, si F actúa en él mismo mediante multiplicación por la izquierda.*

Demostración. Si tenemos un grupo libre de rango 2, tiene generadores libres a, b. Si p es alguno de éstos: $a^{\pm 1}$, $b^{\pm 1}$, sea $W(p)$ el conjunto de elementos de F cuya representación como una palabra (con letras de la forma a, b, a^{-1}, b^{-1}) comienza por la izquierda con p. Tendremos que $F = \{1\} \cup W(a) \cup W(b) \cup W(a^{-1}) \cup W(b^{-1})$ ya que las palabras de F comienzan por algún generador libre o por algún inverso de éstos, y como es un grupo, contiene a la unidad; además, esos conjuntos son disjuntos debido a que toda palabra distinta de 1 empieza por alguna de esas 4 posibilidades, y no puede estar en dos conjuntos porque no puede empezar con dos letras diferentes.

Se cumple que:

$$W(a) \cup aW(a^{-1}) = F,$$
$$W(b) \cup bW(b^{-1}) = F,$$

ya que si $h \in F - W(a)$, entonces $a^{-1}h \in W(a^{-1})$ porque empieza con a^{-1} y h no empieza con a, y además $h = a(a^{-1}h) \in aW(a^{-1})$, análogamente para b. $\square$

La demostración anterior puede ser modificada para que los conjuntos cubran a todo F y no solo a $F - \{1\}$, una demostración así está en [Her, p. 130].

Se dirá que un grupo es paradójico, cuando se refiere a la acción mediante translación por la izquierda.

Si se tiene G un grupo que actúa sobre un conjunto X, la relación $\sim$ en X definida por ([DumFoo, p. 45]):

$$a \sim b \Leftrightarrow a = gb \; \text{ para algún } g \in G$$

es una relación de equivalencia. Para cada $x \in X$, la clase de equivalencia de x bajo la relación $\sim$ es llamada la órbita de x bajo la acción de G, o simplemente la G-órbita de x. Dichas órbitas bajo G particionan al conjunto X.

Mediante este teorema que esta en [Wag, p. 11], y que requiere del Axioma, una descomposición paradójica de un grupo, puede ser llevada a un conjunto donde el grupo actúa, siempre y cuando actúe sin puntos fijos no triviales, refiriéndonos a que ningún elemento del grupo $g \in G$ aparte de la identidad, cumple que $g(a) = a$, para algún $a \in X$.

Proposición 3.5.2. *Si G es paradójico y actúa sobre X sin puntos fijos no triviales, entonces X será G-paradójico. Por lo tanto, X será F-paradójico, siempre que F sea un grupo libre de rango 2 que actúe sobre X sin puntos fijos no triviales.*

Demostración. Supongamos que los conjuntos A_i, $B_j \subset G$ y los elementos g_i, $h_j \in G$ son los que se dicen en la definición de paradójico. Debido al Axioma de Elección, existe un conjunto M el cuál contiene exáctamente un elemento de cada G-órbita en X (M es un conjunto de representantes de la relación $\sim$).

El conjunto $\{g(M) : \; g \in G\}$ será una partición de X. Es una familia disjunta: supongamos que $a \in g_1(M)$ y que $a \in g_2(M)$, con $g_1 \neq g_2$ eso significa que existen b, $c \in M$ tales que $a = g_1(b) = g_2(c)$, por lo tanto $a = g_1 b = g_2 c$, lo que nos dice que $a \sim b$ y que $a \sim c$, pero por la definición de M, se debe cumplir que $a = b = c$, lo cual es una contradicción, ya que $a = g_1(a) = g_2(a)$, indicando que en G hay al menos un elemento distinto de la unidad que fija un punto $a \in X$. Para ver que $\bigcup\{g(M):\ g \in G\} = X$, de la definición de M, éste consta de elementos de $[a]$ la clase de equivalencia de a, para todo $a \in X$, por lo tanto, $\bigcup\{g(M):\ g \in G\} \subset \bigcup[a] = X$, la última igualdad por que la relación $\sim$ es de equivalencia; para el lado contrario, si $x \in X$, entonces $x \in [a]$ para algún $a \in X$, por lo tanto, será un elemento de la clase de equivalencia de a, y como x se relaciona con todos los elementos de dicha clase de equivalencia por definición de ésta, se relaciona con el representante que elije M, digamos m, lo que dice que $x = gm$ para algún $g \in G$, por lo tanto $x \in g(M)$, entonces $x \in \bigcup\{g(M):\ g \in G\}$. Así se demuestra que es una partición.

Definamos estos dos conjuntos:

$$A_i' = \bigcup\{g(M):\ g \in A_i\} \quad \text{y}$$

$$B_i' = \bigcup\{g(M):\ g \in B_i\},$$

se cumple que $\{A_i\} \cup \{B_i\}$ son disjuntos a pares (por la definición de A_i y B_j), y juntándolo con el hecho de que $\{g(M):\ g \in G\}$ es una partición, se tiene que son disjuntos a pares $\{A_i'\} \cup \{B_i'\}$.

Demostraremos que $X = \bigcup g_i A_i' = \bigcup h_j B_j'$. Primero debemos de recordar que como G es paradójico, $G = \bigcup g_i A_i = \bigcup h_j B_j$, que lo ocuparemos para cuando encontremos un $g_k \in G$, que denotaremos por $g_k = g_n g_k'$. Se cumple:

$$\begin{aligned}
x \in X \quad &\Leftrightarrow \quad x \in g_k(M) \text{ para algún } g_k \in G \\
&\Leftrightarrow \quad (x \in g_k(M)) \wedge (g_k \in g_n A_n) \\
&\Leftrightarrow \quad (x \in g_k(M)) \wedge (g_k = g_n g_k') \text{ con } g_k' \in A_n \\
&\Leftrightarrow \quad (x \in g_n g_k'(M)) \wedge (g_k' \in A_n) \\
&\Leftrightarrow \quad (x \in g_n(g_k'(M))) \wedge (g_k' \in A_n) \\
&\Leftrightarrow \quad x \in g_n(\bigcup\{g(M):\ g \in A_n\}) \\
&\Leftrightarrow \quad x \in \bigcup g_i(A_i').
\end{aligned}$$

Lo que acabamos de encontrar, son dos familias de conjuntos que cumplen con la definición de que X sea G-paradójico.

La segunda parte del teorema, se sigue debido al teorema anterior, ya que cualquier grupo libre de rango 2 será paradójico. □

Aunque es un resultado importante, su limitación es el hecho de que requiere una condición fuerte: no dejar puntos fijos. Esta limitación la eliminaremos, para concluir lo deseado. Pero la demostración nos da la idea de qué grupos son paradójicos. Utilizaremos el hecho de que los subgrupos de un grupo actúan en el grupo entero mediante traslación por la izquierda, sin puntos fijos no triviales, dando como consecuencia que[17]:

(3.5.1) Un grupo con un subgrupo paradójico es paradójico. Por lo tanto, cualquier grupo con un subgrupo libre de grado 2 es paradójico.

[17] Por depender del anterior, depende del Axioma.

En particular, los grupos libres no abelianos serán paradójicos. Pero hay muchísimas clases más variadas de grupos paradójicos, en [Wag, p. 12] está planteado un teorema que afirma que hay grupos paradójicos que no tienen subgrupos de cualquier rango.

Hay que aclarar varias cosas, que aunque no se ocuparán en su definición precisa, se requieren para entender unos términos, estas definiciones están en [HofKun, p. 6]. Una matriz de mxn sobre el campo F, con n, $m \in \mathbb{N}$, es una función A del conjunto de pares de enteros $(i,\ j)$, $1 \leq i \leq m$, $1 \leq j \leq n$ al campo F. Las entradas de la matriz A son los escalares $A(i,j) = a_{ij}$, y se da la representación matricial[18] de A de la forma:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Un subconjunto S de un grupo G será llamado independiente, si S es un conjunto libre de generadores de H, el subgrupo de G generado por S, por lo tanto, $|S|$ será el rango del grupo libre H.

En la página 15, Wagon da un ejemplo efectivo de un subgrupo libre de rango 2, para el grupo SO_3, que consta de las matrices ortogonales y reales de nxn con determinante 1. Concluyendo que si $n \geq 3$, se obtiene un subgrupo libre de rango 2 para SO_n. El ejemplo que proporciona, consta de dos rotaciones independientes, ϕ y ρ, sobre los ejes z y x a través del origen, respectivamente, las cuales estan en $\mathbb{R}^3$. Ellas son rotaciones sobre el ángulo arccos$\frac{1}{3}$, su representación matricial es:

$$\phi^{\pm 1} = \begin{pmatrix} \frac{1}{3} & \mp\frac{2\sqrt{2}}{3} & 0 \\ \pm\frac{2\sqrt{2}}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

y

$$\rho^{\pm 1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & \mp\frac{2\sqrt{2}}{3} \\ 0 & \pm\frac{2\sqrt{2}}{3} & \frac{1}{3} \end{pmatrix}$$

Cada elemento de un subgrupo libre de rotaciones F de $\mathbb{R}^3$, podría fijar una linea recta en $\mathbb{R}^3$, así que no se podría aplicar la proposición (3.5.2), pero hay una forma de evadirlo. Si tenemos un grupo libre de rotaciones F en la esfera unitaria (denotada por $S^2 = \{(x,y,z) :\ x^2+y^2+z^2 = 1\}$), toda rotación distinta de la identidad en F tiene dos puntos fijos en S^2, los cuales están en la intersección del eje de rotación con la esfera (cuando giras la esfera respecto a un diámetro, los puntos del diámetro que intersectan la esfera no se mueven). Llamemos D a la colección de dichos puntos: como F es contable, D será contable porque son dos puntos de la aplicación de cada $g \in F$.

Supongamos que $p \in S^2 - D$ y que $g \in F$, se debe tener que $g(P) \in S^2 - D$, debido a que si estuviera fijo $g(P)$ por la rotación h (i.e. h(g(P))=g(P)), se cumpliría que $g^{-1}hg(P) = g^{-1}(h(g(P))) = g^{-1}(g(P)) = P$, con $g^{-1}hg \in F$ por ser grupo, significando que también estaría fijo P. Conclusión: si quitamos todos los puntos que el grupo fija, nos quedamos con un conjunto $S^2 - D$ en el que el grupo actúa sin puntos

[18] Un arreglo rectangular con m filas y n columnas, debido a que dicho arreglo es una correspondencia biyectiva, la representación matricial es más usada por comodidad.

fijos no triviales. Así podemos obtener el siguiente resultado[19] porque aplica la proposición (3.5.2), ya que tenemos un subgrupo libre de rango 2 sin puntos fijos no triviales:

Teorema 3.5.2 (Paradoja de Hausdorff)**.** *Existe un subconjunto contable D de S^2 tal que $S^2 - D$ es SO_3-paradójico.*

El hecho de que sea D contable y por lo tanto mucho menor en tamaño que la esfera completa, nos acerca a la paradoja de Banach-Tarski, sin embargo, falta quitar a D de éste teorema para concluír que S^2 es SO_3-paradójica.

Se necesitan más definiciones:

Definición 3.5.5. *Si G actúa sobre X y $A,\ B \subset X$, se dirá que A y B son G-equidescomponibles*[20] *si A y B pueden ser cada uno particionados en el mismo número finito de respectivas piezas G-congruentes. Matemáticamente:*

$$A = \bigcup_{i=1}^{n} A_i, \qquad B = \bigcup_{i=1}^{n} B_i,$$

con $\emptyset = A_i \cap A_j = B_i \cap B_j$ si $1 \leq i < j \leq n$, y además, existen $g_1, \ldots, g_n \in G$ tales que para cada $1 \leq i \leq n$ se tiene que $g_i(A_i) = B_i$. En dicho caso, la notación usada para dicha relación será $A \sim_G B$.

De hecho, la relación $\sim_G$ es una relación de equivalencia (lo afirma [Wag, p. 23]). De la misma manera que antes, cuando G sea el grupo de todas las isometrías en X, se quitará el subíndice G de $\sim_G$.

Con esta nueva definición, se puede replantear:

Proposición 3.5.3. *Se dice que E es G-paradójico ssi E contiene subconjuntos disjuntos A, B tales que $A \sim_G E$ y también $B \sim_G E$.*

Demostración. E es G-paradójico ssi existen $n,\ m \in \mathbb{N}, A_i,\ B'_j \subset E$ y $g_i,\ h_j \in G$ tales que $E = \bigcup g_i(A_i) = \bigcup h_j(B'_j)$, para aclarar a lo que queremos llegar usamos los mismos nombres solo para identificarlos, pero son dos renombramientos diferentes: $A = \bigcup A_i$ y $E = \bigcup g_i(A_i) = \bigcup B_i$ y de la misma manera $A = \bigcup B'_j$ y $E = \bigcup h_j(B'_j) = \bigcup B_j$ ssi $A \sim_g E$ y $B \sim_g E$. □

Se obtiene así el resultado:

Proposición 3.5.4. *Supongamos que G actúa en X y que E y E_1 son subconjuntos de X G-equidescomponibles. Si E es G-paradójico, también lo será E_1.*

Demostración. Daremos la forma intuitiva de demostrarlo: La hipótesis dice que $E \sim_G E_1$, y que E es G-paradójico, por la proposición anterior existen $A,\ B$ disjuntos, tales que $E \sim_G A$ y $E \sim_g B$. Ya que tenemos una relación de equivalencia, $A \sim_G E_1$ y $B \sim_G E_1$. Por la proposición anterior, pero aplicándola de regreso, se tiene que E_1 es G-paradójico. □

Banach se dio cuenta de que la demostración del Teorema de Schröder-Bernstein (visto éste en la p. 104), puede ser aplicada a cualquier relación de equivalencia que satisfaga dos propiedades abstractas, y aplica a la G-equidescomponibilidad. Se utilizará en esta sección la notación de equipotencia $\preceq$. $A \preceq B$ significará que A es G-equidescomponible con un subconjunto de B.

Dos hechos importantes:

[19] También depende del Axioma.

[20] En inglés, G-equidecomposable. También se les suele llamar finitamente G-equidescomponibles, o G-congruentes por piezas.

1. Si se tiene una $g \in G$ con G actuando sobre un conjunto X, siempre se cumple que $A \sim_G g(A)$ con $A \subset X$ y g biyectiva, ya que para la definición de G-equidescomponibilidad, se toman las piezas G-congruentes como una partición de un solo elemento, i.e., $A = \bigcup_{i=1}^{1} A_i$ y $B = g(A) = \bigcup_{i=1}^{1} g(A_i)$.

2. Si se tienen $A_1 \cap A_2 = \emptyset = B_1 \cap B_2$, con $A_1 \sim B_1$ y con $A_2 \sim B_2$, se cumplirá que $A_1 \cup A_2 \sim B_1 \cup B_2$, debido a que $A_1 = \bigcup A_i^1$, $B_1 = \bigcup B_i^1$, $A_2 = \bigcup A_j^2$ y $B_2 = \bigcup B_j^2$, con $g_i(A_i^1) = B_i^1$ y $h_j(A_j^2) = B_j^2$ y también:

$$\emptyset = A_i^1 \cap A_k^1 = B_i^1 \cap B_k^1 = A_j^2 \cap A_l^2 = B_j^2 \cap B_l^2,$$

luego,

$$A_1 \cup A_2 = \bigcup A_i^1 \cup \bigcup A_j^2 = \bigcup (A_i^1 \cup A_j^2)$$

y

$$B_1 \cup B_2 = \bigcup B_i^1 \cup \bigcup B_j^2 = \bigcup (B_i^1 \cup B_j^2),$$

y se cumplen las dos condiciones de equidescomponibilidad:

$$\begin{aligned}
(A_i^1 \cup A_j^2) \cap (A_k^1 \cup A_l^2) &= ((A_i^1 \cup A_j^2) \cap A_k^1) \cup ((A_i^1 \cup A_j^2) \cap A_l^2) \\
&= ((A_i^1 \cap A_k^1) \cup (A_j^2 \cap A_k^1)) \cup ((A_i^1 \cap A_l^2) \cup (A_j^2 \cap A_l^2)) \\
&= (\emptyset \cup (A_j^2 \cap A_k^1)) \cup ((A_i^1 \cap A_l^2) \cup \emptyset) \\
&= (A_j^2 \cap A_k^1) \cup (A_i^1 \cap A_l^2) \\
&= \emptyset \qquad \text{debido a que } A_1 \cap A_2 = \emptyset.
\end{aligned}$$

Al reemplazar en las ecuaciones anteriores A por B se obtiene la condición para $B_1 \cup B_2$. Para encontrar las g_i´s que pide la definición, debemos encontrar g'_{ij} tal que $g'_{ij}(A_i^1 \cup A_j^2) = B_i^1 \cup B_j^2$; por el teorema (1.5.2), se tiene que definiendo:

$$g'_{ij}(A_i^1) = g_i(A_i^1) \text{ y } g'_{ij}(A_j^2) = h_j(A_j^2),$$

$$\begin{aligned}
\Rightarrow g'_{ij}(A_i^1 \cup A_j^2) &= g'_{ij}(A_i^1) \cup g'_{ij}(A_j^2) \\
&= g_i(A_i^1) \cup h_j(A_j^2) \\
&= B_i^1 \cup B_j^2.
\end{aligned}$$

lo cual es posible ya que dichos conjuntos son disjuntos, por lo tanto $A_1 \cup A_2 \sim B_1 \cup B_2$.

Con estos dos resultados, podemos demostrar el siguiente importante

Teorema 3.5.3 (Teorema de Banach-Schröder-Bernstein). *Supongamos que G actúa en X, y que A, $B \subset X$. Si $A \preceq B$ y $B \preceq A$, se tendrá que $A \sim_G B$. Haciéndonos concluir que $\preceq$ es un orden parcial de las $\sim_G$-clases en $\mathscr{P}(X)$.*

Demostración. Por hipótesis se tienen $A_1 \subset A$ y $B_1 \subset B$, con $A \sim B_1$ y $B \sim A_1$, por el primer punto del resultado anterior, existen funciones biyectivas $f : A \mapsto B_1$ y $g : A_1 \mapsto B$ tales que $\forall C \subset A, D \subset A_1$ se cumple $(C \sim f(C)) \wedge (D \sim g(D))$.

Definimos una sucesión $C_{n+1} = g^{-1}f(C_n)$, con $C_0 = A - A_1$, y un conjunto $C = \bigcup_{n=0}^{\infty} C_n$. Demostraremos que $g(A-C) = B - f(C)$.

$$\begin{aligned}
y \in g(A-C) &\Leftrightarrow \exists x \in A - C :\ g(x) = y \\
&\Leftrightarrow x \in A,\ x \notin C \\
&\Leftrightarrow x \in A,\ \forall i \in \mathbb{N} \cup \{0\} :\ x \notin C_i \\
&\Leftrightarrow x \in A,\ x \notin A - A_1,\ \forall i \in \mathbb{N} :\ x \notin C_i \\
&\Leftrightarrow x \in A,\ x \notin A - A_1,\ \forall i \in \mathbb{N} :\ x \notin C_i \\
&\Leftrightarrow x \in A_1,\ \forall i \in \mathbb{N} \cup \{0\} :\ x \notin g^{-1}f(C_i) \\
&\Leftrightarrow x \in A_1,\ \forall i \in \mathbb{N} \cup \{0\} :\ g(x) \notin f(C_i) \\
&\Leftrightarrow x \in A_1,\ \forall i \in \mathbb{N} \cup \{0\} :\ f^{-1}g(x) \notin C_i \\
&\Leftrightarrow x \in A_1,\ f^{-1}g(x) \notin C \\
&\Leftrightarrow x \in A_1,\ g(x) \notin f(C) \\
&\Leftrightarrow y = g(x) \in B,\ y =\notin f(C) \\
&\Leftrightarrow y \in B - f(C).
\end{aligned}$$

Varias implicaciones se deben a que las funciones g y f son biyectivas, y debido a su otra propiedad, $A - C \sim g(A-C) = B - f(C)$ y $C \sim f(C)$. Por lo tanto, por ser disjuntos $A - C$ y C, aparte de que son disjuntos $B - f(C)$ y $f(C)$, se puede aplicar el segundo punto del resultado anterior a este teorema para la unión de conjuntos disjuntos; además se cumple la propiedad que nos ayuda a «cancelar» $(A-C) \cup C = A$, obteniendo:

$$\begin{aligned}
(A-C) \cup C &\sim (B - f(C)) \cup f(C) \\
&\Rightarrow A \sim B,
\end{aligned}$$

que es lo que se quería demostrar. □

Como corolario:

Corolario 3.5.1. *Un subconjunto $E \subset X$, será G-paradójico ssi existen conjuntos disjuntos A, $B \subset E$ tales que $A \cup B = E$ y $A \sim_G E \sim_G B$.*

Demostración. E es G-paradójico ssi existen A, $B \subset E$ con $A \cap B = \emptyset$ tales que $A \sim_G E \sim_G B$ (por la proposición (3.5.3)), ssi existen A, $B \subset E$ con $B \subset A^c$ que cumplen $A \sim_G E \sim_G B$, ssi existen A, $B \subset E$, que cumplen que $E \sim_G B \subset E \cap A^c = E - A \subset E$ y que $A \sim_G E \sim_G B$, ssi existen A, $B \subset E$, que satisfacen $E \sim_G E - A$ y que $A \sim_G E \sim_G B$ (por el teorema anterior), ssi existen A, $B \subset E$, que cumplen que $B \sim_G E - A$ y que $A \sim_G E \sim_G B$, ssi existen A, $B \subset E$, con $B = E - A$, con $A \cap B = \emptyset$ y con $A \cup B = E$, además de cumplir $A \sim_G E \sim_G B$. □

Como ya se vio, los grupos libres de rango 2 causan paradojas cuando actúan sin puntos fijos no triviales sobre un conjunto, por ello, se requiere una demostración por absorción. Intuitivamente dice que si se tiene un conjunto sin varios puntos equidescomponible, los puntos no cuentan mucho en la equidescomponibilidad, y que pueden ser absorbidos para concluir que el conjunto mismo es equidescomponible.

Pero antes, necesitamos una definición y su resultado para grupos libres ([DumFoo, p. 20]):

Definición 3.5.6. *Si se tiene un grupo G y un $x \in G$, el orden de x es el natural más pequeño tal que $x^n = 1$, y se dice que x es de orden n, cuando dicho natural no existe, el orden de x se dice infinito.*

El ejercicio 4 de [DumFoo, p. 220], dice que todo elemento de un grupo libre que no sea la identidad, tiene orden infinito, y esto se ocupará en:

Teorema 3.5.4. *Sea un subconjunto $D \subset S^2$ contable, se tendrá que S^2 y $S^2 - D$ son SO_3-equidescomponibles, solamente usando dos piezas.*

Demostración. Ya en la p. 130, se vio que SO_3 tiene un subgrupo de rango 2 (incluso dando las matrices). Podemos encontrar de ahí un elemento (una rotación) que no sea la identidad, de orden infinito. Lo que se busca es una rotación de la esfera, llamada ρ (la rotación), que haga que los conjuntos D, $\rho(D)$, $\rho(D)^2$, $\ldots$ sean disjuntos a pares. Si la encontramos, se tendrá que si $D_1 = \bigcup\{\rho^n(D) \ : n \in \mathbb{N} \cup \{0\}\} = D \bigcup\{\rho^n(D) : \ n \in \mathbb{N}\}$, al calcular

$$
\begin{aligned}
y \in \rho(D_1) &\Leftrightarrow \exists d_1 \in D_1 : y = \rho(d_1) \\
&\Leftrightarrow \exists k \in \mathbb{N} \cup \{0\} : \exists d \in D : y = \rho(d_1),\ d_1 = \rho^k(d) \\
&\Leftrightarrow \exists k \in \mathbb{N} \cup \{0\} : \exists d \in D : y = \rho(\rho^k(d)) \\
&\Leftrightarrow \exists k \in \mathbb{N} \cup \{0\} : \exists d \in D : y = \rho^{k+1}(d) \\
&\Leftrightarrow \exists k \in \mathbb{N} : \exists d \in D : y = \rho^k(d) \\
&\Leftrightarrow y \in D_1 - D.
\end{aligned}
$$

Ya que son disjuntos D y $\rho^n(D)$, con $n \in \mathbb{N}$. Se concluye que $D_1 = D \cup \rho(D_1)$, con ésto en mente, por el primer punto del resultado anterior al Teorema de Banach-Schröder-Bernstein, y la propiedad de la p. 14 para «cancelar» conjuntos cuando están contenidos:

$$
\begin{aligned}
S^2 &= (S^2 - D_1) \cup D_1 \\
&\sim (S^2 - D_1) \cup \rho(D_1) \\
&= (S^2 \cap D_1^c) \cup \rho(D_1) \\
&= (S^2 \cap (D \cup \rho(D))^c) \cup \rho(D_1) \\
&= (S^2 \cap D^c \cap (\rho(D))^c) \cup \rho(D_1) \\
&= (S^2 \cap D^c - \rho(D)) \cup \rho(D_1) \\
&= S^2 - D,
\end{aligned}
$$

encontrando que $S^2 \sim S^2 - D$, lo que se quería demostrar.

Ahora solo falta encontrar dicha ρ. Sea L una línea que pasa por el origen y que no toca al conjunto D, dicha existe porque D es contable, y existirán infinitos no numerables puntos que no estén en D. Sea A el conjunto de todos los ángulos θ, tales que para algún $n > 0$ y algún punto del espacio $P \in D$, $\tau(P)$ se encuentre en D, con τ la rotación alrededor de L de ángulo $n\theta$; lo que hace dicho A, es contener ángulos que hagan que las rotaciones de puntos de la esfera que estén en D, vuelvan a caer en D. Ya que a cada ángulo de A le podemos asociar un natural (a saber, el $n \in \mathbb{N}$ que aparece en $n\theta$), A es contable; por ello, existe un ángulo $\theta' \notin A$ (porque hay muchos más) y una rotación ρ alrededor de L con dicho angulo θ'. Como ese ángulo no está en A, para cualquier punto P en D, $\rho(P) \notin D$, pero tampoco se cumple para cualquier natural n, $\rho^n(P) \in D$, ya que el ángulo de éste último es $n\theta'$ porque la rotación suma el ángulo al aplicarla varias veces. Con todo lo anterior, si $n > 0$ se cumple que $\rho^n(D) \cap D = \emptyset$, y por el teorema (1.5.1), el cual aplica debido a que ρ es una rotación y por lo tanto biyectiva, si $1 \leq m < n$ $\emptyset = \rho^{n-m}(D) \cap D \Rightarrow \emptyset = \rho^n(D) \cap \rho^m(D)$, encontrando la rotación buscada. $\square$

El siguiente teorema es el resultado buscado de esta sección, depende del Axioma de Elección debido a que depende de resultados que lo hacen, y es una consecuencia directa de varios resultados vistos:

Teorema 3.5.5 (Paradoja de Banach-Tarski)**.** *S^2 es SO_3-paradójica, así como cualquier esfera centrada en el origen.*

Demostración. Primero recurrimos a la paradoja de Hausdorff, de la p. 131, que nos asegura la existencia de un conjunto contable $D \subset S^2$, cumpliendo que $S^2 - D$ es SO_3-paradójico, dicho conjunto D consta de puntos fijos de rotaciones, pero el teorema anterior indica que S^2 y $S^2 - D$ son SO_3-equidescomponibles, y al satisfacer las hipótesis de la proposición (3.5.4), la cuál dice que al tener dos conjuntos G-equidescomponibles con uno G-paradójico, el otro también es G-paradójico, concluyendo que S^2 es SO_3-paradójico.

Para demostrar que no solo la esfera unitaria, sino cualquier esfera es SO_3-paradójica, se debe notar que nunca se ocupó el radio de la esfera en las demostraciones anteriores. $\square$

También se puede demostrar ([Wag, p. 27-28]) que cualquier esfera es G_3-paradójica (G_n es el grupo de todas las isometrías de $\mathbb{R}^n$), y $\mathbb{R}^3$ es paradójico.

El resultado del teorema anterior, es más desconcertante que la paradoja de Hausdorff, se interpreta diciendo que una pelota, puede ser descompuesta en un número finito de pedazos, y reacomodada únicamente mediante rotaciones, para formar dos pelotas, y si ésto se hace de nuevo, incluso se pueden formar cien millones de pelotas del mismo tamaño, iguales a la original.

Al aplicar rotaciones a cuerpos, se preserva el volumen, debido a que son isometrías, y eso es lo que le hace ganarse a éste teorema el nombre de paradoja, sin embargo, no hay volumen que preservar si se tiene un conjunto sin volumen, y de hecho, unas de las piezas en la descomposición de la paradoja son no-Lebesgue medibles.

La siguiente consecuencia es una generalización de la anterior, y su demostración es por construcción e intuitiva, pero no se dará:

Teorema 3.5.6. *Si A y B son dos conjuntos acotados de $\mathbb{R}^3$, cada uno de los cuales tiene un interior*[21] *no vacío, se tendrá que A y B son equidescomponibles.*

3.6. TOPOLOGÍA Y ANÁLISIS

Empezaremos a dar definiciones fundamentales, tomadas del libro [Dug].

Definición 3.6.1. *Sea X un conjunto. Una topología, o una estructura topológica en X es una familia τ de subconjuntos de X que satisface:*

1. *La unión arbitraria de miembros de τ es un miembro de τ.*
2. *Cada intersección finita de miembros de τ es un miembro de τ.*
3. *Los conjuntos X y $\emptyset$ son miembros de τ.*

Definición 3.6.2. *A la pareja (X,τ) donde X es un conjunto y τ es una topología en X se le llama espacio topológico.*

[21] Término visto en la p. 136.

Generalmente se habla de que τ es una topología para el espacio X en vez de hablar del espacio topológico (X,τ). Cuando no hay necesidad de especificar explícitamente a τ, se habla del espacio X, para distinguirlo de solamente un conjunto.

Los elementos del espacio X se llaman puntos, y los elementos de la topología τ se llaman abiertos.

Definición 3.6.3. *Si tenemos un espacio (X,τ), una vecindad de un punto $x \in X$ es cualquier abierto que contiene a x, y se suele escribir U_x o también $U(x)$.*

Para generar a los conjuntos abiertos, se pueden requerir menos abiertos que los que tiene la topología, teniendo así una base.

Definición 3.6.4. *Sea (X,τ) un espacio topológico. A una subfamilia $\mathscr{B} \subset \tau$ se le llama base para τ, si cada conjunto abierto, o lo que es lo mismo, cada elemento de τ, es la unión de elementos de $\mathscr{B}$.*

Al conjunto $\mathscr{B}$ también se le llama base para el espacio X, y a sus elementos se les llama abiertos básicos de la topología τ.

Este teorema está en [Dug, p. 64]:

Teorema 3.6.1. *Sea $\mathscr{B} \subset \tau$, son equivalentes:*

i) *$\mathscr{B}$ es una base para τ*

ii) *Para cada abierto $U \in \tau$ y cada elemento x de U, existe un elemento V de $\mathscr{B}$ tal que $x \in V \subset U$.*

Supondremos que el espacio X está dado, y ocuparemos varias definiciones, ésta es para caracterizar a los análogos de los abiertos:

Definición 3.6.5. *Un conjunto $A \subset X$ es cerrado si $A^c = \mathscr{C}_X A$ es un conjunto abierto.*

Definición 3.6.6. *Sea $A \subset X$. Un punto $x \in X$ se dirá que es adherente a A (en inglés adherent) si cada vecindad de x contiene al menos un punto de A, el cuál puede ser x mismo si $x \in A$. Al conjunto $\overline{A} = \{x \in X | \forall U(x) : \; U(x) \cap A \neq \emptyset\}$ de todos los puntos de X que son adherentes a A se le llama cerradura de A (en inglés closure), en donde se ocupa el hecho de que $U(x)$ es una vecindad de x.*

Otra forma de caracterizar a la cerradura de un conjunto, es decir que es el cerrado más pequeño que contiene al conjunto.

En [Dug, p. 69] está:

Teorema 3.6.2. *Para todo conjunto A, se cumple que:*

1. *$A \subset \overline{A}$.*
2. *A es cerrado ssi $A = \overline{A}$.*
3. *Si $A \subset B$, entonces $\overline{A} \subset \overline{B}$.*

Definición 3.6.7. *Sea $A \subset X$. El conjunto $A^o = \bigcup\{U | (U \text{ es abierto}) \wedge (U \subset A)\}$ es el interior de A, es decir, el abierto más grande contenido en A.*

El siguiente teorema está en [Dug, p. 65]:

Teorema 3.6.3. *Sea $\mathscr{B} \subset \tau$ una base para τ, un conjunto U será abierto (i.e. $U \in \tau$) ssi para cada $x \in U$ existe un $V \in \mathscr{B}$ tal que $x \in V \subset U$.*

Definición 3.6.8. *Sea $D \subset X$, con X un espacio, se dirá que D es denso en X si $\overline{D} = X$.*

Otra forma de caracterizar a los conjuntos densos es:

Teorema 3.6.4. *Son equivalentes:*

i) *D es denso en X.*

ii) *Cada abierto básico no vacío en X, contiene un elemento de D.*

ii) *El complemento de D tiene interior vacío.*

Una definición importante es:

Definición 3.6.9. *Sean (X, τ_X) y (Y, τ_Y) dos espacios. Una función $f : X \mapsto Y$ se dice continua si la imagen inversa de todo abierto en Y, es abierta en X.*

Se cumple que la composición de funciones continuas es continua, como era de esperarse.

3.6.1. Teorema de Tychonoff

Definición 3.6.10. *Si tenemos (X, τ) un espacio topológico, una colección $\mathscr{C}$ de subconjuntos de X se llama cubierta abierta de X, si $\mathscr{C} \subset \tau$ y si $\bigcup\{C : C \in \mathscr{C}\} = X$.*

Definición 3.6.11. *Si tenemos una cubierta abierta $\mathscr{C}$ de X, una subcubierta abierta $\mathscr{C}_1$ de $\mathscr{C}$ es una subcolección de $\mathscr{C}$ tal que $\bigcup\{C_1 : C_1 \in \mathscr{C}_1\} = X$.*

La definición más importante para esta sección es:

Definición 3.6.12. *Un espacio topológico (X, τ) se llamará compacto, si para toda cubierta abierta de X, se puede extraer una subcubierta abierta finita.*

Si se tiene $\beta \in \mathscr{A}$, y $\prod_{\alpha \in \mathscr{A}} X_\alpha$, la función:

$$\pi_\beta : \prod_{\alpha \in \mathscr{A}} X_\alpha \mapsto X_\beta$$

dada por $c \mapsto c(\beta)$ o lo que es lo mismo $\{a_\alpha\} \mapsto a_\beta$, es la proyección en el β-ésimo factor, y por el teorema 9 de [Dug, p. 23], dicha función es suprayectiva.

Dado un $a \in \mathscr{A}$ y un subconjunto $C_a \subset X_a$, se denota a $\pi_a^{-1}(C_a)$ por $\langle C_a \rangle$, el cuál denota el producto de los X_i´s excepto el a-ésimo factor, que es C_a, y se dice que es como una rebanada (en inglés, slice) en el producto $\prod_{i \in I} X_i$. Cuando se tienen $a_1, \ldots, a_n \in \mathscr{A}$ y subconjuntos $C_{a_i} \subset X_{a_i}$ con $i \in \{1, 2, \ldots, n\}$, al subconjunto $\langle C_{a_1} \rangle \cap \cdots \cap \langle C_{a_n} \rangle = \pi_{a_1}^{-1}(C_{a_1}) \cap \cdots \cap \pi_{a_n}^{-1}(C_{a_n})$ se le denota por $\langle C_{a_1}, \ldots, C_{a_n} \rangle$.

Con esta definición, el siguiente teorema dice que ([Dug, p. 24]):

Teorema 3.6.5. *En el producto $\prod_{i \in I} X_i$ se cumple que*

1. $\prod_{i \in I} C_i = \cap \langle C_i \rangle$,
2. $\langle C_i \rangle^c = \langle C_i^c \rangle$.

Definiciones importantes:

Definición 3.6.13. *Sea $\{X_i | i \in I\}$ una familia de espacios topológicos, y τ_i la topología de X_i para todo $i \in I$. La topología del producto cartesiano $\prod_{i \in I} X_i$ es la que tiene como subbase a los subconjuntos $\langle U_b \rangle = \pi_b^{-1}(U_b)$, donde b recorre todo I y U_b recorre todos los elementos de τ_b.*

Una forma de ver los conjuntos abiertos básicos, es que son de la forma[22]:

$$U_{a_1} \times \cdots \times U_{a_n} \times \prod \{X_i | i \neq a_1, \ldots, a_n\} = \bigcap_1^n \pi_{a_i}^{-1}(U_{a_i}) = \langle U_{a_1}, \ldots, U_{a_n} \rangle,$$

donde los conjuntos U_{a_i} son abiertos en X_i.

En [Dug, p. 223] esta un teorema que nos servirá después:

Teorema 3.6.6. *Son equivalentes:*

1. *X es compacto.*
2. *La propiedad de la intersección finita: Para cada familia $\{A_i | i \in I\}$ de subconjuntos de X que son cerrados en X, y que cumplen que $\bigcap_i A_i = \emptyset$, existe una subfamilia finita $A_{a_1}, \ldots, A_{a_n}$ que cumple lo mismo, i.e., $\bigcap_1^n A_{a_i} = \emptyset$.*

Teorema 3.6.7. *Son equivalentes:*

i) Teorema de Tychonoff: El producto de espacios compactos es compacto.

ii) El Axioma de Elección.

Demostración. $i) \Rightarrow ii)$ Tomemos cualquier familia $(X_i)_{i \in I}$ de conjuntos no vacíos, y denotemos por a un elemento tal que $a \notin \bigcup_{i \in I} X_i$ (existe, debido a que la unión de conjuntos es un conjunto, y no es la clase de todos los conjuntos). Formaremos conjuntos compactos. Sea $Y_i = X_i \cup \{a\}$ y sea su topología $\tau_i = \{\emptyset, Y_i, \{a\}\}$, los conjuntos (Y_i, τ_i) son compactos ya que la única cubierta abierta de Y_i, es el mismo Y_i, porque $Y_i \cup \{a\} = Y_i$, y esa única cubierta abierta de Y_i, tiene una subcubierta abierta finita (a saber, ella misma), por lo tanto, son compactos. De la hipótesis, concluimos que el conjunto $Y = \prod_{i \in I}(Y_i, \tau_i)$ es compacto. Se cumple que $\forall i \in I$ pasa que $\langle X_i \rangle = \pi_i^{-1}(X_i) \neq \emptyset$, ya que como dijimos antes, la función es suprayectiva[23], se concluye que es no es vacío ese conjunto. Además, $\langle X_i \rangle$ es cerrado porque sacando el complemento y utilizando el teorema 3.6.5:

$$\begin{aligned} (\pi_i^{-1}(X_i))^c &= \langle X_i \rangle^c \\ &= \langle X_i^c \rangle \\ &= \langle \{a\} \rangle \end{aligned}$$

que es un abierto en Y para toda $i \in I$ debido a que $\{a\}$ es abierto.

Con estos dos hechos, podemos demostrar que la familia $X = \{\langle X_i \rangle | i \in I\}$ tiene la propiedad de la intersección finita, la intersección

$$\langle X_{a_1}, X_{a_2}, \ldots, X_{a_n} \rangle = X_{a_1} \times \cdots \times X_{a_n} \times \prod \{Y_i | i \in I,\ i \neq a_1, \ldots, a_n\}$$

[22] En realidad, nos ocuparemos solo de las propiedades del producto, y no de la manera en que la multiplicación se hace, es decir, no ocuparemos el hecho de que el producto cartesiano no es conmutativo, pero para ilustrar la forma del producto, se escribe de esa manera.

[23] Para una función suprayectiva f, se cumple que $f^{-1}(B) \neq \emptyset$.

es distinta del vacío para todas las elecciones de $a_i \in I$ con n finito, debido a que como cada $\langle X_i \rangle$ es distinto del vacío, sea n un natural, $a_i \in I$, y sea $c_i \in \langle X_i \rangle$ con $i \in \{a_1, \ldots, a_n\}$, definimos para $i \in I$ la función

$$c(i) = \begin{cases} c_i(i) \in X_i & \text{si } i \in \{a_1, \ldots, a_n\} \\ c_{a_{n+1}}(i) \in Y_i & \text{si } i \neq a_1, \ldots, a_n \end{cases}$$

que manda a elementos de los conjuntos X_i si se evalúa en el índice de éstos, y mandar a elementos de Y_i si se evalúa en un número que no esté en el subconjunto finito de I. Por lo tanto X tiene la propiedad de la intersección finita.

Por eso, reformularemos lo que cumple la familia $X \subset Y$: si toda subfamilia finita de ésta tiene intersección no vacía, se tendrá que la familia entera tiene intersección no vacía porque Y es compacto y por lo que dice el teorema 3.6.6, i.e. $\bigcap_{i\in I} \langle X_i \rangle \neq \emptyset$, pero por el teorema 3.6.5

$$\bigcap_{i\in I} \langle X_i \rangle = \prod_{i\in I} X_i,$$

sabemos por el teorema 3.1.2, que la afirmación de que el producto cartesiano de cualquier familia no vacía de conjuntos no vacíos sea no vacío, es equivalente al Axioma.

ii) $\Rightarrow$ *i*) Sea $\{X_i | i \in I\}$ una familia de espacios, donde cada X_i es compacto. Demostraremos que $X = \prod_{i\in I} X_i$ es compacto. Si $X = \emptyset$, éste será compacto porque cualquier cubierta abierta que lo contenga, tendrá una subcubierta abierta finita (cualquier elemento de la cubierta) que lo contenga. Si no es el caso, tomemos una familia $\mathscr{B}$ de conjuntos cerrados en X que cumpla con la propiedad de la intersección finita, al menos una existe porque a la familia $\{X\}$, al intersectarle un número finito de sus elementos (el único), esa intersección es distinta del vacío. Demostrando que $\bigcap \mathscr{B} \neq \emptyset$, y por el teorema 3.6.6, X es compacto, al ser $\mathscr{B}$ arbitrario. Aplicaremos[24] M7 al conjunto de todas las familias $\mathscr{C}$ de subconjuntos de X que cumplen que $\mathscr{B} \subset \mathscr{C}$ y que tienen la propiedad de la intersección finita. El conjunto de todas esas familias, satisface las hipótesis de M7, que volvemos a enunciar:

M7. Para todo conjunto X y toda propiedad de carácter finito P, existe un subconjunto de X $\subset$-maximal que tiene la propiedad P.

Por lo tanto, sea $\mathscr{U}$ dicha familia maximal con la propiedad de la intersección finita. Se cumple que:

$$\bigcap_{U\in\mathscr{U}} \overline{U} \subset \bigcap_{B\in\mathscr{B}} B,$$

por ello, basta con demostrar el hecho de que $\bigcap_{U\in\mathscr{U}} \overline{U} \neq \emptyset$, y por lo tanto, el conjunto que lo contiene tampoco será igual al vacío. Escogimos las cerraduras de los elementos de $\mathscr{U}$ porque se requieren conjuntos cerrados en las hipótesis del teorema 3.6.6, y por el teorema 3.6.2 las cerraduras son cerrados, hay que notar que los elementos de $\mathscr{U}$ no necesariamente lo son.

Sea $\pi_i : X \mapsto X_i$ la proyección sobre el conjunto X_i (la proyección en el i-ésimo factor). La familia de conjuntos para un $i \in I$ fijo $\{\pi_i(U) | U \in \mathscr{U}\}$ tiene la propiedad de la intersección finita, porque si escogemos cualesquier conjuntos $U_1, U_2, \ldots, U_n \in \mathscr{U}$, y como la familia $\mathscr{U}$ tiene la propiedad de la intersección finita, se sigue que $\bigcap_1^n U_j \neq \emptyset$, por lo que para toda $j \in \{1, \ldots, n\}$ se cumple $\pi_i(U_j) \neq \emptyset$, y por el teorema 1.5.2 de la p. 16, $\emptyset \neq \pi_i(\bigcap_1^n U_j) \subset \bigcap_1^n \pi_i(U_j)$. Lo anterior demuestra que $\{\pi_i(U) | U \in \mathscr{U}\}$ tiene la propiedad de la intersección finita.

[24]En la p. 93 se vio que es equivalente al Axioma.

El hecho de que X_i sea compacto, otra vez el teorema 3.6.6 y el párrafo anterior, nos ayudan a afirmar que $\bigcap_{U\in\mathscr{U}} \overline{\pi_i(U)} \neq \emptyset$, y por ello, podemos encontrar un $x_i \in X_i$ tal que $\forall U \in \mathscr{U}$ se tiene que $x_i \in \overline{\pi_i(U)}$. El Axioma de Elección nos ayuda a formar al elemento $x = (x_i)_{i\in I} \in X$ que tiene como i-ésimo factor a x_i, vamos a demostrar que $x \in \bigcap_{U\in\mathscr{U}} \overline{U}$.

Sea un $A \subset X$ tal que $A \cap U \neq \emptyset$, para toda $U \in \mathscr{U}$, eso implica que $A \in \mathscr{U}$, ya que si no fuera cierto, sería un subconjunto propio $\mathscr{U} \subset \mathscr{U} \cup \{A\}$, tendría la propiedad de la intersección finita porque $\mathscr{U}$ la tiene, y si agregamos a cualquier intersección finita A, como toca a todos los U´s, la intersección seguiría siendo distinta del vacío, y el conjunto $\mathscr{U} \cup \{A\}$ contradeciría la maximidad de $\mathscr{U}$.

Tomemos un entorno básico de x, por la definición 3.6.13, éste debe ser de la forma $A = \bigcap_{i\in F} \pi_i^{-1}(G_i)$, donde $F \subset I$ es finito, y cada G_i es abierto en X_i. Por lo demostrado en el párrafo antes del anterior $\forall U \in \mathscr{U}$ se tiene que $x_i \in \overline{\pi_i(U)}$, y por la definición de cerradura de un conjunto 3.6.6, para esos abiertos se cumple $G_i \cap \pi_i(U) \neq \emptyset$, eso implica que $\emptyset \neq \pi_i^{-1}(G_i) \cap U$, y como dicha condición se cumple para todo $U \in \mathscr{U}$, por el resultado del párrafo anterior $\pi_i^{-1}(G_i) \in \mathscr{U}$, y eso se cumple para toda $i \in F$ porque fue arbitraria. Debido a que $\mathscr{U}$ tiene la propiedad de la intersección finita, la intersección finita de conjuntos de $\mathscr{U}$ está ahí mismo, i.e., $A \in \mathscr{U}$, implicando que el entorno básico A de x se intersecta con todo elemento de $\mathscr{U}$, diciendo que $x \in \overline{U}$, para todo $U \in \mathscr{U}$, que es lo que se quería demostrar. □

3.6.2. Teorema de Hahn-Banach

Si tenemos un espacio vectorial V sobre un campo F:

Definición 3.6.14. *Una función p definida sobre un espacio vectorial V sobre el campo de los reales y con valores en los reales es llamada funcional lineal sobre V si $p(rx+sy) = rp(x)+sp(y)$ se cumple para todos los $r,\ s \in \mathbb{R}$ y todos los $x,\ y \in V$.*

Definición 3.6.15. *Una funcional p sobre V es sublineal si $p(x+y) \leq p(x)+p(y)$ para todos los $x,\ y \in V$ y si $p(rx) = rp(x)$ para todos los reales no negativos r y todos los $x \in V$.*

Teorema 3.6.8 (Teorema de Hahn-Banach)**.** *Sea p una funcional sublineal en un espacio vectorial V, y sea f_0 una funcional lineal definida en un subespacio $V_0 \subset V$ tal que $f_0(v) \leq p(v)$ para todo $v \in V_0$. Entonces, existe una funcional lineal f definida en V que cumple con que $f_0 \subset f$ y que $f(v) \leq p(v)$ para toda $v \in V$.*

Demostración. Llamemos a F el conjunto de todas las funcionales lineales g con dominio algún subespacio $W \subset V$ que satisfagan $f_0 \subset g$ y que $g(v) \leq p(v)$ para toda $v \in W$, F no es vacío porque $f_0 \in F$ por hipótesis. Demostraremos que la funcional lineal requerida f es un elemento maximal del conjunto $(F, \subset)$.

Para verificar que las hipótesis de M3 se satisfacen[25], sea $\emptyset \neq F_0 \subset F$ que esté linealmente ordenado por la relación $\subset$ (es decir, un nido). Como casi siempre, tomemos a $g_0 = \bigcup F_0$ que demostraremos será una cota $\subset$-superior en F_0 siempre que $g_0 \in F$. El rango de g_0 es $\mathbb{R}$, debido a que es el rango de cada una de las funciones que la componen, además, $f_0 \subset g_0$ porque si no, existiría un $x \in f_0$ tal que no existe $g \in F_0$ que satisfaga $x \in g$ porque f_0 no está en la unión, lo que nos diría que $f_0 \not\subset g$, que sería una contradicción a la definición de F_0.

Tenemos que F está linealmente ordenado por $\subset$, por ello, las funciones en F están contenidas unas en otras, traduciéndose en que si tenemos dos funciones, el dominio de una estará contenido en el de otra, por lo que también están ordenadas linealmente por $\subset$. Ahora, demostraremos que la unión de un conjunto de subespacios de V linealmente ordenados por $\subset$ es un subespacio de V: si $c \in \mathbb{R}$ y $\alpha,\ \beta \in dom\ g_0$, tenemos

[25] En la página 93 se demuestra que es equivalente al lema de Zorn, que es equivalente al Axioma por la página 98.

que demostrar que $c\alpha+\beta \in dom\ g_0$, se sigue que existen $g_2,\ g_1 \in F_0$ con $\alpha \in dom\ g_2$ y $\beta \in dom\ g_1$ por definición de unión, y como es un orden lineal F_0, sin pérdida de generalidad $g_2 \subset g_1$, por ello, $\alpha,\ \beta \in dom\ g_1$, y como $dom\ g_1$ es un subespacio de V, $c\alpha+\beta \in dom\ g_1$, por lo tanto $c\alpha+\beta \in dom\ g_0$. Así vemos que el conjunto $dom\ g_0 = \bigcup_{g\in F_0} dom\ g$ (que es unión de subespacios) es un subespacio de V.

Demostraremos que g_0 es una funcional lineal: sean $c \in \mathbb{R}$ y $\alpha,\ \beta \in dom\ g_0$, de la misma forma que el párrafo anterior, existirán $g_2,\ g_1 \in F_0$ con $\alpha \in dom\ g_2$ y $\beta \in dom\ g_1$, y como es un orden lineal F_0, sin pérdida de generalidad $g_2 \subset g_1$, por ello, $\alpha,\ \beta,\ c\alpha+\beta \in dom\ g_1$, y como $g_1 \in F_0$, g_1 es una funcional lineal y

$$g_0(c\alpha+\beta) = g_1(c\alpha+\beta) = cg_1(\alpha)+g_1(\beta) = cg_0(\alpha)+g_0(\beta),$$

demostrando que es una funcional lineal. Lo último que necesitamos, es ver que si $\alpha \in dom\ g_0$ y $\alpha \in dom\ g$ con $g \in F_0$, se cumple por ésto último que

$$g_0(\alpha) = g(\alpha) \leq p(\alpha),$$

tomando a $W = dom\ g_0$ en la definición de F. Todo esto demuestra que $g_0 \in F$, y que g_0 es una cota $\subset$-superior para F_0, debido a que no hay otra función que contenga a g_0 y que esté en F_0. Aplicando M3 que volvemos a enunciar:

M3. Si todo subconjunto no vacío que sea un nido, de un conjunto X no vacío tiene su unión en X, entonces X tendrá un elemento $\subset$-maximal,

concluimos que existe f un elemento maximal de $(F,\subset)$, tal que $f_o \subset f$ y $f(\alpha) \leq p(\alpha)$ para toda $\alpha \in dom\ f$. Lo único que falta demostrar es que $dom\ f = V$, y habremos acabado. Si se diera la inclusión propia en $dom\ f \subset V$, sea $u \in V - dom\ f$, y sea W el subespacio de V generado por el vector u y el conjunto $dom\ f$; si $w \in W \Rightarrow w = \alpha + cu$ con $\alpha \in dom\ f$ y $c \in \mathbb{R}$, definiendo

$$f_b(w) = f_b(\alpha+cu) = f(\alpha)+cb,$$

obtenemos una funcional lineal sobre W porque si $d \in \mathbb{R}$ y $x,\ z \in W$ con $x = x_1 + c_x u$ y $z = z_1 + c_z u$,

$$\begin{aligned} f_b(dx+z) &= f(dx_1+z_1)+(dc_x+c_z)b \\ &= d(f(x_1)+c_x b)+f(z_1)+c_z b \\ &= df_b(x)+f_b(z). \end{aligned}$$

Las propiedades de f_b son que $f \subset f_b$ y la contención es propia porque $dom\ f \subset W$ el espacio generado por u y $dom\ f$ con $u \notin dom\ f$. Lo único que nos falta es encontrar un $b \in \mathbb{R}$ que cumpla:

$$f_b(\alpha+cu) = f(\alpha)+cb \leq p(\alpha+cu)$$

para cualesquiera $\alpha \in dom\ f$ y $c \in R$. Cuando $c = 0$ se cumple por definición de f. Ese b debe satisfacer que:

1. $\forall c > 0$ y $\alpha \in dom\ f$ sea cierto que $f(\alpha)+cb \leq p(\alpha+cu)$ y

2. $\forall c > 0$ y $\beta \in dom\ f$ sea cierto que $f(\beta)+(-c)b \leq p(\beta+(-c)u)$.

Esas condiciones son lo mismo que

$$f(\beta)-p(\beta-cu)\leq cb\leq p(\alpha+cu)-f(\alpha),$$

y multiplicando la desigualdad por c^{-1} que es positivo y ocupando que p y f cumplen con la segunda propiedad de las funcionales lineales se obtiene que

$$f(\frac{1}{c}\beta)-p(\frac{1}{c}\beta-u)\leq b\leq p(\frac{1}{c}\alpha+u)-f(\frac{1}{c}\alpha)$$

se debe cumplir para todos los $\alpha,\ \beta\in dom\ f$ y $c>0$, pero si $\gamma,\ \psi\in dom\ f$ por la propiedad de f y por ser p funcional sublineal:

$$f(\gamma)+f(\psi)=f(\gamma+\psi)\leq p(\gamma+\psi)=p(\gamma-u+\psi+u)\leq p(\gamma-u)+p(\psi+u),$$

lo cual es equivalente a decir que $f(\gamma)-p(\gamma-u)\leq p(\psi+u)-f(\psi)$. Tomando los conjuntos

$$A=\sup\{f(\gamma)-p(\gamma-u)|\gamma\in dom\ f\}$$

y

$$B=\text{ínf}\{p(\psi+u)-f(\psi)|\psi\in dom\ f\}$$

y eligiendo cualquier $A\leq b\leq B$ se concluye la demostración. □

3.6.3. Conjuntos que no son Lebesgue Medibles

Demostraremos la existencia de un conjunto[26] no medible, y lo definiremos como lo hace Bartle en su libro de integración, pero no nos meteremos con definir la integral, ni que la existencia de la integral es equivalente a que las sumas superiores e inferiores de Riemman coinciden al sacarles ínfimo y supremo; no necesitaremos esas definiciones, y lo que se requiera lo pondré explícitamente. Ocuparemos al espacio el conjunto de los reales extendidos, que es $\overline{\mathbb{R}}=\mathbb{R}\cup\{\infty\}\cup\{-\infty\}$. Solamente en esta subsección ocuparemos la notación $|A|$ para otra característica que la de la cardinalidad del conjunto A. La bibliografía para ésta subsección es el libro [Bar].

Para subconjuntos de un conjunto compacto, es equivalente que un conjunto sea medible y que un conjunto sea integrable, pero para subconjuntos generales de $\mathbb{R}$ hay que hacer una distinción. La función característica de un $A\subset\mathbb{R}$ la definimos en la página 16 como:

$$\chi_A(x)=\begin{cases}1 & \text{si } x\in A,\\ 0 & \text{si } x\in A^c.\end{cases}$$

En algunos libros se define $\chi_A(\pm\infty)=0$.

Definición 3.6.16. *Se dirá que $A\subset\mathbb{R}$ es integrable si su función característica es integrable en $\overline{\mathbb{R}}$. Cuando sucede, el número $\int_{-\infty}^{\infty}\chi_A$ es llamado la medida de Lebesgue, o más corto, la medida de A. Se suele denotar por $|A|$ o por $\lambda(A)$.*

Al conjunto de todos los subconjuntos integrables de $\mathbb{R}$ se le denota por $I(\mathbb{R})$.

Definición 3.6.17. *Un subconjunto $A\subset\mathbb{R}$, diremos que es Lebesgue medible, o mas breve, medible, si $A\cap I$ es integrable para todo conjunto compacto $I\subset\mathbb{R}$.*

[26] En realidad, por la construcción, infinitos conjuntos no numerables.

A la familia de todos los conjuntos medibles de $\mathbb{R}$ se le denota por $M(\mathbb{R})$.

Definición 3.6.18. *La medida de Lebesgue, o la medida, se define como la función* $\lambda : M(\mathbb{R}) \mapsto [0,\infty]$ *dada por* $\lambda(A) = |A|$ *si* A *es un subconjunto integrable de* $\mathbb{R}$*, mientras que se define como* $\lambda(A) = \infty$ *si* A *es un subconjunto medible pero no integrable de* $\mathbb{R}$*.*

Propiedades importantes son:

Teorema 3.6.9. *1.* $I(\mathbb{R}) \subset M(\mathbb{R})$*.*

2. *Todo intervalo acotado de* $\mathbb{R}$ *está en* $I(\mathbb{R})$*, y por lo tanto, en* $M(\mathbb{R})$*.*

3. *Todo conjunto abierto en* $\mathbb{R}$ *está en* $M(\mathbb{R})$*.*

4. *Si tenemos* $a < b$*, se cumple que* $|(a,b)| = \lambda((a,b)) = b - a$*.*

5. *Si se cumple que* $A \subset B$*, entonces* $|A| \leq |B|$*.*

6. *Supongamos que tenemos una familia de conjuntos no vacíos disjuntos a pares* $\{A_n | n \in \mathbb{N}\}$*, los cuales están en* $I(\mathbb{R})$*, entonces* $|\bigcup_{n\in\mathbb{N}} A_n| = \sum_{n=1}^{\infty} |A_n|$*.*

Cuando tenemos un subconjunto $A \subset \mathbb{R}$ y a un $r \in \mathbb{R}$, al conjunto $A_r = \{r + x : \ x \in A\}$ se le llama el r-traslado de A, y se cumple que ([Bar, p. 115]):

Teorema 3.6.10 (Teorema de la Invarianza)**.** *i) Si tenemos* $A \in I(\mathbb{R})$ *y* $r \in \mathbb{R}$*, entonces* $A_r \in I(\mathbb{R})$ *y* $|A_r| = |A|$*.*

ii) Si tenemos $A \in M(\mathbb{R})$ *y* $r \in \mathbb{R}$*, entonces* $A_r \in M(\mathbb{R})$ *y también* $\lambda(A_r) = \lambda(A)$*.*

Estamos listos para demostrar la existencia de un subconjunto de $\mathbb{R}$ no medible. De la construcción del conjunto se puede concluir que cualquier conjunto Lebesgue medible E cumpliendo que $\lambda(E) > 0$ contiene un subconjunto que es no Lebesgue medible.

Teorema 3.6.11. *Existe un subconjunto* $V \subset (0,1)$ *que es no Lebesgue medible.*

Demostración. Definiremos una relación sobre los reales por:

$$x \sim y \Leftrightarrow x - y \in \mathbb{Q}.$$

Veamos que es una relación de equivalencia: si $x \in \mathbb{R}$, $x - x = 0 \in \mathbb{Q}$; si $x \sim y$, $x - y \in \mathbb{Q}$, entonces $-(x-y) = y - x \in \mathbb{Q}$, entonces $y \sim x$; si $x \sim y$ y también $y \sim z$, $x - z = (x-y) + (y-z) \in \mathbb{Q}$. Ésto nos ayuda a concluir por el teorema 1.4.1 que la relación divide a $\mathbb{R}$ en clases de equivalencia disjuntas. Cada clase de equivalencia será de la forma $Q_x = \{x + q : \ q \in \mathbb{Q}\}$ para algún $x \in \mathbb{R}$, ya que al restar dos elementos de Q_x, se tendrá una resta de racionales, que es racional. Si tomamos una clase de equivalencia fija, debido a que los racionales son densos en $\mathbb{R}$, ésta tendrá intersección no vacía con cada conjunto abierto no vacío, debido a que éste contendrá infinitos racionales.

En el teorema 3.1.3, se demuestra que es equivalente al Axioma de Elección:

1) Si $\mathscr{A}$ es un conjunto que cumple con que cada uno de sus elementos son no vacíos y cualesquier dos elementos no tienen elementos en común (una familia disjunta de conjuntos no vacíos), entonces, existirá un conjunto C que contiene exactamente un elemento de cada miembro de $\mathscr{A}$, i.e., para todo $A \in \mathscr{A}$, se cumple para algún $x \in A$ que $C \cap A = \{x\}$.

Podemos utilizar al conjunto de representantes $V \subset (0,1)$ que intersecta a cada clase de equivalencia distinta en exactamente un punto. Demostraremos que suponer que V es medible nos lleva a una contradicción del teorema de la invarianza.

Si V es medible, es integrable por las propiedades del teorema 3.6.9, y también su r-trasladado para cualquier $r \in \mathbb{Q}$ por el teorema de la invarianza, además, $|V_r| = |V|$. Tomemos a $r,\ s \in \mathbb{Q}$, con $r \neq s$, la intersección cumple que $V_r \cap V_s = \emptyset$, porque si ésta fuera no vacía, habrían $v_1,\ v_2 \in V$ tales que $r + v_1 = s + v_2$ por ser de la forma de los trasladados, por lo que $v_1 - v_2 = s - r \in \mathbb{Q}$ (porque r y s son racionales), lo anterior implica que $v_1 \sim v_2$. $v_1 \neq v_2$ porque si fueran iguales, también lo serían r y s contradiciendo la hipótesis. La definición de V nos dice que tiene un solo elemento de cada clase de equivalencia, pero v_1 y v_2 están en la misma clase y son diferentes, contradiciendo la definición de V. Por eso, la familia $\{V_r :\ r \in \mathbb{Q}\}$ de conjuntos integrables, es disjunta a pares.

Tomemos todos los racionales entre el 0 y el 2, i.e., $C = \mathbb{Q} \cap (0,2)$, C es numerable por ser un subconjunto del conjunto numerable $\mathbb{Q}$. Demostraremos que:

$$(1,2) \subset \bigcup_{r \in C} V_r \subset (0,3).$$

Tomando a cualquier $x \in \mathbb{R}$, existe un único $v_x \in V$ tal que $x \sim v_x$ (ya que pertenece a una sola clase de equivalencia al ser una partición el conjunto de las clases de equivalencia); nombremos $r = x - v_x \in \mathbb{Q}$ que es racional porque se relacionan los dos. Si dicho $x \in (1,2)$, como $v_x \in V \subset (0,1)$, entonces $r \in (0,2)$ porque se resta desde cero hasta uno, lo que demuestra que $x \in V_r$ para algún $r \in C$, concluyendo que $(1,2) \subset \bigcup_{r \in C} V_r$. Como $V \subset (0,1)$, para cada $r \in (0,2)$, como se suma a lo más dos, $V_r \subset (0,3)$ y se sigue que $\bigcup_{r \in C} V_r \subset (0,3)$.

Ahora la conclusión: Debido a que son disjuntos a pares $\{V_r :\ r \in C\}$, que la medida de Lebesgue es contablemente aditiva (propiedad del teorema 3.6.9), la medida preserva la contención (quinta propiedad del teorema 3.6.9) y que la medida de un intervalo abierto es la resta de los puntos extremos:

$$1 = |(1,2)| \leq \sum_{r \in C} |V_r| \leq |(0,3)| \leq 3.$$

Por el teorema de la invarianza $|V_r| = |V|$, por lo que todos los términos de la suma de arriba son iguales a $|V|$. Si el valor es mayor que cero ($|V| > 0$), la suma diverge porque C es infinito numerable, lo cual contradice el hecho de que la suma es menor o igual a 3; si el valor es cero ($|V| = 0$), la suma es igual a 0, lo cual contradice el hecho de que es mayor o igual a 1.

La conclusión es de que el conjunto V no es Lebesgue medible. $\square$

3.6.4. El problema de la medida

Esta subsección está muy relacionada con la anterior, y aquí se demostrará usando el Axioma que no existe una generalización de "longitud de un intervalo" para conjuntos de cualquier tipo de números reales. El objetivo es contar con una función $\mu : \mathscr{P}(\mathbb{R}) \mapsto [0,\infty) \cup \{\infty\}$ y que cumpla:

1. $\mu([a,b]) = b - a$ para cualesquier $a,\ b \in \mathbb{R}$ con $a < b$.

2. $\mu(\emptyset) = 0$ y $\mu(\mathbb{R}) = \infty$.

3. Se cumple la aditividad contable de μ, i.e., si tenemos una familia de subconjuntos disjuntos mutuamente excluyentes de $\mathbb{R}$, entonces se cumple que

$$\mu\left(\bigcup_{n=0}^{\infty} A_n\right) = \sum_{n=0}^{\infty} \mu(A_n).$$

4. Invarianza de la traslación bajo μ. Si tenemos $a \in \mathbb{R}$, y $A \subset \mathbb{R}$, y denotamos $A+a = \{x+a|x \in A\}$, entonces $\mu(A+a) = \mu(A)$.

De estas propiedades, se puede concluir que la función es monótona[27] y finitamente aditiva[28].

Teorema 3.6.12. *No existe una función $\mu : \mathscr{P}(\mathbb{R}) \mapsto [0,\infty) \cup \{\infty\}$ que cumpla con las 4 propiedades principales mencionadas arriba.*

Demostración. Definiendo la misma relación de equivalencia que la demostración de la existencia de un conjunto no Lebesgue medible, mediante una forma del Axioma de Elección sea X un conjunto de representantes de la relación. Como la relación particiona a los reales, se cumple que:

$$\mathbb{R} = \bigcup\{X + r|r \in \mathbb{Q}\},$$

y dicha unión es una familia disjunta a pares. Se tiene que $\mu(X) > 0$, ya que si no, si fuera cierto que $\mu(X) = 0$, por la propiedad 4, la invarianza bajo traslaciones hace que $\mu(X+q) = 0$ para todo $q \in \mathbb{Q}$, por lo que resultaría por la aditividad contable (propiedad 3) que:

$$\mu(\mathbb{R}) = \sum_{q\in\mathbb{Q}} \mu(X+q) = 0,$$

lo cual es una contradicción a la propiedad 2. Por lo anterior y por la aditividad contable existe un cerrado $[a,b]$ tal que $\mu(X \cap [a,b]) > 0$, llamando $Y = X \cap [a,b]$ entonces:

$$\bigcup_{q\in\mathbb{Q}\cap[0,1]} (Y+q) \subset [a,b+1],$$

ya que si z está en el lado izquierdo de la contención, $z = y+q$ para algunos $y \in Y$ y $q \in \mathbb{Q} \cap [0,1]$; de la definición de Y se cumple que y es un elemento en el conjunto de representantes que además es mayor o igual que el elemento a y menor o igual que el elemento b, por lo que se cumple que $z \in [a,b+1]$. El lado izquierdo de la contención es la unión de infinitos conjuntos disjuntos (porque X lo es al ser conjunto de representantes) de la forma $Y+q$ con medida $\mu(Y+q) = \mu(Y) > 0$, por lo que el lado izquierdo tiene medida infinito y el lado derecho tiene medida $b+1-a$ por la propiedad 1. ☐

El teorema nos dice que debemos de relajar requerimientos de μ si trabajamos con el Axioma. Para el análisis matemático, es muy conveniente decir que μ solo esté definida sobre algunos subconjuntos de $\mathbb{R}$. Que su dominio sea una σ-álgebra es lo adecuado.

Definición 3.6.19. *Si tenemos S un conjunto no vacío, a una colección $\mathscr{A} \subset \mathscr{P}(S)$ se le llama σ-álgebra de subconjuntos de S si:*

1. *$\emptyset \in \mathscr{A}$ y $S \in \mathscr{A}$.*

2. *Si $X \in \mathscr{A}$, se cumple que $X^c = S - X \in \mathscr{A}$.*

3. *Si $X_n \in \mathscr{A}$ para toda $n \in \mathbb{N}$, pasa que*

$$\bigcup_{n=0}^{\infty} X_n \in \mathscr{A} \ \ y \ \ \bigcap_{n=0}^{\infty} X_n \in \mathscr{A}.$$

[27] Una función monótona cumple que si $x \leq y \Rightarrow f(x) \leq f(y)$, y también para la desigualdad contraria.

[28] Una función finitamente aditiva es la que cumple la aditividad finita.

Definición 3.6.20. *Una medida σ-aditiva en una σ-álgebra llamada $\mathscr{A}$ de subconjuntos de S es una función $\mu : \mathscr{A} \mapsto [0,\infty) \cup \{\infty\}$ con:*

1. $\mu(\emptyset) = 0,\ \mu(S) > 0.$
2. *Si $X \in \mathscr{A}$, se cumple que $X^c = S - X \in \mathscr{A}$.*
3. *Si $\{X_n\}_{n=0}^{\infty}$ es una familia de elementos de $\mathscr{A}$ mutuamente disjuntos, entonces:*

$$\mu\left(\bigcup_{n=0}^{\infty} X_n\right) = \sum_{n=0}^{\infty} \mu(X_n).$$

A los elementos del conjunto $\mathscr{A}$ de la definición anterior, se les llama conjuntos μ-medibles.

Como resultado importante: si tenemos cualquier μ que sea una medida σ-aditiva en una σ-álgebra $\mathscr{A}$ de subconjuntos de $\mathbb{R}$, medida en la que se cumplan las propiedades 1 y 4 de las primeras mencionadas, entonces tendremos la existencia de conjuntos de números reales que no son μ-medibles. La demostración es directa del teorema 3.6.12.

3.6.5. Teorema de Categoría de Baire

Las definiciones requeridas son:

La primera de [Dug, p. 181]:

Definición 3.6.21. *Una métrica, o una función de distancia en un conjunto Y es una función $d : Y \times Y \mapsto \mathbb{R}^+ \cup \{0\}$ con las siguientes propiedades:*

1. *$d(x,y) \geq 0$ para todos los $x,\ y \in Y$.*
2. *$d(x,y) = 0$ ssi $x = y$.*
3. *$d(x,y) = d(y,x)$ para todos los $x,\ y \in Y$.*
4. *$d(x,z) \leq d(x,y) + d(y,z)$ para todos los $x,\ y,\ z \in Y$.*

Si en la definición anterior al punto dos, los sustituimos por la implicación de regreso, se llama espacio pseudométrico (o también se llama écart en inglés). Al valor $d(x,z)$ se le llama la distancia entre x y z.

Un espacio métrico es una pareja (X,d) con d una métrica sobre el conjunto X.

Al conjunto $B_d(x,r) = \{y \in Y | d(y,x) < r\}$ se le llama la d-bola de radio r y centro x. Cuando se sobreentiende la métrica, se escribe $B_r(x)$ a la bola centrada en x y de radio r positivo.

Teorema 3.6.13. *La familia $\{B_d(y,r) | y \in Y, r > 0\}$ de todas las d-bolas de Y sirve como una base para una topología en Y.*

En esta sección se definirán los espacios de Baire, que contienen a clases muy importantes de espacios como son los espacios compactos de Hausdorff y los espacios métricos completos[29], se saca la información de [Mun, p. 295].

[29] No se dará la definición de espacio completo, no se ocupará más adelante, sino sólo propiedades que cumplen éstos.

Definición 3.6.22. *Un espacio X diremos que es un espacio de Baire si se cumple que para toda familia contable de conjuntos cerrados $\{A_n\}_{n\in\mathbb{N}}$ que tienen cada uno interior vacío en X, su unión tiene interior vacío en X.*

R. Baire ocupó el término categoría para casi lo que acabamos de definir. Un subconjunto A de un espacio X lo llamó de primera categoría si estaba contenido en la unión contable de conjuntos cerrados de X con interior vacío en X, de otra manera se le llamaría de segunda categoría. Con esto, se define:

Definición 3.6.23. *Un espacio X es un espacio de Baire si y solo si todo abierto no vacío en X es de segunda categoría.*

Lo siguiente es tomado de [Ivo1, p. 96-97].

Definición 3.6.24. *Si tenemos a M un espacio métrico y a $C \subset M$, el diámetro de C es:*

$$d(C) = \sup\{d(x,y)|x,\ y \in C\} \in [0,\infty].$$

Definición 3.6.25. *Un espacio Y de Hausdorff se dirá que es regular si para cualquier punto $x \in Y$ y cualquier vecindad U_x de x, existe una vecindad V_x de x tal que $x \in V_x \subset \overline{V_x} \subset U_x$.*

Un teorema muy parecido al de la propiedad de la intersección finita, porque tiene que ver con él es:

Teorema 3.6.14. *Sea M un espacio métrico completo. Para toda familia decreciente de cerrados $\{C_n\}_{n\in\mathbb{N}}$ en M y que son no vacíos, cumpliendo que $\text{lím}_n\ d(C_n) = 0$, la intersección de la familia es no vacía.*

Teorema 3.6.15 (Teorema de Categoría de Baire)**.** *En cualquier espacio métrico completo, se cumple que la intersección de una familia numerable de abiertos densos es un conjunto denso*[30].

Demostración. Tomemos a un espacio métrico completo M, y a $G = \bigcap_{n=1}^{\infty} G_n$ la intersección numerable de abiertos G_n que son densos en M. Demostraremos que G se intersecta con cualquier abierto, que son de la forma $G_r(x)$, con $x \in M$ y $r \in \mathbb{R}^+$, y el conjunto G sería denso (son equivalentes: D es denso en X, cada abierto básico no vacío en X, contiene un elemento de D, véase 3.6.4).

El hecho de que G_1 es denso en M, nos hace encontrar un $x_1 \in G_1 \cap G_r(x)$, la intersección finita de abiertos es abierto, por lo que $G_1 \cap G_r(x)$ es abierto, y con eso, existirá un $r_1 > 0$ el cuál podemos escoger menor a $r/2$, y que cumpla $\overline{B_{r_1}(x_1)} \subset G_1 \cap G_r(x)$. Para el siguiente conjunto, sabemos por hipótesis que G_2 es abierto y denso en M y que $B_{r_1}(x_1)$ es abierto, por construcción, encontramos $x_2 \in G_2 \cap B_{r_1}(x_1)$ y una vecindad de x_2 llamada $B_{r_2}(x_2)$, tal que su cerradura está contenida en la intersección $G_2 \cap B_{r_1}(x_1)$ y de radio r_2 menor a $r/3$. Se construye inductivamente una sucesión de puntos $\{x_n\}_{n\in\mathbb{N}}$ de M, y una sucesión de puntos $\{r_n\}_{n\in\mathbb{N}}$ en $\mathbb{R}^+$, tales que:

$$\overline{B_{r_n}(x_n)} \subset G_n \cap B_{r_{n-1}}(x_{n-1}),$$

lo cuál podemos hacer por el Axioma de Denumerable, o el de Elecciones Dependientes, tomando a x_n en cada paso, dependiente de los anteriores, y por la forma de la construcción[31], a cada conjunto $B_{r_n}(x_n)$.

Demostraremos que se cumplen las hipótesis del teorema[32] anterior para la familia $\{\overline{B_{r_n}(x_n)}\}_{n\in\mathbb{N}}$:

[30]Para la demostración con espacios compactos de Hausdorff, solo se dará la idea, que es casi análoga a la aquí dada.

[31]Si tuviéramos un espacio compacto de Hausdorff, la regularidad y la definición 3.6.25 nos ayudarían a encontrar dicha secuencia de cerrados.

[32]En el caso de espacios de Hausdorff compactos, se ocupa la propiedad de la intersección finita, la que ya vimos antes del teorema de Tychonoff, de ésta misma sección.

Por el teorema 3.6.2, un conjunto está contenido en su cerradura:

$$\overline{B_{r_n}(x_n)} \subset G_n \cap B_{r_{n-1}}(x_{n-1}) \subset B_{r_{n-1}}(x_{n-1}) \subset \overline{B_{r_{n-1}}(x_{n-1})},$$

por lo que es una familia decreciente de cerrados. Los números positivos r_n satisfacen:

$$0 \leq \lim_{n\to\infty} r_n \leq \lim_{n\to\infty} \frac{r}{n+1} = 0,$$

implicando que:

$$\lim_{n\to\infty} d(\overline{B_{r_n}(x_n)}) = 0.$$

El teorema nos afirma que existe un $y \in M$ tal que $\forall n \in \mathbb{N}$ se tiene que $y \in \overline{B_{r_n}(x_n)}$, resultando que:

$$\begin{aligned}
\emptyset &\neq \bigcap \overline{B_{r_n}(x_n)} \\
&\subset \bigcap (G_n \cap B_{r_{n-1}}(x_{n-1})) \\
&\subset (\bigcap G_n) \cap B_{r_1}(x_1) \\
&= G \cap B_{r_1}(x_1) \\
&\subset G \cap \overline{B_{r_1}(x_1)} \\
&\subset G \cap G_1 \cap B_r(x) \\
&= G \cap B_r(x),
\end{aligned}$$

concluyendo que G intersecta a todo abierto del espacio, que es lo que se quería demostrar. □

Como aplicación del teorema de Baire, tenemos que si $f_n : [0,1] \mapsto \mathbb{R}$ es una secuencia de funciones continuas que convergen[33] a $f(x)$ para toda $x \in [0,1]$, la función f no necesariamente es continua, pero se cumple que no puede ser discontinua en todos lados. Ejemplos de aplicaciones pero que no demostraremos:

Esta en [Mun, p. 297]:

Teorema 3.6.16. *Sea X un espacio y sea (Y,d) un espacio métrico. Sean las funciones $f_n : X \mapsto Y$ continuas y tales que $\forall x \in X : f_n \mapsto f$, donde la función f cumple que $f : X \mapsto Y$. Si el espacio X es un espacio de Baire, el conjunto de puntos en los que f es continua es denso en X.*

El siguiente se encuentra en [Mun, p. 300]:

Teorema 3.6.17. *Sea $h : [0,1] \mapsto \mathbb{R}$ una función continua. Dado cualquier $\varepsilon > 0$, existe una función $g : [0,1] \mapsto \mathbb{R}$ que cumple $|h(x) - g(x)| < \varepsilon$ para toda $x \in [0,1]$, y g es continua, pero es diferenciable en ningún punto.*

El resultado anterior que depende del Teorema de Categoría de Baire, dice que encontraremos una función continua muy cercana a la función dada, pero que no es diferenciable en algún punto de su dominio, gráficamente podríamos imaginar que la función g tiene muchos picos, y que no es «suave» en algún lado.

Otro ejemplo se encuentra en [Ivo1, p. 114]:

Teorema 3.6.18. *El conjunto $\mathbb{Q}$ de todos los racionales como subconjunto de $\mathbb{R}$, no se puede expresar como la intersección numerable de abiertos de $\mathbb{R}$.*

[33] No daremos la definición de convergencia de funciones.

3.7. VARIOS

La primera sección se trata de un concepto fundamental para las matemáticas, mientras que las siguientes se ocupan de propiedades de los reales. En ellas, trabajaremos en el conjunto $\mathbb{R}$.

Definiremos conceptos de varias maneras y mediante formas del Axioma, veremos que son equivalentes dichas definiciones, el paso crucial cuando trabajamos con sucesiones, es decir mediante el Axioma de Elecciones Denumerables, que $\prod_{n\in\mathbb{N}} X_n \neq \emptyset$ si $X_n \neq \emptyset$, con lo que podemos encontrar una función f tal que $f(n) \in X_n$.

3.7.1. ¿Qué es finito?

Para esta subsección tenemos que recordar que para conjuntos A y B la notación $|A| \leq |B|$ significa que existe una inyección de A a B, mientras que $|A| = |B|$ significa que existe una biyección entre ellos.

Algunas definiciones equivalentes que se definen completamente mediante el Axioma Denumerable son:

Definición 3.7.1. *Un conjunto X se dice que es Dedekind-infinito o solo D-infinito, si existe un subconjunto propio Y de X con $|Y| = |X|$, si no se cumple, se dice que X es D-finito o Dedekind-finito.*

Teorema 3.7.1. *Son equivalentes:*

i) *X es un conjunto D-infinito.*

ii) $|X| = |X| + 1$.

iii) $\aleph_0 \leq |X|$.

Demostración. $iii) \Rightarrow ii)$ Denotemos a $f : \aleph_0 \mapsto X$ una inyección dada por la hipótesis, y sea $a \notin X$ que existe por ser X un conjunto y no la clase de todos los conjuntos. Demostraremos que la función $g : X \mapsto X \cup \{a\}$ dada por:

$$g(x) = \begin{cases} a & \text{si } x = f(0), \\ f(n) & \text{si } x = f(n+1), \\ x & \text{cualquier otro caso,} \end{cases}$$

es una biyección, y habremos terminado.

Es inyectiva: como f es una inyección, si $x \neq y \Rightarrow f(x) \neq f(y)$; luego, si $x \neq y$ con $x,\ y \in X$, entonces si $x = f(0)$, $a = g(x) \neq f(n)$ y también $a = g(x) \neq y$; si no son x ni y elementos de la imagen de f, $x = g(x) \neq g(y) = y$; si solo uno está en la imagen de f, por ejemplo $x = f(n+1)$, entonces $f(n) = g(f(n+1)) \neq y$ porque $y \notin f(X)$. Solo queda demostrar que si $x = f(n+1)$ y $y = f(m+1)$ con $x \neq y$, entonces $f(n) = g(x) \neq g(y) = f(m)$: de la hipótesis se sigue que $n \neq m$ (si fueran iguales entonces $x = y$) y como f es inyectiva, $f(n) \neq f(m)$.

g es suprayectiva: cualquier elemento de X está en la imagen de f (los $x = f(n+1)$), mientras que los otros casos no lo están. El valor que falta es a, pero $g(f(0)) = a$, por lo tanto $g(X) = X \cup \{a\}$.

$ii) \Rightarrow i)$ Sea a un elemento tal que $a \notin X$ y sea $f : X \mapsto X \cup \{a\}$ una biyección que afirma la hipótesis existe. La función

$$f^{-1}|_X :\ X \mapsto X - \{f^{-1}(a)\}$$

es una biyección de X a un subconjunto propio de él. Es propio ya que existe un único $x \in X$ tal que $f(x) = a$, por lo tanto $X - \{f^{-1}(a)\} = X - \{x\} \subset X$. $f^{-1}|_X$ es suprayectiva porque como $f(X) = X \cup \{a\}$ se tiene que $f^{-1}|_X(X) = X - \{x\}$. Es inyectiva porque f lo es.

$i) \Rightarrow iii)$ Tenemos que encontrar una función inyectiva, con dominio el conjunto de los naturales y contradominio X. Sea $f : X \mapsto X$ una inyección sobre un conjunto propio de X (que existe por ser X D-infinito). Tomemos un elemento $z \in X - f(X)$ el cuál existe por ser subconjunto propio. Sea $g : \mathbb{N} \mapsto X$ la función $g(0) = z$ y $g(n+1) = f(g(n))$, de esta definición se sigue que:

$$\begin{aligned} g(n+1) &= f(g(n)) = f(f(g(n-1))) \\ &= f^2(g(n+1-2)) = f^{n+1}(g(0)) = f^{n+1}(z). \end{aligned}$$

Tomemos $n \neq m$, sin pérdida de generalidad sea $m = k + n$ con $k \geq 1$, y demostraremos que $f^n(z) = g(n) \neq g(m) = f^m(z)$. Escribiendo $f^m(z) = f^{n+k}(z) = f^n(f^k(z))$ y teniendo en cuenta que $z \notin f(X)$ y que $f^k(z) \in f(X)$, se cumple que $z \neq f^k(z)$, y como f es inyectiva en X, $f(z) \neq f(f^k(z))$, llegando a que $f^n(z) \neq f^n(f^k(z)) = f^m(z)$ que es lo que se quería demostrar. $\square$

El otro concepto de conjunto finito es:

Definición 3.7.2. *Un conjunto X es finito, si cada subconjunto no vacío de $\mathscr{P}(X)$ contiene un elemento minimal con respecto al orden de inclusión $\subset$. Los conjuntos que no son finitos, son llamados infinitos.*

La siguiente proposición nos ayuda a ver que son equivalentes el concepto de «finitez» acabado de definir, y el que comúnmente conocemos: X es finito ssi $|X| = n$ para algún $n \in \mathbb{N}$.

Proposición 3.7.1. *Son equivalentes:*

i) X es finito.

ii) Si tenemos una subfamilia $\mathscr{U} \subset \mathscr{P}(X)$ que satisface:

a) *$\emptyset \in \mathscr{U}$ y*

b) *$A \in \mathscr{U}$ y también $x \in X$ implican $(A \cup \{x\}) \in \mathscr{U}$,*

se concluye que $X \in \mathscr{U}$.

Demostración. $i) \Rightarrow ii)$ Hay que demostrar que si X es finito, y que $\mathscr{U}$ cumple esas dos condiciones, tendremos que $X \in \mathscr{U}$. Tomemos a la subfamilia $\mathscr{B} = \{X - A | A \in \mathscr{U}\}$ del conjunto potencia, la cuál tiene un elemento minimal que llamaremos B (existe porque la definición de finitez lo afirma). Tenemos así que $B = X - A$ con $A \in \mathscr{U}$, por lo tanto, $A = X - B$ es un conjunto maximal para $\mathscr{U}$ (le estamos quitando a X el conjunto más pequeño de inclusión); de aquí, se tiene por la segunda condición de $\mathscr{U}$, que si $x \in X$, $(A \cup \{x\}) \in \mathscr{U}$, pero dicho conjunto sería un conjunto maximal para $\mathscr{U}$ contradiciendo la hipótesis de A. Por lo tanto, se debe cumplir que para cualquier $x \in X$, $A \cup \{x\} = A$, o lo que es lo mismo $X = A$.

$ii) \Rightarrow i)$ Tomemos a la subfamilia $\mathscr{U}$ del conjunto potencia de X, que consta de todos los subconjuntos finitos de X. $\emptyset \in \mathscr{U}$ porque cumple la definición por vacuidad, mientras que si $A \in \mathscr{U}$ y $x \in X$, entonces cada subfamilia de la familia $\mathscr{P}(A)$ tiene un elemento minimal, tomando una subfamilia arbitraria $\mathscr{A}_0 \subset \mathscr{P}(A)$ con elemento minimal A_0 tal que $\forall B \in \mathscr{A}_0$, se tiene que $A_0 \subset B$, por lo que $\forall B \in \mathscr{A}_0$ se cumple que $A_0 \cup \{x\} \subset B \cup \{x\}$, demostrando que el conjunto $A_0 \cup \{x\}$ es un elemento minimal de la subfamilia $\mathscr{A}_0 \cup \{x\}$. Como dicha subfamilia fue arbitraria, se concluye que toda subfamilia de la familia $\mathscr{P}(A \cup \{x\})$ tiene un elemento minimal, por lo tanto, es finito $A \cup \{x\}$ y además por la definición $A \cup \{x\} \in \mathscr{U}$.

Como $\mathscr{U}$ cumple las dos propiedades de $ii)$, su conclusión dice que $X \in \mathscr{U}$, por lo tanto, X es finito. $\square$

Proposición 3.7.2. *Si X, Y son finitos, también lo es $X \cup Y$.*

Demostración. Sea $\mathscr{U}$ cualquier familia no vacía de $\mathscr{P}(X \cup Y)$: por ser X finito, el conjunto $\mathscr{B} = \{C \cap X | C \in \mathscr{U}\}$ tiene un elemento minimal que llamaremos B_0, debido a que la familia $\mathscr{P}(X)$ tiene un subconjunto minimal A_0 con $A_0 \subset A$ para toda $A \subset X$, si $B \subset Y$:

$$A_0 \subset A_0 \cup B \subset A \cup B = C \subset X \cup Y,$$

con C arbitrario, por lo que $B_0 = A_0 \cap C \subset X \cap C$ para todo $C \subset X \cup Y$. De casi la misma manera, $\mathscr{C} = \{C \cap Y | C \in \mathscr{U}, A \cap X = B\}$ tiene un elemento minimal D_0, porque Y es finito, así encontramos un elemento minimal $B_0 \cup C_0$ de $\mathscr{U}$, por lo tanto, $X \cup Y$ es finito. □

Proposición 3.7.3. *Uniones finitas de conjuntos finitos, son finitas.*

Demostración. Sea $\mathscr{M}$ un conjunto finito de conjuntos finitos, demostraremos que $\bigcup \mathscr{M}$ es finito. Sea:

$$\mathscr{U} = \{\mathscr{B} \subset \mathscr{M} | \text{con} \bigcup \mathscr{B} \text{ finito}\},$$

como $\emptyset \subset \mathscr{M}$, entonces $\emptyset \in \mathscr{U}$. Siendo $\mathscr{B} \in \mathscr{U}$ y $A \in \mathscr{M}$ se cumple que tanto A (por la definición de $\mathscr{M}$) como $\bigcup \mathscr{B}$ (por la definición de $\mathscr{U}$) son finitos. Por la proposición anterior, $\bigcup \mathscr{B} \cup A$ es finito, por lo cual $\bigcup \mathscr{B} \cup A \in \mathscr{U}$, y al ser A arbitrario, se satisfacen las dos condiciones de la equivalencia de finitez de la proposición 3.7.1, concluyéndose que $\mathscr{M} \in \mathscr{U}$. Por la definición de $\mathscr{U}$ se tiene que $\bigcup \mathscr{M}$ es finito. □

Proposición 3.7.4. *Si tenemos al conjunto X finito, entonces $\mathscr{P}(X)$ será finito.*

Demostración. Sea la familia:

$$\mathscr{U} = \{A \subset X | \text{ con } \mathscr{P}(A) \text{ finito}\},$$

de la misma manera que antes, $\emptyset \in \mathscr{U}$, y tomando un $x \in X$ (que es finito) y $B \in \mathscr{B} \subset \mathscr{P}(\mathscr{P}(A))$, dicho $\mathscr{B}$ tiene un elemento minimal B_0 porque $\mathscr{P}(A)$ es finito, y se cumple que $\forall B \in \mathscr{B}$ se tiene $B_0 \subset B$, uniendo $\{x\}$, $\forall B \in \mathscr{B}$ se tiene que $B_0 \cup \{x\} \subset B \cup \{x\}$, resultando que toda subfamilia de $\mathscr{P}(\mathscr{P}(A \cup \{x\}))$ tiene un elemento minimal, por lo que $\mathscr{P}(A \cup \{x\})$ es finito, por eso $A \cup \{x\} \in \mathscr{U}$. De la conclusión del segundo inciso de la proposición 3.7.1, se tiene que $X \in \mathscr{U}$ y por lo tanto $\mathscr{P}(X)$ es finito.

□

Proposición 3.7.5. *La imagen de un conjunto finito, es finita.*

Demostración. Sea X un conjunto finito. Y sea $f : X \mapsto Y$ una función suprayectiva ($f(X) = Y$), demostraremos que si $\mathscr{U}$ es cualquier subfamilia no vacía de $\mathscr{P}(Y)$, entonces tiene un elemento minimal. Sea

$$\mathscr{B} = \{f^{-1}(A) | A \in \mathscr{U}\},$$

dicha familia es un subconjunto de $\mathscr{P}(X)$ y además es no vacía porque f es suprayectiva[34] y $\mathscr{U}$ es no vacía. Se cumple que $\mathscr{B} \subset \mathscr{P}(X)$ y porque X es finito, ésta tiene un elemento minimal B_0 tal que $\forall B \in \mathscr{B}$ se tiene que $B_0 \subset B$, por eso $\forall B \in \mathscr{B}$ se cumple que $f(B_0) \subset f(B)$, por lo que $f(B_0)$ es un elemento minimal de $\mathscr{U}$. □

La siguiente proposición relaciona al concepto de D-finito con el de finito.

Proposición 3.7.6. *Todo conjunto finito es D-finito.*

[34] Para funciones suprayectivas las imágenes inversas de conjuntos no vacíos son no vacías.

Demostración. Demostraremos que si no se cumple la conclusión, no se cumple la hipótesis, que es el equivalente lógico a lo que queremos demostrar. Sea X D-infinito, por lo que existe una función inyectiva $f : \mathbb{N} \mapsto X$. Consideremos la familia:

$$\mathscr{U} = \{\{f(m)|m \geq n\}|n \in \mathbb{N}\} = \{A_n|n \in \mathbb{N}\},$$

dicha familia de subconjuntos de X aunque es no vacía, no contiene un elemento minimal, ya que $A_{n+1} \subset A_n$ para toda $n \in \mathbb{N}$, i.e., es una familia decreciente, y por ser f inyectiva, ningún conjunto es igual, i.e. $A_n \neq A_m$ si $n \neq m$, al haber encontrado una familia sin elemento minimal, se tiene que X es infinito. □

Para el siguiente teorema[35] hay que recordar que $|A| \leq^* |B|$ significa que existe una función suprayectiva $f : B \mapsto A$.

Teorema 3.7.2. *Son equivalentes:*

i) $\aleph_0 \leq^* |X|$.

ii) $\mathscr{P}(X)$ *es D-infinito.*

Teorema 3.7.3. *Son equivalentes:*

i) X es finito ssi es D-finito.

ii) Uniones D-finitas de conjuntos D-finitos son D-finitas.

iii) Imágenes de conjuntos D-finitos son D-finitas.

iv) El conjunto potencia de todo conjunto D-finito es D-finito.

v) Para cada conjunto X se cumple que $\aleph_0 \leq |X|$ *o que* $|X| \leq \aleph_0$.

Demostración. $i) \Rightarrow ii)$ Es lo que afirma la proposición 3.7.3.

$ii) \Rightarrow i)$ Sea X un conjunto D-finito y $f : X \mapsto Y$ una función suprayectiva. El conjunto $Y = f(X) = \bigcup_{x \in X}\{f(x)\}$ es una unión D-finita de conjuntos D-finitos, y por la hipótesis, será Y D-finito.

$iii) \Rightarrow iv)$ Demostraremos que si se cumple $iii)$, y tenemos un conjunto con conjunto potencia D-infinito, entonces el propio conjunto será D-infinito. Sea $\mathscr{P}(X)$ un conjunto D-finito, el teorema anterior nos ayuda a afirmar que existe una función suprayectiva $f : X \mapsto \mathbb{N}$, como el conjunto $\mathbb{N}$ es D-infinito (por ejemplo, existe una biyección entre los naturales y los números pares), tenemos la negación de la conclusión de $iii)$, que la imagen de X es un conjunto D-infinito, por lo que se cumple la negación de su hipótesis, i.e., X es D-infinito.

$iv) \Rightarrow i)$ Por la proposición 3.7.6, solo falta demostrar que todo conjunto D-finito es finito, o lo que es lo mismo, que todo conjunto infinito es D-infinito. Sea $f : \mathbb{N} \mapsto \mathscr{P}(\mathscr{P}(X))$ definida por $f(n) = \{A \subset X||A| = n\}$. Esta función es inyectiva, porque si $n \neq m$ y si $A \in f(n)$ entonces $|A| = n \neq m \Rightarrow A \notin f(m)$, por lo que $f(n) \neq f(m)$. Teniendo una inyección de los naturales a $\mathscr{P}(\mathscr{P}(X))$, el teorema 3.7.1 aplica y $\mathscr{P}(\mathscr{P}(X))$ es D-infinito, la hipótesis $iv)$ dice que $\mathscr{P}(X)$ es D-infinito, y por la misma hipótesis, se tiene que X es D-infinito.

$i) \Rightarrow v)$ Supongamos que las definiciones de finito y D-finito son equivalentes. Supongamos que tenemos X D-infinito, por el teorema 3.7.1 $\aleph_0 \leq |X|$. Ahora supongamos que X no es D-infinito, que

[35] La demostración está en [Her, p. 47].

por hipótesis es lo mismo que decir que no es infinito, y por definición, sería finito. Se cumpliría así que $|X| = n \in \mathbb{N}$, y por lo tanto $|X| \leq \mathbb{N}$ porque existe una inyección de n a los naturales. Se concluye que $\aleph_0 \leq |X|$ o que $|X| \leq \aleph_0$.

$v) \Rightarrow i)$ Si tenemos un conjunto X, supongamos que $\aleph_0 \leq |X|$, por el teorema 3.7.1 sabemos que X será D-infinito, y por la proposición 3.7.6 X será infinito. Ahora, si $|X| \leq \aleph_0$ y X es D-finito, hay que demostrar que X es finito: como X es D-finito, por el teorema 3.7.1 $\aleph_0 \not\leq |X|$, lo que nos dice junto con la otra hipótesis que $|X| < \aleph_0$, por lo tanto existe un $n \in \aleph_0$ tal que $|X| = n$, diciendo que X es finito. $\square$

Los siguientes dos resultados (el segundo que depende del Axioma de Elecciones denumerables) nos ayudarán para demostrar la equivalencia de definiciones.

Teorema 3.7.4. *Todo conjunto que tiene un subconjunto denumerable, es D-infinito.*

Demostración. Tomando un conjunto A que cumpla la hipótesis, sea $B \subset A$ denumerable, por lo tanto, existe una biyección $f : B \mapsto \mathbb{N}$. Definiendo a g mediante:

$$g(x) = \begin{cases} f^{-1}(f(x)+1) & \text{si } x \in Dom\ f, \\ x & \text{si } x \in A - Dom\ f. \end{cases}$$

Lo que se acaba de definir, es una función tal que si $x \in Dom\ f$, se tiene que $f(x) = n$ para un único natural n, por lo que se puede escribir de la forma $x = x_n$, y será $g(x) = g(x_n) = f^{-1}(n+1) = x_{n+1}$. Con ello, al ser $Dom\ f = B \subset A \Rightarrow Dom\ g = A$ y la función identidad en $A - Dom\ f$, ahí es inyectiva; si $x \in A - Dom\ f$ y $z_n \in Dom\ f$ con $x \neq z$, se cumple que $x = g(x) \neq g(z_n) = z_{n+1}$; y también será inyectiva en $Dom\ f$, porque si $x_n \neq x_m$, entonces $f(x_n)+1 = n+1 \neq m+1 = f(x_m)+1$ ya que f es biyectiva, por lo cual $g(x_n) = f^{-1}(n+1) \neq f^{-1}(m+1) = g(x_m)$ (porque f es biyectiva), encontrando que g es inyectiva. El rango de la función es $g(A) = A - \{f^{-1}(0)\}$ porque no se puede tener un $n \in \mathbb{N}$ tal que $n+1 = 0$ para que $g(x) = f^{-1}(0)$.

La conclusión es que existe una biyección g entre A y un subconjunto propio de A que es $A - \{f^{-1}(0)\}$, por lo tanto A es D-infinito. $\square$

Teorema 3.7.5. *Si un conjunto es infinito, entonces tendrá un subconjunto denumerable.*

Demostración. Sea A un conjunto infinito, y sea f una función de elección dada por el Axioma. Definimos a $g : \mathbb{N} \mapsto A$ como:

$$g(n) = \begin{cases} f(A) & \text{si } n = 0, \\ f(A - \{g(k) :\ k \leq n-1\}) & \text{cualquier otro caso.} \end{cases}$$

La idea de la función es asignar 0 al primer elemento x_0 escogido de A, asignar 1 al elemento x_1 escogido de $A - \{x_0\}$, y así sucesivamente. Como siempre que $f(B) \in B$, la función g es inyectiva, porque la definición dice $g(n+1) \notin \{g(k) :\ k \leq n\}$, y por lo tanto si $m \neq n$, $g(m) \neq g(n)$.

Suponiendo que existe un n tal que $\varnothing = A - \{g(k) :\ k < n\}$, tendriamos una biyección $|A| = n$, contradiciendo que A es infinito. Por lo tanto, $\forall n \in \mathbb{N}$ se tiene que $A - \{g(k) :\ k < n\} \neq \varnothing$, llegando a que $|g(\mathbb{N})| = \aleph_0$, pero $g(\mathbb{N}) \subset A$, conluyendo lo pedido. $\square$

Podemos ya demostrar el teorema más importante de esta subsección, que dice:

Teorema 3.7.6. *Asumiendo que se cumple el Axioma de Elecciones Denumerables, las definiciones de conjunto finito y D-finito son equivalentes.*

Demostración. Sabemos por la proposición 3.7.6 que todo conjunto finito es D-finito. Solo falta demostrar que todo conjunto D-finito es finito, o lo que es lo mismo, que todo conjunto infinito es D-infinito.

Por el teorema anterior, si tenemos un conjunto infinito, éste tendrá un subconjunto denumerable, y por el anterior a ese, si tenemos un subconjunto denumerable de un conjunto, éste será D-infinito. □

Después de estos resultados, podemos concluir que al no asumir el Axioma Denumerable, podría pasar en algún sistema axiomático que:

1. Uniones D-finitas de conjuntos D-finitos pueden ser D-infinitas.
2. El conjunto potencia de un conjunto D-finito podría ser D-infinito.
3. Un conjunto D-infinito puede ser la imagen de un conjunto D-finito. Incluso, cualquier $\aleph_\alpha$ no importando cuán grande, puede ser la imagen de un conjunto D-finito.
4. Los números cardinales de conjuntos D-finitos podrían no satisfacer el principio de inducción.
5. Existan modelos en donde hay conjuntos D-finitos y al mismo tiempo infinitos.
6. Existan modelos en donde los conjuntos son finitos ssi son D-finitos, pero no se satisface el Axioma de Elecciones Contables.

3.7.2. ¿Qué es continuo?

Requerimos la definición de convergencia de una sucesión.

Definición 3.7.3. *Si se tiene una sucesión en los reales $\{x_n\} \subset \mathbb{R}$, se dice que la sucesión converge al punto $x \in X$ y se escribe $\lim x_n = x$, si se cumple que $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$ tal que $\forall n \geq N$, se cumple que $|x - x_n| < \varepsilon$.*

Se tienen los siguientes dos conceptos:

Definición 3.7.4.

1. *Una función $f : \mathbb{R} \mapsto \mathbb{R}$ se dice continua en $x \in Dom\ f$ ssi $\forall \varepsilon > 0, \exists \delta > 0$ tal que si $|x - y| < \delta$ entonces $|f(x) - f(y)| < \varepsilon$.*
2. *Una función $f : \mathbb{R} \mapsto \mathbb{R}$ se dice secuencialmente continua en $x \in Dom\ f$ ssi para toda $\{x_n\}$ que converga a x se tiene que $\{f(x_n)\}$ converge a $f(x)$.*

Una función que sea continua en todo su dominio, se dirá continua. Una función que sea secuencialmente continua en todo su dominio, se dirá secuencialmente continua.

Teorema 3.7.7. *Asumiendo que se cumple el Axioma de Elecciones Denumerables, las definiciones anteriores son equivalentes.*

Demostración. $1) \Rightarrow 2)$ Si tenemos que f es continua en x, entonces $\forall \varepsilon > 0, \exists \delta > 0$ tal que si $|x - y| < \delta$ entonces $|f(x) - f(y)| < \varepsilon$. Suponiendo una sucesión cualquiera $\{x_n\}$ que converga a x (existe por lo menos $\{x_n\} = x$), hay que demostrar que $\{f(x_n)\}$ converge a $f(x)$. Como $\lim x_n = x$, entonces $\forall \delta$, $\exists N \in \mathbb{N}$ tal que si $n \geq N$ se tiene que $|x_n - x| < \delta$.

Sea $\varepsilon > 0$ arbitrario, existirá $\delta > 0$ de la definición de convergencia de la sucesión, y por ello existirá $N \in \mathbb{N}$, tal que $|x_n - x| < \delta$ para toda $n \geq N$, y por la definición de continuidad, se cumple que $|f(x) - f(x_n)| < \varepsilon$ para toda $n \geq N$, por lo tanto, se cumple la definición de convergencia y $\lim f(x_n) = f(x)$.

$2) \Rightarrow 1)$ Demostraremos que si no se cumple $1)$ entonces no se cumple $2)$, que es el equivalente lógico.

Si f no es continua en x, $\exists \varepsilon > 0$ tal que $\forall \delta > 0$ existe un a tal que $|x-a| < \delta$ pero $|f(x) - f(a)| \geq \varepsilon$, en particular, para cada $n \in \mathbb{N}$ podemos elegir un x_n tal que $|x_n - x| < \frac{1}{n}$ y $|f(x_n) - f(x)| \geq \varepsilon$, i.e., $x_n \in (-\frac{1}{n} + x, x + \frac{1}{n})$. Por el Axioma de Elecciones Denumerables, dicha sucesión creada $\{x_n\}$ converge a x debido a que la longitud del intervalo decrece al tender n a infinito, sin embargo, se cumple que $|f(x_n) - f(x)| \geq \varepsilon$ porque f no es continua en x, se sigue que no se cumple que $\lim f(x_n) = f(x)$, debido a que estamos negando la definición de la convergencia.

$\square$

3.7.3. ¿Qué es cerrado?

Definición 3.7.5. *Si se tiene un conjunto A, un punto x estará en la cerradura de A ssi:*

1. *Toda vecindad de x intersecta a A.*
2. *Para alguna sucesión $\{x_n\} \subset A$, se cumple que $x = \lim x_n$.*

La primera definición se refiere a conjuntos cerrados, mientras que la segunda a conjuntos secuencialmente cerrados.

Teorema 3.7.8. *Asumiendo que se cumple el Axioma de Elecciones Denumerables, las definiciones anteriores son equivalentes.*

Demostración. $1) \Rightarrow 2)$

Sea $B_n = (-\frac{1}{n} + x, x + \frac{1}{n})$ para toda $n \in \mathbb{N}$, entonces B_n es una vecindad de x y por hipótesis $\emptyset \neq A \cap B_n$. Tomando a cualquier elemento $x_n \in A \cap B_n$ que el Axioma de Elecciones Contables nos permite elegir, formamos una sucesión $\{x_n\}$ que demostraremos converge a x.

Como $x_n \in B_n = (-\frac{1}{n} + x, x + \frac{1}{n})$, se tiene que $\forall n \in \mathbb{N}$:

$$-\frac{1}{n} + x < x_n < x + \frac{1}{n} \Leftrightarrow -\frac{1}{n} < x_n - x < \frac{1}{n} \Leftrightarrow |x_n - x| < \frac{1}{n},$$

y tomando a un N tal que[36] $\frac{1}{N} < \varepsilon$ se cumplirá que $\forall \varepsilon > 0, \exists N \in \mathbb{N}$ tal que $\forall n \geq N$ se tiene que $|x_n - x| < \frac{1}{n} < \frac{1}{N} < \varepsilon$ por lo que $\{x_n\} \subset A$ y $\lim x_n = x$.

$2) \Rightarrow 1)$

Por hipótesis, existe $\{x_n\} \subset A$ tal que $\forall \varepsilon > 0, \exists N \in \mathbb{N}$ tal que $\forall n \geq N$ se cumple $|x_n - x| < \varepsilon$.

Sea V_x una vecindad arbitraria de x, queremos demostrar que $\emptyset \neq A \cap V_x$. El conjunto $\mathscr{B} = \{B_{\frac{1}{n}}(x) : n \in \mathbb{N},\, x \in \mathbb{R}\}$ con $B_{\frac{1}{n}}(x) = \{y : |x - y| < \frac{1}{n}\}$ es una base de $\mathbb{R}$, y al ser V_x abierto (esto por la definición de vecindad), se tiene que $\exists m \in \mathbb{N}$ tal que $B_m = (-\frac{1}{m} + x, x + \frac{1}{m}) \subset V_x$ (por el teorema 3.6.3). Al cumplirse para toda ε, sea $\varepsilon = \frac{1}{m}$, entonces $\exists N \in \mathbb{N}$ tal que $\forall n \geq N$:

$$\begin{aligned}
|x_n - x| < \frac{1}{m} \quad &\Leftrightarrow \quad x_n \in (-\frac{1}{m} + x, x + \frac{1}{m}) \\
&\Leftrightarrow \quad x_n \in (-\frac{1}{m} + x, x + \frac{1}{m}) \cap A \\
&\subset \quad V_x \cap A \\
&\Rightarrow \quad V_x \cap A \neq \emptyset,
\end{aligned}$$

[36]Dicho N siempre existe debido a que $\lim_{n\to\infty} \frac{1}{n} = 0$.

usando el hecho de que $\{x_n\} \subset A$.

$\square$

3.7.4. Una función muy extraña

Llamaremos a una $f : \mathbb{R} \mapsto \mathbb{R}$ aditiva si se cumple que

$$f(x+y) = f(x) + f(y)$$

para todo $x,\ y \in \mathbb{R}$. Dicha ecuación se llama la ecuación de Cauchy. Denotaremos a las funciones aditivas y continuas de la forma $f(x) = ax$ para toda $x \in \mathbb{R}$ por f_a, con $a \in \mathbb{R}$ fija.

Sea f una función aditiva, y sea $f(1) = a$. Se cumple que:

$$f(2) = f(1) + f(1) = a \cdot 2\,, \qquad f(3) = f(2) + f(1) = a \cdot 3,$$

aplicando inducción, si $f(n) = a \cdot n$, entonces $f(n+1) = f(n) + f(1) = a \cdot n + a = a \cdot (n+1)$, por lo que $f(n) = a \cdot n$ si $n \in \mathbb{N} - \{0\}$.

Calculando $f(0) = f(0) + f(0) = 2 \cdot f(0)$ por lo que $f(0) = 0$. Además, $f(-n) + f(n) = f(0) = 0$ por lo que $f(-n) = -f(n) = -a \cdot n$ si $n \in \mathbb{N}$. Vemos que:

$$a = f(1) = f\left(\sum_{k=1}^{n} \frac{1}{n}\right) = \sum_{k=1}^{n} f\left(\frac{1}{n}\right) = n \cdot f\left(\frac{1}{n}\right),$$

concluyendo que $f\left(\frac{1}{n}\right) = a \cdot \frac{1}{n}$. Además, si $\frac{m}{n} \in \mathbb{Q}$ es racional:

$$f\left(\frac{m}{n}\right) = f\left(\sum_{k=1}^{m} \frac{1}{n}\right) = \sum_{k=1}^{m} f\left(\frac{1}{n}\right) = m \cdot f\left(\frac{1}{n}\right) = a \cdot \frac{m}{n}.$$

Todo lo anterior demuestra que si $r \in \mathbb{Q}$, se tiene que $f(r) = a \cdot r$, por lo que podríamos pensar que si $x \in \mathbb{R}$ se cumpliría que $f(x) = a \cdot x = f_a(x)$, i.e. que cualquier función aditiva es de la forma f_a para algún a. Si suponemos que f es continua, se cumple que $f(x) = a \cdot x = f_a(x)$ para todo real x, la pregunta es: ¿existen soluciones no continuas de la ecuación de Cauchy?.

Dichas soluciones no continuas no se pueden encontrar en ZF, existen modelos en ZF en los que todas las soluciones de la ecuación de Cauchy son continuas, pero en ZFC hay más soluciones de las que creeríamos:

Teorema 3.7.9. *Existen $2^{\aleph_0}$ soluciones continuas f a la ecuación de Cauchy y existen $2^{(2^{\aleph_0})}$ soluciones no continuas a dicha ecuación en ZFC.*

Demostración. Para cada $r \in \mathbb{R}$, la función $f : \mathbb{R} \mapsto \mathbb{R}$ definida por $f(x) = r \cdot x$ es una solución continua de la ecuación de Cauchy, y no hay más.

Para ZFC, si consideramos el espacio vectorial $\mathbb{R}$ sobre el campo $\mathbb{R}$, dicho espacio tiene una base de Hamel B por el teorema 3.4.1 p. 116, como subconjunto de $\mathbb{R}$ se cumple que $|B| = 2^{\aleph_0}$, y por el teorema de Hahn-Banach p. 140 podemos extender de manera única una función lineal $B \mapsto \mathbb{R}$ a una función lineal $\mathbb{R} \mapsto \mathbb{R}$ y se cumple que:

$$|\mathbb{R}^B| = |\mathbb{R}|^{|B|} = (2^{\aleph_0})^{2^{\aleph_0}} = 2^{\aleph_0 \cdot 2^{\aleph_0}} = 2^{2^{\aleph_0}}$$

son el número de soluciones a la ecuación de Cauchy, y como solo hay $2^{\aleph_0}$ que son continuas, por el teorema 3.3.6 de la p. 109 hay $2^{2^{\aleph_0}}$ no continuas. $\square$

Teorema 3.7.10. *Existe una función aditiva que no es de la forma $f = f_a$ para toda $a \in \mathbb{R}$ si suponemos que se cumple el Axioma de Elección.*

Demostración. Sea B una base de Hamel para $\mathbb{R}$, y elijamos a un $x_0 \in B$ fijo, definiendo:

$$f(x_0) = \begin{cases} r_k & \text{si } x = \sum_{i=1}^{n} r_i x_i \text{ y además } x_0 = x_k \text{ con } k \in \{1, \ldots, n\}, \\ 0 & \text{cualquier otro caso.} \end{cases}$$

Si tenemos que $x,\ y \in \mathbb{R}$, hay 3 casos, considerando que escribimos a $x_i,\ y_i \in B$ y a $r_i,\ s_i \in \mathbb{Q}$:

1. Si las representaciones de las combinaciones lineales de x y y no tienen a x_0, tampoco su suma, y se cumple que $f(x+y) = 0 = f(x) + f(y)$.

2. Si solo uno tiene representación donde x_0 sucede, sea $x = \sum_{i=1}^{n} r_i x_i$ con $x_0 = x_k$ y además $y = \sum_{i=1}^{m} s_i y_i$, luego, $x+y$ tendrá a x_0 en la k-ésima suma, y no puede pasar que $s_k = r_k$ porque sería $x_k = y_k = x_0$ contrario a que solo aparece en la representación de x, por lo que $f(x+y) = r_k = f(x) + f(y) = r_k + 0$.

3. Si se cumple que $x = \sum_{i=1}^{n} r_i x_i$ con $x_0 = x_k$ y además $y = \sum_{i=1}^{m} s_i y_i$ con $x_0 = y_j$, se cumplirá que $x+y$ tendrá como sumando a $(r_k + s_j)x_0$ y se dará que $f(x+y) = r_k + s_j = f(x) + f(y)$.

Lo anterior demuestra que f es aditiva. Debemos de tener en cuenta que, como B es una base, $0 \notin B$ y que B es infinito. De la definición y representación del mismo x_0, se cumple que $f(x_0) = 1$, y al ser independientes los elementos de B, se cumple que si $x_1 \neq x_0$ y $x_1 \in B$, $f(x_1) = 0$ al no estar x_0 en la representación de x_1.

La conclusión: si suponemos que se llega a cumplir para algún $a \in \mathbb{R}$ que $f = f_a$, se cumple:

$$f(x_0) = 1 = a \cdot x_0$$

diciendo que $a \neq 0$, mientras que:

$$f(x_1) = 0 = a \cdot x_1$$

diciendo que $a = 0$, llegando a una contradicción. □

Dichas funciones no solo son muchas en $\mathbb{R}$, también están casi por todos lados en el plano $\mathbb{R}^2$:

Teorema 3.7.11. *Si f es una función no continua y que satisface las ecuaciones de Cauchy, se tiene que su gráfica $G(f) = \{(x, f(x)) | x \in \mathbb{R}\}$ es densa en $\mathbb{R}^2$.*

Demostración. Sea $(x,y) \in \mathbb{R}^2$ arbitrario, demostraremos que si U es una vecindad de (x,y) cualquiera, se cumple que la intersección de U y de $G(f)$ es no vacía, y por el teorema 3.6.4 $G(f)$ será denso.

Por la propiedad de f, existen $a \neq 0$ y $b \neq 0$ tales que los números $\psi = \frac{f(a)}{a}$ y $\xi = \frac{f(b)}{b}$ son diferentes, ya que la función f no es de la forma f_a para algún real, y por eso $\frac{f(x)}{x}$ no es constante para x. Los vectores $u = (a, f(a)) = a(1, \psi)$ y $v = (b, f(b)) = b(1, \xi)$ son linealmente independientes en el espacio vectorial $\mathbb{R}^2$, por que ψ y ξ son diferentes, y al tener dos vectores linealmente independientes de ese espacio de dimensión 2, tenemos una base para $\mathbb{R}^2$. Los escalares que se necesitan para representar cualquier elemento (x,y) como combinación lineal de los elementos de la base son:

$$p = \frac{y - \xi x}{a(\psi - \xi)} \qquad \text{y} \qquad q = \frac{\psi x - y}{b(\psi - \xi)},$$

con $(x,y) = p\cdot u + q\cdot v$, ya que:

$$\begin{aligned} p\cdot u + q\cdot v &= \frac{y-\xi x}{a(\psi-\xi)}\cdot a(1,\psi) + \frac{\psi x - y}{b(\psi-\xi)}\cdot b(1,\xi) \\ &= \frac{1}{\psi-\xi}\cdot((y-\xi x, \psi y - \psi\xi x) + (\psi x - y, \xi\psi x - \xi y)) \\ &= \frac{1}{\psi-\xi}\cdot((\psi-\xi)x, (\psi-\xi)y). \end{aligned}$$

Sabemos que $\mathbb{Q}^2$ es denso en $\mathbb{R}^2$, ya que cualquier abierto en $\mathbb{R}^2$ contiene un punto con las dos coordenadas racionales (porque $\mathbb{Q}$ es denso en $\mathbb{R}$). El vector $p\cdot u + q\cdot v$ depende continuamente de p y q, por lo que existen números racionales r y s tales que $r\cdot u + s\cdot v \in U$ al tomar valores cercanos a (x,y). Con esto, se tiene que:

$$r\cdot u + s\cdot v = (ra+sb, rf(a)+sf(b)) = (ra+sb, f(ra+sb)),$$

y de ahí, se deduce que dicho punto pertenece a la gráfica de f, i.e., $r\cdot u + s\cdot v \in U\cap G(f)$. □

También pasa que:

Teorema 3.7.12. *Las funciones no continuas que satisfacen la ecuación de Cauchy, son no medibles.*

Demostración. Sea f una tal función. Sean $a\neq 0$ y $b\neq 0$ tales que $f(a)=1$ y $f(b)=0$, si no se cumple, se pueden tomar como la demostración pasada, unos puntos tales que $\frac{f(a)}{a}\neq\frac{f(b)}{b}$, y se reemplazaría a f por g, con $g = (bf(a)-af(b))^{-1}(bf(x)-f(b)x)$.

Para todo $n\in\mathbb{Z}$, definamos

$$A_n = f^{-1}[n,n+1)$$

y elijamos cada q_n racional tal que $|na-q_nb| < \frac{1}{2}$ (recuérdese que estamos suponiendo que se cumple el Axioma). Definiendo a $B_0 = A_0\cap[-\frac{1}{2},\frac{3}{2}]$ y para todo $n\neq 0$ definamos:

$$B_n = \{x+na-q_nb|\ \ x\in B_0\},$$

Si $x\in(A_n\cap[0,1])$, entonces $f(x)\in[n,n+1)$ y $x\in[0,1]$, por lo que si $y = x-(na-q_nb)$:

$$f(y) = f(x)+f(-na+q_nb) = f(x)-nf(a)+q_nf(b) = f(x)-n,$$

y se tendrá que $f(y)\in[0,1)$, y entonces $y\in f^{-1}[0,1)$. Como $x\in[0,1]$ y $|na-q_nb|<\frac{1}{2}$, se tiene que $y = x-(na-q_nb)\in[-\frac{1}{2},\frac{3}{2}]$, por lo tanto, si $x\in A_n\cap[0,1]$, entonces $y = x-(na-q_nb)\in A_0\cap[-\frac{1}{2},\frac{3}{2}] = B_0$, y volteando $x = y+(na-q_nb)\in B_n$, la consecuencia es que:

$$(A_n\cap[0,1])\subset B_n\subset[-1,2],$$

de lo anterior, se sigue que:

$$[0,1]\cap\bigcup_{n\in\mathbb{Z}} A_n = \bigcup_{n\in\mathbb{Z}}([0,1]\cap A_n)\subset\bigcup_{n\in\mathbb{Z}} B_n\subset[-1,2].$$

Demostraremos que los B_n son disjuntos. Supongamos que $x+na-q_nb = y+ma-q_mb$, con $x,\ y\in B_0$, entonces $x,\ y\in A_0$ y por eso $f(x),\ f(y)\in[0,1)$. Se sigue que:

$$
\begin{aligned}
x-y+a(n-m)+b(q_m-q_n) &= 0 \\
\Rightarrow 0 = f(0) &= f(x-y)+(n-m)f(a)+(q_m-q_n)f(b) \\
\Rightarrow 0 &= f(x)-f(y)+n-m,
\end{aligned}
$$

por las propiedades de x y de y, pasa que $f(x)-f(y)\in(0,1)$, mientras que $n,\ m\in\mathbb{Z}$. Para que la suma de éstos tres sea cero, se requiere que $f(x)=f(y)$, pero eso implicaría que $n=m$, por lo que necesariamente $B_m=B_n$, demostrando que los B_n son disjuntos. También son congruentes, si μ denota su medida, para todo entero $\mu(B_n)=\mu(B_0)$, y ocupando la aditividad infinita de la medida de Lebesgue:

$$1=\mu([0,1])=\sum_{n\in\mathbb{Z}}\mu(B_0)\leq\mu([-1,2])=3,$$

entonces, B_0 y también A_0 son no medibles, porque si un sumando es cero, la suma en total sería cero, y si algún sumando es positivo, la suma en total sería infinita. ☐

Acabamos también de demostrar que:

Teorema 3.7.13. *En ZFC, existen $2^{2^{\aleph_0}}$ funciones no medibles, que satisfacen la ecuación de Cauchy, y existen $2^{2^{\aleph_0}}$ subconjuntos no medibles de $\mathbb{R}$.*

Demostración. El teorema anterior nos generó conjuntos y funciones no medibles, la cardinalidad de éstos resulta del teorema 3.7.9. ☐

3.7.5. Uniones contables

Este teorema se ocupa varias veces en la carrera, y no necesita del Axioma si es efectiva la numeración del conjunto, sin embargo, cuando no tenemos el ejemplo de dicha numeración, demostraremos usando el Axioma de Elecciones Contables que:

Teorema 3.7.14. *La unión contable de conjuntos contables, es contable.*

Demostración. Sea S un conjunto contable, cuyos elementos son contables, y sea $A=\bigcup S$, demostraremos que A es contable. Debido a que S es contable, existe una sucesión inyectiva $\{A_n\}_{n\in\mathbb{N}}$ tal que podemos escribir $S=\{A_n\}_{n\in\mathbb{N}}$, además, para cada $n\in\mathbb{N}$ el conjunto A_n es contable, por lo que también existe una sucesión con rango A_n.

Escogemos una sucesión para cada $n\in\mathbb{N}$, es decir, sea S_n el conjunto de todas las sucesiones cuyo rango es A_n, y sea F una función de elección sobre el conjunto $\{S_n\}_{n\in\mathbb{N}}$. F tomará un $s_n=F(S_n)$ para cada n. Con todas las sucesiones $s_n=\{a_n(k)\}_{k\in\mathbb{N}}$ elegidas para cada n, se obtiene una función f biyectiva de $\mathbb{N}\times\mathbb{N}$ a A haciendo $f(n,k)=a_n(k)$, y como podemos siempre hacer a los conjuntos disjuntos, si $n\neq m$, entonces $f(n,k)\in A_n$ y $f(m,l)\in A_m$ entonces $a_n(k)\neq a_m(l)$. Como $s_n(\mathbb{N})=A_n$, se tendrá que $f(\mathbb{N},\mathbb{N})=A$; al ser $\mathbb{N}\times\mathbb{N}$ contable, A su imagen y f biyección, se concluye que A es contable. ☐

Corolario 3.7.1. *El conjunto $\mathbb{R}$ de todos los números reales, no es la unión contable de conjuntos contables.*

Demostración. El conjunto $\mathbb{R}$ es incontable, como demostró Cantor por su método de diagonalización, por lo que si fuera la unión contable de conjuntos contables, sería contable por el teorema anterior, contradiciendo que es incontable. ☐

3.7.6. Sobre sombreros

El Axioma de Elección también puede ser empleado para resolver acertijos:

Un número infinito numerable de prisioneros están colocados sobre una línea que representa a los números naturales, mirando cada uno de ellos hacia la dirección positiva, es decir, ve a los infinitos prisioneros que tiene enfrente, y que están numerados de uno en uno a partir del uno. A todos los prisioneros se les coloca un sombrero blanco o negro y éste no puede ver de que color es. Luego se le pregunta a cada prisionero por el color del sombrero que trae puesto, comenzando por el prisionero que está en la posición del primer natural. Ningún prisionero puede oír lo que los anteriores dicen, ni si aciertan o no. En caso de dar una respuesta correcta, el prisionero es liberado, si no la da, se le encierra para siempre. La pregunta es: en caso de poder acordar una estrategia entre todos de antemano para responder, ¿existe una estrategia en la que se pueda asegurar cuántos se salvarán con seguridad? y ¿cuál es dicha estrategia?.

El hecho de que ningún prisionero puede obtener información de los prisioneros que están a su espalda, hace parecer que no es posible idear una estrategia, y se podría creer que solo podemos tratar de atinarle. Pero una forma del Axioma nos ayuda a resolver el problema, haciendo sorprendente el hecho de que sin escuchar las respuestas de los anteriores, un número infinito de prisioneros acierten su color.

Llamaremos cero al color de sombrero blanco y uno al negro. Al tener puestos todos sus sombreros, obtenemos una representación de sucesiones infinitas de ceros y de unos. Definiremos una relación de equivalencia entre sucesiones de ceros y unos: dos sucesiones serán equivalentes si son iguales excepto en un número finito de términos, i.e., si existe un natural n tal que $\forall m \geq n$ se tiene que $x_m = y_m$. Claramente toda sucesión se relaciona consigo misma, y si una se relaciona con otra, también se cumple al revés, por eso la relación es simétrica y reflexiva; para ver que es transitiva, si x_n se relaciona con y_n y ésta se relaciona con z_n, donde las dos primeras son iguales a partir de n y las dos últimas a partir de m, tomando a $k = \text{máx}\,(n,m)$, se cumple que x_p y z_p son iguales para todo p mayor o igual que k.

Aquí es donde los prisioneros usan el Axioma de Elecciones Contables, para elegir un representante de cada clase de equivalencia formada, memorizando cada una de esas sucesiones representantes. Al haberles colocado los sombreros a todos, el prisionero verá un número infinito de términos de una sucesión excepto a los anteriores que son finitos, reconocerá la clase de equivalencia en la que está y podrá recordar el representante de dicha clase de equivalencia. Conociendo el representante, elige de entre el cero y el uno el valor que tiene la sucesión en el número natural sobre el que está parado dicho prisionero, y el problema está resuelto: la sucesión que se representa a partir del orden actual de los sombreros debe estar en una clase de equivalencia, y por lo tanto es equivalente a la sucesión representante que están recitando los prisioneros, por lo que tras un número finito, dichas sucesiones son iguales, y todos esos prisioneros se salvarán.

Cabe señalar que se podría generalizar el problema a un número arbitrariamente grande de colores de sombreros, y de la misma forma se seguiría cumpliendo lo antes dicho.

3.7.7. Sobre ejemplos efectivos

En este ejemplo daremos una definición para la cuál existe una clasificación, pero un ejemplo de ella no se puede dar. No lo demostraré pero daré las razones por las que sucede.

Clasificaremos a todos los subconjuntos de los números naturales como «pequeños» y «no pequeños», definiendo la palabra pequeño que signifique:

1. Cualquier conjunto con un elemento o ningún elemento se llamará pequeño.

2. Cualquier unión de dos conjuntos pequeños es pequeña.

3. Un conjunto es pequeño si y solo si su complemento no es pequeño.

Estos son ejemplos que satisfacen cualesquiera dos de las características de pequeño:

1. Definiendo «pequeño» que signifique «finito», claramente se satisfacen las dos primeras reglas, pero la tercera no, porque los números pares y los impares no son finitos, pero son complemento uno del otro.

2. Diciendo que un conjunto es «pequeño» si el número 1 no es un elemento de dicho conjunto. Se satisfacen las dos últimas reglas, porque si dos conjuntos no tienen al 1, tampoco la unión, y si un conjunto lo contiene, su complemento no. Pero el conjunto $\{1\}$ es «no pequeño» mediante esta caracterización, sin embargo, tiene solamente un elemento, no se cumple la primera regla.

3. Diciendo que un conjunto es «pequeño», si contiene a lo más a uno de los tres siguientes elementos: 1, 2, 3. La primera regla se satisface debido a que si el conjunto contiene a uno de esos números, tiene un elemento y se llamará pequeño, si el conjunto no tiene elementos, tiene a lo más a alguno de esos tres números; la tercera regla se cumple porque si tiene a alguno de esos números, su complemento tendrá a los otros dos, diciendo que es «no pequeño»; sin embargo, no se satisface la segunda regla, porque los conjuntos «pequeños» $\{1\}$ y $\{2\}$, tienen como unión al conjunto «no pequeño» $\{1,2\}$ debido a que tiene dos elementos de los seleccionados.

La pregunta importante es: ¿existe un esquema de clasificación que satisfaga las 3 reglas?. Se cumple que dicho esquema de clasificación existe pero un ejemplo de dicho esquema no puede ser dado, es decir, un ejemplo de dicho esquema ¡no existe!. No estoy diciendo que aún no encontramos un ejemplo tal, lo que afirmo es que las demostraciones de existencia son intrínsecamente no constructivas, y no pueden ser reemplazadas por demostraciones constructivas (como cuando el Axioma afirma existe un conjunto no Lebesgue medible, o un conjunto de representantes de una clase de equivalencia), por lo que algún ejemplo de un esquema de clasificación que satisfaga las 3 reglas nunca podrá ser dado. La demostración de esto sale del alcance de esta tesis, pero para demostrar la existencia de dicho esquema de clasificación, se pueden llamar a «grandes» los elementos de algún ultrafiltro no principal sobre los naturales, y llamando a sus complementos «pequeños». Con dicha clasificación, cualquier conjunto más grande de un conjunto «grande» también lo es; el contrario es: si tenemos una clasificación de conjuntos «grandes» y «pequeños», los conjuntos «grandes» no necesariamente forman un ultrafiltro no principal, pero los conjuntos más grandes que los «grandes» si lo hacen. Ya se demostró en la página 125 que la existencia de ultrafiltros no principales se sigue del Axioma; pero probar que la demostración de la existencia es no constructiva (refiriéndonos por «ejemplo» algo cuya existencia pueda ser demostrada usando ZF+DC), es mucho más difícil.

3.7.8. Todas los conjuntos son interesantes

Veremos una aplicación del Axioma generalizando una aplicación del principio de inducción[37].

He aquí algunas características de los números naturales: la obsesión que tienen los matemáticos por los números primos y las distintas propiedades que éstos tienen los hacen interesantes, los números cuyos divisores excepto él mismo suman a dicho número son interesantes, los números que satisfacen la ecuación $a^2+b^2=c^2$ también lo son. Demostraré que todos los números naturales son «interesantes».

[37] Debo esta aplicación del Axioma a la idea en el conjunto de los números naturales que me compartió Alejandro Betanzos Gómez, mejor amigo de la carrera.

Supongamos el contrario, que existen números naturales los cuales no son interesantes, y tomemos al conjunto de todos ellos, dicho conjunto no es vacío, porque si lo fuera, todos los números serían interesantes contrario a la hipótesis, entonces, tenemos un conjunto no vacío de números naturales que cumplen con la propiedad de no ser interesantes. En dicho conjunto tenemos el orden natural de los naturales, del cuál obtenemos el natural menor de entre todos ellos, así podemos encontrar el menor número natural que no es interesante, i.e., ¡el primer natural que no es interesante!, dicha propiedad lo convierte en interesante, contradiciendo el hecho de que no lo es, por lo tanto, todos los números naturales son interesantes.

La demostración anterior se basa en una característica de todos los subconjuntos de los naturales: el buen orden. Aceptando el Axioma de Elección, el cuál ya se vio en la página 96 ser equivalente al Principio del Buen Orden: para todo conjunto se puede encontrar una relación que lo bien ordene. Generalizando:

Supongamos que tenemos un conjunto cualquiera no vacío, veremos que dicho conjunto solo contiene elementos interesantes. Si solo tuviera un elemento, éste es el elemento más interesante del conjunto, por ser el único, por ser el menor y mayor elemento del conjunto, la intersección y unión del conjunto (también se le podrían encontrar otras propiedades interesantes al elemento). Ahora supongamos que hay un conjunto que tuviera elementos no interesantes, tomaríamos al conjunto de todos los elementos no interesantes (es un conjunto debido a que consideramos un subconjunto de un conjunto), dicho conjunto tendría un buen orden por el Principio del Buen Orden, por lo que encontraríamos un elemento menor en dicho conjunto de elementos no interesantes, y al igual que antes, se convertiría en interesante, por lo que cualquier conjunto no solo contiene elementos interesantes, sino que también él mismo es interesante, ¡porque todos sus elementos son interesantes!.

El argumento anterior adquiere mayor validez al recordar que se puede encontrar un elemento menor para cualquier buen orden del conjunto, en particular, para el buen orden elegido por el Principio del Buen Orden, por lo que no es relativo el resultado del buen orden que se asigne al conjunto.

Si se admitiera una generalización del Axioma, por ejemplo el Axioma de Elección Global, se concluiría que todas las clases son interesantes.

Capítulo 4

CONCLUSIONES

El uso del Axioma de Elección para las matemáticas de hoy es esencial, principalmente por sus consecuencias, pero también porque es un axioma que históricamente ha estado presente desde el comienzo de las matemáticas modernas que aplican a conjuntos generales.

La controversia del Axioma de Elección nos ayudó a ser más meticulosos durante una demostración, la intuición y lo que parece una sencillez en generalizar algún resultado llevaron a veces a cometer errores grandes, como la suposición de la existencia de conjuntos que no necesariamente existen.

Al tratar de ser más precavidos, muchos se dieron cuenta de las carencias que había respecto a términos como: definición, existencia, propiedad definida de un conjunto, incluso el término conjunto había estado en una posición débil debido a que no existía definición convincente sobre él.

Pasaron los años, y se construyeron paradojas que hacían imperante la necesidad de una definición satisfactoria de conjunto.

Surgieron problemas como la Hipótesis del Continuo y el Principio del Buen Orden, que ganaron público durante la proclamación de los 23 problemas de Hilbert. Al resaltarse dichos problemas, los matemáticos se empezaron a encontrar con paredes para demostrarlos, como lo eran elecciones de elementos en conjuntos arbitrarios, o sumas infinitas de cardinales.

Con todo esto, hubo quienes dedicaron su vida a darle una axiomatización formal a las matemáticas, creyendo que éstas eran una ciencia que abarcaba todas las «verdades», y que cualquier proposición, el sistema podría decidir si es verdadera o falsa.

El Axioma de Elección se incluyó como un axioma en el sistema axiomático de Zermelo, y con ello, la comunidad matemática respondió con muchas críticas sobre su axiomatización, dando lugar a modificaciones como el sistema de Zermelo y Fraenkel ZFC, y el sistema de Gödel, Bernays y von Neuman NBG.

Esos sistemas se creían «buenos» en el sentido de que satisfacían muchos de los requerimientos matemáticos del momento, evitando incluso paradojas, sin embargo, Gödel demostró que no solo ZFC y NBG son incompletos o inconsistentes, también demostró que cualquier sistema axiomático finito es inconsistente o incompleto, resolviendo la duda de si alguna axiomatización existiría que demostrara que cualquier teorema es verdadero o falso.

La comunidad matemática que aceptó el Axioma diversificó las áreas de aplicación de éste, y profundizó los alcances que un axioma puede tener, de tal manera que, en libros modernos no se mencionaba explícitamente el uso del Axioma. La comunidad que rechazó el Axioma, encontró resultados parecidos o mejorados con axiomas alternos, sin embargo, pagaban el precio de debilitar hipótesis con obtener resultados no tan prominentes como con el Axioma.

Al recorrer la tesis el lector pudo haber encontrado teoremas que haya utilizado antes, sin haberse percatado antes del uso de una forma del Axioma. Encontrar la manera en que se utiliza y necesita una forma del Axioma es fundamental al hablar de un teorema en matemáticas formales, ya que los resultados que uno obtiene dependen del sistema axiomático que uno utilice. Aquí se mostró la manera en que varios teoremas dependen del Axioma, que a veces es evitable dicha dependencia, que a veces es indispensable, y que en la carrera se enseña a utilizar indiscriminadamente generalizaciones y teoremas sin buscar sus fundamentos, en gran parte porque los libros no lo mencionan, pero en mayor parte porque las matemáticas son menos axiomáticas. En la tesis se enseñó que existen diferentes modelos y universos matemáticos, que uno elije los axiomas que mejor se ajusten al modelo que se quiere crear, y que tengan mejor significado los resultados obtenidos del modelo.

REFERENCIAS

[Avi] AVIGAD, Jeremy. Forcing in proof theory [en línea]. [citado 07 de Septiembre de 2009] Disponible en Internet: http://www.andrew.cmu.edu/user/avigad/Papers/forcing.pdf.

[Bar] BARTLE, Robert G.. A Modern Theory of Integration. American Mathematical Society. 2001.

[BirMac] BIRKHOFF, Garrett; MAC LANE, Saunders. A Survey in Modern Algebra. Macmillan Publishing Co., Inc. 1977.

[Cle] CLEMENS, John Daniel. Descriptive Set Theory, Equivalence Relations, and Classification Problems in Analysis. Dissertation for the degree of Doctor of Philosophy. 2001.

[Dug] DUGUNJI, James. Topology. Allyn and Bacon Inc, 1966.

[DumFoo] DUMMIT, David S.; FOOTE, Richard M.. Abstract Algebra. John Wiley and Sons, Inc. Third Edition. 2004.

[End] ENDERTON, Herbert B.. Elements of Set Theory. Academic Press, Inc.. 1977.

[Fer] FERREIRÓS, José. Labyrinth of Thought, A History of Set Theory and its Role in Modern Mathematics. Birkhäuser. 2007.

[Fri] FRIEDBERG, Stephen H.; INSEL, Arnold J.; SPENCE, Lawrence E.. Álgebra Lineal. Publicaciones Cultural, S.A. Primera Edición. 1982.

[GivHal] GIVANT, Steven; HALMOS, Paul. Introduction to Boolen Algebras. Springer. 2009.

[Hal] HALMOS, Paul R.. Naive Set Theory. D. Van Nostrand Company, Inc. 1964

[Her] HERRLICH, Horst. Axiom of Choice. Springer. 2006.

[Her1] HERRLICH, Horst. Axiom of Choice [en línea]. [citado 24 de Julio de 2009] Disponible en Internet: http://extracoder.com/genesis/0004.html.

[HofKun] HOFFMAN, Kenneth; KUNZE, Ray. Linear Algebra. Prentice-Hall, Inc.. Second Edition. 1971.

[HrbJec] HRBACEK, Karel; JECH, Thomas. Introduction to Set Theory. Marcel Dekker, Inc. Third Edition. 1999.

[Ivo] IVORRA, Carlos. Lógica y Teoría de Conjuntos [en línea]. [citado 17 de Julio de 2007] Disponible en Internet: http://www.uv.es/ivorra/Libros/Logica.pdf

[Ivo1] IVORRA, Carlos. Análisis Matemático [en línea]. [citado 17 de Julio de 2007] Disponible en Internet: http://www.uv.es/ivorra/Libros/Analisis.pdf

[Moo] MOORE, Gregory H.. Zermelo´s Axiom of Coice, its Origins, Development, and Influence. *Springer Studies in the History of Mathem. and Physic. Sci,* 8, 1982.

[Moo1] MOORE, Gregory H.. Zermelo´s Axiom of Coice, its Origins, Development, and Influence [en línea]. [citado 17 de Septiembre de 2009] Disponible en Internet: http://extracoder.com/genesis/0071.html.

[Mun] MUNKRES, James. Topology. 2nd Edition.

[RubRub] RUBIN, H.; RUBIN, J. E.. Equivalents of the axiom of choice, II. North Holland Studies in Logic and Foundation of Math. 1985.

[RubRub1] RUBIN, H.; RUBIN, J. E.. Equivalents of the axiom of choice, II [en línea]. [citado 17 de Septiembre de 2009] Disponible en Internet: http://avaxhome.ws/ebooks/Equivalents_of_the_Axiom_of_Choice_II_Rubin.html.

[Sie] SIERPIŃSKI, Waclaw. Cardinal and Ordinal Numbers. Monografie Matematyczne, tom 34, Polska Akademia Nauk, Second Edition, 1965.

[Tou] TOURLAKIS, George. Lectures in Logic and Set Theory, Volume 2: Set Theory. Cambbridge University Press. 2003.

[Wag] WAGON, Stan. The Banach-Tarski Paradox. Cambridge University Press. 1985.

Apéndice A

BIBLIOGRAFÍA

En este anexo proporciono las bibliografías de todas las referencias que hago en mi tesis pero que no consulté directamente, si no que lo encontré citado en algún libro.

[AckHil] ACKERMANN, Wilhelm; HILBERT, David. Grundzüge der theoretischen Logik. Springer. Berlin. 1928.

[AleUry] ALEXANDROFF, Paul; URYSOHN, Paul. Mémoire sur les espaces topologiques compacts. NA 14. 1929.

[ArtSch] ARTIN, Emil; SCHREIER, Otto. Algebraische Konstruction reeller Körper. MSH 5. 1927.

[BanTar] BANACH, Stefan; TARSKI, Alfred. Sur la décomposition des ensembles de points en parties respectivement congruentes. FM 6. 1924.

[BeFr] BERNAYS, P.;FRAENKEL, A.A.. Axiomatic Set Theory. North Holland, 1958.

[Ber1] BERNAYS, Paul. A System of Axiomatic Set Theory. Part I. JSL 2. 1937.

[Bet] BETTAZZI, Rodolfo. Gruppi finiti ed infiniti di enti. 1896. AT 31, 506-512.

[Bor] BOREL, Emile. Leçons sur la théorie des fonctions. Paris. 1914.

[Bor1] BOREL, Emile. Quelques remarques sur les principes de la théorie des ensembles. MA 60. 1905.

[Bor2] BOREL, Emile. La philosophie mathématique et l´infini. RM 14. 1912.

[Bou] BOURBAKI, Nicolas. Eléments de mathématique. Première partie: Les structures fondamentales de l´analyse. Livre I, Théorie des ensembles. Paris Hermann. 1939.

[Can] CANTOR, Georg. Gesammelte Abhandlungen, Berlin 1932.

[Can1] CANTOR, Georg. Über unendliche, lineate Punktmannichfaltigkeiten V. MA 21, 1883.

[Car] CARTAN, Henri. Théorie des filtres. CRP 205. 1937.

[Car1] CARTAN, Henri. Filtres et ultrafiltres. CRP 205. 1937.

[Coh] COHEN, Paul J.. A minimal Model for Set Theory. BAMS 69. 1963.

[Fra] FRAENKEL, A. A.. L´axiome du choix. Revue Phil. de Louvain 50 (1950).

[Fra1] FRAENKEL, A. A.. Zehn Vorlesungen über die Grundlagen der Mengenlehre, Leipzig und Berlin, 1927.

[Fra2] FRAENKEL, Abraham. Zu den Grundlagen der Cantor-Zermeloschen Mengenlehre. MA 86. 1922.

[Fre] FRÉCHET, Maurice. Sur quelques points du calcul fonctionnel. CMP 22. 1906.

[Goo] GOODSTEIN, R. L.. Existence in mathematics. In *Logic and Foundations of Mathematics*, pages 70-82. Wolters-Noordhoff Publ., Groningen, 1968.

[God] GÖDEL, Kurt. Die Vollstandigkeit der Axiome des logischen Funktionenkalkuls. MMP 37. 1930.

[God1] GÖDEL, Kurt. Über formal unentscheidbare Sӑtze der Principia Mathematica und verwandter Systeme I. MMP 38. 1931.

[God2] GÖDEL, Kurt. The Consistency of the Axiom of Choice and of the Generalized Continuum Hypothesis. PAS 24. 1938.

[God3] GÖDEL, Kurt. Consistency-Proof for the Generalized Continuum-Hypothesis. PAS 25. 1939.

[God4] GÖDEL, Kurt. The Consistency of the Axiom of Choice and of the Generalized Continuum-Hypothesis with the Axioms of Set Theory. Annals of Mathematics Studies, No 3. Princeton University Press. 1940.

[HalLev] HALPERN, J. D.; LEVY, A.. The Boolean prime ideal theorem does not imply the axiom of choice. Proceedings of the Symposium in Pure Math. of the A.M.S. (1967) Vol XIII, Part I. 1971.

[Har] HARDY, G. H.. The continuum and the second number class. *Proc. London Math. Soc. Series 2*, 4:10-17, 1906.

[Hau] HAUSDORFF, Felix. Die Graduierung nach dem Endverlauf. SW 61. 1909.

[Hau1] HAUSDORFF, Felix. Grundzüge der Mengenlehre (Leipzig de Gruyter). 1914. Reprinted (New York Chelsea, 1965).

[Hau2] HAUSDORFF, Felix. Grundzüge der Mengenlehre (Leipzig de Gruyter). Second edition. 1927.

[Hau3] HAUSDORFF, Felix. Zur Theorie der linearen metrischen Raume. JM 167. 1932.

[Hes] HESSENBERG, Gerhard. Grundbegriffe der Mengenlehre. Göttingen Vandenhoeck and Ruprecht. 1906

[Kru] KRULL, Wolfgang. Algebraische Theorie der Ringe 1. MA 88. 1923.

[KinWag] KINNA, W. ; WAGNER, K.. Uber eine Abschwächung des Auswahlpostulates. FM 42. 1955.

[Leb] LEBESGUE, Henri. A proppos de quelques travaux mathématiques récents. EM (2) 17. 1971.

[Leb1] LEBESGUE, Henri. Les controverses sur la théorie des ensembles et la question des fondements. Gonseth 1941. 1941.

[Lev] LEVY, A.. Axioms of Multiple Choice. FM 50. 1962.

[LinMos] LINDENBAUM, Adolf; MOSTOWSKI, Andrzej. Über die Unabhangigkeit des Auswahlaxioms und einiger seiner Folgcrungen. CVR 31. 1938.

[LinTar] TARSKI, Alfred; LINDENBAUM, Adolf. Communication sur les recherches de la théorie des ensembles. CVR 19. 1926.

[MycSte] MYCIELSKI, Jan; STEINHAUS, Hugo. A Mathematical Axiom Contradicting the Axiom of Choice. AP 10. 1962.

[Neu] VON NEUMANN, John. Die Axiomatisierung der Mengenlehre, MZ 27, 669-752.

[Pea] PEANO, G.. Additione. Revista di Mat. III. 8, 1906.

[Pea1] PEANO, G.. Sur une courbe qui remplit toute une aire plane, Math. Ann. 36, 1890.

[Poi] POINCARÉ, Henri. La logique de l´infini. SC 12. 1912.

[Rus] RUSSELL, Bertrand. Sur les axiomes de l´infini et du transfini. *Soc. math. France, Comptes rendues des séances,* 1911. Translated in Grattan-Guinnes 1977.

[Rus1] RUSSELL, Bertrand. On Some Difficulties in the Theory of Transfinite Numbers and Order Types. LMS (2) 4.

[RusWhi] RUSSELL, Bertrand; WHITEHEAD, Alfred North. Principia Mathematica. Cambridge: Cambridge University Press, vol 1. 1910

[RusWhi1] RUSSELL, Bertrand; WHITEHEAD, Alfred North. Principia Mathematica. Cambridge: Cambridge University Press, vol 2. 1912

[RusWhi2] RUSSELL, Bertrand; WHITEHEAD, Alfred North. Principia Mathematica. Cambridge: Cambridge University Press, vol 3. 1913

[Sch] SCHRÖDER, Ernst. Ueber zwei Definitionen dei Endlichkeit und G Cantor´sche Sätze, «NAL 71,», 303-362.

[Sko] SKOLEM, Thoralf. Einige Bemerkungen zur axiomatischen Begründung der Mengenlehre, *Den femte skandinaviska matematikerkongressen, Redogörelse*. Helsinki Akademiska Bokhandeln. 1923.

[Sto] STONE, Marshall. The Theory of Representations for Boolean Algebras. TAMS 40. 1936.

[Tar] TARSKI, Alfred. Sur quelques théorèmes qui équivalent à l´axiome du choix. FM 5. 1924.

[Tar1] TARSKI, Alfred. Ein Überdeckungssatz für endliche Mengen nebst einigen Bemerkungen über die Definitionen der Endlichkeit. FM 30. 1938.

[Tyc] TYCHONOFF, Andrei. Uber die topologische Erweiterung von R*ă*umen. MA 102. 1930.

[Ury] URYSOHN, Paul. Uber die M*ă*chtigkeit der zusammenh*ă*ngenden Mengen. MA 94. 1925.

[Zer] ZERMELO, Ernst. Über Stufen der Quantifikation und die Logik des UNendlichen. DMVA 41. 1932.

[Zer1] ZERMELO, Ernst. Untersuchungen über die Grundlagen der Mengcnlehre I. MA 65. 1908. Traducido por van Heijenoort 1967.

[Zer2] ZERMELO, Ernst. Über den Begrff der Definitheit in der Axiomatik. FM 14. 1929.

Índice alfabético