UNIVERSIDAD SAN FRANCISCO DE QUITO
USFQ

Colegio de Ciencias e Ingenierías

The Power to Choose

The Influence of the Axiom of Choice in Mathematics

Pablo José Padilla Ortiz

Matemáticas

Trabajo de titulación presentado como requisito para la obtención del título de

Matemático

Quito, Diciembre de 2023

UNIVERSIDAD SAN FRANCISCO DE QUITO - USFQ

Colegio de Ciencias e Ingenierías

HOJA DE CALIFICACIÓN DE TRABAJO DE TITULACIÓN

The Power to Choose

The Influence of the Axiom of Choice in Mathematics

Pablo José Padilla Ortiz

Calificación: ____________

Nombre del Profesor: John R. Skukalek,

Título Académico: Ph.D.

Firma: ______________________________

Quito, Diciembre de 2023

©Derechos de Autor

Firma del estudiante: ________________________________

Nombres y apellidos:	Pablo José Padilla Ortiz
Código:	00211548
Cédula de identidad:	1105709925
Lugar y fecha:	Quito, diciembre de 2023

ACLARACIÓN PARA PUBLICACIÓN

Nota: El presente trabajo, en su totalidad o cualquiera de sus partes, no debe ser considerado como una publicación, incluso a pesar de estar disponible sin restricciones a través de un repositorio institucional. Esta declaración se alinea con las prácticas y recomendaciones presentadas por el Committee on Publication Ethics COPE descritas por Barbour et al. (2017) Discussion document on best practice for issues around theses publishing, disponible en `http://bit.ly/COPETheses`.

UNPUBLISHED DOCUMENT

Note: The following capstone project is available through Universidad San Francisco de Quito USFQ institutional repository. Nonetheless, this project – in whole or in part – should not be considered a publication. This statement follows the recommendations presented by the Committee on Publication Ethics COPE described by Barbour et al. (2017) Discussion document on best practice for issues around theses publishing available on `http://bit.ly/COPETheses`.

Karma Police, arrest this man

He talks in maths.

-Radiohead

Resumen

El Axioma de elección es un principio fundamental en la teoría de conjuntos de Zermelo-Fraenkel. Postula que dada una colección de conjuntos no vacíos, es posible elegir un elemento de cada conjunto en la colección. Este trabajo tiene como objetivo proporcionar una referencia informativa detallada acerca del Axioma de Elección y su impacto en el ámbito matemático. Se presenta una discusión sobre la historia del axioma y el razonamiento de por qué es controversial. Se detallan también demostraciones de equivalencia entre el Axioma de Elección y varias proposiciones formuladas en diferentes áreas de la matemática como el teorema de Tychonoff, lemma de Zorn, teorema del Buen-Ordenamiento, la existencia de bases en espacios vectoriales y otros más. Se presentan también consecuencias del Axioma de Elección y versiones más débiles del mismo.

Abstract

The Axiom of Choice is a fundamental principle in Zermelo-Fraenkel set theory. It postulates that given a collection of non-empty sets, it is possible to choose one element from each set in the collection. This work aims to provide a detailed informative reference about the Axiom of Choice and its impact on mathematics. A discussion of the history of the axiom and the reasoning for why it is controversial is presented. Proofs of equivalence between the Axiom of Choice and several propositions formulated in different areas of mathematics such as Tychonoff's theorem, Zorn's lemma, Well-Ordering theorem, the existence of bases in vector spaces and others are also detailed. The consequences of the Axiom of Choice and weaker versions of it are also presented.

Agradecimientos

Quiero agradecer a John Skukalek, quien me ha orientado y apoyado desde mi primer día en la universidad y quien me ha ayudado durante el proceso de comprensión y lectura sobre el tema de este proyecto. Gracias a Nicola di Teodoro, Tere Matos, Julio Ibarra, David Hervas y demás profes que me han brindado su tiempo para guiarme en mi camino como estudiante. Agradezco también a mi familia, en especial a mis padres, quienes con su sacrificio me han dado los recursos necesarios para estudiar y han procurado siempre mi bienestar. A mis hermanos y a toda la gente que me preguntó qué estoy haciendo de trabajo final, gracias por su disposición para escucharme hablar durante dos largas horas ininterrumpidas sobre este tema que no tiene nada que ver con sus vidas, admiro mucho su paciencia, por ustedes pude organizar mejor mis ideas y finalmente materializarlas en este manuscrito. Agradezco en especial a Nico Montesinos, Nico Cardozo, Patricio Jiménez y Erick Argüello sin quienes probablemente yo no estaría aquí.

Me siento en deuda con toda la gente que ha aportado a mi crecimiento intelectual, personal y artístico en la universidad. Gracias a Adrián Vasconez, Manu Larrea, Jorge García, Ricardo Mayorga y Pablo Mendieta por expandir los horizontes de mi pensamiento y darme ideas para reflexionar. A mis principales fuentes de apoyo emocional: José Palacios, Juliana Proaño, Santiago Vela, Tefa Coronel, Eddy Torres, Carla Torres, María del Carmen Salazar, Pedro Ramírez, José Ocampo y Vale Vaca.

A mis seguidores/as en @maths.posting, quienes con su apoyo y entusiasmo han sido la inspiración detrás de esta tesis. A Charly García, Luis Alberto Spinetta, Diego Lorenzini y todos quienes con su música han formado parte de mis largas noches de trabajo. Por último, y no por eso menos importante, a mis amigos y amigas... mi segunda familia. Desearía poder poner los nombres de todos ustedes, pero lamentablemente no caben en el margen de esta página. Ustedes saben quiénes son, los amo incondicionalmente y agradezco todo lo que han hecho por mi. Gracias a toda la gente que ha formado parte de mi vida en los últimos meses por su motivación.

Contents

List of Figures

Introduction

In the process of obtaining a scientific truth, we subject our hypotheses to various tests to gauge their validity and obtain results with different degrees of certitude. These hypotheses offer glimpses into a reality, whether tangible or intangible (in the case of mathematics). We aim for a certain level of certainty through evidence that some conjecture is true. If it holds true, we seek a broader principle encapsulating this truth, linking it to established knowledge via chains of inference. These truths, in turn, might stem from validated conjectures or from assumed statements that form the bedrock of our knowledge, known as axioms. Axioms serve as the foundational pillars for any system of knowledge, whether it's a pseudoscience or the rigorous structures of mathematics. Every belief or knowledge we claim is built from some axioms, statements so fundamental to our systems that we (knowingly or unknowingly) assume they are true without question.

The Axiom of choice, or AC as it will be commonly referred to throughout this work, is an axiom of set theory that postulates the possibility to select elements from each and every set belonging to a collection of non-empty sets. This postulate might seem obvious (possibly even uninteresting) at first. One might think *'why can't we just pick any element from a set if we already know it is not empty?'* And that is true, as long as we work on a finite setting. Without the Axiom, we can sometimes specify a choice, if $\varphi(x)$ was a statement like "x is the smallest number in the set", then a rule for choosing elements from a collection of finite subsets of $\mathbb{Z}$ would be: "from the given set, choose element x if it makes $\varphi(x)$ a true statement." Whenever we can construct a rule φ for selecting elements, there is no need for the Axiom of Choice. However, when the family of sets is infinite and it does not have a clear way to construct a rule for selecting elements, we cannot write down infinitely many instructions for choosing the elements from each set, that is when the Axiom is required.

The reason I chose this topic for my capstone project is that it combines many branches of math, specially the ones I find most interesting. It sparks my curiosity and fills me with joy to find exciting connections between the most fundamental parts of mathematics, like set theory and more complex ones that are built on top of that first layer of complexity, like topology and analysis. Throughout my studies at the university, I have seen results related to the Axiom of Choice in different courses: in Calculus and Real Analysis we use the equivalence between the sequential definition of continuity and the ε, δ definition. In Topology we use it to show Tychonoff's theorem, in Logic and set Theory we show that the countable union of countable sets is countable, in Linear Algebra and Functional Analysis we show that every vector space has a basis and so on.

Mathematicians, usually associated with the sureness of their findings, have rarely clashed as fervently over a core idea as they have with the Axiom of Choice. However, the absence of this Axiom would significantly reshape mathematics today. Surprisingly enough, the acceptance of the Axiom of Choice leads also to some strange results which have been a subject of debate. This axiom's distinctive feature lies in its ability to assert the existence of specific mathematical entities without the need to explicitly define them, similar to a magic trick, bringing objects to our world from an unknown source. The acceptance of the Axiom of Choice yields perplexing consequences that run counter to intuition, exemplified by the Banach-Tarski Paradox and the existence of unmeasurable sets of real numbers. Mathematicians have argued over the validity of those results, since they defy any physical law and are purely ideal concepts. The mathematicians that advocate for the denial of the Axiom are mostly constructivists, they adhere to a particular approach in mathematics that emphasizes the need for proofs to be constructive. They typically require that mathematical objects or proofs not only demonstrate the existence of something but also provide a method for constructing or finding that something.

This perspective contrasts with more traditional mathematical approaches that might accept the existence of objects without providing an explicit method for their construction.

What I find particularly interesting is that if the constructivist criticisms of the Axiom persisted, the entire subject would transform into a mere set of procedures, like an algorithm. In essence, the Axiom symbolizes the profound shifts—mathematical, philosophical, and psychological—that occurred when mathematicians earnestly delved into the study of infinite sets.

The purpose of this project is to explore the Axiom of Choice and the implications of accepting it and denying it in different branches of mathematics. The main objective is to make an informative reference and to highlight the use of the Axiom in some important results. It is intended that the reader acquires knowledge about the various mathematical and historical topics where the Axiom of Choice is applied, finding detailed demonstrations and understanding the history of the subject. Due to the wide diversification of the subject, in most of the important results we will cite the text where the information was obtained for future references. In the first chapter we will give some definitions and mention some results as well as define the axiomatic system ZF, which will be relevant on the proofs given in the subsequent sections. We will also mention some history related to the Axiom, its first appearance, the controversy and acceptance. Next, in chapter 2, we will define the axiom formally and prove some of the equivalent statements to it, as well as other non equivalent axioms that can serve as alternatives. In chapter 3, some applications of the Axiom in other parts of mathematics are mentioned, like in algebra, set theory, topology, analysis and measure theory. Most of the proofs are omitted in this chapter, since they are rather lengthy and do not add significatively to the purpose of showing the implications of the Axiom. In the fourth chapter, we'll explore how the Axiom aligns with the rest of the axioms in Zermelo-Fraenkel set theory. Finally, in chapter 5, we will take a look at mathematics without the Axiom of Choice.

Chapter 1

Preliminaries

> *But as one cannot apply infinitely many times an arbitrary rule by which one assigns to a class A an individual of this class, a determinate rule is stated here.*

Giuseppe Peano

1.1 Mathematical Background

In this section we will give some basic definitions, define the notation we will be working with and state important results in order to make the comprehension of the subsequent chapters more clear. I will assume that the reader has some familiarity with the common symbols used in logic and set theory and refer the interested reader to an introductory text in the bibliography [15]. It is also assumed that the reader has familiarity with the definition of a vector space, a topological space, metric spaces, rings, fields and ideals. We will be dealing with sets mostly, it is difficult to give an exact definition of a set, it is used as a synonym to collection, group or conglomerate. The philosophical and psychological implications of the definition of a set are beyond the scope of this work. The objects that compose a set are called elements, and we write $x \in S$ to mean x is a member of (or belongs to) the set S. We will define (almost) everything in terms of sets. The formal rules that determine what is possible to construct from sets are called the Axioms of ZF (with or without C) and are presented in Chapter 4.

Definition 1.1.1 Given two sets x, y, the *ordered pair* (x, y) is the set

$$\{x, \{x, y\}\}$$

Note that now we have the property $(x, y) = (a, b)$ if and only if $x = a$ and $y = b$.

Definition 1.1.2 Given two sets X, Y, the *Cartesian Product* $X \times Y$ is the set

$$X \times Y \coloneqq \{(x, y) \mid (x \in X) \wedge (y \in Y)\}.$$

Definition 1.1.3 A *function* $f\colon X \to Y$ from the set X to the set Y is a subset of $X \times Y$ such that

- $(\forall x \in X)(\exists y \in Y)((x, y) \in f)$
- $(x, y_1), (x, y_2) \in f \Rightarrow (y_1 = y_2)$

In other words, for any $x \in X$, there exists a unique $y \in Y$ such that $(x, y) \in f$. We write $f(x) = y$ to mean $(x, y) \in f$. [15]

Definition 1.1.4 A function $f\colon X \to Y$ is *injective* or *one-to-one* if

$$(a, y), (b, y) \in f \Rightarrow (a = b).$$

The function f is *surjective* or *onto* if

$$(\forall y \in Y)(\exists x \in X)((x, y) \in f).$$

And f is *bijective* if it is *injective* and *surjective*.

Definition 1.1.5 A *collection of sets* is a set whose members are sets.

In order to avoid Russell's Paradox, it is not possible for a set to be a member of itself. This actually follows from the Axiom of Regularity on ZF.

Definition 1.1.6 A *class* is a collection of sets that can be unambiguously defined by a property that all its members share.

A class is not necessarily a set. But every set is a class.

Definition 1.1.7 An *indexing function* of S by a set I is a function $x\colon I \to S$. The element $x(i)$ is the image under x of the element i and is denoted by x_i. We often write the images as $\langle x_i \rangle_{i \in I}$ or $(x_i\colon i \in I)$. When S is a collection of sets indexed by I, we call it an *indexed family of sets*.

Definition 1.1.8 A *relation* $\mathcal{R}$ between the sets A and B is a subset of $A \times B$. If $A = B$, then the relation is called a *binary relation*.

If $\mathcal{R}$ is a relation between A and B and $a \in A$ and $b \in B$, we write $a\mathcal{R}b$ to mean $(a, b) \in \mathcal{R}$. This will be important since it might strike as odd to write $(x, y) \in <$ instead of $x < y$ for the relation $<$, although it is completely valid.

Definition 1.1.9 ([15]) A binary relation $\mathcal{R}$ on a set X is said to be

- *Reflexive* if for all $x \in X$, $(x, x) \in \mathcal{R}$.
- *Irreflexive* if for all $x \in X$, $(x, x) \notin \mathcal{R}$.
- *Symmetric* if for all $x, y \in X$, $(x, y) \in \mathcal{R}$ implies $(y, x) \in \mathcal{R}$.
- *Antisymmetric* if for all $x, y \in X$, $(x, y) \in \mathcal{R}$ and $(y, x) \in \mathcal{R}$ implies $x = y$.
- *Asymmetric* if for all $x, y \in X$, if $(x, y) \in \mathcal{R}$, then $(y, x) \notin \mathcal{R}$.
- *Transitive* if for all $x, y, z \in X$, if $(x, y) \in \mathcal{R}$ and $(y, z) \in \mathcal{R}$, then $(x, z) \in \mathcal{R}$.

Definition 1.1.10 A *partial order* on a non-empty set X is a binary relation $\leq$ that is reflexive, antisymmetric and transitive. The pair $(X, \leq)$ is called a *partially ordered set* or a *poset*.

Definition 1.1.11 A *strict (partial) order* on a non-empty set X is a binary relation $<$ on X that is asymmetric and transitive.

The next definitions are valid for strict orders as well.

Definition 1.1.12 If $(X, \leq)$ is a poset, then two elements $x, y \in X$ are *comparable* if $x \leq y$ or $y \leq x$.

Definition 1.1.13 A *total order, linear order, chain* or *entire relation* of X is a partial order in which any two elements in X are comparable. The pair $(X, \leq)$ is called a *totally ordered set*.

Definition 1.1.14 Let $(X, \leq)$ be an ordered set. An element $x \in X$ is called *least element* of X under $\leq$ if for all $y \in X$, $x \leq y$.

Definition 1.1.15 An *upper bound* for a set $G \subseteq X$ is an element $x \in X$ such that $a \leq x$ for all $a \in G$.

Note that it need not be the case that $x \in G$.

Definition 1.1.16 A *maximal element* x of X is an element $x \in X$ such that $x < a$ for no $a \in X$.

Definition 1.1.17 Let $\mathcal{R}$ be a relation on X and $\mathcal{L}$ a relation on Y. We say that $(X, \mathcal{R})$ and $(Y, \mathcal{L})$ are *isomorphic* if there exists a bijection $f \colon X \to Y$ such that if $(x_1, x_2) \in \mathcal{R}$ then $(f(x_1), f(x_2)) \in \mathcal{L}$ and conversely. We denote this by $(X, \mathcal{R}) \cong (Y, \mathcal{L})$.

Definition 1.1.18 A *well-order* on a set X is a total order $\leq$ that has the property that every non-empty subset Y of X has a least element with respect to the restriction of $\leq$ to Y.

Definition 1.1.19 A set X is *transitive* if for every $A \in X$, we have $A \subset X$.

Definition 1.1.20 A set α is an *ordinal number* if

- α is transitive.

- α is a well-ordered set under $\in$.

The last definition can be equivalently formulated as in [15].

Definition 1.1.21 (Von Neumann verison) An ordinal number is a well-ordered set $(X, <)$ such that $\alpha = \{x \in X \mid x < \alpha\}$ for all $\alpha \in X$, i.e., every element of X is equal to the set of all its predecessors.

We can see how the ordinals are built starting from the empty set. Let α be an ordinal, then it has a least element x. If there existed an element $a \in x$, then $a \in \alpha$ (since α is transitive and x is consequently a subset of α) which contradicts the condition that x is the least element. So $x = \varnothing$. Here we assumed x to be the least element and concluded it was empty, moreover $\varnothing$ is vacuously an ordinal, so the ordinals start at $\varnothing$, and it is the smallest element of any non-empty ordinal number. Now, if $\alpha \neq \{\varnothing\}$, then $\alpha \setminus \{\varnothing\}$ has a smallest element b which cannot be empty. If $s \in b$ then $s \notin \alpha \setminus \{\varnothing\}$, otherwise b would not be the least element, and also $s \in X$, since $b \in X$, thus $b = \{\varnothing\}$. Then the second smallest ordinal is $\{\varnothing\}$. If we continue this process, we find that the next smallest ordinal is $\{\varnothing, \{\varnothing\}\}$ and so on. The finite ordinals can be identified with each natural number starting from zero as follows

$$
\begin{aligned}
&0 = \varnothing \\
&1 = \{\varnothing\} = \{0\} \\
&2 = \{\varnothing, \{\varnothing\}\} = \{0, 1\} \\
&3 = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\} = \{0, 1, 2\} \\
&4 = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}\} = \{0, 1, 2, 3\} \\
&\vdots
\end{aligned}
$$

In this sense, every natural number is an ordinal. The first infinite ordinal is the set of all natural numbers $\mathbb{N}$ and its denoted by ω.

Proofs for the following theorems can be found in the first chapter of [15].

Theorem 1.1.1 *Every well-ordered set is isomorphic to a unique ordinal.*

Theorem 1.1.2

- *If α is an ordinal, then $\alpha \cup \{\alpha\}$ is an ordinal.*
- *The union of a set of ordinals is an ordinal.*

There are three types of ordinals

- The *first type* is the zero ordinal $0 = \varnothing$.
- The *second type* is the one formed by *successor ordinals*. If α is an ordinal, then $\alpha \cup \{\alpha\}$ is also an ordinal, called the successor of α or $\alpha + 1$. It is the smallest ordinal that is greater than α.
- A non-zero ordinal λ is called *limit ordinal* if it is the union of all predecessors

$$\lambda = \bigcup_{\alpha<\lambda} \alpha$$

  Successor ordinals are not limit ordinals.

Definition 1.1.22 For any set X, the *Hartogs number* of X, denoted by $\mathfrak{H}(X)$, is the least ordinal λ such that there is no injective function from λ to X.

Theorem 1.1.3 $\mathfrak{H}(X)$ *exists for all* X.

Proof: See Chapter 7 in [16]. □

Theorem 1.1.4 *The class* Ord *of ordinals is not a set. It can be regarded as an ordered class.*

Proof: Suppose Ord is a set. It is clearly transitive and well ordered, so we have $\mathsf{Ord} \in \mathsf{Ord}$; but it is greater than all of its members, which leads to a contradiction. □

The next theorem will be used to show the equivalence between AC and the Well Ordering Theorem WO in the next section. A proof can be found in Chapter 4 of [16].

Theorem 1.1.5 (Transfinite Recursion Theorem) *Given a function* $G\colon \mathsf{Set} \to \mathsf{Set}$*, where* Set *is the class of all sets. There exists a unique transfinite sequence* $F\colon \mathsf{Ord} \to \mathsf{Set}$ *such that*

$$F(\alpha) = G(\{(a, b) \in F \mid a \in \alpha\}).$$

Definition 1.1.23 Two sets X and Y are *equipotent* or *have the same cardinality* if there exists a bijective function $f\colon X \to Y$.

Definition 1.1.24 A *cardinal* α is an ordinal such that for every $\beta < \alpha$ (the relation is equivalent to $\in$ or $\subset$ for ordinals), then there is no bijection between β and α, i.e., they are not equipotent.

With this definition we can consider all natural numbers (finite ordinals) as cardinals, the first infinite cardinal would be ω (the set of all natural numbers) and $\omega + 1$ would not be a cardinal, since there is a bijection with ω.

Definition 1.1.25 The *cardinality* of a well-ordered set X is the unique cardinal that is equipotent to X and is denoted by $|X|$.

Thus, if α is a cardinal, then $|\alpha| = \alpha$.

The crucial part of the use of AC in set theory is that with AC we can show (and we will, in the next section) that every set can be well-ordered, thus, every set is equipotent to some cardinal. Cantor was the one who introduced the *aleph* notation for cardinals. It is a way to index the cardinals using ordinals and transfinite recursion.

Definition 1.1.26 The *aleph sequence* is the sequence of infinite cardinals indexed using ordinals as follows

- $\aleph_0 = \omega$ is the smallest infinite cardinal.
- For all ordinals α, $\aleph_{\alpha+1}$ is the smallest cardinal greater than $\aleph_\alpha$.

- If λ is a limit ordinal, then

$$\aleph_\lambda = \bigcup_{\beta<\lambda} \aleph_\beta.$$

We now have $|X| \leq |Y|$ if and only if there is an injection $f\colon X \to Y$. This means (also) that $|X|$ is an ordinal and a subset of $|Y|$.

Theorem 1.1.6 (Schröder-Bernstein Theorem) *For any sets X, Y, if $|X| \leq |Y|$ and $|Y| \leq |X|$, then $|Y| = |X|$.*

1.2 Some History

In the evolution of mathematics, mathematicians shifted between exploring their foundational assumptions and the outcomes or structures derived from those assumptions. Frequently, these underlying assumptions were left unspecified, only becoming explicit after extensive use and reflection, which lead to an explicit formulation of them. We will now see how the use of choices of sets when constructing certain mathematical objects or proving theorems led in the course of a century to Zermelo's Axiom of Choice.

On 24 Septermber 1904, Ernst Zermelo proposed a version of the Axiom of Choice in a letter sent to David Hilbert. In this letter, Zermelo proposed a proof for the Well-Ordering Theorem from what was latter called the Axiom of Choice by Zermelo Himself. This was the first ever explicit mention of AC in the history of mathematics and it was translated from [2] in [22] as follows:

(Axiom VI (from [2])) *If T is a set whose elements are all sets that different from 0 and mutually disjoint, its union $\mathfrak{S}T$ includes at least one subset S_1 having one and only one element in common with each element of T.*

(Axiom of Choice)

We can also express this axiom by saying that it is always possible to choose a single element from each *element $M, N, R, \ldots$ of T and to combine all the chosen elements, $m, n, r, \ldots$, into a set S_1.*[1]

During that period, Zermelo viewed the Axiom of Choice as formalizing a presumption that many mathematicians had already assumed without explicit declaration [10]. When Zermelo presented his proof of the Well-Ordering Theorem using the Axiom of Choice, mathematicians split in constructivists and non-constructivists. The constructivists asserted that mathematics is a creation of the human mind, thus it has to align with what is reasonable to our minds. Conversely, the non-constructivists perceived mathematics as inherently existent in nature. For this group, discovering math concepts that might be logically sound but seemingly nonsensical was acceptable, mirroring the idea that not all natural discoveries inherently make sense upon revelation. Zermelo, Hilbert, Haussdorf, Hadamard and Keyser were believers of the latter ideology. While Borel, Baire and Lebesgue were some of the constructivist mathematicians that opposed the Axiom, but it should be noted that certain critics of the axiom had inadvertently employed this method before without noticing it.

In order to address the critics of his work, Zermelo published a second formulation of AC and a proof of the Well-Ordering Theorem in 1908. Some time later that year, he published his well known Axiomatic system, which included the Axiom of Choice. This axiomatic system, with the contributions of Abraham Fraenkel would latter become the pilar of modern set theory, the axiomatic system is known as ZFC and ZF if one decides to work without the Axiom of Choice. All those controversies brought Zermelo some fame and a position as a professor in Göttingen.

It is important to note that the formulation of AC arises as a consequence of another statement proposed by Cantor in 1882, which is the Well-Ordering Principle. Cantor believed that every set can be well-ordered and it was more like a *law of thought* than a proposition that required a proof. This principle deals with the 'enumeration' of elements of infinite sets, it extends what we might call as a labeling of each element in a given set. In order to do that, he defined the ordinal numbers, which serve as an extension of $\mathbb{N}$. Ordinal numbers are impor-

[1] The original article states that this second equivalent formulation is justified in [21].

tant, since they allow us to define cardinal numbers, which generalize the notion of the *size* of a set. All finite ordinals (the natural numbers) are cardinals. If every set can be well-ordered (as a result of the Axiom of Choice), then every set is bijectable with an ordinal and thus has a cardinal number.

Even as the practical value of AC became evident, the uncertainties about its validity persisted and were strengthened by the emergence of specific and notably counter-intuitive outcomes associated with it. Like the *paradoxical* decomposition of the sphere, known as the Banach-Tarski Paradox. The debate of whether the Axiom was true or not was settled when in 1935 Kurt Gödel established the relative consistency of AC with the axioms of ZF in the course of his attempt to show the Continuum hypothesis. And in 1963 Paul Cohen showed the independence of AC from the standard axioms of set theory. Over the following years, up until the present day, mathematicians have been using different formulations of AC in many branches of mathematics and many important results have been shown with the help of this, seemingly trivial, assumption.

Chapter 2

Statement, Equivalences and Alternatives

> *“The Axiom of Choice is obviously true, the Well-Ordering Principle is obviously false, and nobody knows about Zorn’s Lemma.”*
>
> Jerry Bona [17]

In this section we will look at the formal statement of AC, different ways of formulating it, some equivalent statements and some weaker versions of AC.

2.1 The Statement and Other Equivalent Choice Principles

There are multiple equivalent ways of stating the Axiom of Choice, in specific contexts, some of them are more useful than others. The first formulation of AC was proposed by Zermelo in a letter to David Hilbert [21] as mentioned in the previous section.

(Original Statement AC) *Given a family T of non-empty sets, there is a function f which assigns to each member A of T an element $f(A)$ of A.*

We will refer to this first formulation as AC. The function f here is called a *choice function*. Sometimes it is easier to work with disjoint sets in the family, that is when the following formulation is relevant.

(Disjoint-family Axiom of Choice dAC) *A family $\mathcal{F}$ of non-empty disjoint sets has a choice function.*

Theorem 2.1.1 *The two previous statements are equivalent.*

Proof: AC implies dAC trivially.

To show that dAC implies AC, suppose dAC and let $\mathcal{F}$ be a collection of non-empty sets. For all $X \in \mathcal{F}$, let $S_X = \{(x, X) \mid x \in X\}$ (which is non-empty since X is non-empty). Note that $S_X \cap S_Y = \varnothing$ if and only if $X \neq Y$, since the elements in S_X and S_Y have different coordinates if and only if the sets X, Y are not equal. Now, the collection $\mathcal{C} = \{S_X \mid X \in \mathcal{F}\}$ is a collection of pairwise disjoint non-empty sets. By dAC there is a choice function f for $\mathcal{C}$ such that $f(S_X) \in S_X$ for every $X \in \mathcal{F}$. Now, define a function $h\colon \mathcal{F} \to \mathcal{C}$ by $h(X) = S_X$ and a function $\pi_1\colon \cup\mathcal{C} \to \cup\mathcal{F}$ by $\pi_1(a, X) = a$. Then, for any $X \in \mathcal{F}$ we have $(\pi_1 \circ f \circ h)(X) = \pi_1(f(S_X))$ which is an element of X since $f(S_X) \in S_X$ and the first coordinate of this is an element of X. Hence $\pi_1 \circ f \circ h$ is the desired choice function on $\mathcal{F}$. □

In the previous section we defined the Cartesian Product for two sets and then defined what a function is in terms of the product. We can extend this product to an arbitrary family of sets. In the finite case, for a family with n sets, we can define their Cartesian Product $X_1 \times \cdots \times X_n$ as the set of all ordered n-tuples

$$X_1 \times \cdots \times X_n \coloneqq \{(x_1, \ldots, x_n) \mid x_1 \in X_1 \wedge \ldots \wedge x_n \in X_n\}.$$

Where an n-tuple is defined inductively starting from the case $n = 2$ by the rule

$$(x_1, \ldots, x_n) = ((x_1, \ldots, x_{n-1}), x_n) \coloneqq \{(x_1, \ldots, x_{n-1}), \{(x_1, \ldots, x_{n-1}), x_n\}\}.$$

We can also define the Cartesian Product for a family of sets indexed by a set I. Given an indexed family $\langle X_i \rangle_{i \in I}$, the *cartesian product* $\prod_{i \in I} X_i$ is the set of all choice functions for the family, i.e., functions $f\colon I \to \bigcup_{i \in I} X_i$ such that $f(i) \in X_i$ for all $i \in I$. We can think of this as generalizing the notion of an ordered tuple even further than for the finite case. Elements of this product are of the form $(f(i)\colon i \in I)$.

Note that if for some $j \in I$ we have $X_j = \varnothing$, then the cartesian product is empty, since we cannot choose any element in that set. But the converse of this statement is not entirely trivial and, in fact, is another form of the Axiom of Choice. This form is what is usually referred to as the Multiplicative Axiom.

(Multiplicative Axiom) *The Cartesian Product of a non-empty collection of non-empty sets is non-empty.*

It is easy to see that this axiom is equivalent to AC, since AC implies that the cartesian product of the family of non-empty sets is non-empty (there exists a choice function by AC). And the Multiplicative Axiom implies that at least one choice function exists for the family indexed by itself. The following is another way of stating the Multiplicative Axiom for indexed families of pairwise disjoint sets.

(Transversal Axiom) *For every family $\mathcal{F}$ of non-empty pairwise disjoint sets, there exists a set $C \subseteq \cup\mathcal{F}$ such that for every $X \in \mathcal{F}$, $C \cap X$ has exactly one element. The set C is called the set of representatives.*

Finally, we present the indexed version of AC.

(Indexed Axiom of Choice) *Let $\mathcal{F} = \langle X_i \rangle_{i \in I}$ be a collection of non-empty sets indexed by I. There exists a function $f\colon I \to \bigcup_{i \in I} X_i$ such that $f(i) \in X_i$*

Theorem 2.1.2 *The following are equivalent:*

1. AC.
2. *Transversal Axiom.*
3. *Indexed Axiom of Choice.*

Proof: 1. implies 2. Let $\mathcal{F}$ be a family of non-empty pairwise disjoint sets. Let f be the choice function for the family, then the set $C = \{f(X) \mid X \in \mathcal{F}\}$ is a set of representatives of $\mathcal{F}$. Note that $f(X) \in C \cap X$ but $f(Y) \notin C \cap X$ for $Y \neq X$, thus $C \cap X = \{f(X)\}$ for all $X \in \mathcal{F}$.
2. implies 3. Let $\mathcal{F} = \langle X_i \rangle_{i \in I}$ be a collection of non-empty sets indexed by I. Let

$$C_i \coloneqq \{(i,x) \mid x \in X_i\},$$

note that $i \neq j$ implies $C_i \cap C_j = \varnothing$. Consider now the family $\mathcal{C} \coloneqq \{C_i \mid i \in I\}$, then it is a family of non-empty pairwise disjoint sets and there exists a set of representatives f such that for all i there is a unique x such that $(i,x) \in f \cap C_i$, this means that f is a function $f \colon I \to \bigcup_{i \in I} X_i$ such that $f(i) \in X_i$ for all $i \in I$.
3. implies 1. Let $\mathcal{F}$ be a family of non-empty sets, we can index it by setting $I = \mathcal{F}$ and letting $X_i = i$ for all $i \in I$. The family $\langle X_i \rangle_{i \in I}$ is an indexed family of non-empty sets now, hence it has a choice function f such that $f(i) \in X_i = i$ for all $i \in I = \mathcal{F}$. $\square$

From now on, whenever we use AC we will be using any of its previously mentioned equivalent forms. Those forms have a distinctive character to them, mainly that they allow us to choose one element from each set in a family. That is why it is somewhat surprising that the next statement is equivalent to AC.

(Axiom of Multiple Choice (AMC)) *For every family $\{X_i \mid i \in I\}$ of nonempty sets there exists a family $\{F_i \mid i \in I\}$ of non-empty finite sets such that $F_i \subseteq X_i$.*

This axiom is equivalent to AMC under the axiomatic system of ZF (see Section 4.1 for the definition of ZF), but it is not equivalent in other axiomatic systems such as ZFA, which is the same as ZF but with modifications to allow the existence of *atoms*, which are objects that are not sets.

Theorem 2.1.3 *In* ZF, AC *is equivalent to* AMC.

AC implies AMC trivially by letting $F_i = \{f(i)\}$ where f is the choice function on the family. But the proof that AMC implies AC requires a sequence of implications which we will prove after defining some other principles.

Definition 2.1.1 Given a poset X, an *antichain* A is a subset of X such that any two elements in A are not comparable.

Definition 2.1.2 An antichain is *maximal* if it is not a proper subset of any other antichain.

(Antichain Principle, A) *Every partially ordered set has an antichain.*

(L) *Any totally ordered set can be well-ordered.*

It might be usefull to remind the reader that a totally ordered set under some order might not necessarily be well-ordered under the same order. Take for example $\mathbb{R}$ with the usual order $<$, it is totally ordered (any two elements are comparable), but it is not well-ordered under this relation (the subset $(-\infty, 0)$ does not contain a least element under $<$).

Definition 2.1.3 The *power set* of a set X, denoted $\mathcal{P}(X)$ is the set of all subsets of X.

The existence of the power set is assumed in the axioms of ZF.

(Power set, P) *The power set of any well-ordered set is well-orderable.*

We can now state our plan, which holds for ZF and ZFA.

Theorem 2.1.4 AC $\Rightarrow$ AMC $\Rightarrow$ A $\Rightarrow$ L $\Rightarrow$ P.

Proof: AC $\Rightarrow$ AMC is trivial.
AMC $\Rightarrow$ A, let $(S, <)$ be a partially ordered set. Consider the family of sets $\mathcal{P}(S) \setminus \{\varnothing\}$. There exists a function f such that for all $X \in \mathcal{P}(S)$, $f(X) \subseteq X$ non-empty and finite. Let g be defined by: $g(X)$ is the set of all minimal elements of $f(X)$ under $<$. Thus, $g(X)$ is a non-empty finite antichain. We now construct a maximal antichain in S:

- $A_0 = g(S)$.
- For every ordinal α, $A_\alpha = g(X)$ where X is the set of all $x \in S$ that are incomparable with every element of $\bigcup_{\beta<\alpha} A_\beta$.

The union of all A_α is a maximal antichain.

$\mathsf{A} \Rightarrow \mathsf{L}$, let $(T, <)$ be totally ordered. To show that it can be well ordered, it is sufficient to define a choice function f on $\mathcal{P}(T)$ and define the least element on any subset of T to be the element chosen by f. Let

$$\mathcal{S} := \{(X, x) \mid (X \in \mathcal{P}(T)) \wedge (x \in X)\}.$$

It can be shown that $\prec$ defines a partial order on $\mathcal{S}$ by

$$(X, x) \prec (Y, y) \Leftrightarrow (X = Y) \wedge (x < y).$$

By A, $(\mathcal{S}, \prec)$ has a maximal antichain C. This defines a choice function on $\mathcal{P}(T)$ since each set in the first coordinate of the elements of C occurs exactly once, and the second coordinate can be the choice.

$\mathsf{L} \Rightarrow \mathsf{P}$, If X is well-ordered, then $\mathcal{P}(X)$ can be totally ordered using the order $A \prec B$ if and only if the least element in $(A \setminus B) \cup (B \setminus A)$ is in B. It can be shown that this order is a total order. By L, $\mathcal{P}(X)$ can be well-ordered. $\square$

To complete the implications, we have

Theorem 2.1.5 *In* ZF, $\mathsf{P} \Rightarrow \mathsf{AC}$.

To show that every family of sets admits a choice function, we first need to construct all sets in stages and if we show that at every stage they all can be well-ordered, then we can put a choice function in them by choosing the least element. In the following construction, attributed to Zermelo, every family of sets appears at some stage of the construction. His idea was to construct all sets in well-ordered stages. At every stage there are no sets which are members of themselves, so it avoids Russell's Paradox[1]. The construction is indexed by the ordinals and is as follows

- $V_0 = \varnothing$.
- For any ordinal α, $V_{\alpha+1} = \mathcal{P}(V_\alpha)$.
- For limit ordinals λ, $V_\lambda = \bigcup_{\alpha<\lambda} V_\alpha$.

We can see that this procedure gives rise to all sets in stages.

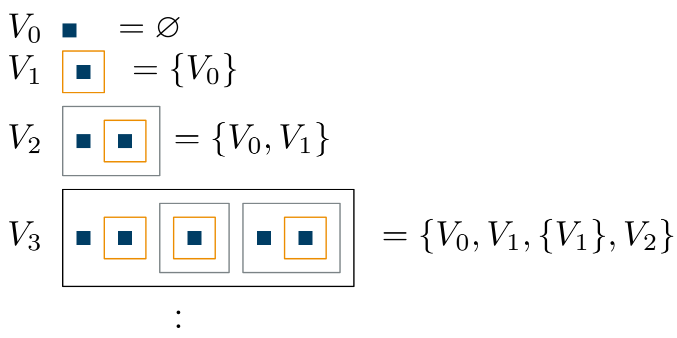

Figure 2.1: Stages of the construction of the sets for $\alpha = 1, 2, 3$.

The sets V_α are called *ranks* or *stages*. The rank of a set x, denoted by $\mathrm{rank}(x)$ is the smallest ordinal α such that $x \in V_{\alpha+1}$. Every set is contained in V_α for some α, and the number of elements at each stage grows rapidly. In an intuitive sense, it is like[2] expressing

$$\mathsf{Sets} = \bigcup_{\alpha \in \mathsf{Ord}} V_\alpha.$$

Now we can use this to prove the theorem.

Proof: (of Theorem 2.1.5) Suppose that the power set of every well-ordered set can be well-ordered, i.e., the power set of every ordinal is well-orderable. We can proceed by showing that V_α is well-orderable for every α since given any set X, we have $X \in V_\alpha$ for some α and

[1]The existence of a set that contains all sets which are not members of themselves.

[2]Not a real mathematical expression, just notation to help grasp the concept

since V_α is transitive, $X \subseteq V_\alpha$ so X is also well-orderable. Since for all α, $V_{\alpha+1} = \mathcal{P}(V_\alpha)$, a well-ordering on V_α induces a well-ordering on the next stage by assumption.

We use transfinite induction. V_0 is well ordered by the well ordering $\varnothing$ (trivially well ordered). Now, if there is a well-order $\prec$ for some ordinal α on V_α, there is a unique ordinal β that is bijective with V_α and a unique isomorphism $g\colon (V_\alpha, \prec) \to (\beta, \in)$. By P, $\mathcal{P}(\beta)$ is well-orderable and induces an order in $V_{\alpha+1}$ since they are bijective, that is, with g we can define a bijection $G\colon V_{\alpha+1} \to \mathcal{P}(\beta)$ by

$$G(X) = g[X] \coloneqq \{g(x) \mid x \in X\}$$

for every $X \in V_{\alpha+1}$. This function G induces a well-order of $V_{\alpha+1}$ by the rule $X \prec Y$ iff $G(X) < G(Y)$.

Now we have to show that V_α is well-orderable for all limit ordinals α. We can try to 'patch' the well-orderings that appear at previous stages to form a new one for V_α by ordering sets as they appear in V_α. For $X, Y \in V_\alpha$ (α limit ordinal), we define $X < Y$ iff

- $\mathrm{rank}(X) < \mathrm{rank}(Y)$, or
- $\mathrm{rank}(X) = \mathrm{rank}(Y) = \beta$ and $X <_{\beta+1} Y$ under the well-order $<_{\beta+1}$ of $V_{\beta+1}$.

This makes clear that for any $S \subseteq V_\alpha$, if λ is the least ordinal such that S has an element of rank λ, then the least element under $<_{\lambda+1}$ is the least element of S.

The problem now is to define what well order to use. Our assumption was that there exists a well-order of each $\mathcal{P}(X)$, not that there is a way of choosing one. Without a way of choosing an order, the reasoning would be circular. We need a sequence of well orderings with respect to which we can define a well order on V_α. We do this by finding an ordinal large enough such that the cardinality of each V_β for $\beta < \alpha$ is less than that ordinal and then using P to induce an order.

Given α a limit ordinal, let $\kappa = \mathfrak{H}(V_\alpha)$ (thus not injective to any subset of it). Let $<_T$ be a well-ordering of $\mathcal{P}(\kappa)$, we define a sequence of well-orders $\langle <_\beta \rangle_{\beta_\alpha}$ so that $<_\beta$ is a well-order for V_β for all $\beta < \alpha$.

- $<_0 = \varnothing$.
- If $\lambda < \alpha$ is a limit ordinal, we define $<_\lambda$ by: For $X, Y \in V_\lambda$, $X <_\lambda Y$ iff
  - $\operatorname{rank}(X) < \operatorname{rank}(Y)$, or
  - $\operatorname{rank}(X) = \operatorname{rank}(Y) = \beta$ and $X <_{\beta+1} Y$.
- From $<_\beta$ we can define $<_{\beta+1}$ as follows: Let η be the unique ordinal such that there is an isomorphism

$$f \colon (V_\beta, <_\beta) \to (\eta, \in).$$

  Since η is bijective with V_β, κ is not injective with η, thus $\eta < \kappa$, so the order $<_T$ in $\mathcal{P}(\kappa)$ well-orders $\mathcal{P}(\eta)$. This induces a well-order $<_{\beta+1}$ on $V_{\beta+1}$. □

It is important to remark that the last equivalence only holds if one assumes the axioms of ZF. In ZFA, it has been shown [6] that AMC, and thus P, is weaker than AC.

2.1.1 The Finite Version

We started this chapter with different ways of formulating AC, but so far we have not said anything about what makes it fundamentally different from the finite version, which can be proved from the axioms of ZF alone, so this is not an axiom, it is a theorem.

Theorem 2.1.6 (Finite Collection Choice) *There exists a choice function for every finite collection of non-empty sets.*

Proof: Let $\mathcal{S}$ be a collection of non-empty sets. We use (finite) induction on the number of sets in the family. For the base case, suppose $\mathcal{S}$ has one non-empty set X. By the definition of non-empty, there exists some $x \in X$, hence there exists a choice function given by $f(X) = x$. Now, suppose every collection of n sets has a choice function. Let $\mathcal{S}$ be a collection of $n + 1$

non-empty sets, fix $X \in \mathcal{S}$, then $\mathcal{S} \setminus \{X\}$ is a collection of n sets, thus it has a choice function $f \colon \mathcal{S} \to \cup\mathcal{S}$ such that $f(A) \in A$ for all $A \in \mathcal{S}$. Now, since X is non-empty, there exists some $x \in X$, thus there exists a function $\tilde{f} \colon \mathcal{S} \to \cup\mathcal{S}$ given by

$$\tilde{f}(A) = \begin{cases} f(A), & \text{if } A \neq X, \\ x, & \text{if } A = X. \end{cases}$$

This function is a choice function on $\mathcal{S}$, since for every $A \in \mathcal{S}$, $f(A) \in A$. □

The reader might wonder why this proof cannot be generalized for every possible collection of choices. The answer lies in the following subsection.

2.1.2 A Common Missconception

We make decisions everyday, we are used to the idea of choosing something from a vast collection of things. From choosing what dish to eat from a menu to choosing what we ultimately want to pursue in life, the possibilities are always there and we just have to grab one of them. The idea behind having a choice is simple: if there are things, there is at least one of them, so grab one. Then one might think: why is AC so debated when it is obvious? If you have multiple possible choices for something, each of which is possible to make, then it should be reasonable to assume that you can make all of them.

In practice, our imagination often limits these choices to a finite number, which indeed feels manageable. The problem arises when we deal with an infinite number of choices. And still, one might think of a rule for choosing elements, but what if there was no clear rule?

Here are some examples in which some form of choice is made to illustrate the idea of why we need to postulate it as an axiom.

Example 2.1.1 *Let $\mathcal{X} = \{\{a, b\} \mid a, b \in \mathbb{R}\}$, then $\mathcal{X}$ is a family of sets of two real numbers. We can* *easily define a choice function for this set by specifying the element we want to pick:*

$$f(\{a,b\}) = \min\{a,b\}$$

The idea of having a choice function on every family of nonempty sets does not challenge our intuition. However, consider the following example:

Example 2.1.2 *Let* $\mathcal{X} = \{\{A,B\} \mid A, B \subseteq \mathbb{R}\}$, *then* $\mathcal{X}$ *is a family of sets that consist of sets that contain two subsets of* $\mathbb{R}$.

Note that the set $\{A, B\}$ is not ordered, it would give exactly the same to write $\{B, A\}$. Writing an explicit choice function for this family of sets is an impossible task. Although we know that each set in the family is non-empty, one cannot specify which element to choose from each set. One needs the Axiom of Choice to show that such a function exists.

To put things more formally. Asserting that a set X is non-empty is stating that the formula $\phi(X) \coloneqq (\exists x)(x \in X)$ is true. The statement that this formula is true implies that we can find a *witness* or an example for which it holds true, that is: X is not empty, thus there is an element in X, and we can give it a name and use it in other parts of a proof or a statement. This is an example of what is known as *Existential Instantiation*.

In logic, we can use the symbol $\wedge$ to create another formula (a well formed formula), this symbol expresses conjunction, it means “and”. Now, if $X_1, \dots, X_n$ are non-empty sets, then this is also a formula in first order logic

$$\phi(X_1) \wedge \cdots \wedge \phi(X_n).$$

This is what allows us to choose an element from each set by using existential instantiation finitely many times. However, it is not possible in logic to express the conjunction of infinitely many formulas. So we cannot use infinitely many times the existential instantiation. This is why The Axiom is needed.

2.2 Equivalents

The following are equivalents to AC that are not explicitly stated in terms of Choice functions or cartesian products of sets.

2.2.1 The Well Ordering Theorem

This theorem was the primary motivation for Zermelo to formulate the Axiom of Choice.

(The Well Ordering Theorem WO) *Every set can be well-ordered.*

This means that for every set X, there exists a relation $\preceq$ such that every subset of X has a least element under $\preceq$. This can be easily seen with the natural numbers[3] $\mathbb{N}$. The relation $<$ well-orders $\mathbb{N}$. The smallest element is 0 since it is contained (as a set) in every $n = \{0, 1, \ldots, n-1\}$, then the order is

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

There is always a smallest element in every subset of $\mathbb{N}$. In fact, this is called the Well-Ordering Principle of $\mathbb{N}$ and its known to be equivalent to the Principle of mathematical induction[4].

But the relation $<$ as we commonly know it (less than) does not induce a well-order on $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$. For example, in $\mathbb{Z}$, the set $\{-n \mid n \in \mathbb{N}\}$ does not have a least element under $<$. What we can do is create a new relation $<_{\mathbb{Z}}$ as follows, for every $x, y \in \mathbb{Z}$

$$\begin{cases} x <_{\mathbb{Z}} y, & \text{if } |x| < |y| \\ -x <_{\mathbb{Z}} x, & \text{if } 0 < x. \end{cases}$$

Under this order we obtain

$$0 <_{\mathbb{Z}} -1 <_{\mathbb{Z}} 1 <_{\mathbb{Z}} -2 <_{\mathbb{Z}} \ldots <_{\mathbb{Z}} -n <_{\mathbb{Z}} n <_{\mathbb{Z}} -(n+1) <_{\mathbb{Z}} n+1 <_{\mathbb{Z}} \ldots$$

[3] Including zero.

[4] See [20].

Which is a well order on $\mathbb{Z}$. For $\mathbb{Q}$ the process is a bit more complicated, but we can construct a bijection $f\colon \mathbb{N} \to \mathbb{Q}$ using Theorem 2 in [23] and then use this bijection to define a well-ordering $p <_{\mathbb{Q}} q$ iff $f^{-1}(p) < f^{-1}(q)$. In this sense, every set that is bijectable with $\omega = \mathbb{N}$ is well-orderable. However, no relation has been found that transforms $\mathbb{R}$ into a well-ordered set. But with the help of the axiom of choice, we can find a bijection between every set and a well-ordered set and use that to define a well-order on every set.

Theorem 2.2.1 $\mathsf{AC} \Leftrightarrow \mathsf{WO}$.

Proof: We begin by showing that AC implies WO. Let X be a set. We want to construct a bijection form an ordinal λ to X. We proceed by Transfinite Recursion. Let Y be a set not contained in X and let $f\colon \mathcal{P}(X) \setminus \{\varnothing\} \to X$ be a choice function. We can define a function g on the class of all ordinals as

$$g(\alpha) = \begin{cases} f(X \setminus \{g(\beta) \mid \beta < \alpha\}), & \text{if } X \setminus \{g(\beta) \mid \beta < \alpha\} \neq \varnothing, \\ Y, & \text{if } X \setminus \{g(\beta) \mid \beta < \alpha\} = \varnothing. \end{cases}$$

Note that $g(\varnothing) = g(0) \in X$ and

$$g(1) = \begin{cases} f(X \setminus \{g(0)\}), & \text{if } X \setminus \{g(0)\} \neq \varnothing, \\ Y, & \text{if } X \setminus \{g(0)\} = \varnothing. \end{cases}$$

Transfinite recursion guarantees that g is defined for all ordinals. What g does is exhaust all the elements in X, when there are no more elements in X, then $g(\alpha) = Y$.

Observe that for $\alpha < \eta$, if $g(\eta) \neq Y$, then $g(\eta) \in X \setminus \{g(\beta) \mid \beta < \eta\}$, which means that $g(\eta) \notin \{g(\beta) \mid \beta < \eta\}$ and in particular, since $\alpha < \eta$, we have $g(\alpha) \neq g(\eta)$. This implies that X gets exhausted at some $\lambda < \mathfrak{H}(X)$, since we would get an injection from $\mathfrak{H}(X)$ to X if $g(\alpha) \neq Y$ for every $\alpha < \mathfrak{H}(X)$ and that contradicts the definition of $\mathfrak{H}(X)$.

Let λ be the least element of the set $\{\alpha \mid g(\alpha) = Y\}$. Clearly g restricted to λ is injective. We can conclude that g restricted to λ is surjective if we show $\{g(\beta) \mid \beta < \lambda\} = X$. Clearly we have $\{g(\beta) \mid \beta < \lambda\} \subseteq X$. If $\{g(\beta) \mid \beta < \lambda\} \subset X$, then $X \setminus \{g(\beta) \mid \beta < \lambda\} \neq \varnothing$ and $g(\lambda) \neq Y$, which contradicts the definition of λ. Thus, the restriction g to λ is a bijection between λ and X.

The second part of the implication is easier. Assuming WO, let $\mathcal{S}$ be a collection non-empty sets. then $\cup\mathcal{S}$ can be well-ordered by a relation $\prec$, thus, for every $X \in \mathcal{S}$, there is a least element in X. We can now define a choice function $f\colon \mathcal{S} \to \cup\mathcal{S}$ by $f(X) = x$, where x is the least element in X under $\prec$. $\square$

2.2.2 Zorn's Lemma

(Zorn's Lemma (ZL)) *If $(X, <)$ is a poset with the property that every chain in X has an upper bound, then X contains a maximal element.*

Theorem 2.2.2 $\mathsf{AC} \Leftrightarrow \mathsf{ZL}$.

Proof: We begin by showing that AC implies ZL. We proceed by contradiction. Let X be a poset that satisfies the hypotheses of ZL and suppose X has no maximal element. By assumption X is non-empty, since it is vacuously true that $X = \varnothing$ has a maximal element in X. Let $f\colon \mathcal{P}(X) \setminus \{X\} \to X$ be a choice function and define a map $g\colon \mathsf{Ord} \to X$ setting

- $g(0) = f(X)$;
- $g(\alpha + 1) = f(\{x \in X \mid x > g(\alpha)\})$;
- If λ is a limit ordinal, $g(\lambda) = f(U)$ where U is the set of upper bounds for $\{g(\alpha) \mid \alpha < \lambda\}$.

Note that if $\{x \in X \mid x > g(\alpha)\} = \varnothing$, then $g(\alpha) \in X$ would be a maximal element, so it is non-empty. Also $\{g(\alpha) \mid \alpha < \lambda\}$ is a chain since $\alpha < \beta$ implies $g(\alpha) < g(\beta)$. Thus, we have created an injection from Ord to X, which is impossible.

For the other implication, i.e. $\mathsf{ZL} \Rightarrow \mathsf{AC}$, consider a collection $\mathcal{S}$ of non-empty sets. Let

$$\mathcal{F} = \{f\colon\ \mathrm{dom}(f) \to \cup\mathcal{S} \mid (\mathrm{dom}(f) \subseteq \mathcal{S}) \wedge (\forall X \in \mathrm{dom}(f))\, (f(X) \in X)\}\,.$$

In other words, $\mathcal{F}$ is the set of choice functions whose domain is a subset of $\mathcal{S}$. The collection $\mathcal{F}$ is non-empty, since by existential instantiation there exists a non-empty $Y \in \mathcal{S}$ and also $(\exists y)(y \in Y)$ since Y is non-empty. Thus, the function (regarded as a set of pairs) $f = \{(Y, y)\} \in \mathcal{F}$. Note that $(\mathcal{F}, \subset)$ is a poset. Let $\mathcal{F}_0$ be a chain (totally-ordered subset of $\mathcal{F}$) and $f_0 = \cup \mathcal{F}_0$, then f_0 is a function in $\mathcal{F}$ since if $(X, y_1), (X, y_2) \in f_0$, we have $(X, y_1) \in F_1, (X, y_2) \in F_2$ for some F_1 and F_2 in $\mathcal{F}_0$. Since $\mathcal{F}_0$ is totally ordered under inclusion, we have either $F_1 \subset F_2$, $F_2 \subset F_1$ or $F_1 = F_2$. In the last case, we obtain $y_1 = y_2$ immediately since F_1 is a function. For the other two cases, both pairs will be elements of either F_1 or F_2, and since those are functions, we obtain $y_1 = y_2$. Note also that f_0 is an upper bound for $\mathcal{F}_0$. Since $\mathcal{F}_0$ was an arbitrary chain in $\mathcal{F}$, we conclude that every chain in $\mathcal{F}$ has an upper bound. By ZL, there exists a maximal element F in $\mathcal{F}$. If $\mathrm{dom}(F) \neq \mathcal{S}$, then $\mathcal{S} \setminus \mathrm{dom}(F) \neq \varnothing$ and there exists some non-empty $A \in \mathcal{S} \setminus \mathrm{dom}(F)$, by existential instantiation, we can select some $a \in A$ and form the function $\overline{F} = F \cup \{(A, a)\}$, which contradicts the maximality of F. Thus, $\mathrm{dom}(F) = \mathcal{S}$ and F is a choice function. □

2.2.3 Tuckey's Lemma

Definition 2.2.1 A family of non-empty sets $\mathcal{S}$ is said to have *finite character* if $X \in \mathcal{S}$ if and only if every finite subset of X is in $\mathcal{S}$.

(Tukey's Lemma) *If $\mathcal{S}$ is a collection of non-empty sets. If $\mathcal{S}$ has finite character, then $\mathcal{S}$ has a maximal element with respect to set inclusion.*

Theorem 2.2.3 AC *is equivalent to Tukey's Lemma.*

Proof: AC implies Tukey's Lemma:

If $\mathcal{S}$ is a collection of non-empty sets and has finite character, then it is partially ordered under $\subset$. Let $\mathcal{C}$ be a chain in $\mathcal{S}$. Let $X = \cup \mathcal{C}$. Now, let $B \subset X$ be finite. For every $x \in B$ there exists a set $A_x \in \mathcal{C}$ such that $x \in A_x$. Since $\mathcal{C}$ is a chain and B is finite, the collection $\{A_x \mid x \in B\}$ is finite and consequently has a maximal element $M \in \mathcal{C}$. Now, for any $x \in B$, we have

$x \in A_x \subseteq M$, thus $B \subseteq M \in \mathcal{S}$ and since $\mathcal{S}$ has finite character, $B \in S$. Thus every chain has an upper bound. By ZL, $\mathcal{S}$ has a maximal element.

Tukey's Lemma implies AC:

We can repeat a similar argument to the one we used for the second part of the proof of ZL. Consider a collection $\mathcal{S}$ of non-empty sets. Let

$$\mathcal{F} = \{f\colon \operatorname{dom}(f) \to \cup\mathcal{S} \mid (\operatorname{dom}(f) \subseteq \mathcal{S}) \wedge (\forall X \in \operatorname{dom}(f))\,(f(X) \in X)\}\,.$$

This is the set of choice functions whose domain is a subset of $\mathcal{S}$. The collection $\mathcal{F}$ is non-empty, by the same argument as in the mentioned proof. Note that $(\mathcal{F}, \subset)$ is a poset and has finite character, since the finite subsets of a choice function are choice functions themselves. By Tukey's Lemma, $\mathcal{F}$ has a maximal element F. We can apply the same argument as we did for ZL to show that $\operatorname{dom}(F) = \mathcal{S}$, thus it is a choice function on $\mathcal{S}$ and AC holds. □

2.2.4 Every Vector Space Has a Basis

We begin this subsection by recalling that a *basis* (or *Hamel Basis*) in a vector space V is a subset $B \subset V$ such that every element of V can be expressed uniquely (ignoring the elements with zero-coefficients) as a linear combination of a finite subset of B.

The following result highlights the importance of the Axiom of Choice in the development of modern algebra. Namely, that the AC implies that every vector space has a basis[5], and if we suppose that every vector space has a basis, then we can show the Axiom of Choice. The original version of this result was published in [11].

Theorem 2.2.4 *In* ZF, *the assertion that every vector space has a basis is equivalent to* AC.

Proof: For the first part of the implication, we will assume that every vector space has a basis and use that to show AC.

Consider a family $\{X_i \mid i \in I\}$ of non-empty sets. Without loss of generality, suppose the sets

[5] The finite-dimensional case is immediate and does not require choice.

X_i are pairwise disjoint and define $X := \bigcup\{X_i \mid i \in I\}$. Now, adjoin all the elements in X as indeterminates to some arbitrary field K, forming the polynomial ring $K[X]$,

$$K[X] := \bigcup\{K[A] \mid A \subset X \ \wedge \ A \text{ finite}\},$$

which is an integral domain and let $K(X)$ be the field of fractions of $K[X]$, which is the field of rational functions whose "variables" are in X.

If $h \in K(X)$ is a monomial, the i-degree of h, denoted as $\deg_i(h)$, is the sum of the exponents of the elements in X_i in that monomial. Let $\mathcal{H}(i,n) \subset K(X)$ be the set of rational functions in $K(X)$ that are the quotient of two polynomials such that all monomials in the denominator have the same i-degree m and all monomials in the numerator have the same i-degree $n+m$. It is easy to show that

$$H := \bigcap_{i \in I} \mathcal{H}(i,0)$$

is a subfield of $K(X)$. Thus $K(X)$ can be regarded as a vector space over H. Let V be the subspace spanned by X. By assumption, there exists a basis B of V. Now, for all $i \in I$ and $x \in X_i$, we can express x as an H-linear combination of elements in B

$$x = \sum_{b \in B} \alpha_b(x) \cdot b, \tag{2.1}$$

where $\alpha_b(x) \in H$ is zero for all but finitely many $b \in B$. Similarly, if $y \in X_i$, then

$$y = \sum_{b \in B} \alpha_b(y) \cdot b. \tag{2.2}$$

If we multiply equation (2.1) by $y/x \in H$, then

$$y = \sum_{b \in B} \left(\frac{y}{x}\right) \alpha_b(x) \cdot b. \tag{2.3}$$

Since B is a basis, we conclude that $(y/x)\alpha_b(x) = \alpha_b(y)$ for all $b \in B$, thus

$$\frac{\alpha_b(x)}{x} = \frac{\alpha_b(y)}{y}$$

which shows that the expression $\alpha_b(x)/x \in K(X)$ depends only on $i \in I$ and not the element $x \in X_i$, we can therefore denote it as α_{bi}. Since $\alpha_b(x) \in H$, we have $\alpha_{bi} \in \mathcal{H}(i, -1)$ and $\alpha_{bi} \in \mathcal{H}(j, 0)$ for all $j \neq i$. Therefore, when α_{bi} is written as a quotient of polynomials in reduced form, some variables of X_i must occur in the denominator. Let F_i be the set of those members of X_i that occur in the denominator of α_{bi} for some $b \in B$ (such that $\alpha_b \neq 0$). Then for every $i \in I$, F_i is a nonempty finite subset of X_i, which is exactly what the axiom of multiple choice requires. Since AMC is equivalent to AC, the result follows.

For the second part of the implication, we start with a non-trivial vector space V and let $\mathcal{M}$ be the set of all linearly independent subsets of V. Since V is non-trivial, there is some $v \neq 0 \in V$, so $\{v\} \in \mathcal{M}$. The relation $\subset$ defines a partial order on $\mathcal{M}$. Every chain $\mathcal{C}$ in $\mathcal{M}$ has an upper bound, which is the union of all the members of $\mathcal{C}$. By ZL, $\mathcal{M}$ has a maximal element B.

We now that B is a basis for V. Let $W = \operatorname{span}(B)$, then W is a subspace of W. If $W \neq V$, then for any $v \in V \setminus W$, the set $B \cup \{v\}$ is a linearly independent set that contains B, which cannot happen due to the maximality of B, thus $W = V$ and every vector space has a basis. $\square$

2.2.5 Surjective Functions Have Right Inverses

Theorem 2.2.5 AC *is equivalent to the statement: Every surjective function has a right inverse.*

Proof: Let $f\colon X \to Y$ be a surjection. Then, for all $y \in Y$, the set $f^{-1}(\{y\})$ is non-empty. Using AC, for all $y \in Y$ we can choose an element $g(y) \in f^{-1}(\{y\})$, let g be the function $Y \to X$ defined by these choices. Now, for any $y \in Y$,

$$(f \circ g)(y) = f(g(y)) = y.$$

Thus g is a right inverse of f.

For the other implication, let $\mathcal{F}$ be a collection non-empty sets. For any $X \in \mathcal{F}$, define

$$S_X = \{(x, X) \mid x \in X\}.$$

Consider the collection $\mathcal{C} = \{S_X \mid X \in \mathcal{F}\}$ and the relation $f \subseteq \cup\mathcal{C} \times \mathcal{F}$ consisting of pairs of the form

$$((x, X), X) \in f \quad \text{i.e.} \quad f(x, X) = X.$$

The relation f is clearly a function, since if $((a, A), A), ((c, C), C)$ in f and $(a, A) = (c, C)$ then $A = C$. It is clearly surjective. Now, by our assumption, there exists a right inverse of f, that is, a function $g\colon \mathcal{F} \to \cup\mathcal{C}$ such that for all $X \in \mathcal{F}$,

$$f(g(X)) = X.$$

This implies that $g(X) = (x, X)$ for some $x \in X$. The function that takes the element $x \in X$ that occurs in $g(X)$ is a choice function in $\mathcal{F}$, thus AC holds. $\square$

2.2.6 Tychonoff's Theorem

To prevent undue lengthiness in the project, for this section we have opted not to provide an exhaustive list of definitions and theorems necessary to show the main equivalence within this section. Instead, we present the main definitions and recommend readers refer to [12] for a comprehensive understanding of these foundational concepts in topology.

Definition 2.2.2 Let I be a set and $(X_i, \mathcal{T}_i)$ be a topological space for each $i \in I$. The *product space*, denoted by $\prod_{i\in I}(X_i, \mathcal{T}_i)$ is the set $\prod_{i\in I} X_i$ with the topology $\mathcal{T}$ that has as its basis the family

$$\left\{ \prod_{i\in I} O_i \;\middle|\; O_i \in \mathcal{T}_i \text{ and } O_i = X_i \text{ for all but a finite number of } i \right\}.$$

The topology $\mathcal{T}$ is called the *product topology* or *Tychonoff Topology*.

Recall that a topological space X is compact if every open cover[6] of X has a finite subcover.

Theorem 2.2.6 (Tychonoff's Theorem) *Let* $\{(X_i, \mathcal{T}_i) \mid i \in I\}$ *be a family of topological spaces. Then* $\prod_{i\in I}(X_i, \mathcal{T}_i)$ *is compact if and only if each* $(X_i, \mathcal{T}_i)$ *is compact.*

Theorem 2.2.7 $\mathsf{AC} \Rightarrow$ *Tychonoff's Theorem.*

This is a well known implication that uses Zorn's Lemma. For a proof, the reader can check [12], Chapter 10 in [24], or any other Point Set Topology textbook. The focus of this subsection will be the reverse implication, which was first noted by Kelley in 1950 [3].

Theorem 2.2.8 *Tychonoff's Theorem implies* AC.

Proof: Let $\{X_i \mid i \in I\}$ be a family of non-empty sets. We want to show that the cartesian product of the family is non-empty.
By the Axiom of Regularity and the Pair Set Axiom[7] we know that for each $i \in I$ the set $\{X_i\}$ is disjoint from X_i (i.e., $X_i \cap \{X_i\} = \varnothing$ or equivalently $X_i \notin X_i$). Let $F_i = X_i \cup \{X_i\}$ and define a topology $\mathcal{T}_i$ on F_i by $\mathcal{T}_i = \{\varnothing, \{X_i\}, X_i, F_i\}$. Since this is finite, $(F_i, \mathcal{T}_i)$ is compact. By Tychonoff's Theorem, the product space $\prod_{i\in I}(F_i, \mathcal{T}_i)$ is compact. We will use this fact to show that $\prod_{i\in I} X_i$ is non-empty.

Consider the coordinate-projection map $\pi_j \colon \prod_{i\in I} F_i \to F_j$, defined by

$$\pi_j(\langle x_i \rangle_{i\in I}) = x_j.$$

This map is continuous (see Proposition 10.1.5 in [24]). Thus, for every $j \in I$, the set

$$\pi_j^{-1}[\{X_j\}] = \left\{ \langle x_i \rangle_{i\in I} \;\middle|\; (x_j = X_j) \wedge (\forall i \neq j)(x_i \in X_i) \right\}$$

[6]A family of open sets whose union is equal to X.
[7]See Chapter 4.

is open in $\prod_{i\in I}(F_i,\mathcal{T}_i)$. Consider the collection

$$\mathcal{C} := \left\{\pi_i^{-1}\left[\{X_i\}\right] \;\middle|\; i\in I\right\}.$$

We will show that it does not have a finite subcover of $\prod_{i\in I} F_i$. Let

$$\left\{\pi_{i_1}^{-1}\left[\{X_{i_1}\}\right],\ldots,\pi_{i_n}^{-1}\left[\{X_{i_n}\}\right]\right\}$$

be a finite subcollection of $\mathcal{C}$. We can use the Finite Choice Principle to choose an element $p_{i_k}\in X_{i_k}$ for $k\in\{1,\ldots,n\}$ and let $\langle a_i\rangle_{i\in I}\in\prod_{i\in I}F_i$ be defined by

$$a_i = \begin{cases} p_{i_k}, & \text{if } i\in\{i_1,\ldots,i_n\}, \\ X_i, & \text{otherwise.}\end{cases}$$

Then for all $j\in\{i_1,\ldots,i_n\}$, $\pi_j(\langle a_i\rangle_{i\in I})\notin\{X_j\}$, which implies that

$$\langle a_i\rangle_{i\in I}\notin\bigcup_{k=1}^{n}\pi_{i_k}^{-1}\left[\{X_{i_k}\}\right].$$

Since $\prod_{i\in I}(F_i,\mathcal{T}_i)$ is compact, $\mathcal{C}$ cannot be an open cover of $\prod_{i\in I}F_i$, thus

$$\left(\prod_{i\in I}F_i\right)\setminus\left(\bigcup_{i\in I}\pi_i^{-1}\left[\{X_i\}\right]\right)\neq\varnothing.$$

That is, there exists some $\langle y_i\rangle_{i\in I}\in\prod_{i\in I}F_i$ such that $y_i\in F_i$ and $y_i\neq X_i$, hence $y_i\in X_i$ for every $i\in I$. This implies that the cartesian product is non-empty, thus AC holds. $\square$

One could think that in constructing a subcollection of $\mathcal{C}$ and then putting an enumeration of the members of that subcollection in the above proof, we have used AC, but this is not true, since we can think of this process as using existential instantiation[8] finitely many times on the collection $\mathcal{C}$ without the element we choose at each step in the construction.

[8]See Section 2.1.2 in this chapter.

2.3 Other Equivalents to The Axiom of Choice

So far, we've focused on the most popular and widely applied equivalent forms of AC. Nevertheless, The Axiom of Choice has numerous equivalent statements. Some of them are more convenient for some purposes than others. Various versions of the Axiom of Choice are prevalent in nearly all mathematical disciplines, manifesting in diverse forms. That is why mathematicians are interested in finding equivalent statements in other branches of math, some of them worry they might be using one form of AC unknowingly. This matter is so important that there is a whole book dedicated to equivalent statements to AC. The following equivalent forms and more can be found in the book by H. Rubin and J. Rubin [4] and another version by H. Rubin [13].

Theorem 2.3.1 (Tarski's Theorem About Choice [1]) AC *is equivalent to the statement that for any infinite set X, there is a bijection between X and $X \times X$.*

Theorem 2.3.2 ([1]) *Any two sets have comparable cardinalities.*

In abstract algebra we also have an equivalence.

Definition 2.3.1 A *maximal ideal* of a ring R is an ideal I such that for any ideal J that contains I, either $J = I$ or $J = R$.

Theorem 2.3.3 (Krull's Theorem) *Every ring with identity has a maximal ideal.*

A proof that ZL implies Krull's theorem can be found in chapter 6 of [15] and a proof for the converse statement (and the next statement) is given in [9].

The following is a result in graph theory. To get an idea of the meaning of the statement, some definitions are given.

Definition 2.3.2 A graph is a *connected graph* if for each pair of vertices, there exists at least one single path which joins them.

Definition 2.3.3 A *tree* is a connected graph without cycles. A spanning tree is a tree T of a graph G that contains all the vertices of G.

Theorem 2.3.4 *Every connected graph has a spanning tree.*

2.4 Alternatives: Weaker Forms of The Axiom of Choice

In this section we show some alternative formulations to the Axiom of Choice that are implied by it, but do not imply AC. The idea is that we can work in ZF without AC but with one of those alternatives and in doing so, we might be able to prove some theorems (which we will review in the next section), but avoid some controversial or paradoxical results that come along witht the addition of AC to ZF.

According to [19], we can weaken the original statement of AC by doing one of the following:

1. Restrict the size of the family of sets (or of the indexing set I).
2. Restrict the properties of the sets in the family.
3. Change the axiom so that we can choose some set of finitely many elements at each stage and not one particular element.

This project does not delve into demonstrating that the alternative axioms presented in this section do not imply AC. The omission is intentional, since a thorough exploration of this topic is extensive and exceeds the defined scope of the project. The interested reader can review Chapter 8 in [6].

The first form of weakening AC gives us the Principle of Finite Choice and the Axiom of Countable Choice CC. The second form gives us

(Axiom of Choice for Families of Finite Sets (AC(fin))) *Every family of non-empty finite sets admits a Choice Function.*

This statement is weaker than AC (it is implied by AC but it is not equivalent to it). The second procedure also gives us the Axiom of Choice for Well-Orderable Sets.

The third form of weakening AC gives us AMC, which is equivalent in ZF but not in general. It also gives us

(Kinna–Wagner Selection Principle) *For every family $\mathcal{F} = \{X_i \mid i \in I\}$ of sets each of which has at least 2 elements, there is a family $\mathcal{C} = \{Y_i \mid i \in I\}$ of non-empty sets such that Y_i is a proper subset of X_i.*

We can combine the above mentioned procedures to form other axioms weaker than AC, like

(Countable Choice in $\mathbb{R}$) *The family $\{X_n \mid n \in \mathbb{N}\}$ of subsets of $\mathbb{R}$ admits a choice function.*

Some other examples and equivalences can be found in [19].
A weaker version of AC that is not obtained by following our procedure but is closely related is the following.

2.4.1 Axiom of Dependent Choice

We know that AC validates limitless independent choices as long as they are independent from each other, mathematicians sometimes made an infinity of arbitrary selections such that a given choice depended on those previously made. Paul Bernays proposed this axiom in 1942 as a weakened form of the Axiom useful in analysis [7]. This form is now known as the Axiom of Dependent Choice:

(Axiom of Dependent Choice.) *Let R be an entire binary relation on a non-empty set X, then there exists a sequence $\langle x_n \rangle_{n \in \mathbb{N}}$ in X such that $(x_n, x_{n+1}) \in R$ for all $n \in \mathbb{N}$.*

Theorem 2.4.1 AC $\Rightarrow$ DC.

Proof: Let X be a non-empty set and R an entire binary relation on X. Since R is entire, for any $x \in X$ there exists some $y \in X$ such that $(x, y) \in R$, therefore the set

$$r(x) := \{y \in X \mid (x, y) \in R\}$$

is non empty. Now consider the indexed family of sets $\mathcal{R} = \langle r(x) \colon x \in X \rangle$. By AC, there exists a choice function $f \colon \mathcal{R} \to \cup\mathcal{R}$ such that $f(r(x)) \in r(x)$ for every $x \in X$, this means $(x, f(r(x))) \in R$. For simplicity, let $h(x) = f(r(x))$.

We can define a sequence $\langle x_n \rangle_{n\in\mathbb{N}}$ with $x_n = h^n(x)$ for all $n \in \mathbb{N}$, where $h^n(x)$ denotes the composition of h with itself n times. This sequence has the property that $(x_n, x_{n+1}) \in R$ for all $n \in \mathbb{N}$ since by construction $(x_n, f(r(x_n))) \in R$. □

2.4.2 Axiom of Countable Choice

(Axiom of Countable Choice CC) *Any countable collection of non empty sets has a choice function.*

The axiom of dependent choice implies the axiom of countable choice and is strictly stronger.

Theorem 2.4.2 DC ⇒ CC.

Proof: Let $\langle X_n \rangle_{n\in\mathbb{N}}$ be a family of sets. Since we want to use DC, we have to create an entire relation. To do this, let

$$X = \bigcup_{n\in\mathbb{N}} \{(a, n) \mid a \in X_n\}.$$

And define a relation $R \subseteq X \times X$ by the rule

$$((a, n), (b, m)) \in R \iff m = n + 1.$$

Now, R is entire since for any $(a, n) \in X$ and any $b \in X_{n+1}$ we have $(b, n + 1) \in X$ and $((a, n), (b, n + 1)) \in R$. By DC, there exists a sequence $\langle \mathbf{p}_n \rangle_{n\in\mathbb{N}}$ in X such that $(\mathbf{p}_n, \mathbf{p}_{n+1}) \in X$ for all $n \in \mathbb{N}$.

Letting $\mathbf{p}_n = (p_n, k_n)$ for all n, and by the definition of R, it follows that $k_{n+1} = k_n + 1$. By induction it can be shown that all numbers k_n are of the form $k_n = k_0 + n$, this implies that $p_n \in X_{N+n}$ for all n. □

2.4.3 Axiom of Choice for Well-Orderable Sets

(ACWO) *Every family of non-empty well-orderable sets has a choice function.*

Theorem 2.4.3 AC $\Rightarrow$ ACWO.

Proof: Since every family of non-empty sets has a choice function, then so does a family of non-empty well-ordered sets. $\square$

Chapter 3

Accepting the Axiom of Choice

> “*Though the Axiom of Choice is responsible for many beautiful results, it is equally responsible for the existence of several dreadful monstrosities — unwelcome and unneeded.*”

Horst Herrlich [19]

Here we will look at some of the desirable consequences and controversial results that are derived from incorporating the axiom of choice to the axioms of ZF. There is a debate on whether the disadvantages of AC are balanced by the advantages of accepting the Axiom of Choice as true. Most of the results in this section will be presented with a reference for a proof due to their length.

3.1 Consequences in Set Theory

The main results in ZF set theory related to the acceptance of AC are the implications about cardinality of sets. We can extend the notion of cardinality presented in Chapter 1 to all sets. What this mainly implies is that the ordering induced by the cardinality $|X|$ is a linear ordering, i.e., every two sets have comparable cardinalities.

Theorem 3.1.1 AC *implies that for every set X there is a unique cardinal equipotent to X.*

Proof: The finite case follows easily, from the fact that every finite set with n elements is bijective with $\{0, 1, \ldots, n-1\}$. For the infinite case, assuming AC, X can be well-ordered, therefore it is equipotent to some infinite ordinal and hence to a cardinal. $\square$

Theorem 3.1.2 *For any two sets X, Y we have $|X| < |Y|$, $|Y| < |X|$ or $|Y| = |X|$.*

Proof: Assuming AC, and by the previous theorem, we have $|X| = \alpha$, $|Y| = \beta$ for some unique cardinals α, β which are comparable, so one of these holds $\alpha < \beta$, $\beta < \alpha$ or $\alpha = \beta$. The result follows. $\square$

Theorem 3.1.3 *Every infinite set has a countable subset.*

Proof: If X is an infinite set, we can well-order it and write it as a succession of terms in a transfinite sequence $\langle x_\alpha \rangle_{\alpha<\lambda}$ where λ is an ordinal bijective with X. Since X is infinite, it is at least countable, thus $\lambda \geq \omega$ and the initial string $\{x_\alpha \mid \alpha < \omega\}$ is a countable subset of X. $\square$

This proposition can actually be proved only using the Axiom of Countable Choice. But avoiding the full Axiom complicates the proof.
Another relevant result implied by the countable Axiom of Choice is:

Theorem 3.1.4 *The union of a countable collection of countable sets is countable.*

Proof: Let $\mathcal{F}$ be a countable collection of countable sets and let $F = \cup\mathcal{F}$. Without loss of generality, assume that the sets in $\mathcal{F}$ are disjoint. Since $\mathcal{F}$ is countable, there is an enumeration of the elements of $\mathcal{F}$, thus we can write it as $\mathcal{F} = \{X_n \mid n \in \mathbb{N}\}$. For all $n \in \mathbb{N}$, X_n is countable and we can also enumerate it, so the elements in X_n can be written as

$$X_n = \{x_n(k) \mid k \in \mathbb{N}\} = \{x_n(0), x_n(1), x_n(2) \ldots\}$$

We can define a bijection $f \colon \mathbb{N} \to \cup\mathcal{F}$ by the construction

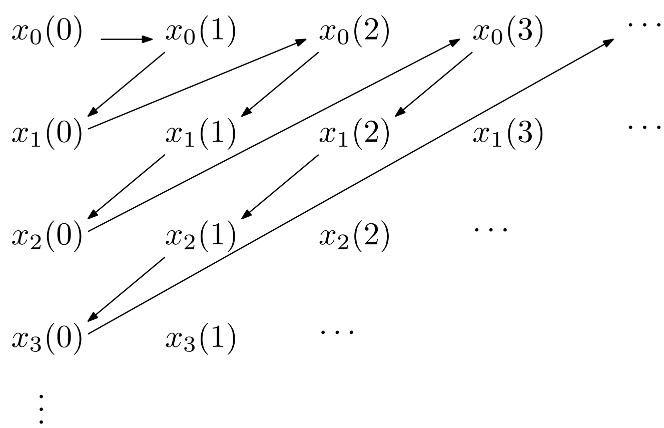

Figure 3.1: Construction that defines the bijection f.

which means $f(0) = x_0(0)$, then $f(1) = x_0(1)$, $f(2) = x_1(0)$ and so on. The inverse of this function is given by

$$f^{-1}(x_m(k)) = \frac{(m+k)(n+k+1)}{2} + (m+1).$$

Which implies that $\cup\mathcal{F}$ is countable. $\square$

Note that the use of the Axiom of Countable Choice is subtle in this proof. There could be many ways of enumerating the elements in X_n, so for every X_n we have to choose one of those enumerations. This is not the case for the enumeration of the members of $\mathcal{F}$, since we can use existential instantiation to choose one enumeration from the set consisting of all enumerations of the elements of $\mathcal{F}$.

Another important result that follows from AC is the fact that the two main definitions of finiteness are equivalent.

Definition 3.1.1 (Peano) A set S is *Peano-finite* if there exists some $n \in \mathbb{N}$ such that there is a bijective function $f \colon S \to \{0, 1, \ldots, n-1\}$.

Definition 3.1.2 (Dedekind) A set S is *Dedekind-infinite* if there is a bijection from S to a proper subset A of S. A set is *Dedekind-finite* if no such bijection exists.

A set S which is Peano-finite is always Dedekind-finite. To show this, suppose that there exists a bijection $f\colon S \to \{0, 1, \ldots, n-1\}$ for some n and that S was bijective with a proper subset $K \subset \{0, 1, \ldots, n-1\}$, then K would be bijective with $\{0, 1, \ldots, n-1\}$ under some function g, let σ be a permutation so that $\sigma(g(k)) = k$ for all $k \in K$, this composition is the identity bijection, so it has an inverse and it implies $\{0, 1, \ldots, n-1\} = K$ which contradicts the fact that K was a proper subset.

The converse, however, relies on the axiom of countable choice. We want to show that every Peano-infinite set is Dedekind-infinite. Let X be a set and define $\mathcal{G}\colon \mathbb{N} \to \mathcal{P}(X)$ so that $\mathcal{G}(n)$ is the collection of subsets of X that are Peano-finite bijective with $\{0, 1, \ldots, n-1\}$. Clearly $\mathcal{G}(n)$ is nonempty for all $n \in \mathbb{N}$ since X is infinite and if there is a smallest n such that $\mathcal{G}(n) = \varnothing$, then for any subset of X, the set is not bijective with $\{0, 1, \ldots, n-1\}$, in particular, X is not, thus X has to be Peano-finite bijective with $\{0, 1, \ldots, n-2\}$, otherwise, if there was a bijection between X and $\{0, 1, \ldots, n\}$, it would induce a bijection from a subset of X to $\{0, 1, \ldots, n-1\}$. We now can form a family of non-empty sets $\{\mathcal{G}(n) \mid n \in \mathbb{N}\}$. By the Countable Axiom of Choice, there exists a choice function f such that $f(\mathcal{G}(n)) \in \mathcal{G}(n)$ for all $n \in \mathbb{N}$. Take

$$C = \bigcup_{n \in \mathbb{N}} f(\mathcal{G}(n)).$$

Then C is a countable union of Peano-finite sets, thus it is countable and we can write it as $C = \{c_n \mid n \in \mathbb{N}\}$. Now, note that $h\colon X \to X \setminus \{c_0\}$ defined by

$$h(x) = \begin{cases} x, & \text{if } x \notin C, \\ c_{n+1}, & \text{if } x = c_n, \end{cases}$$

is a bijection. Thus X is Dedekind-infinite.

This proposition was assumed by most mathematicians before the foundational crisis of mathematics, but it is in fact impossible to prove without the Axiom of Countable Choice.

3.2 Algebra

We have already seen equivalent statements to AC in algebra like

Theorem 3.2.1 *In ZFC, every ring with identity has a maximal ideal.*

Theorem 3.2.2 *In ZFC, every vector space has a basis.*

We now turn to some weaker statements implied by AC.

Definition 3.2.1 The *algebraic closure* of a field F is an extension field $\overline{F}$ of F such that $\overline{F}$ is the smallest algebraically closed field that contains F, i.e., every non-constant polynomial has a root and every element of $\overline{F}$ is a root of a polynomial in $F[x]$.

Theorem 3.2.3 *Every field has a unique algebraic closure.*

The proof of this fact relies heavily on ZL and the reader can check it on page 12 of [6].

The following is an example of the application of the fact that assuming AC, every vector space has a basis.

Theorem 3.2.4 *The additive groups* $(\mathbb{R}, +)$ *and* $(\mathbb{C}, +)$ *(or* $(\mathbb{R}^2, +)$*) are isomorphic.*

Proof: Consider $\mathbb{R}$ as a vector space over $\mathbb{Q}$. By AC we know that this vector space has a basis B (it is uncountable since $\mathbb{R}$ is uncountable). The set $K = (B \times \{0\}) \cup (\{0\} \times B)$ is a basis of $\mathbb{R}^2$ over $\mathbb{Q}$ and has the same cardinality as B, thus there exists a bijection g between B and K. We now define a bijection $f : \mathbb{R} \to \mathbb{R}^2$. If $x \in \mathbb{R}$, there is a unique linear combination

$$x = \sum_{b \in B} \alpha_b \cdot b$$

with $\alpha_b = 0$ for all but finitely many b. We can define f as

$$f(x) = \sum_{b \in B} \alpha_b \cdot g(b)$$

this is an isomorphism of the additive groups. $\square$

3.3 Topology

Probably the most important (weaker) implication of AC in topology is the fact that the two common notions of continuity are equivalent.

Theorem 3.3.1 CC *implies that if A is a subset of a metric space (X, d) and $x \in X$, the following are equivalent: x is in the closure of A if*

1. *every neighborhood of x intersects A.*
2. *there exists a sequence $\langle x_n \rangle_{n\in\mathbb{N}}$ of points in A such that $\lim_{n\to\infty} x_n = x$.*

Proof: 1. implies 2.) Let x be in the closure of A. Consider the open ball around x,

$$B_{1/n}(x) := \left\{ y \in X \;\middle|\; d(x, a) < \frac{1}{n} \right\}.$$

By 1. $B_{1/n}(x) \cap A \neq \varnothing$ and by CC we can choose $x_n \in B_{1/n}(x) \cap A$ for every $n \in \mathbb{N}$, this is a sequence of points in A that converges to x.
2. implies 1.) Does not require the Axiom. A proof can be found in [24] or [12]. □

Definition 3.3.1 A metric space S is separable if it has a countable dense subset.

Theorem 3.3.2 *The Countable Axiom of Choice implies that every subspace of a separable metric space is separable.*

For a proof, see [6].

3.4 Analysis

The following is an implication of CC.

Theorem 3.4.1 *The sequential definition of continuity is equivalent to the ε, δ definition. That is, a function $f : \mathbb{R} \to \mathbb{R}$ is continuous at a point x if*

1. $(\forall \varepsilon > 0)(\exists \delta > 0)(\forall y)(|x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon)$.

2. $\lim_{n\to\infty} x_n = x \Rightarrow \lim_{n\to\infty} f(x_n) = f(x)$.

Proof: 1. implies 2.) does not require the Axiom. We skip the proof and refer the reader to [20] or any other introductory real analysis textbook.
2. implies 1.) Suppose 2. holds but 1. is false, then there exists some $\varepsilon > 0$ such that $\forall\delta > 0$ there exists some y such that $|x - y| < \delta$ but $|f(x) - f(y)| \geq \varepsilon$. Given that ε, define $\delta = 1/n$ and the set $A_n = \{y \mid |x - y| < \delta\}$. By CC we can choose $x_n \in A_n$ such that $|f(x) - f(x_n)| \geq \varepsilon$. Then we have $x_n \to x$ but $f(x_n) \to f(x)$ is false, which is a contradiction. Thus 1. holds. $\square$

The following theorem is called the Hahn-Banach Theorem and it essential in functional analysis. It is strictly weaker than the Axiom of Choice. A proof can be found in section 4.2-1 in [14] and uses ZL.

Definition 3.4.1 A function $f\colon V \to \mathbb{R}$ on a vector space V over $\mathbb{R}$ is a *linear functional* on V if

$$f(a\mathbf{u} + b\mathbf{v}) = af(\mathbf{u}) + bf(\mathbf{v})$$

for all $a, b \in \mathbb{R}$ and $\mathbf{u}, \mathbf{v} \in V$.

Definition 3.4.2 A function ρ on a vector space V over $\mathbb{R}$ is a *sublinear functional* on V if $\rho(\mathbf{u} + \mathbf{v}) \leq \rho(\mathbf{u}) + \rho(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v} \in V$ and $\rho(a\mathbf{u}) = a\rho(\mathbf{u})$ for all $a \in \mathbb{R}$ and $\mathbf{u} \in V$.

Theorem 3.4.2 (Hahn-Banach Theorem) *Let X be a real vector space and ρ a sublinear functional on X. Let f be a linear functional which is defined on a subspace Y of X and satisfies $f(x) \leq \rho(x)$ for all $x \in Y$. Then there exists a linear functional $\overline{f}$ such that $f \subseteq \overline{f}$ and $\overline{f}(x) \leq \rho(x)$ for all $x \in X$.*

Theorem 3.4.3 *Every nontrivial Hilbert space has a orthonormal basis.*

For a proof, see Theorem 4.1-8 in [14].

3.5 Measure Theory

For this particular branch in math, AC has been problematic since it allows us to use mathematical objects which are not possible to define explicitly and it has problematic consequences that defy intuition. In this section we will construct an object whose entire existence depends on AC.

Definition 3.5.1 A *partition* P of a set X is a family of non-empty sets that cover X and are disjoint. I.e.,

- For all $A \in P$, $A \neq \varnothing$.
- For all $A, B \in P$ such that $A \neq B$, $A \cap B = \varnothing$.
- $\cup P = X$.

It should be noted that every equivalence relation on X induces a partition with the set of equivalence classes and, conversely, every partition induces an equivalence relation on X.

Definition 3.5.2 A *coset* of $\mathbb{Q}$ in $\mathbb{R}$ is a set of the form

$$x + \mathbb{Q} = \{x + q \mid q \in \mathbb{Q}\}$$

where $x \in \mathbb{R}$.

It can be shown that the cosets of $\mathbb{Q}$ partition $\mathbb{R}$. In particular:

- if $x, y \in \mathbb{R}$ and $x - y \in \mathbb{Q}$, then $x + \mathbb{Q} = y + \mathbb{Q}$.
- if $x, y \in \mathbb{R}$ and $x - y \notin \mathbb{Q}$, then $(x + \mathbb{Q}) \cap (y + \mathbb{Q}) = \varnothing$.

Note that $\mathbb{Q}$ is dense in $\mathbb{R}$ (see [24]). We will now give a brief insight into the *Lebesgue measure*.

The Lebesgue measure, $\mu\colon \mathcal{P}(\mathbb{R}) \to \mathbb{R}_{\geq 0} \cup \{\infty\}$, is a way to assign a 'size' of a set of real numbers, we will not go into full detail of its definition. All we require is that

- The measure of a bounded subset of $\mathbb{R}$ (if it exists) is a non-negative real number.
- If $A \subseteq B$, then $\mu(A) \leq \mu(B)$.
- The measure of an interval (a,b) (with or without the endpoints) is its length, $\mu((a,b)) = b - a$.
- Countable additivity: for all sets for which the measure is defined, if A_k is measurable for all $k \in \mathbb{N}$, and $A_i \cap A_j = \varnothing$ when $i \neq j$, then

$$\mu\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mu(A_k).$$

- Unnafected by translation: for all measurable sets S, $\mu(S) = \mu(t+S)$, where $t \in \mathbb{R}$.

Since $\{x\} = [x,x]$, we have $\mu(\{x\}) = 0$ for all singleton sets. This fact, together with countable additivity, implies that $\mu(S) = 0$ for all countable subsets S of $\mathbb{R}$.

Definition 3.5.3 A *Vitali set* is a set $V \subseteq [0,1]$ that contains a unique point from each coset of $\mathbb{Q}$ in $\mathbb{R}$.

Theorem 3.5.1 AC *implies the existence of Vitali sets.*

Proof: Apply AC to each equivalence class $(x+\mathbb{Q}) \cap [0,1]$. Note that CC does not work in this case, since there are uncountably many equivalence classes. □

We would have shown the following

Theorem 3.5.2 *In ZFC, there exist non measurable bounded sets of real numbers.*

If we show that

Theorem 3.5.3 *Vitali sets are non-(Lebesgue) measurable.*

For that we begin by showing

Theorem 3.5.4 *Let V be a Vitali set, then the sets in $\{q+V \mid q \in \mathbb{Q}\}$ are pairwise disjoint and $\mathbb{R} = \bigcup_{q\in\mathbb{Q}}(q+V)$.*

Proof: Suppose $x \in (q+V) \cap (p+V)$ for $p, q \in \mathbb{Q}$. Then $x = q + v_1 = p + v_2$ for some $v_1, v_2 \in V$. Thus $v_1 = x - q$ y $v_2 = x - p$ so that $v_1, v_2 \in x + \mathbb{Q}$.
But V has at most one element of each coset of $\mathbb{Q}$, thus $v_1 = v_2$ y $p = q$. Now, for all $x \in \mathbb{R}$ there exists some $v \in V$ such that $v \in x + \mathbb{Q}$. Therefore $v = q + x$ for some $q \in \mathbb{Q}$, thus $x = v - q$ from which we conclude $x \in -q + V$. □

Theorem 3.5.5 *Let $V \subseteq [0,1]$ be a Vitali set, let $S = \mathbb{Q} \cap [-1,1]$ and define*

$$U = \bigcup_{q\in S}(q+V).$$

Then

$$[0,1] \subseteq U \subseteq [-1,2].$$

Proof: Since $V \subseteq [0,1]$, we know that $q+V \subseteq [-1,2]$ for all $q \in [-1,1]$. Therefore, $U \subseteq [-1,2]$. To show that $[0,1] \subseteq U$, let $x \in [0,1]$. Since V is a Vitali set, there exists some $v \in V$ such that $v \in x + \mathbb{Q}$. Thus, $v = x + q$ for some $q \in \mathbb{Q}$. Since $v, x \in [0,1]$, we have $q = v - x \in [-1,1]$. Consequently, $q \in S$ and $x \in -q + V$, which shows that $x \in U$. □

We can now show that Vitali sets are non-measurable.

Proof: (of Theorem 3.5.3) Let V be a Vitali set. Suppose it is measurable. Let $S = \mathbb{Q} \cap [-1,1]$ and $U = \cup_{q\in S}(q+V)$. U is a countable union of measurable sets, thus measurable. By our previous result,

$$[0,1] \subseteq U \subseteq [-1,2]$$

thus $1 \leq \mu(U) \leq 3$. But

$$\mu(U) = \mu\left(\bigcup_{q\in S}(q+V)\right) = \sum_{q\in S}\mu(q+V) = \sum_{q\in S}\mu(V).$$

If $\mu(V) = 0$, then $\mu(U) = 0$. If $\mu(V) > 0$, then $\mu(U) = \infty$. In both cases we reach a contradiction to $1 \leq \mu(U) \leq 3$. $\square$

The existence of non-measurable sets was the starting point for the construction of other paradoxical decompositions like the Banach-Tarski paradox, informally, the statement is

Theorem 3.5.6 (Banach-Tarski paradox) *It is possible to partition $\{x \in \mathbb{R}^3 \mid \|x\| \leq 1\}$ (the unit closed ball in $\mathbb{R}^3$) into a finite number of pieces, which can be re-assembled by rigid motions to form two unit spheres.*

For a proof of this seeminly 'absurd' result, see [6]. Such an unexpected conclusion challenges our intuitive understanding of 'space' and 'volume'. As a consequence, mathematicians have debated about the philosophical and logical implications of the Axiom of Choice, contemplating its role in shaping the boundaries of mathematical reality and its intuitive appeal.

3.6 Logic

Finally, if one wishes to exclude AC from the rest of the axioms of set theory, one has to sacrifice one of the most useful propositions in logic.

(Law of Excluded Middle) *For every proposition p, 'either p or $\neg p$' is a true statement.*

Theorem 3.6.1 (Diaconescu's theorem) AC *implies the law of excluded middle.*

For a proof, see [8]. Rejecting the law of excluded middle has concerning consequences: Proofs by contradiction may not work anymore, but the discussion of this matter falls beyond the scope of this manuscript.

Chapter 4

The Axiom of Choice and ZF

> *“Ideas can be works of magic; they are in a chain of development that may eventually find form. All ideas need not be made physical.”*
>
> Jeremy Millar, Sentences on Magic (After Sol Lewitt), 2009

When considering the criterion of convenience for accepting the Axiom of Choice, its extensive use across various mathematical branches in the last 50 years strongly suggests its validity. However, acknowledging the Axiom's nonconstructive nature, which makes it less intuitively evident than other axioms, prompts the question of its formal consistency. Does introducing the axiom to the existing set of axioms in set theory result in a contradiction? Godel's work in 1939 resolved this concern, confirming that the axiom of choice is consistent within the framework of axiomatic set theory.

Some people might be compelled to use AC due to its equivalences and desirable results, as shown in Chapter 2 and Chapter 3. While it would be ideal if we could find a proof for the Axiom of Choice using more fundamental and universally accepted principles, it has become evident that this is highly unlikely due to the Axiom of Choice's unique nature. Our next best course of action is to demonstrate its compatibility with other mathematical principles, thereby establishing its consistency within the framework of ZF.

In this chapter we will see how the Axiom of Choice fits with the other axioms of set theory.

4.1 Axioms of Set Theory

Zermelo's presentation of the Well-Ordering Theorem, utilizing the Axiom of Choice, was succeeded by his 1908 publication introducing a set of axioms. He intended to formalize set theory and establish a solid foundation for mathematics. Zermelo's axioms sought to encapsulate the intuitive concept of a set while avoiding paradoxes such as Russell's Paradox. The axiomatic system proposed by Zermelo was then enhanced and through the contributions made by Fraenkel, Skolem, and von Neumann in the following years.

The nine axioms of ZF (Zermelo-Fraenkel) are

1. *Axiom of Extension* If two sets have the same members, they are equal, more formally:

$$(\forall x)(x \in A \leftrightarrow x \in B) \to (A = B).$$

2. *Empty Set Axiom* There exists a set with no elements:

$$(\exists x)(\forall y)\neg(y \in x).$$

3. *Axiom of Pairing* If A and B are sets, then there is a set $\{A, B\}$ whose only members are A and B.

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

4. *Axiom of Union* If X is a set, there is a set $\cup X$ whose elements are the members of the members of X.

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

5. *Power Set Axiom* If X is a set, there exists a set $\mathcal{P}(X)$ that contains all the subsets of X.

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

6. *Axiom of Infinity* There is a set A such that $\varnothing \in A$ and if $x \in A$, then $\{x \in A\}$.

$$(\exists x)((\varnothing \in x) \wedge (\forall y)((y \in x) \rightarrow (\exists z)((y \in z) \wedge (z \in x)))).$$

7. *Axiom of Specification* Let ϕ be a formula in the language of set theory with one variable x and let A be a set. There exists a set B that contains all members of A that satisfy ϕ.

$$(\forall a)(\exists b)(\forall x)((x \in b) \leftrightarrow ((x \in a) \wedge \phi(x))).$$

8. *Axiom of Regularity* If $A \neq \varnothing$, then there exists $x \in A$ such that $x \cap A = \varnothing$.

$$(\forall x)((x \neq \varnothing) \rightarrow (\exists y)((y \in x) \wedge (y \cap x = \varnothing))).$$

It should be noted that the Axiom of Pairing and the Axiom of Regularity imply that no set can be a member of itself. The with the addition of the Axiom of Choice to ZF we obtain ZFC.

4.2 Consistency and Independence

A concern might rise from adopting the Axiom of Choice into our axiomatic system. If it has so many unwanted consequences, should we expect it not give rise to contradictions in our system?

We say that an axiomatic system is inconsistent whenever we can deduce some statement ϕ and its negation $\neg\phi$ within the system; if no such statement exists, then the system is consistent. In 1938, Gödel proved that if ZF is consistent, then ZFC is consistent, thus establishing the relative consistency of ZFC. Due to the second incompleteness theorem, it is impossible to show that ZF is consistent from ZF alone. The same happens with ZFC. This result has placed AC in almost the same position as the parallel postulate in geometry. One can work in a setting in which the Axiom is false or true and obtain interesting results.

The independence of the Axiom of Choice from other axioms within ZF was later demonstrated by Paul Cohen in 1963. Utilizing a method known as forcing, Cohen constructed a model of ZF where a set of real numbers cannot be well-ordered, implying a model in which both ZF and the negation of AC coexist. Cohen's work solidified the idea that, assuming the consistency of ZF, the addition of the negation of AC to ZF is also consistent.

The results of Gödel and Cohen show that AC is independent from ZF, that is, we cannot prove nor refute AC.

Chapter 5

Negating the Axiom of Choice: A Discussion

> *A formal system in which $\exists x G(x)$ is provable, but which provides no method for finding the x in question, is one in which the existential quantifier fails to fulfill its intended function.*

-R.L Goodstein [5]

5.1 Reasons to Object the Axiom of Choice

There are two main reasons, each accompanied by its respective arguments, that have been presented as valid points to either negate or not incorporate AC in ZF or some other axiomatic system. The first reason has to do with what we consider to be intuitive and the second one deals with the concept of existence in mathematics. In this section we provide a brief explanation of both.

5.1.1 Intuition

There are many results obtained from the Axiom of Choice that seem *obviously* true, such as the union of countably many countable sets is countable or that the definitions of continuity are equivalent. The Axiom itself seems obvious in some instances. And given that those results seem true without question and have desirable consequences, when working with a system of axioms for set theory, it should be reasonable to want those consequences to be true in that system. So one should better include AC in that set of axioms.

This argument is similar in style to affirming the consequent, that is, drawing the conclusion that if the consequent is true, then the antecedent must be true too. *If the Axiom of Choice allows us to have those desirable results, and those are intuitively true, then the Axiom must be true*. But it can also (and more rightfully according to [18]) be seen as an Inference to the Best Explanation, which describes a situation in which the results that we want to obtain, in turn, provide an essential part for believing that the assumption of the Axiom is a reasonable thing to do.

However, the initial appeal to intuition can be challenged by the observation that when we find the Axiom of Choice obvious, it often stems from an intuition grounded in the objects of our real-world experiences. While it might be intuitive to imagine a choice function selecting an element from each box in a collection of infinitely (but countably) many boxes containing distinct items, this intuition may rely on imagining the objects existing in a physical space, where we can assign a coordinate to each point and thus come up with a rule to order the elements in each box and pick the one that is the 'smallest' under that order. Doing so without the very physical nature of the setting takes away the intuition, and the Axiom becomes difficult to justify or visualize.

The Banach-Tarski Paradox, the existence of non-measurable sets and similar results highlight the nonconstructive nature of the Axiom of Choice, as it allows for the existence of mathematical objects with counterintuitive properties that cannot be visualized or realized in the physical world. This clash between mathematical formalism and our intuitive grasp of space and quantity is a starting point for the philosophical debate about the foundations and limits of set theory and the role of the Axiom of Choice in shaping mathematical reality. It poses the question of what really counts as being *mathematically intuitive*.

5.1.2 Constructivism

In mathematics, constructivists and intuitionists deny the existence of all things that cannot be explicitly "constructed". That is, there is no clear rule that allows us to build those objects. The constructivists idea of "existence" is very strict, and it can lead them to extremes such as the abolishment of the law of excluded middle. So they deny AC since incorporating it would automatically imply that law. Under such strictness, the existential quantifier $(\exists x)G(x)$ only makes sense when one can explicitly construct such an x that has property G. So the Axiom of Choice is naturally false in this setting, since it poses the existence of a choice function for any collection of non-empty sets, without ever mentioning a specific set of instructions for the construction of such a function. Take for example the existence of Vitali sets, which require us to make infinitely (uncountably) many choices, since we cannot create infinitely many rules due to the limitations of logic, we have to rely on the Axiom to construct the set. Vitali sets exist as a consequence of a non-constructive assumption, so in constructive mathematics, it would be impossible to have a non-measurable set of real numbers. In a constructive world, the only things that are true are those which can be computed.

While what we have presented in this section doesn't constitute an assertion that the axiom of choice is untrue, it does serve a rhetorical purpose akin to an argument against it. By challenging the axiom of choice's presumed obviousness, it lays the groundwork for presenting a contrasting perspective.

5.2 Negation of the Axiom of Choice: Consequences

In order to contrast the discussion of the previous chapter, we present here some not-so-desirable consequences of negating the Axiom of Choice. Some of those consequences would have a great impact on the way we do mathematics. Working in ZF with the assumption that AC is false or with any other axiom that allows us to disprove AC, we can sometimes show that

- There is a set that has a partition into strictly more equivalence classes than the original set has elements.
- The definitions of continuity (sequential and ε, δ) are not equivalent.
- There is an infinite set of real numbers without a countably infinite subset.
- The real numbers are a countable union of countable sets. [6]
- There exists a vector space with no basis.
- There is a field with no algebraic closure.
- There is a vector space with two bases of different cardinalities. [19]
- There exists a model of ZF with the negation of AC such that every subset of $\mathbb{R}$ is measurable and The Banach-Tarski paradox does not hold.

For proofs of these facts, see [6].

Chapter 6

Final Remarks

When working on tasks like solving differential equations, studying manifolds, or finite-order groups, the Axiom of Choice is unlikely to play a role. Nevertheless, a significant portion of contemporary mathematics revolves around abstract infinite structures, which leads us to delve deeper into the foundational aspects of set theory. The axiom of choice is essential not only in foundational theories such as set theory and model theory but also in various modern disciplines like point set topology, algebra, functional analysis, and measure theory.

Even though the Axiom of Choice brings about some undesirable consequences, as seen in Chapter 3, it is still consistent with the other axioms of ZF. It does not give rise to contradictions, it only defies our intuition and it makes us question what does it mean to be constructible or what existence means in the context of mathematics. This means we can comfortably accept the Axiom of Choice as a foundational principle for mathematics. But in my opinion, the intuitive nature of this axiom tends to disappear when one thinks about it deeply, so there is no reason whatsoever to believe the truth or the falsity of the axiom. However, it should be noted that most of modern mathematics take AC for granted and many mathematicians use it, even without noticing it. All this work demonstrates how useful or convenient AC is. But the question of the truth of AC is left unanswered.

The debate over the Axiom of Choice reminds us that objectivity is elusive. Our theories are built from assumptions that can be modified. There is room for discussion about the use of mathematics in certain aspects of science, since it is subject to the arbitrary categories that we defined in order to make those mathematical applications. Nonetheless, there is joy is in the exploration of the consequences of these assumptions. Theories are nothing but playgrounds for grown up kids.

Bibliography

[1] Alfred Tajtelbaum-Tarski. “Sur quelques théorèmes qui équivalent à l’axiome du choix”. In: *Fundamenta Mathematicae* 5.1 (1924), pp. 147–154.

[2] Ernst Zermelo. “Über Grenzzahlen Und Mengenbereiche: Neue Untersuchungen Über Die Grundlagen der Mengenlehre”. In: *Fundamenta Mathematicæ* 16 (1930), pp. 29–47.

[3] J. Kelley. “The Tychonoff product theorem implies the axiom of choice”. eng. In: *Fundamenta Mathematicae* 37.1 (1950), pp. 75–76. URL: `http://eudml.org/doc/213229`.

[4] Herman Rubin and Jean E. Rubin. *Equivalents of the axiom of choice*. Vol. 34. Studies in Logic and the Foundations of Mathematics. North-Holland, 1963.

[5] R. L. Goodstein. “Existence in mathematics”. en. In: *Compositio Mathematica* 20 (1968), pp. 70–82. URL: `http://www.numdam.org/item/CM_1968__20__70_0/`.

[6] Thomas Jech. *The Axiom of Choice*. London: North-Holland Publishing Company, 1973.

[7] Paul Bernays. “A System of Axiomatic Set Theory**Almost unchanged reproduction of seven papers under the same general title. Reprinted by permission of the publisher, the Association for Symbolic Logic, from The Journal of Symbolic Logic, Volumes 2 (1937) 66-77; 6 (1941) 1-17; 7 (1942) 65-89, 133-146; 8 (1943) 89-106; 13 (1948) 65-79; 19 (1964) 81-96.” In: *Sets and Classes on The Work by Paul Bernays*. Ed. by Gert H. Müller. Vol. 84. Studies in Logic and the Foundations of Mathematics. Elsevier, 1976, pp. 1–119. DOI: `https://doi.org/10.1016/S0049-237X(08)70892-9`. URL: `https://www.sciencedirect.com/science/article/pii/S0049237X08708929`.

[8] N. Goodman and J. Myhill. “Choice Implies Excluded Middle”. In: *Zeitschrift fur mathematische Logik und Grundlagen der Mathematik* 24.25-30 (1978), pp. 461–461. DOI: 10.1002/malq.19780242514.

[9] Wilfrid Hodges. “Krull Implies Zorn”. In: *Journal of The London Mathematical Society-second Series* (1979), pp. 285–287. URL: https://api.semanticscholar.org/CorpusID:122446732.

[10] Gregory H. Moore. “The Prehistory of the Axiom of Choice”. In: *Zermelo’s Axiom of Choice: Its Origins, Development, and Influence*. New York, NY: Springer New York, 1982, pp. 5–84. ISBN: 978-1-4613-9478-5. DOI: 10.1007/978-1-4613-9478-5_2. URL: https://doi.org/10.1007/978-1-4613-9478-5_2.

[11] Andreas Blass. “Existence of bases implies the axiom of choice”. In: 1984. URL: https://api.semanticscholar.org/CorpusID:116499095.

[12] James Dugundji. *Topology*. Allyn and Bacon, 1987.

[13] Herman Rubin and Jean E. Rubin. “Equivalents of the Axiom of Choice, II”. In: *Journal of Symbolic Logic* 52.3 (1987), pp. 867–869. DOI: 10.2307/2274372.

[14] Erwin Kreyszig. *Introductory functional analysis with applications*. Wiley, 1989.

[15] Peter; J. Cameron. *Sets, Logic and Categories*. 1st ed. Springer Undergraduate Mathematics Series. Springer London, 1999.

[16] Karel Hrbacek and Thomas Jech. *Introduction to set theory*. Marcel Dekker, 1999.

[17] Steven G. Krantz. “The Axiom of Choice”. In: *Handbook of Logic and Proof Techniques for Computer Science*. Boston, MA: Birkhäuser Boston, 2002, pp. 121–126. ISBN: 978-1-4612-0115-1. DOI: 10.1007/978-1-4612-0115-1_9. URL: https://doi.org/10.1007/978-1-4612-0115-1_9.

[18] THOMAS FORSTER. “THE AXIOM OF CHOICE AND INFERENCE TO THE BEST EXPLANATION”. In: *Logique et Analyse* 49.194 (2006), pp. 191–197. ISSN: 00245836, 22955836. URL: http://www.jstor.org/stable/44084824 (visited on 11/26/2023).

[19] Horst Herrlich. “Disasters with Choice”. In: *Axiom of Choice*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 117–136. ISBN: 978-3-540-34268-7. DOI: 10.1007/3-540-34268-0_5. URL: https://doi.org/10.1007/3-540-34268-0_5.

[20] Michael Spivak. *Calculus*. Fourth. Publish or Perish, 2008.

[21] “Zermelo 1908a”. In: *Ernst Zermelo - Collected Works/Gesammelte Werke: Volume I - Set Theory, Miscellanea / Band I - Mengenlehre, Varia*. Ed. by Heinz-Dieter Ebbinghaus, Craig G. Fraser, and Akihiro Kanamori. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 120–159. ISBN: 978-3-540-79384-7. DOI: 10.1007/978-3-540-79384-7_5. URL: https://doi.org/10.1007/978-3-540-79384-7_5.

[22] Ulrich Felgner. “Zermelo 1908b”. In: *Ernst Zermelo - Collected Works/Gesammelte Werke: Volume I - Set Theory, Miscellanea / Band I - Mengenlehre, Varia*. Ed. by Heinz-Dieter Ebbinghaus, Craig G. Fraser, and Akihiro Kanamori. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 160–229. ISBN: 978-3-540-79384-7. DOI: 10.1007/978-3-540-79384-7_6. URL: https://doi.org/10.1007/978-3-540-79384-7_6.

[23] Aimeric Malter, Dierk Schleicher, and Don Zagier. “New Looks at Old Number Theory”. In: *The American Mathematical Monthly* 120.3 (2013), pp. 243–264. ISSN: 00029890, 19300972. URL: https://www.jstor.org/stable/10.4169/amer.math.monthly.120.03.243 (visited on 11/15/2023).

[24] Sidney A. Morris. *Topology without tears*. University of New England, 2023.