

The College of Wooster

Open Works

Senior Independent Study Theses

2022

The Infinity Conundrum: Understanding Topics In Set Theory And The Continuum Hypothesis

Sabrina Grace Helck

The College of Wooster, shelck22@wooster.edu

Follow this and additional works at: <https://openworks.wooster.edu/independentstudy>



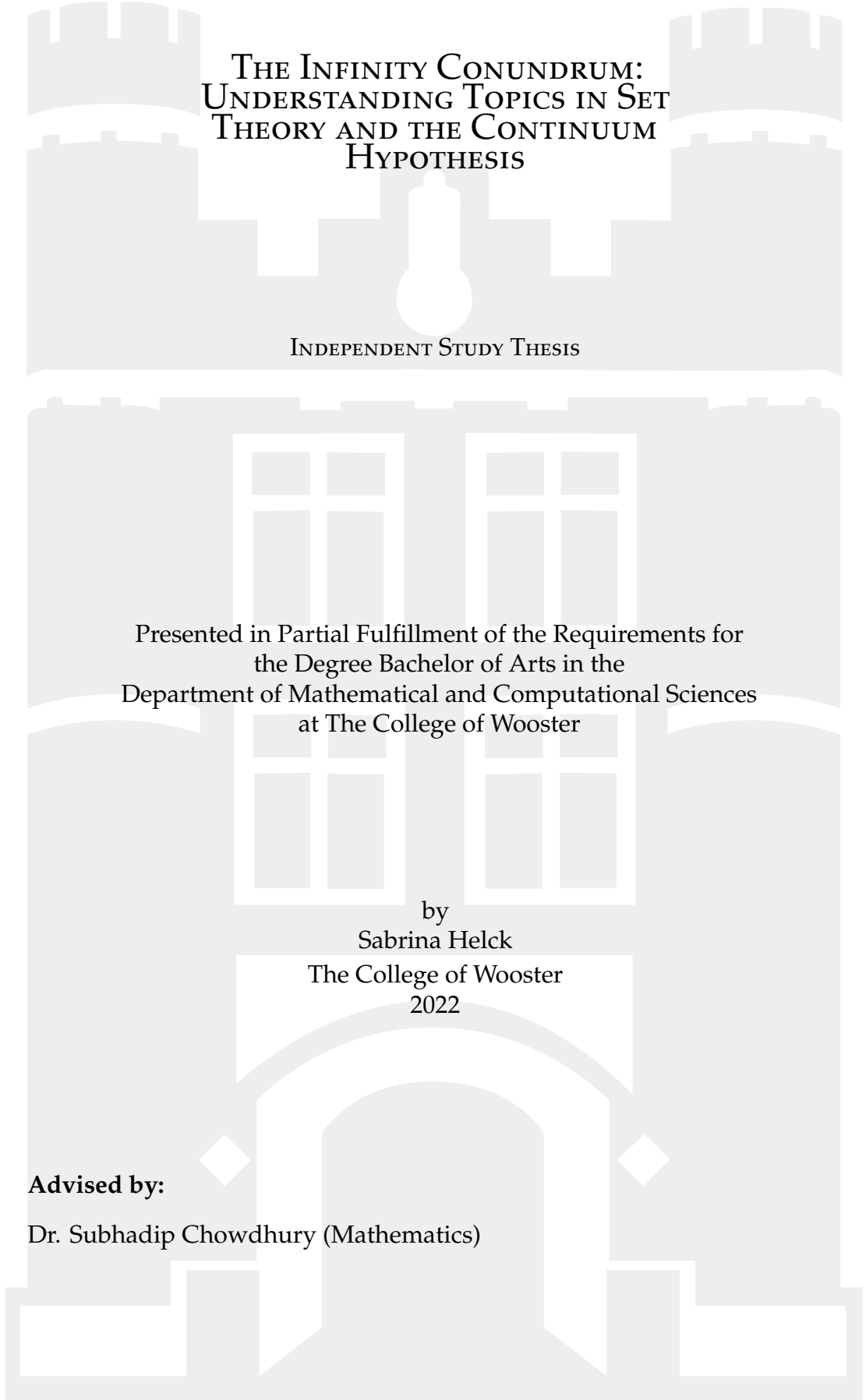
Part of the [Set Theory Commons](#)

Recommended Citation

Helck, Sabrina Grace, "The Infinity Conundrum: Understanding Topics In Set Theory And The Continuum Hypothesis" (2022). *Senior Independent Study Theses*. Paper 9828.

This Senior Independent Study Thesis Exemplar is brought to you by Open Works, a service of The College of Wooster Libraries. It has been accepted for inclusion in Senior Independent Study Theses by an authorized administrator of Open Works. For more information, please contact openworks@wooster.edu.

© Copyright 2022 Sabrina Grace Helck



THE INFINITY CONUNDRUM:
UNDERSTANDING TOPICS IN SET
THEORY AND THE CONTINUUM
HYPOTHESIS

INDEPENDENT STUDY THESIS

Presented in Partial Fulfillment of the Requirements for
the Degree Bachelor of Arts in the
Department of Mathematical and Computational Sciences
at The College of Wooster

by
Sabrina Helck

The College of Wooster
2022

Advised by:

Dr. Subhadip Chowdhury (Mathematics)



THE COLLEGE OF
WOOSTER

© 2022 by Sabrina Helck

ABSTRACT

This project is concerned with articulating the necessary background in order to understand the famous result of the undecidability of the continuum hypothesis. The first chapter of this independent study discusses the foundations of set theory, stating fundamental definitions and theorems that will be used throughout the remainder of the project. The second chapter focuses on ordinal and cardinal numbers which will directly relate to the final chapter. First, there is a clear explanation of the notion of order and what it means for a set to be well-ordered. Then ordinal numbers are defined and some properties are listed and proved. The second half of this chapter discusses cardinal numbers. Similarly, they are defined and some of their properties are stated. Some arithmetic rules surrounding cardinal numbers are discussed as an extension to those properties. The next chapter is concerned with Zermelo-Fraenkel set theory and the axiom of choice (ZFC) which introduces the idea of set theoretic systems and models. All nine axioms are listed and expanded upon. Additional focus is put on the axiom of choice and its equivalent statements. The final chapter states the continuum hypothesis, as well as the weak continuum hypothesis and the generalized continuum hypothesis. Some additional background of inner models is discussed for subsequent proof. Kurt Gödel proved that the continuum hypothesis could not be proven false within ZFC. The outline for this proof is discussed to reflect its main points. Paul Cohen proved that the continuum hypothesis could not be proven true within ZFC, although this

is not discussed as extensively. With this last chapter, the end result becomes clear that the continuum hypothesis is independent of ZFC.

ACKNOWLEDGMENTS

I would like to take the opportunity to thank my advisor, Dr. Subhadip Chowdhury, for his encouragement and patience throughout this process. There were many topics I struggled to understand, but his support and dedication to my independent study helped me through it. I would also like to thank my other mathematics professors who initially sparked my interest in the subject and encouraged me to pursue it as a major. I am grateful that they have kept me motivated these past four years, and I carried over that motivation to this project.

I want to thank Coach Dennis Rice, for creating a great team environment and supporting my endeavors. I would also like to thank my friends on campus, and especially my housemates who always encouraged me and were there to make me laugh whenever I needed some stress relief. Finally, I would like to thank my parents for their never-ending care and support over these past four years and throughout this independent study process.

CONTENTS

Abstract	v
Acknowledgments	vii
Contents	ix
List of Figures	xi
CHAPTER	PAGE
1 Introduction	1
2 Countable and Uncountable Sets	5
2.1 Finite Sets	5
2.2 Infinite Sets	7
3 Ordinal Numbers and Cardinal Numbers	29
3.1 Order and Ordinal Numbers	29
3.2 Cardinal Numbers	38
4 Zermelo-Fraenkel Set Theory	45
4.1 Zermelo-Fraenkel Axioms	46
4.2 Axiom of Choice	55
5 The Continuum Hypothesis	65
5.1 Background	65
5.2 Provability	67
5.3 Subsequent Inquiries and Implications	72
Afterword	75
References	77

LIST OF FIGURES

Figure		Page
2.1	Process of constructing the sequence between sets A and B for the Cantor-Bernstein Theorem.	13
2.2	Visual process of listing the rational numbers [4].	17
2.3	Construction of the Cantor Set [17]	19
2.4	Graph of $\tan\left(\pi x - \frac{\pi}{2}\right)$ using Desmos Graphing Calculator	22
3.1	Order diagram of division relations on the set S	31
3.2	Order diagram of division relations on the set S'	32
3.3	How to determine the initial segment of a set.	34
3.4	Mapping from A to $f(A)$ [7]	41
4.1	The element b is a predecessor of a , so it must also be in B	59
4.2	Any c where cRa must map to c contained in the set of all d where dSb	60
4.3	We know that $C \subset A$. Then $\langle B, S \rangle$ is an initial segment of $\langle A, R \rangle$, and a is the least element of $B \cap C$	61
5.1	Visual representation of the cumulative hierarchy of sets [3].	68

CHAPTER 1

INTRODUCTION

Infinity has been a topic that mathematicians have pondered for decades and became an important subject discussed in the field of set theory. Set theory is the basis for many famous results in mathematics. It provides us with the tools needed to understand fascinating concepts such as infinity. And with this understanding, we can ask more questions and investigate new topics and problems such as the continuum hypothesis. The continuum hypothesis is a fascinating problem, and it combines many different set theoretic objects such as cardinal and ordinal numbers to form a problem that tells us a lot about the nature of mathematical systems. The path to resolving the continuum hypothesis also shows how much the field of set theory has evolved since it was first introduced.

Set theory originated from German mathematician Georg Cantor in the late 1800s when up to this point there was only the discovery of finite sets. Cantor was the first to introduce infinite sets and the idea of countable and uncountable infinities. He constructed some fundamental definitions such as subsets, power sets, and cardinality. His discovery of countable and uncountable sets presented new ideas about different sizes of infinity, leaving mathematicians to ask more questions such as how many different sizes of infinity exist. Turns out, there are an infinite number of different sizes of infinity.

Not long after Cantor's discovery of different sizes of infinity, he formulated his

continuum hypothesis initially stating that there exists no infinity of a size that is in between that of countable and uncountable infinity. In other words, there is no set less than an uncountable set but greater than a countable set. He later generalized this, to state it for all sizes of infinity, beyond just the two countable and uncountable infinite sets. The continuum hypothesis was attempted to be proven or disproven, but no one was successful. For this reason, it was placed first on Hilbert's list of problems which was published in 1900. Hilbert's list of problems is a famous list of 23 influential mathematical problems that were unsolved at the time. The fact that the continuum hypothesis was placed first on the list emphasizes its significance.

Having introduced Cantor's role in the field of set theory, it is important to note that his original set theory was not perfect and revealed issues and paradoxes such as Russell's Paradox. This does not at all dismiss his work as he is one of the most influential mathematicians in the field. However, these issues did need to be resolved with a more rigorous system of set theory. Mathematicians Ernst Zermelo and Abraham Fraenkel did just that by proposing a system of axioms that would eliminate those paradoxes known as the Zermelo-Fraenkel axioms in the early 1900s. Another important axiom, the axiom of choice, was added later to construct the general model of set theory we use today referred to as ZFC. Within this system, mathematicians were able to find the conclusive results of the continuum hypothesis. It was proven that it cannot be proven true and proven that it cannot be proven false within the Zermelo-Fraenkel axioms.

The famous result of the continuum hypothesis is that it is unprovable in ZFC. In 1940 Kurt Gödel proved that it cannot be proven false within ZFC, and in 1963 Paul Cohen proved that it cannot be proven true within ZFC. This is a compelling result that requires much background to fully understand. We will begin with the foundations of set theory leading us to explore topics of infinity and cardinality. Expanding on this, we will talk about different types of numbers called ordinal and

cardinal numbers. The last piece of background we will go over are the axioms in ZFC and their relevance to the topic at hand. We will see that most of these topics are intertwined. For this reason, our final discussion on the continuum hypothesis will be cumulative tying in all the prior concepts. We will be able to clearly state the hypothesis, consider its implications, and gain a more sound understanding of how Gödel and Cohen arrived at the result of unprovability.

CHAPTER 2

COUNTABLE AND UNCOUNTABLE SETS

The objective of this chapter is to introduce important concepts in set theory related to cardinality and infinity. First, we want to understand some baseline definitions introduced by Georg Cantor in the 1800s. This collection of definitions that we will cover in this chapter is often referred to as naive set theory. When we use the term naive, we mean that although these definitions are true, they have not been established in a formal system. We will see in a later chapter that this can lead to some negative consequences.

2.1 FINITE SETS

Simply speaking, a **set** is a collection of things such as numbers, letters, ordered pairs, functions, etc. The things inside the sets are called **elements**. Sets are typically denoted by listing the elements in closed brackets separated by commas such as $A = \{1, 2, 3\}$ which is the set A with elements 1, 2, 3. We can write $1 \in A$ and $4 \notin A$ since 1 is an element of A and 4 is not. A set is **finite** if it contains a finite number of elements such as $\{2, 4, 6, 8\}$. A set is **infinite** if it contains an infinite number of elements. For infinite sets, there will be several sets used throughout this chapter that we must define. We have $\mathbb{N} = \{1, 2, 3, \dots\}$ denotes the set of natural numbers. \mathbb{Z} denotes the set of all integers $\{\dots -2, -1, 0, 1, 2, \dots\}$. Then $\mathbb{Q} = \{\frac{p}{q} : p, q \in \mathbb{Z} \text{ and } q \neq 0\}$

which denotes the set of rational numbers. Finally, \mathbb{R} denotes the set of real numbers. There is also the **empty set** denoted \emptyset that contains no elements. Thus, $\emptyset = \{\}$.

If every element of A is an element of another set B then A is a **subset** of B denoted $A \subseteq B$. This would make B a **superset** of A . We would call A a **proper subset** of B denoted $A \subset B$ if A is still a subset of B , but there are elements in B that are not contained in A . Formally, we define subsets as follows,

Definition 2.1.1. *A set A is a subset of B if for any $x \in A$, $x \in A \Rightarrow x \in B$.*

It is important to note that the empty set \emptyset will always be a subset of any set. If a finite set contains n elements then it has 2^n subsets. Two subsets are said to be equal if they both contain the same elements. However, as sets become larger, it may become difficult to compare every single element. Instead, we can use subsets to determine this equality.

Theorem 2.1.2. *If $A \subseteq B$ and $B \subseteq A$ then $A = B$.*

Another type of set is a **power set** which is the set of all the subsets of a set. As mentioned earlier, if a set A contains n elements, then it has 2^n subsets. Thus, $\mathcal{P}(A)$ must contain 2^n elements.

Definition 2.1.3. *Given a set A , the power set of A denoted $\mathcal{P}(A)$ is the set of all subsets of A . $\mathcal{P}(A) = \{X : X \subseteq A\}$.*

Now we can discuss some operations on sets such as the union and intersection between two sets. The **union** between two sets is the set of elements that are in either set. The **intersection** of two sets contains the elements that the two sets have in common. If two sets are **disjoint**, they share no elements in common. Thus, the intersection of two disjoint sets equals the empty set.

Definition 2.1.4. 1. *The union of sets A and B denoted $A \cup B$ is $A \cup B = \{x : x \in A \text{ or } x \in B\}$.*

2. The intersection of sets A and B denoted $A \cap B$ is $A \cap B = \{x : x \in A \text{ and } x \in B\}$.

Another operation that we can perform on sets is the cartesian product. The **cartesian product** of two sets A and B denoted $A \times B$ results in a set of ordered pairs.

Definition 2.1.5. *The cartesian product of two sets A and B denoted $A \times B$ results in another set defined as $A \times B = \{(a, b) : a \in A, b \in B\}$.*

Let's suppose $A = \{1, 2\}$ and $B = \{2, 4\}$, then the cartesian product would be $A \times B = \{(1, 2), (1, 4), (2, 2), (2, 4)\}$.

Next, we want to introduce the idea of **cardinality**. For finite sets, we can define cardinality to be the number of elements in the set. We can say the cardinality of our set $A = \{1, 2\}$ denoted $|A|$ is 2 since there are two elements in the set. Thus, $|A| = 2$. Also, $A \cup B = \{1, 2, 4\}$, so $|A \cup B| = 3$. The cardinality of the empty set \emptyset equals 0. Knowing the definition of power set, we can also state the cardinality of any power set.

Theorem 2.1.6. *If $|A| = n$ for some $n \in \mathbb{N}$, then $|\mathcal{P}(A)| = 2^n$.*

The difference between cardinality of finite sets and infinite sets is that cardinality of finite sets will give us some natural number. However, with an infinite set, we determine its cardinality by comparing it to another infinite set which we will now discuss.

2.2 INFINITE SETS

Cardinality for finite sets is as simple as counting each element in the set. However, when sets get larger and larger the process of counting the number of elements becomes more difficult. It becomes even more complicated when we are dealing with infinite sets. To resolve this issue, we need to define cardinality differently

when discussing infinite sets. To do this we must introduce functions and what it means for a function to be one-to-one and onto.

A **function** is a relation or mapping between two sets such as $f : A \rightarrow B$. In this case we would call the set A the **domain** of the function and B the **codomain**. The **range** is all of the possible outputs of f or the set $\{f(a) : a \in A\}$. The formal definition of a function is stated below:

Definition 2.2.1. *A function f from A to B is a relation $f \subseteq A \times B$ from A to B , satisfying the property that for each $a \in A$ the relation f contains exactly one ordered pair of form (a, b) .*

We can have $f(a) = b$ for some $a \in A$ and $b \in B$ resulting in the ordered pair (a, b) . Then we would call a the **pre-image** of b , and we would call b the **image** of a .

Definition 2.2.2. *Consider the function $f : A \rightarrow B$.*

1. *If $X \subseteq A$, the image of X is the set $f(X) = \{f(x) : x \in X\} \subseteq B$*
2. *If $Y \subseteq B$, the pre-image of Y is the set $f^{-1}(Y) = \{x \in A : f(x) \in Y\} \subseteq A$*

Understanding the fundamentals of functions, we can now define what it means for a function to be **injective**, **surjective**, and **bijective**.

Definition 2.2.3. 1. *A function is **injective** (one-to-one) if for all $a_1, a_2 \in A$, $a_1 \neq a_2$ implies $f(a_1) \neq f(a_2)$.*

2. *A function is **surjective** (onto) if for all $b \in B$, there exists an $a \in A$ such that $f(a) = b$.*

3. *A function is **bijective** if it is both injective and surjective.*

Another way of saying a function is injective is saying that it is one-to-one which may be more intuitive. No two elements in the domain can map to the same element

in the range. Each element in the range can only be mapped to by one element in the domain, hence the phrase one-to-one. Another way of saying a function is surjective is saying that it is onto. In other words, every element in the codomain must be mapped to by some element in the domain. Finally, another way of saying a function is bijective is by saying there exists a one-to-one correspondence between the sets. Each element in the domain maps to exactly one element in the codomain, and the mapping of all the elements in the domain gives us the entire codomain.

The definitions of injection and surjection can be used to compare cardinalities of sets. If there are two sets A and B and $|A| \leq |B|$, then there exists an injective function $f : A \rightarrow B$. If $|A| \geq |B|$, then there exists a surjective function $g : A \rightarrow B$. Finally, if $|A| = |B|$, then there exists a bijective function $h : A \rightarrow B$. We have defined cardinality for finite sets to be the number of elements in the sets. Thus, two finite sets have the same cardinality if they have an equal number of elements. Since this definition does not work for infinite sets, we can use the notion of bijection to determine if two sets have the same cardinality.

Definition 2.2.4. *Two sets A and B have the same cardinality ($|A| = |B|$) if and only if there exists a bijective function $f : A \rightarrow B$.*

When two sets have the same cardinality, we say the two sets are **equinumerous** ($A \approx B$). Thus, we can say a set is finite if and only if it is equinumerous to some unique natural number. This is using the convention that each natural number is the set including all the natural numbers before it. Using this convention in an example, we have that $5 = \{1, 2, 3, 4, 5\}$. This can hold for any subset of \mathbb{N} . If a finite set has a cardinality of k , then it forms a bijection with a subset $\{1, 2, 3, \dots, k\}$ of \mathbb{N} . We know that the number must be unique because we can show that a finite set cannot be equinumerous to a proper subset of itself. To show this, we will first start by stating the Pigeonhole Principle.

Theorem 2.2.5 (Pigeonhole Principle). *If n objects are placed into k pigeonholes, where*

$k < n$, then at least one pigeonhole will have more than one object. In other words, no natural number is equinumerous to a proper subset of itself.

Since finite sets are equinumerous to a natural number, we can use the same logic we use in the pigeonhole principle to show that the natural number must be unique.

Theorem 2.2.6. *A finite set A is equinumerous to a unique natural number.*

Proof. Assume for the sake of contradiction that the natural number is not unique, that is $A \approx n$ and $A \approx m$ where $n \neq m$. Then m must be equinumerous to n , $m \approx n$. Then, by trichotomy either $n < m$, $m < n$, or $n = m$. We already stated that $n \neq m$, so that rules out the last possibility. Then either $n < m$ or $m < n$ meaning one must be a proper subset of the other. However, since $m \approx n$ and m and n are finite, this cannot be possible, so we have found a contradiction from our assumption. Therefore, $n = m$ so the natural number is unique. \square

Finally, we can formally define what it means for a set to be finite or infinite.

Definition 2.2.7. *Consider the set X .*

1. *The set X is finite if there exists a bijection between X and some subset of natural numbers $\{1, 2, 3, \dots, k\}$.*
2. *The set X is infinite if it is not finite.*

If a set is not equinumerous to a natural number then the set is infinite. Using this definition we can now prove sets to be either finite or infinite. For example, we can show that the set $S = \{2, 3, 6, 7, 9\}$ is finite by showing that there exists a bijection between S and some subset of \mathbb{N} . We can define a bijective function from S to the set $\{1, 2, 3, 4, 5\} \subseteq \mathbb{N}$:

$$f(x) = \begin{cases} 1 & x = 2 \\ 2 & x = 3 \\ 3 & x = 6 \\ 4 & x = 7 \\ 5 & x = 9 \end{cases}$$

We now can see that this function is injective since each element in S maps to a unique element in the subset of \mathbb{N} . It is also surjective because it maps to all elements in $\{1, 2, 3, 4, 5\}$. Since it is both injective and surjective, we can conclude that the function is bijective. Thus, we have formed a bijection between the set S and some subset of natural numbers. Therefore S must be finite.

This is one example of how we can form a bijective function between two sets to determine that a set is finite. Now, let's look at an example of how we may determine that a set is infinite such as the set of natural numbers.

Theorem 2.2.8. \mathbb{N} is an infinite set.

Proof. We will proceed with this proof by contradiction. Suppose that \mathbb{N} is finite. Then by definition 2.2.7, there exists a bijection between \mathbb{N} and some subset of natural numbers $K = \{1, 2, 3, \dots, k\} \subseteq \mathbb{N}$. We can define a function $f : \mathbb{N} \rightarrow K$ and begin to attempt to show that f is bijective. Let's define f by mapping elements in \mathbb{N} to elements in K : $f(1) = 1, f(2) = 2, f(3) = 3 \dots$. We can proceed like this up to k , so $f(k-1) = k-1, f(k) = k$. However, let's see what happens when we consider the $f(k+1)$. Notice that all elements in S have already been mapped to. Then $f(k+1)$ must map to an element that has been mapped to by another element by the Pigeonhole Principle. So, we have two elements in \mathbb{N} that map to the same element in S . Therefore, f is not injective. Since f is not injective, f is not bijective.

We have found a contradiction to our original assumption. Thus, \mathbb{N} must be an infinite set. \square

We can also draw additional conclusions that can help us determine whether or not a set is finite such as the theorem below that makes use of proper subsets.

Theorem 2.2.9. *If A is not finite and $A \subset B$, then B is not finite.*

Proof. Suppose for the sake of contradiction, B is finite and A is not finite with $A \subset B$. Then B contains n elements for some $n \in \mathbb{N}$. Then if $A \subset B$, then there exists an injection $f : A \rightarrow B$. However, we can first map n elements in A to n unique elements in B . But since A does not have a finite number of elements n , we can try to map $n + 1$ to some other unique element in B . Then, by the pigeonhole principle, it will have to map to an element that has already been mapped to. Thus, f is not injective which poses a contradiction. Therefore, B cannot be finite. \square

When working with cardinality of infinite sets, we will also find definition 2.2.4 to be very useful. Another useful theorem is the Cantor-Bernstein Theorem which will give us an alternate way to determine if there exists a bijection between two sets [9].

Theorem 2.2.10 (Cantor-Bernstein). *If there are injections $f : A \rightarrow B$ and $g : B \rightarrow A$, then there is a bijection $A \rightarrow B$. In other words $|A| \leq |B|$ and $|B| \leq |A|$ implies $|A| = |B|$.*

Proof. Consider the functions $f : A \rightarrow B$ and $g : B \rightarrow A$ to be injective functions. Beginning with any element $b_1 \in B$, we can construct a sequence $b_1, a_1, b_2, a_2, b_3, a_3, \dots$ alternating elements of A and B such that first, there may or may not exist an $a_1 \in A$ where $f(a_1) = b_1$. If a_1 exists then it is unique since f is injective. Then we let a_1 be the inverse of the image of b_1 under f . Similarly, we choose a b_2 such that $g(b_2) = a_1$ where b_2 is a unique element in B . Then we can choose a_2 such that $f(a_2) = b_2$, and the process continues, $f^{-1}(b_1) = a_1$, $g^{-1}(a_1) = b_2$, $f^{-1}(b_2) = a_2$.

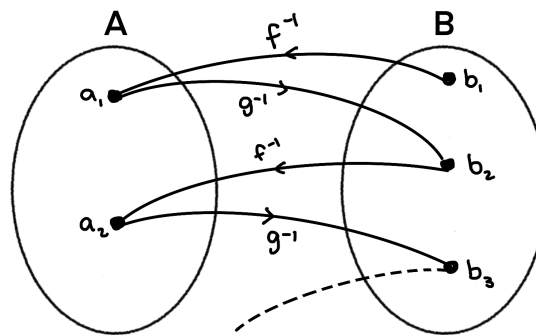


Figure 2.1: Process of constructing the sequence between sets A and B for the Cantor-Bernstein Theorem.

Continuing this process, there are three possible outcomes.

1. The process ends because we reach some $a_n \in A$ where there exists no $b_{n+1} \in B$ such that $g^{-1}(a_n) = b_{n+1}$ (which is possible because g is not necessarily surjective).
2. The process ends because we reach some $b_n \in B$ where there exists no $a_n \in A$ such that $f^{-1}(b_n) = a_n$ (f is also not necessarily surjective).
3. The process continues forever.

The process for each $a \in A$ and $b \in B$ has three different possibilities, so we can first partition the set A into three mutually disjoint subsets that correspond to each possibility.

1. A_A = the set of all $a \in A$ such that the process ends with a a_n
2. A_B = the set of all $a \in A$ such that the process ends with a b_n

3. $A_\infty =$ all $a \in A$ such that the process never ends

We can do the same thing with B

1. $B_A =$ the set of all $b \in B$ such that the process ends with a a_n
2. $B_B =$ the set of all $b \in B$ such that the process ends with a b_n
3. $B_\infty =$ all $b \in B$ such that the process never ends

We want to show that there exist bijections $A_A \rightarrow B_A$, $A_B \rightarrow B_B$, and $A_\infty \rightarrow B_\infty$ to show that there exist bijections between sets A and B .

Let's start by proving $A_A \rightarrow B_A$. We know f is injective, so we want to show the $f(a)$ is in B_A . Suppose $a_n \in A_A$, that is the process applied to a ends in A . Now consider the same process applied to $f(a_n)$ where we know $f(a_n) \in B$. Then the next element in the sequence would be a_n since that would be the preimage of $f(a)$. Then the process would once again be applied to a_n which we had already established ends in A . Thus $f(a) \in B_A$.

Next, we want to show that $f : A_A \rightarrow B_A$ is surjective, that is for all $b \in B_A$, there exists an $a \in A_A$ such that $f(a) = b$. Suppose that $b_n \in B_A$, so the process applied to b ends in A . This must mean that $f^{-1}(b_n)$ exists in A_A because if that wasn't true then b would end in B rather than A . Continuing the process $f^{-1}(b_n) = a_n$. Note that the continuation of the process applied to b must be the same process applied to a , therefore it ends in A . Thus $a_n \in A_A$. We have shown that $f : A_A \rightarrow B_A$ is bijective. The same argument can be followed to show $A_B \rightarrow B_B$ using g^{-1} instead of f^{-1} . The first part of the proof remains exactly the same but proving surjectivity is slightly different. We want to show that for all $b \in B_B$ there exists an $a \in A_B$ such that $g^{-1}(a) = b$. Starting with an arbitrary $b_n \in B_B$, we know that $g(b_n) \in A_B$ exists because if it didn't then the process would end in A rather than B . And $g(b_n) = a_{n-1}$. Then once we continue the process we end up with $g^{-1}(a_{n-1}) = b_n$ which we know ends in B . Thus, $g^{-1}(b_n) \in A_B$.

Finally, we want to show that $f : A_\infty \rightarrow A_\infty$ is bijective. We already know that f is injective by the way it is defined. To show that f is surjective, we want to show that for every $b \in B_\infty$, there exists an a such that $f(a) = b$. This must be true since the sequence never ends. Then this would be the same as starting with $b \in B$, $g(b) = a$. Thus we have shown there is a bijection $A \rightarrow B$.

Thus, we can form a bijective function between each pair of subsets. We can define the bijection $F : A \rightarrow B$ as,

$$F(x) = \begin{cases} f(x) & x \in A_A \\ g^{-1}(x) & x \in A_B \\ f(x) & x \in A_\infty \end{cases}$$

□

The Cantor-Bernstein Theorem holds for both finite and infinite sets. Although it may seem counterintuitive, not all infinite sets have the same cardinality. In this chapter, we will define infinite sets to be either countable or uncountable.

Definition 2.2.11. A set is **countable** if there exists a bijection between that set and \mathbb{N} . If there does not exist such a bijection then the set is **uncountable**.

The cardinality of uncountable sets is larger than countable sets which we will confirm later in this chapter. Although we will not explicitly show it, we could confirm that the integers \mathbb{Z} are countable using a bijective piecewise function $f : \mathbb{N} \rightarrow \mathbb{Z}$ where $f(n) = \frac{n}{2}$ when n is even, and $f(n) = -\frac{n-1}{2}$ when n is odd. Using definition 2.2.11 we can also show that $\mathbb{N} \times \mathbb{N}$ is countable.

Theorem 2.2.12. $\mathbb{N} \times \mathbb{N}$ is countable.

Proof. To show that $\mathbb{N} \times \mathbb{N}$ is countable we want to show that there exists a bijection between $\mathbb{N} \times \mathbb{N}$ and \mathbb{N} . Let's consider the function $f : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ defined

by $f(m, n) = 2^{m-1}(2n - 1)$. To prove that f is bijective, we want to prove that it is both injective and surjective. First, let's prove that f is injective. We want to show that if $f(m, n) = f(k, p)$, then $(m, n) = (k, p)$ or $m = k$ and $n = p$. Let's suppose $f(m, n) = f(k, p)$. Then, $2^{m-1}(2n - 1) = 2^{k-1}(2p - 1)$. By the uniqueness of prime factorization, if two numbers are equal then they must have the same prime factorization or the same amount of each prime factor. Since both $2n - 1$ and $2p - 1$ are odd, there are no 2's in their prime factorization. However, since both numbers $2^{m-1}(2n - 1)$ and $2^{k-1}(2p - 1)$ are equal, and they must have an equal number of 2's in their prime factorization, then $2^{m-1} = 2^{k-1}$. Multiplying both sides by two gives $2^m = 2^k$. So $m = k$, thus f is injective.

Next, we want to show that f is surjective. Consider some $x \in \mathbb{N}$. Since x is a natural number it can be written as a product of some power of 2 multiplied by an odd number since it can either be even or odd. Thus, $x = 2^y(2z + 1)$ for some $y, z \in \mathbb{N} \cup \{0\}$. We want to show that there exists some $(m, n) \in \mathbb{N} \times \mathbb{N}$ such that $f(m, n) = x$. Suppose $m = y + 1$ and $n = z + 1$. Then $x = 2^{m-1}(2n - 1)$ and from how we defined our function $f(m, n) = 2^{m-1}(2n - 1)$. Thus, there exists an $(m, n) = (y + 1, z + 1)$ such $f(m, n) = x$. Therefore, f is surjective.

Since f is both injective and surjective, then f is bijective. Thus, there exists a bijection between $\mathbb{N} \times \mathbb{N}$ and \mathbb{N} . □

We can use the result above and the Cantor-Bernstein theorem to also prove that the rational numbers are countable.

Theorem 2.2.13. \mathbb{Q} is countably infinite.

Proof. Knowing that $\mathbb{N} \times \mathbb{N}$ is countably infinite since it is bijective with \mathbb{N} , we want to define a function $f : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{Q}$ as $f(m, n) = \frac{m}{n}$ where $\frac{m}{n}$ is fully reduced ($\gcd(m, n) = 1$) and show that f is bijective. First, let's show that f is injective.

Suppose $f(a, b) = f(c, d)$. Then $\frac{a}{b} = \frac{c}{d}$. Since both fractions are fully reduced, then $a = c$ and $b = d$. Thus, f is injective.

Next let's show that the function $g : \mathbb{Q} \rightarrow \mathbb{N} \times \mathbb{N}$ is also injective where g is defined as follows: $g(\frac{m}{n}) = (m, n)$ where $\frac{m}{n}$ is fully reduced. Suppose $f(\frac{a}{b}) = f(\frac{c}{d})$. Then $(a, b) = (c, d)$ or $a = c$ and $b = d$. Thus g is injective.

Since both f and g are injective, by the Cantor-Bernstein Theorem, we can state that there is a bijection between $\mathbb{N} \times \mathbb{N}$ and \mathbb{Q} . \square

The image below shows a visual argument for why \mathbb{Q} is countably infinite by listing them out in the following way.

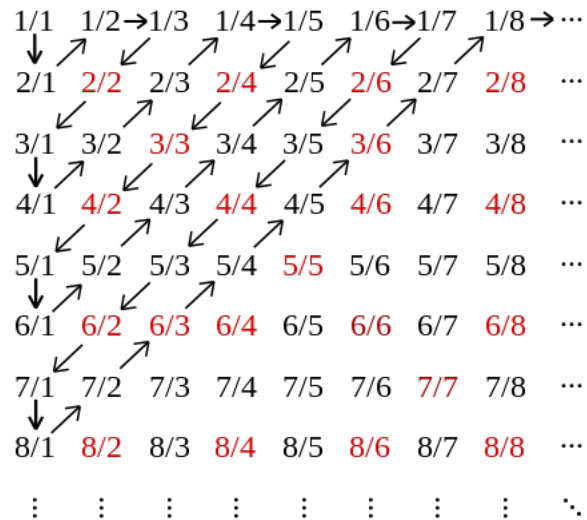


Figure 2.2: Visual process of listing the rational numbers [4].

Now we must understand how we can show that some sets are uncountable such as the set $(0, 1)$. We observe that a set is countable if and only if its elements can be arranged in a list.

Theorem 2.2.14. *The interval $(0, 1)$ is uncountable.*

Proof. This proof follows Cantor's diagonalization argument. Suppose for the sake

of contradiction that the interval $(0, 1)$ is countable, that is we can list out all the numbers in the interval,

$$x_1 = 0.a_1a_2a_3a_4\dots$$

$$x_2 = 0.b_1b_2b_3b_4\dots$$

$$x_3 = 0.c_1c_2c_3c_4\dots$$

$$x_4 = 0.d_1d_2d_3d_4\dots$$

$$\vdots$$

Let's consider the digits on the diagonal, $a_1, b_2, c_3, d_4, \dots$. Now consider an element $y \in (0, 1)$ where $y = 0.y_1y_2y_3y_4\dots$ where $y_1 \neq a_1, y_2 \neq b_2, y_3 \neq c_3, y_4 \neq d_4$, etc. Since each element in our list has at least one digit that is not equal to the corresponding digit in y , we have found an element that is not in our list. Therefore, the interval $(0, 1)$ is uncountable. \square

This proof also shows that $|(0, 1)| > |\mathbb{N}|$ since there is an injection from \mathbb{N} to $(0, 1)$, but no surjection. For this reason, we can conclude that uncountable sets are greater than countable sets. We can also show that the intervals $(0, 1)$, $[0, 1]$, and $[0, 1)$ are in bijection with each other and are all uncountable.

Theorem 2.2.15. *There exists a bijection between $[0, 1)$ and $(0, 1)$. That is, $[0, 1)$ is uncountable.*

Proof. By the Cantor-Bernstein Theorem, we can show a bijection by showing there exists injective functions $f : (0, 1) \rightarrow [0, 1)$ and $g : [0, 1) \rightarrow (0, 1)$. Let's consider the function $f(x) = x$. We can show this is injective by considering $f(a) = f(b) \in (0, 1)$. Then $a = b$ by how we defined f and $a, b \in [0, 1)$. Next, let's consider the function $g : [0, 1) \rightarrow (0, 1)$ and define it as a piecewise function.

$$g(x) = \begin{cases} \frac{1}{2} & x = 0 \\ \frac{x}{2} & x \neq 0 \end{cases}$$

We can show that g is injective by supposing $g(a) = g(b)$. If $g(a) = \frac{1}{2} = g(b)$ we know that $a = b = 0$. Or we have that $\frac{a}{2} = \frac{b}{2}$, and solving for that we get $a = b$. Thus, g is injective. Since we have shown that there exists injective functions f from $(0, 1)$ to $[0, 1)$ and g from $[0, 1)$ to $(0, 1)$, we have proven that there is a bijection between $[0, 1)$ and $(0, 1)$. \square

We would use a similar method to show that $|[0, 1]| = |(0, 1)|$.

We can combine some of the methods we have used to discuss a very fascinating set called the Cantor set and prove that it is uncountable. The Cantor set is the set constructed by starting with all the real numbers on the number line from 0 to 1. Then the middle third is removed and you are left with two lines each of length $1/3$. Then you remove the middle third from each of those lines and you are left with four lines of length $1/9$. This process continues forever and results in the Cantor set shown below.



Figure 2.3: Construction of the Cantor Set [17]

To prove that the Cantor set is uncountable we must first understand what elements are in the Cantor set. This part of the proof uses base-3 numbers. Although we will not encounter them throughout the rest of this project, they are very significant in this proof. To briefly summarize base-3 notation is a way of representing decimal numbers (base-10 numbers) using 0s, 1s, and 2s. It is similar to how binary numbers work when adding powers of 2, but instead, we are adding

powers of three or powers of $\frac{1}{3}$ like $a \times \frac{1}{3^1} + b \times \frac{1}{3^2} + c \times \frac{1}{3^3} + \dots$ where a, b, c can be 0, 1, or 2 [16].

Theorem 2.2.16. *For any element $x \in C$ where C is the Cantor set, the base 3 representation of x contains no 1.*

Proof. For each iteration of the Cantor set, we remove the middle third. For the first iteration, we are removing the interval $(\frac{1}{3}, \frac{2}{3})$. And we know that those numbers must be represented as $\frac{1}{3^1} + \frac{a_2}{3^2} + \frac{a_3}{3^3} + \dots$. Thus, we are removing the numbers that can only be represented as $0.1a_2a_3a_4\dots$ in base 3 leaving numbers that can be represented as $0.0a_2a_3a_4\dots$ and $0.2a_2a_3a_4\dots$. Next, we are removing the middle third in the second iteration. Thus we are removing any number represented by $0.01a_3a_4\dots$ and $0.21a_3a_4\dots$ since we are removing the intervals $(\frac{0}{3^1} + \frac{1}{3^2}, \frac{0}{3^1} + \frac{2}{3^2})$ and $(\frac{2}{3^1} + \frac{1}{3^2}, \frac{2}{3^1} + \frac{2}{3^2})$. Therefore, every element whose second ternary digit is 1 is removed. This leaves elements with ternary representation as $0.00a_3a_4\dots, 0.02a_3a_4\dots, 0.20a_3a_4\dots, 0.22a_3a_4\dots$ in the Cantor set. Then the third step would remove all elements containing a 1 in the third ternary digit, and the fourth step would remove all elements containing a 1 in the fourth ternary digit. At the n th step, we would have removed all elements containing one from the first to the n th ternary digit. Thus, the base 3 representation of any element $x \in C$ contains no 1. \square

Now that we know which elements are in the Cantor set, we can now prove that it is uncountable [16].

Theorem 2.2.17. *The Cantor set is uncountable.*

Proof. We want to show that the Cantor set is uncountable, that is it forms a bijection with the interval $[0,1]$. This proof follows Cantor's diagonalization proof. We have proven above that the Cantor set is made up of any element that contains only 0's and 2's in base 3 representation. We can treat these elements as sequences of 0's and 2's. We can prove this the same way we proved the interval $[0,1]$ was uncountable.

Assume for the sake of contradiction that the set of numbers on the interval $[0,1]$ containing 0 and 2 is countable. Then we can list them as,

$$x_1 = 0.00000\dots$$

$$x_2 = 0.22222\dots$$

$$x_3 = 0.02020\dots$$

$$x_4 = 0.20202\dots$$

$$\vdots$$

Then we can take all the digits on the diagonal to get a number 0.0200. Then we can form a new number y such that none of the digits in y equal the digits on the diagonal $y = 0.2022$. Since each element in the list contains at least one different digit than y , we have found an element that is not in our set. Thus, we cannot list all the possible numbers containing 0 and 2 on the interval $[0,1]$. Therefore, the Cantor set contains uncountably infinite elements. In other words, it is uncountable. \square

For countable sets, we determined they were countable by seeing if there exists a bijection between a given set and a known countable set, such as \mathbb{N} . We can do a similar thing to prove that certain sets are uncountable. If we can prove that there exists a bijection between a certain set A and a known uncountable set such as $(0, 1)$ we can say that set A is also uncountable and $|A| = |(0, 1)|$. We can do this to show that the set of real numbers \mathbb{R} is uncountable.

Theorem 2.2.18. \mathbb{R} is uncountable.

Proof. To show that \mathbb{R} is uncountable, we can show that there exists a bijection between \mathbb{R} and $(0,1)$ since we have just proven $(0,1)$ is uncountable. Let's consider

the function $f : (0, 1) \rightarrow \mathbb{R}$ defined by $f(x) = \tan\left(\pi x - \frac{\pi}{2}\right)$. The graph of this function is below.

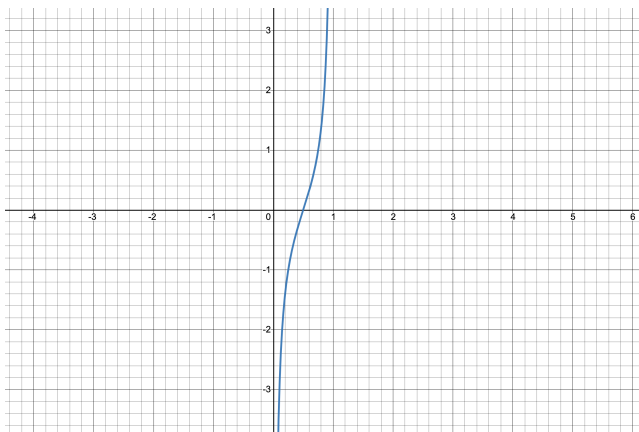


Figure 2.4: Graph of $\tan\left(\pi x - \frac{\pi}{2}\right)$ using Desmos Graphing Calculator

From the graph, we can see that the function should be both injective and surjective, but we are still going to confirm that fact. First, let's show that this function is injective. Suppose we have $a, b \in [0, 1]$ and $f(a) = f(b)$. Then $\tan\left(\pi a - \frac{\pi}{2}\right) = \tan\left(\pi b - \frac{\pi}{2}\right)$. We know in general the tangent function is not one-to-one unless we restrict it to a range, and in this case, our range is $0 < x < 1$ which means $-\frac{\pi}{2} < \pi x - \frac{\pi}{2} < \frac{\pi}{2}$. Since we are taking $\tan\left(\pi x - \frac{\pi}{2}\right)$, that range makes it so that no two numbers can map to the same thing. With that clarification, we can proceed to solve.

$$\begin{aligned}\tan^{-1} \tan\left(\pi a - \frac{\pi}{2}\right) &= \tan^{-1} \tan\left(\pi b - \frac{\pi}{2}\right) \\ \pi a - \frac{\pi}{2} &= \pi b - \frac{\pi}{2}\end{aligned}$$

Solving this the rest of the way we find that $a = b$. Thus, f is injective. Next, we want to show that f is surjective. Let's suppose $y \in \mathbb{R}$. We want to find a $w \in [0, 1]$ such that $f(w) = y$, so $\tan\left(\pi w - \frac{\pi}{2}\right) = y$. Solving for this we get $w = \frac{\tan^{-1} y}{\pi} + \frac{1}{2}$. Note that the range for the inverse tangent function is $(0, 1)$. Therefore, $\frac{1}{2} \leq w \leq \frac{3}{4}$ which

is within $(0, 1)$. Since there exists a $w \in (0, 1)$ such that $f(w) = y$, then f is also surjective. Since f is both injective and surjective, we can conclude that f is bijective. Thus, $|(0, 1)| = |\mathbb{R}|$ implying \mathbb{R} is uncountable. \square

Next, let's look at a generalization that relates two sets in bijection with their power sets. The theorem below holds for both finite and infinite sets.

Theorem 2.2.19. *For two sets A and B , if there exists a bijection $g : A \rightarrow B$, then there exists a bijection $f : \mathcal{P}(A) \rightarrow \mathcal{P}(B)$. In other words $|A| = |B|$ implies $|\mathcal{P}(A)| = |\mathcal{P}(B)|$.*

Proof. We want to show there exists a function f such that f forms a bijection between $\mathcal{P}(A)$ and $\mathcal{P}(B)$. Let's consider the function $f(X) = \{g(x) : x \in X\}$ for all $X \subseteq A$. First, let's show that f is injective. Let $S, T \subseteq A$ and $f(S) = f(T)$. Then, $\{g(x) : x \in S\} = \{g(x) : x \in T\}$. If we take some $s \in S$, then we have $g(s) \in f(S)$. And since $f(S) = f(T)$, then $g(s) \in f(T)$. Then there must exist a $t \in T$ such that $g(s) = g(t)$. And since g is a bijection $s = t$. So $s \in S$ implies $s \in T$. Thus, $S \subseteq T$. We can use the same logic to prove that $T \subseteq S$. Thus, $S = T$ as desired.

Next, let's show that f is surjective. We want to show that for all $Y \subseteq B$, there exists some $X \subseteq A$ such that $f(X) = Y$. Since g is a bijection, g^{-1} exists so we can define it as $X = \{g^{-1}(x) : x \in Y\}$. Let $y \in Y$ which means $g^{-1}(y) \in X$. Then $g(g^{-1}(y)) \in f(X)$, so $y \in f(X)$. We have that $y \in Y$ implies $y \in f(X)$, so $Y \subseteq f(X)$. Now we need to show that $f(X) \subseteq Y$. Let's suppose that $t \in f(X)$. Then there exists an $x \in X$ such that $f(x) = t$. Since, $x \in X$, there must be some $y \in Y$ such that $f^{-1}(y) = x$. Then substituting this back into $f(x) = t$, we get that $f(f^{-1}(y)) = t$. Then, $y = t$ and since $t \in f(X)$, $y \in f(X)$. Thus $f(X) \subseteq Y$. We have that $Y \subseteq f(X)$ and $f(X) \subseteq Y$, so $f(X) = Y$ as desired. Since we have shown that f is both injective and surjective, f is bijective. Therefore, $|\mathcal{P}(A)| = |\mathcal{P}(B)|$. \square

While on the topic of power sets, we should discuss a significant theorem, Cantor's Theorem, We have already established that there are at least two different

types of infinity, countable and uncountable. However, with Cantor's theorem, we can conclude that there are in fact infinitely many different sizes of infinity [6].

Theorem 2.2.20 (Cantor's Theorem). *Given any set A , the set is always going to be strictly less than its power set: $|A| < |\mathcal{P}(A)|$.*

Proof. To show that $|A| < |\mathcal{P}(A)|$, we want to show that the map $f : A \rightarrow \mathcal{P}(A)$ is not surjective. Suppose for the sake of contradiction that f is surjective. Consider the set B where $B \subseteq A$ such that elements in B are not in the image of f . So $B = \{a \in A \mid a \notin f(a)\} \subseteq A$. Since we assumed f is surjective, then there exists an $x \in A$ such that $f(x) = B$. It is either the case that $x \in B$ or $x \notin B$. If $x \in B$, then $x \notin f(x) = B$, so $x \notin B$ which is a contradiction. If $x \notin B$, then $x \in f(x) = B$, so $x \in B$ which is a contradiction for the second case. In both cases, we reach a contradiction, so our assumption is not true. Thus, f is not surjective, so $|A| < |\mathcal{P}(A)|$. \square

Using Cantor's theorem, we can conclude that for any A , the following result hold:

$$|A| < |\mathcal{P}(A)| < |\mathcal{P}(\mathcal{P}(A))| < \dots$$

Cantor's theorem will be prevalent throughout much of this thesis, so it is important that we understand that this holds for both finite and infinite sets. Cantor's theorem establishes that $|\mathbb{N}| < |\mathcal{P}(\mathbb{N})|$. And we previously established $|\mathbb{N}| < |\mathbb{R}|$. Now we are ready to prove the following result [9].

Theorem 2.2.21. $|\mathcal{P}(\mathbb{N})| = |\mathbb{R}|$.

Proof. We want to show that $|\mathcal{P}(\mathbb{N})| = |[0, 1]|$ knowing the fact that $|\mathbb{R}| = |[0, 1]|$. To show this, we must show there exists injections $\mathcal{P}(\mathbb{N}) \rightarrow [0, 1]$ and $[0, 1] \rightarrow \mathcal{P}(\mathbb{N})$.

First, let's show that there exists an injection $f : \mathcal{P}(\mathbb{N}) \rightarrow [0, 1]$. Let $X \subseteq \mathbb{N}$. Then we can construct a decimal expansion $0.a_0a_1a_2\dots$ where

$$a_i = \begin{cases} 0 & i \notin X \\ 1 & i \in X \end{cases}$$

Suppose $f(X) = f(Y) = 0.a_0a_1a_2\dots$. The $i \in X$ if and only if $a_i = 1$ which also means $i \in Y$, so $X = Y$ since both subsets map to the same decimal expansion. So $a_i = 1$ for $f(Y)$. Thus, f is injective.

Next, we want to show that there exists an injection $g : [0, 1) \rightarrow \mathcal{P}(\mathbb{N})$. We must first note that any element of $[0, 1)$ can be expressed uniquely as a decimal in the form,

$$0.n_0n_1n_2\dots, \quad 0 \leq n_k \leq 9$$

And we take $g(0) = \emptyset$. For some given $x \in [0, 1)$ we can write $x = 0.n_0n_1n_2\dots$ as defined above and define $g(x) = \{n_k10^k : k \in \mathbb{N}\}$. Suppose $g(x) = g(y)$ where $x = 0.n_1n_2n_3\dots$ and $y = 0.m_1m_2m_3\dots$. Given some $k \in \mathbb{N}$ we know $m_k10^k \in g(x)$, thus $m_k10^k \in g(y)$ since $g(x) = g(y)$. Then $m_k10^k = m_i10^i$ for some $i \in \mathbb{N}$. Since m_k and n_i are single digit numbers, then $i = k$. So $m_k = n_k$. Thus, $x = y$, so $g : [0, 1) \rightarrow \mathcal{P}(\mathbb{N})$ is injective. Since we have shown injections $f : \mathcal{P}(\mathbb{N}) \rightarrow [0, 1)$ and $g : [0, 1) \rightarrow \mathcal{P}(\mathbb{N})$, we can conclude by the Cantor-Bernstein Theorem that there exists a bijection $\mathcal{P}(\mathbb{N}) \rightarrow [0, 1)$. Thus, it must also follow that there there exists a bijection $\mathcal{P}(\mathbb{N}) \rightarrow \mathbb{R}$ since $|\mathbb{R}| = |[0, 1)|$. Therefore, $|\mathcal{P}(\mathbb{N})| = |\mathbb{R}|$. \square

Before we begin our next chapter which discusses cardinal numbers and ordinal numbers, let's go back to one of the infinite sets that we have become familiar with: the natural numbers \mathbb{N} . We have established that the set of natural numbers is the set of integers greater than or equal to 1, that is $\mathbb{N} = \{1, 2, 3, \dots\}$. Although this is true, we want to redefine what it means for a number to be a natural number in a set theoretic way. There are multiple ways that we could define natural numbers, but the way we will use is from mathematician John von Neuman as it is the most

advantageous and is the standard used by most set theorists. Von Neumann's construction defines each natural number as the set of natural numbers that precedes it [6]. We begin with the convention that 1 is the empty set. With this convention, we can continue to define the natural numbers as follows.

$$1 = \{\} = \emptyset$$

$$2 = \{1\} = \{\emptyset\}$$

$$3 = \{1, 2\} = \{\emptyset, \{\emptyset\}\}$$

$$4 = \{1, 2, 3\} = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$$

$$\vdots$$

We see some interesting properties result from this construction. First notice that each number, which we are now treating as sets, is an element in its successor.

$$1 \in 2 \in 3 \in 4 \dots$$

We could also write the same thing using the sets we defined corresponding to each natural number which may make the result more obvious.

$$\emptyset \in \{\emptyset\} \in \{\emptyset, \{\emptyset\}\} \in \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} \in \dots$$

Another interesting result is that each set is also a subset of the succeeding set.

$$1 \subseteq 2 \subseteq 3 \subseteq \dots$$

$$\emptyset \subseteq \{\emptyset\} \subseteq \{\emptyset, \{\emptyset\}\} \subseteq \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} \subseteq \dots$$

After seeing how we have defined the first few natural numbers, we have a

better understanding, so that we can define the entire set of natural numbers. We have used the word successor several times throughout this discussion which we will formally define in the next chapter. Simply speaking, the successor here is the set that follows immediately after the set before it. With this definition of natural numbers, we are now ready to introduce the notion of ordinal numbers and cardinal numbers which is a topic directly related to infinite sets.

CHAPTER 3

ORDINAL NUMBERS AND CARDINAL NUMBERS

We all have an intuitive sense of what infinity is, and from our initial discussion of infinite sets, it may have become clearer. We have learned that there are different sizes of infinity and have looked in-depth at the two sizes: countable and uncountable. However, as Cantor's theorem had introduced there are an infinite number of sizes of infinity. There are more things we can say about this idea which we will elaborate on through our discussion of ordinal and cardinal numbers. Not only, will this chapter be helpful to that extent, but ordinal and cardinal numbers are directly used in our concluding chapter about the continuum hypothesis, so it is important to keep these definitions in mind. We will begin with discussing something we have made use of, but have yet to formally define. This is the notion of order which will lead us to our discussion of ordinal numbers.

3.1 ORDER AND ORDINAL NUMBERS

Our first step towards understanding ordinal numbers is to understand the notion of ordering and order relations.

Definition 3.1.1. A binary relation on a set A is an **order** relation if it is

1. reflexive: aRa holds for every $a \in A$
2. anti-symmetric: aRb and bRa implies $a = b$ for all $a, b \in A$, if $x \neq y$ and $x < y$, then $y < x$ does not hold
3. transitive: aRb and bRc implies aRc for all $a, b, c \in A$

With this definition, we can now simply define what it means for a set to be ordered.

Definition 3.1.2. An **ordered set** X is a set such that there exists an order relation R that holds in X .

To be more specific, an ordered set can either be partially ordered or totally ordered. These two types of sets require additional definitions of what makes a relation a partial ordering or a total ordering. We will start with partial ordering.

Definition 3.1.3. A **strict partial ordering** is a relation R that satisfies the following two conditions:

1. R is transitive
2. R is irreflexive

Some partial ordering examples are $<$, or \subset . Notice because both are transitive and irreflexive. Both hold for transitivity,

$$a < b < c \Rightarrow a < c$$

$$a \subset b \subset c \Rightarrow a \subset c$$

The relations are also irreflexive, that is the statements $x < x$ and $x \subset x$ are impossible. Next, we must define a total order relation.

Definition 3.1.4. An order relation R is a **total order** relation on a set S if and only if for every pair $x, y \in S$ either xRy or yRx is true.

Let's consider two examples to clarify what we mean by total order. Suppose we have the set $S = \{1, 2, 4, 6, 8, 10, 12, 14, 16\}$ where the order operation on S is an element divides another element. In fig. 3.2, we have an ordering diagram to see how each element is related through our division relation. Each arrow in the figure represents the existence of the relation between the elements. Note, that the relation is transitive. So although there is no direct arrow between some related elements because we are accounting for transitivity. For example, from 1 to 8 we can follow the arrows from 1 to 2 to 4 to 8. Thus, by transitivity, 1 and 8 are related although there is not one direct arrow relating the two.

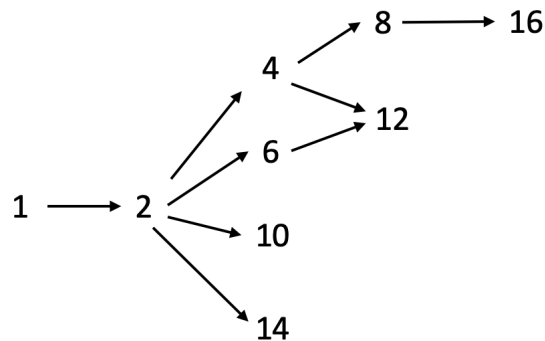


Figure 3.1: Order diagram of division relations on the set S .

However, for a relation to be a total order on a set, we need either xRy or yRx for *all* elements in S . If we look at the elements 4 and 6, neither 4 divides 6 nor 6 divides 4. Thus, our condition for the relation to be a total order fails, and this is the case for several other pairs of elements in S . We can alter this set to create $S' = \{1, 2, 4, 8, 16\}$ with the same relation of elements divide other elements. Now our relation holds, because for any pair of elements $x, y \in S'$, we have that either xRy or yRx hold. Keeping transitivity in mind, we can more clearly see how this is true in fig. 3.2.

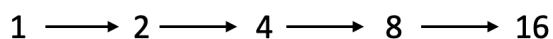


Figure 3.2: Order diagram of division relations on the set S' .

More common examples of total ordering are the relations of magnitude (less than or greater than) on the natural numbers, integers, or real numbers.

We can also use orders to compare two different ordered sets. Similar to how we compared cardinalities of sets, we can compare two sets based on their order relation through order isomorphism defined below.

Definition 3.1.5. *Given sets X and Y , let X be ordered by a relation R and Y be ordered by a relation S . The function $f : X \rightarrow Y$ is order-preserving if for all $n, m \in X$, $(n, m) \in R$ if and only if $(f(n), f(m)) \in S$. There exists an order isomorphism if f is also bijective. Two sets are **order isomorphic** if there is an order isomorphism between them.*

Let us look at an example of two sets that are order isomorphic and understand why using the definition above [9].

Theorem 3.1.6. *The sets $\mathcal{P}(\mathbb{N})$ and $\mathcal{P}(\mathbb{Z})$ both ordered by the subset relation \subseteq are order isomorphic.*

Proof. First, let us show that there exists a bijection between $\mathcal{P}(\mathbb{N})$ and $\mathcal{P}(\mathbb{Z})$. We know that there exists a bijection $g : \mathbb{N} \rightarrow \mathbb{Z}$ as we had established in chapter 2. Both sets are countably infinite and have the same cardinality. We also know from theorem 2.2.19 that if two sets have the same cardinality, then their power sets have the same cardinality. Then if both sets have the same cardinality there must exist a bijection $f : \mathcal{P}(\mathbb{N}) \rightarrow \mathcal{P}(\mathbb{Z})$. Specifically, this function is defined as $f(A) = \{g(x) : x \in A\}$ for all $A \subseteq \mathbb{N}$ as we had also discussed shown chapter 2. Next, we want to show that the two sets are order-preserving. We want to show $A \subseteq B$ if and only if $f(A) \subseteq f(B)$ for all $A, B \in \mathbb{N}$. First, let's suppose $A \subseteq B$ and let $a \in f(A)$.

This means that $a = g(x)$ for some $x \in A$. It also must hold that $x \in B$ since $A \subseteq B$. Since x is also in B , then a is also an element of $f(B)$. So we have that $a \in f(A)$ implies $a \in f(B)$. Thus, $f(A) \subseteq f(B)$ as desired. Now we will show the other direction supposing that $f(A) \subseteq f(B)$ and letting $x \in A$. If $x \in A$ then $g(x) \in f(A)$ meaning $g(x) \in f(B)$ since $f(A) \subseteq f(B)$ as desired. Now we will show the other direction supposing that $f(A) \subseteq f(B)$. Then $g(x) = g(y)$ for some $y \in B$. Finally, since g is a bijective function it must follow that $x = y$ for $x \in A$ and $y \in B$. Thus, $x \in B$. We have that $x \in A$ implies $x \in B$; therefore, $A \subseteq B$. We have shown that the two sets are in bijection with each other and that they are order-preserving. Thus, we can conclude that the two sets are order isomorphic. \square

An important property of order isomorphisms is that they also preserve least and greatest elements. If we had two sets A ordered by R and B ordered by S , and a bijection $f : A \rightarrow B$, then if $x \in A$ is the least element in A , $f(x)$ must be the least element in B in order for the two sets to be order isomorphic. For example, \mathbb{Z} and \mathbb{N} both ordered by the less than relation are not isomorphic because there exists a least element $1 \in \mathbb{N}$, but there does not exist a least element in \mathbb{Z} . The last two definitions we need to cover before our discussion of ordinal numbers are related to well-ordered sets. First, let's understand what makes a set well-ordered.

Definition 3.1.7. *A total ordered set A is called **well-ordered** if every non-empty subset of A has a least element.*

An example of a well-ordered set would be the natural numbers ordered by magnitude. Since there is a least element $1 \in \mathbb{N}$, then any subset will have a least element $n \geq 1$. A set that isn't well-ordered would be the set of integers \mathbb{Z} . For example, let's take the subset $\{\dots - 4, -3, -2, -1\} \subseteq \mathbb{Z}$ to be the set of all negative integers. This set does not contain a least element, and since we found a subset of \mathbb{Z} that does not have a least element, \mathbb{Z} is not well-ordered. However, this example

brings up another important principle in mathematics called the well-ordering principle [9].

Theorem 3.1.8 (Well-Ordering Principle). *Given any set X , there exists a binary relation R on X that makes X a well-ordered set.*

This principle will come up later during our discussion of Zermelo-Fraenkel set theory and the axiom of choice. With the idea of well-ordered sets established, we can look at how to find an initial segment of a well-ordered set. This concept of initial segments will be important when later discussing properties of ordinal numbers [9].

Definition 3.1.9. *Suppose X is a well-ordered set, ordered by a relation R . An **initial segment** of X is found by choosing an element $a \in X$. The initial segment is the subset $\{y \in X : yRa \text{ and } y \neq a\}$. We would denote this set X_a .*

Suppose we have the well-ordered set $X = \{1, 2, 3, 4, 5, 6, 7, 8\}$ where our relation R is magnitude (less than). Then we can form the initial segment $X_5 = \{1, 2, 3, 4\}$ for example. This example is represented in fig. 3.3 where the blue arrows are the relation R .

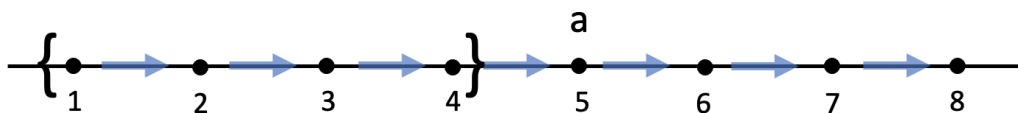


Figure 3.3: How to determine the initial segment of a set.

The element $a \in X$ that we choose will not be in X_a , so we can say that for any $a \in X$, a will be the least element in $X \setminus X_a$. Another important remark to state is that if we have $a, b \in X$, then $X_a \subseteq X_b$.

We are now ready to discuss ordinal numbers. Although we have called it an ordinal *number*, it is actually a set, not a number. Often ordinal numbers are just referred to as ordinals. We will use these two terms interchangeably.

Definition 3.1.10. *An ordinal number is a well-ordered set in which each element is the set of all its predecessors, that is a set A well-ordered by a relation R is an ordinal number if for every $x \in A$, $x = \{y \in A : yRx \text{ and } y \neq x\}$.*

Note the similarities between this definition and definition 3.1.9. The distinction here is that we are not creating a subset from a well-ordered set by choosing some element a as we saw with initial segments. In this case, we are concerned with the whole set, not subsets. We have actually seen an application of this definition already. Recall the way we defined the natural numbers. Following von Neumann's definition, he states that a natural number is the set of natural numbers that precede it beginning with the convention that $1 = \emptyset$. Thus, each natural number is the set of all its predecessors. Then by definition 3.1.10 every natural number is an ordinal number. We call these finite ordinals.

Definition 3.1.11. *For a given ordinal α , the next biggest ordinal is the set $\alpha \cup \{\alpha\}$ and is called the **successor** of α denoted α^+ .*

For the finite ordinals we just discussed, we can rewrite them in terms of successors as shown below.

$$1 = \emptyset, \quad 2 = \emptyset^+, \quad 3 = \emptyset^{++}, \quad 4 = \emptyset^{+++}, \dots$$

There are finite ordinals and limit ordinals which we will discuss shortly. Before we introduce limit ordinals, let's discuss some important properties of ordinal numbers to further our understanding of what we have already discussed, and to later help us make sense of limit ordinals. These properties and their proofs follow from [9].

Theorem 3.1.12. *Given any two well-ordered sets, either the two sets are isomorphic to each other or one is isomorphic to an initial segment of the other.*

We use the term isomorphic here the same way we had defined it in definition 3.1.5. We will omit the proof of this theorem, but this theorem is important to include as it will help us prove some other significant properties of ordinals listed below.

Theorem 3.1.13. *The following about ordinals holds:*

1. *An initial segment of an ordinal is an ordinal, that is every element of an ordinal is an ordinal.*
2. *The order relation on an ordinal is always \subseteq .*
3. *If two ordinals are isomorphic, then they are equal.*
4. *If α and β are ordinals, then either $\alpha = \beta$, $\alpha \in \beta$, $\beta \in \alpha$.*

Proof. We will omit the proof for parts 1 and 3, and establish that they are true.

Proof of 2. Suppose we have an ordinal α with the order relation R on α . We want to show that for any $x, y \in \alpha$, xRy if and only if $x \subseteq y$. Then, from what we stated earlier about initial segments, we have that xRy if and only if $\alpha_x \subseteq \alpha_y$. But since α is an ordinal number, $x = \alpha_x$ and $y = \alpha_y$. Thus, xRy if and only if $x \subseteq y$.

Proof of 4. Suppose α and β are ordinals. Then, by theorem 3.1.12, either α is isomorphic to β , α is isomorphic to an initial segment of β , or β is isomorphic to an initial segment of α . If α is isomorphic to β , then $\alpha = \beta$ by part 3. Now let's suppose that α is an initial segment of β , let's say β_a . Then $\alpha \in \beta$ because α is the set of all predecessors of $a \in \beta$. That set of predecessors must also be in β since β is ordinal, so $\alpha \in \beta$. For the same reason if β is an initial segment of α , then $\beta \in \alpha$. \square

So far, we have only discussed finite ordinals, but there is a different type of ordinal number called a limit ordinal.

Definition 3.1.14. A *limit ordinal* is an ordinal number that is not the successor of any other ordinal.

Let's suppose λ is a limit ordinal, and $\beta \in \lambda$. Now let's consider the successor of β , β^+ . By theorem 3.1.13, then $\beta^+ = \lambda$, $\beta^+ \in \lambda$, or $\lambda \in \beta^+$. We can rule out the last option since β^+ is still a predecessor to λ . We can also rule out the possibility that $\beta^+ = \lambda$ since this would mean λ is the successor to β , but λ is not a successor to any ordinal as we had established. Thus, $\beta^+ \in \lambda$. Then if $\beta^+ \in \lambda$, then $\beta^{++} \in \lambda$, $\beta^{+++} \in \lambda$, $\beta^{++++} \in \lambda$, and so on [6]. The successor for any element in the limit ordinal is contained in the limit ordinal and is never equal to it. We call the first limit ordinal ω which contains all natural numbers. Notice that it is not the successor of any of the numbers (the finite ordinals) because for every number, its successor is another finite ordinal, not a limit ordinal. For any smaller ordinal, there is another ordinal as its successor that is still smaller than ω .

Although we have made a distinction between finite and limit ordinal, the properties from theorem 3.1.13 hold for both. We can further convince ourselves of this by looking at one of the properties such as the order relation on an ordinal is always the subset relation \subseteq . We can recall that for the natural numbers each number is in fact ordered by the subset relation. On the finite ordinal of the third natural number for example, we have the set $3 = \{\emptyset, \{\emptyset\}\} = \{1, 2\}$. Thus, it holds that $\emptyset \subseteq \{\emptyset\} \subseteq \{\emptyset, \{\emptyset\}\}$ or $1 \subseteq 2 \subseteq 3$. With this, we can also see how this would hold for the first limit ordinal. The limit ordinal is the set

$$\omega = \{1, 2, 3, 4, \dots\} = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset, \{\emptyset\}\}\} \dots\}$$

which continues forever. Note that the subset relation still holds for the infinite set $1 \subseteq 2 \subseteq 3 \subseteq 4 \subseteq \dots$. The subset relation will hold on all ordinal numbers, both finite and limit ordinals. This completes our discussion on ordinal numbers, but we will

be able to make some connections between this topic and our next topic: cardinal numbers.

3.2 CARDINAL NUMBERS

Simply speaking cardinal numbers are used to define cardinality of sets and are often referred to just as cardinals. Each set is associated with a unique cardinal number, and if two sets are equinumerous then they are associated with the same cardinal number. However, there are different types of cardinal numbers. The simplest of these would be the cardinal numbers used to denote finite sets. For finite sets, our definition is the same as we had when we introduced cardinality in chapter 2. For finite sets, our cardinal number is the number of elements in that set. This will always be represented by a natural number. Consider the finite set $A = \{w, x, y, z, v\}$. Then the cardinal number used to represent A is 5 the same way that $|A| = 5$. There also exists the cardinal number 0 which is used to represent the empty set.

Finite cardinal numbers are the simplest case. However, there is one other type of cardinal number which will be much more relevant throughout the rest of this project and in our discussion of the continuum hypothesis. These are **infinite cardinals** which are associated with infinite sets. Unlike cardinal numbers for finite sets, we do not consider infinite cardinals to be numbers, nor do we represent them as such. Instead, we will introduce a new symbol called aleph \aleph . When we use this symbol, we do not think of it as a number, but rather just a notation used to represent the size of an infinite set. For example, by convention we would denote \aleph_0 to be the cardinal number associated with \mathbb{N} , that is $|\mathbb{N}| = \aleph_0$ [7]. And we know that if two sets are equinumerous they are associated with the same cardinal number. Since \mathbb{N} is equinumerous to \mathbb{Z} , then we would also associate \mathbb{Z} with the cardinal

number \aleph_0 . Thus, we have cardinal numbers that are either equal to 0 for the empty set, some natural number for finite sets, or some infinite cardinal \aleph_α for infinite sets. And we will also introduce the idea that \aleph_1 is associated with the set of real numbers, that is $|\mathbb{R}| = \aleph_1$. We will not elaborate on this convention at this point as that would require a more in-depth discussion of the continuum hypothesis that we will get to later. However, this convention is necessary to state as we begin discussing some properties of cardinals. We will begin by redefining what it means for a set to be finite, countable, or uncountable in terms of cardinal numbers.

Definition 3.2.1. 1. A set is considered finite if its cardinality is less than \aleph_0 .

2. A set is considered countable if its cardinality is at most \aleph_0 .

3. A set is considered uncountable if its cardinality is at least \aleph_1 .

These definitions reiterate what was discussed earlier, but with our newly defined cardinal numbers. We have discussed ordinal numbers, and have now introduced cardinal numbers. It turns out that the two concepts are closely related. The following theorem regarding the ordering on cardinals will allow us to elaborate on this idea [7].

Theorem 3.2.2. Given a cardinal \aleph_α , there exists a successor cardinal $\aleph_{\alpha+1}$. In other words, the set of cardinals is well-ordered.

We will not prove this as the proof is too complex for the scope of this project. However, let's convince ourselves that the set of cardinals is well-ordered. The smallest cardinal is 0 associated with the empty set, so in the case of the set of finite cardinals, any subset must also have a least element. For the set of infinite cardinals, we still have a least element: \aleph_0 . Thus, if we had a subset of infinite cardinals, we will always have a least element. We can order the cardinals in the following way.

$$0 < 1 < 2 < \dots \text{ (all other finite cardinals) } \dots < \aleph_0 < \aleph_1 < \aleph_2 < \dots$$

If we were to write this in terms of successors we would have to separate them. Similar to ordinals, the first infinite cardinal \aleph_0 is not a successor to any other cardinal. For this reason we would call \aleph_0 a **limit cardinal**.

$$\aleph_0 < \aleph_0^+ < \aleph_0^{++} < \aleph_0^{+++} < \dots$$

Cardinal numbers also follow similar properties to ordinals such as the trichotomy property. For any two cardinals α and β , either $\alpha < \beta$, $\beta < \alpha$, or $\alpha = \beta$.

One other thing that is interesting about infinite cardinal numbers is that they do not follow the same arithmetic rules as natural numbers or real numbers. One may be curious to know what happens if we try to add infinities or multiply them. The answer to these questions can be answered using cardinal arithmetic. These arithmetic rules and their proofs follow from [7]. First, let's define addition on infinite cardinals.

Definition 3.2.3. Let \aleph_α and \aleph_β be infinite cardinals such that $\aleph_\alpha = |A|$ and $\aleph_\beta = |B|$. Then, $\aleph_\alpha + \aleph_\beta = |A \cup B|$.

This definition can be used to prove the next two theorems.

Theorem 3.2.4. Let \aleph be any infinite cardinal. Then,

1. If $n = |A|$ where $n \in \mathbb{N}$, then $n + \aleph = \aleph$
2. $\aleph + \aleph = \aleph$

Proof. Proof of 1. Let there be an infinite set A such that $\aleph = |A|$.

Case 1. Suppose $n = 0$. We have established that this means $n = |\emptyset|$. Then we have

$$0 + \aleph = |\emptyset \cup A| = |A| = \aleph$$

Case 2. Suppose $n > 0$. Since $n < \aleph$, and A is infinite, then there exists a finite set $\{a_1, a_2, \dots, a_n\} \subset A$ such that $|\{a_1, a_2, \dots, a_n\}| = n$. Then we have,

$$n + \aleph = |\{a_1, a_2, \dots, a_n\} \cup A| = |A| = \aleph$$

Proof of 2. Continuing the convention that $\aleph = |A|$ we have,

$$\aleph + \aleph = |A \cup A| = |A| = \aleph$$

□

Theorem 3.2.5. For two infinite cardinals α and β such that $\alpha < \beta$, then $\alpha + \beta = \beta$.

Proof. Let's denote sets A and B such that $\alpha = |A|$ and $\beta = |B|$, so $|A| < |B|$. Then there is an injective function $f : A \rightarrow B$. First let's show that there is a bijective function F such that $F : A \rightarrow f(A)$, in other words $A = f(A)$. We can look at fig. 3.4 to visually understand this mapping.

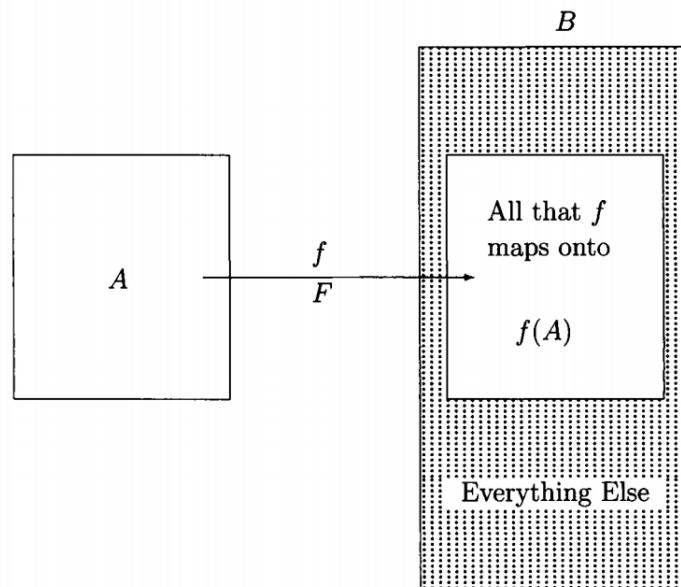


Figure 3.4: Mapping from A to $f(A)$ [7]

Both f and F follow the same rule and conditions but just map into different sets. So f is only injective and not surjective, but we will show that F is bijective. We know F is injective because it follows the same rules as f which we had established is an injective function. To show that F is surjective consider a $y \in f(A)$. Then by definition of $f(A)$ there exists an $x \in A$ such that $F(x) = f(x) = y$. Thus F is surjective. Since F is injective and surjective, then $F : A \rightarrow f(A)$ is bijective. Since $f(A) \subset B$, then $f(A) \cup B = B$. So,

$$\alpha + \beta = |A| + |B| = |f(A)| + |B| = |A \cup B| = |B| = \beta$$

Thus we have shown that the sum of two infinite cardinals is equal to the larger cardinal. \square

The result from theorem 3.2.5 can also be applied to cardinal multiplication. An example of this is considering the two sets \mathbb{R} and \mathbb{N} . We used the convention that $|\mathbb{N}| = \aleph_0$ and $|\mathbb{R}| = \aleph_1$ and $\aleph_0 < \aleph_1$. Then by definition 3.2.3, $\aleph_0 + \aleph_1 = |\mathbb{N} \cup \mathbb{R}|$. The natural numbers are a subset of the real numbers, so the union of the two must equal the real numbers. Thus, $\aleph_0 + \aleph_1 = |\mathbb{N} \cup \mathbb{R}| = |\mathbb{R}| = \aleph_1$. Therefore, theorem 3.2.5 holds. The result from this theorem can also be applied to understand cardinal multiplication.

Definition 3.2.6. For two infinite cardinals α and β such that $\alpha < \beta$, then $\alpha \times \beta = \beta$.

We have introduced addition and multiplication on cardinals, but now we must discuss powers of cardinals which will be most relevant to this project. For finite cardinals, the power rules hold as one would expect. For two cardinal numbers, 5 and 2 associated with two different sets, taking the power would be intuitive, $5^2 = 25$, and would result in another finite cardinal number: 25. We can define this.

Definition 3.2.7. Suppose α and β are cardinals associated with two sets A and B such that $\alpha = |A|$ and $\beta = |B|$. Then $\beta^\alpha = |B^A|$ where B^A is the set of all functions $f : A \rightarrow B$.

What may not be as intuitive is that the same applies to infinite cardinals. However, using this definition for infinite cardinals it is important to be cautious that we are no longer dealing with numbers. In this project, we are only really concerned with using them to denote power sets of infinite sets. Looking back to how we defined power sets if a finite set has cardinality n , its power set has cardinality 2^n . This is still the same for infinite cardinals, but we must think of it differently. If we have a finite cardinal number n associated with a set, then we would get another finite number m associated with its power set. However, let's say we wanted to take the power set of \mathbb{N} . We know that the size of \mathbb{N} is associated with the cardinal number \aleph_0 . Then the power set of the natural numbers, $\mathcal{P}(\mathbb{N})$ would be 2^{\aleph_0} . This is the main extent to which powers of cardinals are used: to denote power sets of infinite sets. Thus, what we have stated is that $|\mathcal{P}(\mathbb{N})| = 2^{\aleph_0}$. We also know that $|\mathcal{P}(\mathbb{N})| = |\mathbb{R}|$, so we can say $|\mathbb{R}| = 2^{\aleph_0}$. Then by Cantor's Theorem, we can state that $\aleph_0 < 2^{\aleph_0}$. We will keep this fact in mind for our discussion of the continuum hypothesis, but will not go any further at this moment.

Every concept we have discussed so far is related to set theory. However, we have yet to define a formal system used for set theory. Throughout this project, we have yet to question the existence of these objects such as ordinals and cardinals. Everything we have covered intuitively makes sense and we take it to be true. However, the existence of these objects can only be guaranteed if we define a formal system of set theory for which they are a part of.

CHAPTER 4

ZERMELO-FRAENKEL SET THEORY

The way we defined sets in the previous chapters gives us a base understanding of some of the concepts we will see throughout this project. However, it is not always sufficient. I would like to raise an issue that may occur with the way we have defined sets using naive set theory in chapter 2 using a famous example called Russell's Paradox. Mathematician Bertrand Russell was the first to realize that this misuse of sets can lead to paradoxical situations. The 'set' constructed in Russell's paradox is defined as follows:

$$A = \{X : X \text{ is a set and } X \notin X\}$$

In other words, A is the set of all sets that do not contain themselves as elements. This holds for the empty set because \emptyset is a set and $\emptyset \notin \emptyset$, thus $\emptyset \in A$. We can also provide an example of a set that is not in A . Consider the set $B = \{\{\{\{\dots\}\}\}\}$ where B can be thought of as a box containing a box, containing a box, and so on where the boxes are endlessly nested inside each other. Then, for this reason, B is the set B itself

$$B = \underbrace{\{\{\{\{\dots\}\}\}\}}_B = \{B\}$$

Thus, $B \in B$, so $B \notin A$. However, what would happen if we asked the question

of whether or not $A \in A$. Then A cannot be in A by the way the set is defined. If $A \notin A$, then A must be in A again by how the set is defined.

This is a paradox within naive set theory. To resolve this, mathematicians Ernst Zermelo and Abraham Fraenkel created a system of axioms in the early 1900s that was structured in a way that would eliminate these paradoxes and contradictions. It was called Zermelo-Fraenkel set theory and consisted of nine axioms and later the axiom of choice. This system is what we use when we discuss set theory. So, up to this point, when we proved theorems about cardinality, sets, cardinals, and ordinals, we were proving them within this system although there was no need to formally state it. We will go through all nine axioms beginning with the axiom of extensionality, and end with a discussion on the axiom of choice.

4.1 ZERMELO-FRAENKEL AXIOMS

Axiom 4.1.1. (*Axiom of Extensionality*) *Two sets are equal if and only if they contain the same elements.*

Formally, we can write the axiom of extensionality as,

$$\forall A, \forall B (A = B \iff \forall x (x \in A \iff x \in B))$$

In this expression, we have $x \in A \Rightarrow x \in B$ and $x \in B \Rightarrow x \in A$ implies $A = B$. This is the same thing as saying $A \subseteq B$ and $A \supseteq B$ implies $A = B$. Therefore, we can conclude that the axiom of extensionality states the same thing as theorem 2.1.2. We can rewrite our formal definition as follows,

$$\forall A, \forall B (A = B \iff A \subseteq B \text{ and } B \subseteq A)$$

It is important to note that it does not matter how many times a single element is listed or the order of the elements. Let's look at the following example.

$$\{1, 1, 2\} = \{1, 2\} = \{2, 1\} = \{2, 1, 2\}$$

All these sets are equal according to the axiom of extensionality since they contain the same elements. Without the axiom of extensionality, we could have multisets which are sets where the number of a single element in the set is used to define the set. If we to consider multisets, then $\{1, 1, 2\} \neq \{1, 2\}$. We would also have ordered sets where the order of elements in the set is used to define the set. If this were the case, then $\{1, 2\} \neq \{2, 1\}$. However, by implementing the axiom of extensionality we do not come across these issues, and those sets are all considered to be equal. The next axiom we will discuss is the empty set axiom, which will require a much shorter explanation.

Axiom 4.1.2. *(The Empty Set Axiom) There is a set called the null set, denoted \emptyset that contains no elements.*

We have discussed the existence of the set in chapter 2. Although this axiom may seem trivial, this axiom is necessary for understanding other Zermelo-Fraenkel axioms and to uphold certain set theory consequences and results.

The next axiom is the pairing axiom which discusses a procedure for constructing sets, but only involving pairs of sets.

Axiom 4.1.3. *(Pairing Axiom) Given any sets X and Y , there exists a set Z such that the elements of Z are X and Y .*

In other words, if X is a set and Y is a set, then the set $\{X, Y\}$ exists. This is an example of the case where $X \neq Y$ which results in sets of pairs of sets. Now, let's consider the case where we have two sets A and B such that $A = B$. In this case, we get singletons or sets that contain a singular element. Thus, there exists a set

C whose only element is the set A . The pairing axiom results in pairs, $\{X, Y\}$, or singletons, $\{X\}$. The pairs are unordered pairs, but as we have seen from the axiom of extensionality $\{X, Y\} = \{Y, X\}$ and we know that the set $\{X, X\}$ will result in the singleton $\{X\}$. The pairing axiom is closely related to the next axiom, the axiom of union as it is necessary to introduce a new notion of the union of sets.

Axiom 4.1.4. (*Axiom of Union*) *If X is a set whose elements are sets, there is a set $\bigcup X$ consisting of all elements of all elements of X .*

When we had first introduced what the union between two sets looks like we defined it as $c \in A \cup B$ if $c \in A$ or $c \in B$. Although this is true, we need to redefine this notion so we can explicitly define the process of constructing these sets. Suppose we have two sets A and B . Our goal is to construct the union of these sets, which is the set $A \cup B$. First, by axiom 4.1.3, we know that the set $\{A, B\}$ exists, so we construct that set. Then, by the axiom of union, we can construct the set $\bigcup\{A, B\}$ whose elements are all elements of A and B which is the set we had intended to form. Thus, combining the pairing axiom and the axiom of union guarantees the union of two sets and provides a process of their formations. We can now formally redefine union from these two axioms.

Definition 4.1.5. *The union of two sets $A \cup B$ is $\bigcup\{A, B\}$ following axiom 4.1.3 and axiom 4.1.4 for sets A and B .*

This axiom is also helpful for proving the existence of ordinal numbers. With the completion of this fourth axiom, we can now construct basic finite sets using singletons, pairs, and unions. We have seen examples already of forming singletons, pairs, and unions of each. However, let's say we had three sets C, D , and E . Then we could define the set $\{C, D, E\}$ as $\bigcup\{\bigcup\{C, D\}, \{E\}\}$. Now, let's say we have more than three sets. Suppose we have n sets for some $n \in \mathbb{N}$, that is we have sets $s_1, s_2, s_3, \dots, s_n$. Then similarly we can define the set $\{s_1, s_2, s_3, \dots, s_n\}$ as $\bigcup\{\bigcup\{s_1, s_2, s_3, \dots, s_{n-1}\}, \{s_n\}\}$ [9].

We have yet to understand how we may construct the intersection between two sets, but that will be discussed in a later axiom. The next axiom discusses the existence of power sets, a notion we have already become somewhat familiar with.

Axiom 4.1.6. (*Power Set Axiom*) *Given any set X there is a power set of X , denoted $\mathcal{P}(X)$ whose elements are all subsets of X .*

Formally, the axiom states,

$$\forall X, \exists \mathcal{P}(X), \forall a (a \in \mathcal{P}(X) \iff a \subseteq X)$$

We introduced power sets when we discussed set theory background, so this axiom does not require much more explanation. If the set X exists then the set $\mathcal{P}(X)$ exists. Since $\mathcal{P}(X)$ is a set, then the set $\mathcal{P}(\mathcal{P}(X))$ also exists. An interesting result of this axiom allows us to create a sequence of power sets which we will come back to shortly.

These first five axioms give us enough information to understand basic set properties and construct simple sets. The remaining axioms follow more intuitive principles about the construction and properties of sets. The next two axioms relate to the comprehension principle which was first introduced by Zermelo and later modified by Fraenkel and Thoralf Skolem [9].

Axiom 4.1.7. (*Axiom of Separation*) *Given a condition ψ and a set X there exists a set Y that contains the elements of X and satisfies the condition ψ .*

We can formally state the axiom as follows,

$$\forall X, \exists Y, \forall a (a \in Y \iff (a \in X \text{ and } a \in \psi(a)))$$

Essentially this axiom is stating that we can create subsets from a given condition or assertion ψ . However, notice that there are an infinite number of assertions that

can be made, so this is not technically one single axiom. For this reason, the axiom of separation is often referred to as an axiom scheme. The condition takes one free variable a , so we would have $\psi(a)$. *The condition must also be expressible as a well-formed formula in Zermelo-Fraenkel set theory.* This means it can be expressible using only the following propositions: equality, $a = b$ or set membership $a \in B$. The formulas must also be made up of logical operations such as quantifiers, implications, and negations. Let's look at two simple examples where we can take a condition and express it as a well-formed formula.

Let A and B be two fixed sets and we start with the set $A \cup B$. Suppose we want to construct the set $A \cap B$. So we want the condition that for any element z , $z \in A \cap B$. However, this formula is not well-formed in Zermelo-Fraenkel set theory. Instead, we want the set which can be constructed from a condition expressed as a well-formed formula, such as the one stated below.

$$\{z \in A : z \in B\}$$

This is stating the same exact thing as $A \cap B$, but we have instead expressed it using a well-formed formula. Next, suppose we wanted to construct the set of the relative complement of B , that is the set of all elements z in A/B . However, using this notation of A/B is again not well-formed in Zermelo-Fraenkel set theory. Instead, we can construct the set by expressing the condition again using a well-formed formula stated below.

$$\{z \in A : z \notin B\}$$

Note, that we correctly stated the condition of $z \in A \setminus B$ only using the allowed notation that characterizes it as a well-formed formula. However, with any of the possible conditions ψ , there is no guarantee that any possible member satisfies that

condition. It may instead be the case that the set contains no elements and we would have the empty set which we know exists by axiom 4.1.2.

This axiom was one of the significant responses to Russell's paradox as it asserted a correct way of constructing a set. However, it was later discovered that the axiom was not sufficient to cover all constructible sets. Let's go back to the interesting result that we introduced in our discussion of the power set axiom. Suppose we have a set X . Then we know the power set of X $\mathcal{P}(X)$ exists by axiom 4.1.6. Then again, by axiom 4.1.6 $\mathcal{P}(\mathcal{P}(X))$ also exists. We can continue constructing these sets as we know they will all exist by the power set axiom, so we would have $X, \mathcal{P}(X), \mathcal{P}(\mathcal{P}(X)) \dots$ and continue with this process forever. Now let's try constructing the sets of all these power sets which would look like $\{X, \mathcal{P}(X), \mathcal{P}(\mathcal{P}(X)) \dots\}$. Intuitively, we should be able to do this; however, with the axioms we have discussed so far, there is actually no guarantee that this is a legitimate set. Axiom 4.1.7 cannot be applied here because this set is not a subset of our starting set X . These types of issues lead us to the next axiom, the axiom of replacement.

Axiom 4.1.8. (*Axiom of Replacement*) Suppose \mathcal{F} is a function that can be expressed as a well-formed formula in Zermelo-Fraenkel set theory. For any set X , there exists a set Y which is the set of all images of each element in X under the function \mathcal{F} .

Roughly speaking, this axiom states that the image of a set under any function must also be a set. Similar to the axiom of separation, the axiom of replacement is also considered to be an axiom schema since there are infinite possibilities for a function \mathcal{F} . We can create a well-formed formula for our function \mathcal{F} the same way we did for our condition ψ in the axiom of separation. However, in axiom 4.1.7, our condition took one variable, so we had $\psi(a)$. In this case, our function is constructed by taking two free variables, so we would have $\mathcal{F}(x, y)$. After discussing both the axiom of separation and the axiom of replacement, we can look at an interesting result that shows the close relation between the two axiom schemes [9].

Theorem 4.1.9. *The axiom of replacement implies the axiom of separation.*

Proof. Let $\psi(y)$ be some well-formed formula as we saw in axiom 4.1.7, and let X be a set. Then the elements in X either satisfy the condition ψ or do not satisfy the condition. If no element satisfies the condition, then we get the empty set which we know exists by axiom 4.1.2. Then let $\mathcal{F}(u, v)$ be the formula $u = v$ and $\psi(u)$. This formula determines the identity function on all u for which $\psi(u)$ holds. If we restrict this function to the set X , then it would be the function on all $u \in X$ such that $\psi(u)$ holds. This gives us the set

$$w = \{v : (\exists u)(u \in X \text{ and } \mathcal{F}(u, v))\}$$

deduced by the axiom of separation which is the exact same set as

$$\{v \in X : \psi(v)\}$$

required by the axiom of replacement. Therefore, we can deduce the same formula from the axiom of separation without the need for the axiom of replacement. \square

This result shows that the axiom of separation is a repetitive axiom in Zermelo-Fraenkel set theory, but it is still included since it describes a more standard way of constructing sets than the axiom of separation.

Going back to the issue we had earlier of the set of power sets $\{X, \mathcal{P}(X), \mathcal{P}(\mathcal{P}(X)) \dots\}$, we can use the axiom of foundation to show that this is a legitimate set. We can state that this set is the image of the set of natural numbers. The function associates every natural number n with $\mathcal{P}^{n-1}(X)$. To more clearly show this, we can create a table of the elements in \mathbb{N} associated with each elements image under this function.

n	Image
1	X
2	$\mathcal{P}(X)$
3	$\mathcal{P}(\mathcal{P}(X))$
\vdots	\vdots

We started with the set \mathbb{N} and defined a function $\mathcal{P}^{n-1}(X)$. In this case our variables would be the set \mathbb{N} and the set X . Then, by the replacement axiom there exists a set containing all images of the elements of \mathbb{N} under this function. Therefore, the set does in fact exist and we have shown the first few elements that would be in the set. However, this example still has an issue. We have used a procedure that allows us to construct an infinite set using another infinite set, the set of natural numbers. However, up to this point we have not yet seen an infinite constructible set, nor have any of our axioms led us from finite sets to infinite sets. We must address this notion of infinite sets before moving on, which leads us to the infinity axiom.

Axiom 4.1.10. (*Infinity Axiom*) *There is a set X that contains the empty set \emptyset and if $a \in X$ then $a \cup \{a\} \in X$ must also be true.*

Such a set is infinite, so this axiom confirms that there exists a set which is infinite. In addition, it states exactly what elements the set contains. The element we start with in the set is the empty set \emptyset , so $\emptyset \in X$. Then as the axiom states, $\emptyset \cup \{\emptyset\}$ must also be in the set X . Since this is also an element in the set, we continue the process. So we get $\emptyset \cup \{\emptyset\} \cup \{\emptyset \cup \{\emptyset\}\} \in X$. This process continues forever and results in the following set.

$$\{\emptyset, \emptyset \cup \{\emptyset\}, \emptyset \cup \{\emptyset\} \cup \{\emptyset \cup \{\emptyset\}\} \dots\}$$

In this case, we were just considering the set containing the empty set, but the

set may contain other elements as well. We have chosen the sequence $\emptyset, \emptyset \cup \{\emptyset\}, \emptyset \cup \{\emptyset\} \cup \{\emptyset \cup \{\emptyset\}\} \dots$. With this sequence it is clear that all the members are distinct which is necessary for us to say that the set is infinite. However we could have also chosen a different sequence of such as $\emptyset, \{\emptyset\}, \{\{\emptyset\}\} \dots$ used by Zermelo and still have an infinite sequence of distinct members. There are many other processes we could use to form the sequence, but the process we worked with was first used by mathematician John von Neumann [9]. The final axiom we will discuss is the axiom of foundation, which follows an intuitive property of set theory.

Axiom 4.1.11. (*Axiom of Foundation*) *If X is a set, there is an $a \in X$ such that $a \cap X = \emptyset$.*

In other words this axiom states that every non-empty set contains an element that is disjoint from that set. The purpose this axiom serves is that it tells us it is impossible to have an infinite chain of sets of which there must be a predecessor contained in each set. We have seen examples of the opposite being true, that is each element is contained in its successor such as the construction of the natural numbers.

$$1 \in 2 \in 3 \in \dots$$

However, this axiom states that the opposite of this construction is impossible. Suppose, we started with a set X_0 . Then we could choose an element $X_1 \in X_0$, then $X_2 \in X_1$, $X_3 \in X_2$, and so on attempting to create the sequence shown below.

$$\dots X_{n+1} \in X_n \in \dots X_3 \in X_2 \in X_1 \in X_0$$

This is impossible by the axiom of foundation and let's convince ourselves why this is the case. Suppose we had the set $A = \{a_1, a_2, a_3, \dots\}$ and assume $a_{n+1} \in a_n$ for all $n \in \mathbb{N}$. Then by the axiom of foundation, A must contain an element b that is disjoint from the set, that is $A \cap b = \emptyset$. If $b \in A$, then $b = a_m$ for some $m \in \mathbb{N}$. And by our original assumption $a_{m+1} \in a_m$. If this assumption is true then a_{m+1} must also

be in A . So, we have that $a_{m+1} \in a_m$, where $a_m = b$, and $a_{m+1} \in A$. Thus, $a_{m+1} \in b \cap A$. However, $b \cap A = \emptyset$, so no elements can be contained within the intersection. This poses a contradiction. Thus, using the axiom of foundation, we can understand that these types of sequences where $\dots a_{m+1} \in a_m \dots a_2 \in a_1$ are impossible [9].

With the completion of this axiom, we now have constructed the model of Zermelo-Fraenkel set theory also referred to as ZF. The next section discusses an axiom that is used as an extension to ZF, the axiom of choice. With the addition of the axiom of choice, we form the model often referred to as ZFC.

4.2 AXIOM OF CHOICE

The axiom of choice allows mathematicians to select elements from a nonempty set to construct other mathematical objects. Some mathematicians argue that the axiom is unnecessary. However, there are some true statements in math that cannot be proved without the axiom of choice. For example, we need the axiom of choice to prove that the union of countably many countable sets is uncountable. The axiom also becomes very useful in our discussion of the continuum hypothesis as we are proving its independence from ZFC. Before we state the axiom of choice we must first define a choice function.

Definition 4.2.1. *A choice function f defined on a collection X of nonempty sets is a function such that for every set $A \in X$, $f(A) \in A$.*

This choice function is what we will use to form new sets. Now we can state the axiom of choice which we will abbreviate as AC.

Axiom 4.2.2 (AC). *For any set X (whose elements are sets) where $\emptyset \notin X$, there exists a choice function f on X that maps each set of X to an element of that set.*

We can see how this may work for a finite set such as $X = \{W, Y, Z\}$ where

$W = \{a, b, c, d, e\}$, $Y = \{1, 2, 3, 4, 5\}$, $Z = \{\alpha, \beta, \gamma, \delta, \sigma\}$. Then we can define a choice function f that maps each set W, Y , and Z to an element in each corresponding set.

$$\{a, b, c, d, e\} \xrightarrow{f(W)} b$$

$$\{1, 2, 3, 4, 5\} \xrightarrow{f(Y)} 5$$

$$\{\alpha, \beta, \gamma, \delta, \sigma\} \xrightarrow{f(Z)} \alpha$$

Note that the axiom of choice is really an intuitive truth as an extension to our Zermelo-Fraenkel axioms. In many of the proofs that we have done so far in set theory have just been ZF, we did not assume the axiom of choice. This is the case for all set theory proofs unless explicitly stating that we are assuming that the axiom of choice holds. There has been much discussion about whether or not the axiom of choice should be accepted, although now most mathematicians do accept it [12]. However, for our purposes of discussing the continuum hypothesis, the axiom of choice becomes very relevant. But first, let's look at some other direct applications related to infinite sets.

One application of the axiom of choice is proving that every infinite set has an infinite countable subset [9]. Informally, the proof follows that given an infinite set A , which can be countable or uncountable, choose an element $a_0 \in A$. Then choose another element $a_1 \in A$ where $a_0 \neq a_1$. Then, choose an $a_2 \in A$ such that $a_2 \neq a_1 \neq a_0$. This continues forever and we form a sequence of distinct elements in A as follows: a_0, a_1, a_2, \dots . Notice how the axiom of choice is necessary for this outline of the proof. The axiom of choice also has many equivalences such as the well-ordering principle as stated in theorem 3.1.8 and Zorn's Lemma which we will state below.

Lemma 4.2.3 (Zorn's Lemma). *Suppose S is a partially ordered set. If every totally ordered subset of S has an upper bound, then S contains a maximal element.*

With these three equivalent statements, we can actually show that each of these statements implies the other. So if we only assumed one of them to be an axiom, it must follow that the other two are also true. Typically there are three implications in the following order and we will prove the last one.

Theorem 4.2.4. 1. *The axiom of choice implies Zorn's Lemma.*

2. *Zorn's lemma implies the well-ordering principle.*

3. *The well-ordering principle implies the axiom of choice.*

Proof of 3. We can prove this implication by showing that we can construct a choice function by means of the well-ordering principle. Given any collection of sets, let's first form the union of all sets in the collection. We know this union can be well-ordered by the well-ordering principle. Then for each set X in the collection we can specify a choice function f where $f(X)$ is the least element of X relative to the order relation. Thus, we were able to specify a choice function f from the well-ordering principle. \square

Although theorem 4.2.4 states the typical direction of these implications, we can prove the opposite direction without going through Zorn's Lemma, although the proof is more extensive.

Theorem 4.2.5. *The axiom of choice implies the well-ordering principle.*

To complete this proof, we will prove several claims that will lead us to the desired result. The proof technique we will be using is from [14]. We will utilize our result from theorem 3.1.12 about isomorphisms between well-ordered sets.

First, consider an arbitrary set X . We want to use the axiom of choice to show that there exists some well-ordering on X . Let's consider a choice function f such that $f(Y) \in X \setminus Y$ where $Y \subset X$ has a non-empty complement. The remainder of this

proof will relate to this function f . With this function, we will define something called an f -string which will be referred to many times throughout the proof.

Definition 4.2.6. *An f -string is a pair $\langle A, R \rangle$ where $A \subseteq X$ and R is a well-ordering on A such that*

$$\forall a \in A, (a = f(\{b \in A \mid bRa\}))$$

As of right now, we do not know for sure if an f -string exists, but if this condition is true we would call that pair an f -string. Before we begin proving anything, let's make a final remark about the commonality of the least element of an f -string.

Remark 4.2.7. *If $\langle A, R \rangle$ is a nonempty f -string and a_0 is the R -least element of A (the least element relative to the relation R), then the least element will be common for all non-empty f -strings.*

$$a_0 = f(\{b \in A \mid bRa_0\}) = f(\emptyset)$$

When we say for all non-empty f -strings, we mean that given the set A for all R that well-orders A , the R -least element of A will always be the same and be equal to $f(\emptyset)$. With this definition and remark established, we are now ready to begin the proof of our first lemma.

Lemma 4.2.8. *An initial segment of an f -string is an f -string.*

Proof. Suppose $\langle B, S \rangle$ is an initial segment of the f -string $\langle A, R \rangle$ and let $a \in B$. Then,

$$a = f(\{b \in A \mid bRa\}) = f(\{b \in B \mid bSa\})$$

Since $\langle B, S \rangle$ is an initial segment of $\langle A, R \rangle$ then the above statement holds since S is just R relative to only $B \subset A$. And we know that $b \in B$ by transitivity since b is the predecessor of some $a \in B$ which we can see more clearly in fig. 4.1.

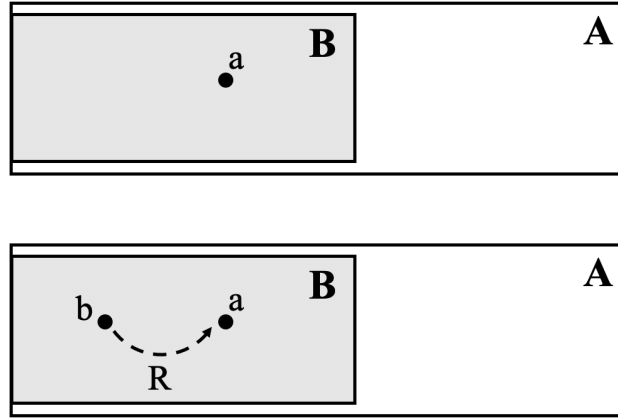


Figure 4.1: The element b is a predecessor of a , so it must also be in B .

Thus, the initial segment is still an f -string. □

We now move on to proving four claims in progressing order that will lead us to the end of the proof.

Claim 4.2.9. *Suppose that $\langle B, S \rangle$ and $\langle A, R \rangle$ be f -strings and let g be an isomorphism between initial segments of the strings. Then, $g(a) = a$ for all a in the domain.*

Proof. For the sake of contradiction, let a be the least element in A under a relation R such that $g(a) \neq a$. Instead, let $g(a) = b$ for some $b \in A$. Then $g(c) = c$ whenever $c \in A$ and cRa . We know that g is an isomorphism, so the following must be true for some $d \in B$.

$$\{d \in B \mid dSb\} = \{g(c) \mid cRa\} = \{c \in A \mid cRa\}$$

Here we have for the set of $d \in B$ where dSb . This must be the same as $g(c)$ where cRa since $g(a) = b$ and we said $g(c) = c$ wherever cRa . So by isomorphism, for wherever dSb , the same must be true to give us cRa as depicted in fig. 4.2.

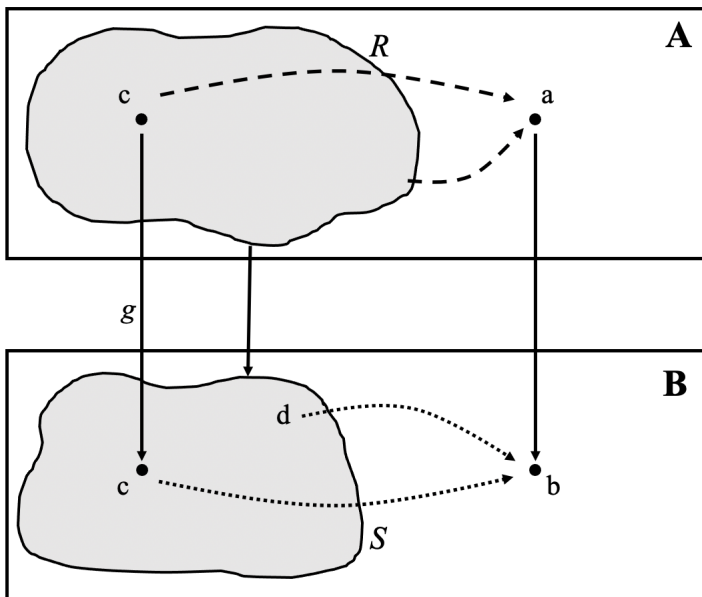


Figure 4.2: Any c where cRa must map to c contained in the set of all d where dSb .

However, since $\langle B, S \rangle$ and $\langle A, R \rangle$ are f -strings it also follows from definition 4.2.6 that,

$$a = f(\{c \in A \mid cRa\}) = f(\{d \in B \mid dSb\}) = b$$

So we have $a = b$. However, this is a contradiction from our original assumption that $a \neq g(a)$ and $g(a) = b$. Therefore, it must be true that $g(a) = a$ for all a in the domain. \square

Claim 4.2.10. *Any two f -strings are either equal, or one is an initial segment of the other.*

Proof. By theorem 3.1.12 the two f -strings are either isomorphic or one is isomorphic to an initial segment of the other. Then by lemma 4.2.8 and claim 4.2.9 the isomorphism will always be the identity function. Therefore, the claim holds. \square

With this claim, we can now form a total order on the f -strings by initial segment ordering. Let A be the union of all B where $\langle B, S \rangle$ is an f -string and let R be the union of all relations S such that $\langle B, S \rangle$ is again an f -string.

Claim 4.2.11. $\langle A, R \rangle$ is an f -string.

Proof. There are two parts to this proof. We want to show that R is a well-ordering on A , and we want to show that $\langle A, R \rangle$ follows our established definition of f -string. Let's first show that R is a well-ordering on A .

We want to show that any arbitrary subset of A contains a least element. Suppose $C \subset A$ where C is nonempty. By its construction, we know that $A = \bigcup B_\alpha$ for all B_α where $\langle B_\alpha, S_\alpha \rangle$ is an f -string. Then there exists an α such that $B_\alpha \cap C \neq \emptyset$. Continuing on, we will just refer to B_α as B and S_α as S . Since $\langle B, S \rangle$ is an f -string, then S is a well-ordering on B . Then S is also a well-ordering on $B_\alpha \cap C$ meaning that there exists some a that is the S -least element of $B_\alpha \cap C$. Recall from claim 4.2.10 that any two f -strings are either equal or one is an initial segment of the other. So, we know $\langle B, S \rangle$ is an f -string and we know $A = \bigcup B$. We also know by the construction of $\langle A, R \rangle$, that S is just R restricted to B . With these two facts we can conclude that $\langle B, S \rangle$ is an initial segment of $\langle A, R \rangle$. So far we have established what is shown in fig. 4.3 for some arbitrary subset $C \subset A$.

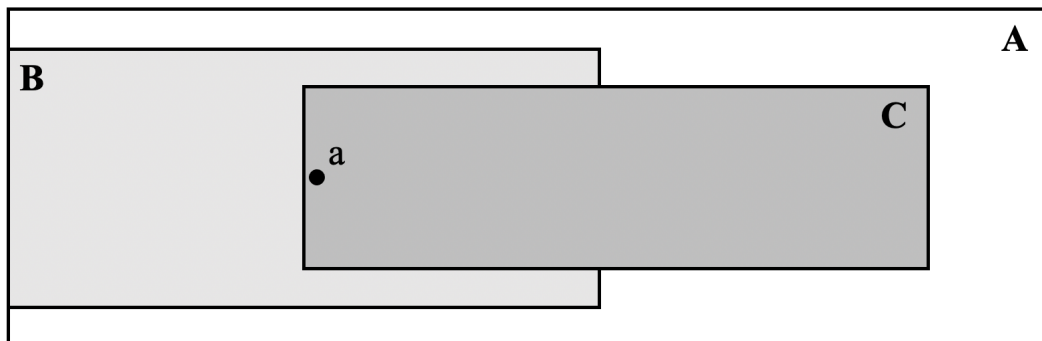


Figure 4.3: We know that $C \subset A$. Then $\langle B, S \rangle$ is an initial segment of $\langle A, R \rangle$, and a is the least element of $B \cap C$.

Now suppose for the sake of contradiction that $d \in C \setminus B$ and dRa where a is still our least element in $B \cap C$. However, since $a \in B$ and $\langle B, S \rangle$ is an initial segment, then dSa which means that $d \in B$. This poses a contradiction since we initially stated

that $d \in C \setminus B$. Thus, aRd for all $d \in C \setminus B$. Thus we have shown that a is the S -least on $B \cap C$ and $C \setminus B$. In other words, a is the R -least element *restricted* to $B \cap C$ and $C \setminus B$. Therefore, a is the R -least element on our arbitrary subset C , so R is a well-ordering on A . Next, we want to show that $\langle A, R \rangle$ is consistent with our definition of f -string in definition 4.2.6.

We know that if $a \in B$, then $a = f(\{b \in B \mid bSa\})$ since $\langle B, S \rangle$ is an f -string. We want to show that $a = f(\{b \in A \mid bRa\})$. Suppose $a \in A$ where $A = \bigcup B_\alpha$. Then there exists a B_β such that $a \in B_\beta$. So we have,

$$a = f(\{b \in B_\beta \mid bS_\beta a\})$$

However, by our construction of $\langle A, R \rangle$ the relation R *restricted* to only B_β is exactly S_β . Thus, we can state the following.

$$a = f(\{b \in B_\beta \mid bRa\})$$

Recall, that this B_β is some B_α in $\bigcup B_\alpha$. And if $b \in B_\beta$ then $b \in \bigcup B_\alpha$, where $\bigcup B_\alpha = A$. Now we can state the following.

$$a = f(\{b \in \bigcup B_\alpha \mid bRa\}) = a = f(\{b \in A \mid bRa\})$$

This is our definition of f -string and we have proven that it holds for $\langle A, R \rangle$. \square

Finally we can state and prove our last theorem that will lead us to the desired result.

Claim 4.2.12. *The set A that we constructed is equal to X . Thus, R is a well-ordering of X .*

Proof. Assume for the sake of contradiction that $A \subset X$. Also let $b = f(A)$, $B = A \cup \{b\}$, and $S = R \cup \{(a, b) \mid a \in A\}$. Then $\langle B, S \rangle$ is also an f -string, thus $b \in A$ by our

construction of A . This is a contradiction since we said that $b \notin A$. Therefore the claim holds. \square

For an arbitrary set X , we have concluded by claim 4.2.12 that there does in fact exist a relation R that well-orders X . Thus, we have proven that the axiom of choice implies the well-ordering principle.

With Zermelo-Fraenkel axioms and the axiom of choice, we have formed our model for ZFC which is the standard model used for set theory. The importance of this model is directly related to our final topic: the continuum hypothesis. The major result we will discuss is that the continuum hypothesis is *independent* of ZFC, and we will elaborate on what exactly it means when a statement is independent of a system.

CHAPTER 5

THE CONTINUUM HYPOTHESIS

5.1 BACKGROUND

What we have discussed so far has been the background required to understand a significant result in mathematics: the continuum hypothesis. There is the weak continuum hypothesis and the generalized continuum hypothesis. The continuum hypothesis was first thought up by Cantor after he discovered two different sizes of infinity: countable and uncountable. We later learned that there are infinitely many sizes of infinity, but this hypothesis originally stems from just those two sizes: $|\mathbb{N}|$ and $|\mathbb{R}|$. This leads us to Cantor's weak continuum hypothesis [13].

Hypothesis 5.1.1 (Weak Continuum Hypothesis). *Any infinite subset of \mathbb{R} either can either be put in one-to-one correspondence with \mathbb{N} or \mathbb{R} .*

In other words, this theorem states that there exists no set S such that,

$$|\mathbb{N}| < |S| < |\mathbb{R}|$$

This was Cantor's weak continuum hypothesis, but with this as a starting point, he was able to state the hypothesis more specifically in terms of cardinal numbers. First, recall from theorem 2.1.6 that for any set X with $|X| = n$, where n is a finite cardinal

number $|\mathcal{P}(X)| = 2^n$. And from Cantor's Theorem, we know that $|X| < |\mathcal{P}(X)|$. Let's look at how this might translate to the continuum hypothesis. We have established in chapter 3 that \aleph_0 is the first infinite cardinal, that is $|\mathbb{N}| = \aleph_0$. We also proved in theorem 2.2.21 that $|\mathcal{P}(\mathbb{N})| = |\mathbb{R}|$. Thus, we can use 2^{\aleph_0} to denote $|\mathcal{P}(\mathbb{N})|$ and $|\mathbb{R}|$, so $\aleph_0 < 2^{\aleph_0}$. These are all conclusions we can draw from what we have already discussed. With these conclusions, we can state the more complete version of Cantor's continuum hypothesis [5].

Hypothesis 5.1.2 (Continuum Hypothesis). *The continuum c is the immediate successor of the first infinite cardinal. In other words,*

$$\aleph_0 < 2^{\aleph_0} = \aleph_0^+ = \aleph_1 = c$$

Thus, $|\mathbb{N}| = \aleph_0$, and $|\mathbb{R}| = \aleph_1 = c$. We could go even further with this hypothesis to discuss the generalized continuum hypothesis. Going back to hypothesis 5.1.1 more generally, we would say that for sets, A and B , if $|A| \leq |B| \leq |\mathcal{P}(A)|$ then either $|B| = |A|$ or $|B| = |\mathcal{P}(A)|$. This is the main idea of the generalized continuum hypothesis; however, we want to understand it in terms of cardinal numbers. We take the infinite cardinals are ordered by magnitude. We can recall the following from chapter 3.

$$\aleph_0 < \aleph_1 < \aleph_2 < \aleph_3 < \dots$$

Then, we can form the generalized continuum hypothesis [7].

Hypothesis 5.1.3 (Generalized Continuum Hypothesis). *For every cardinal number \aleph_α , its successor is the next immediate cardinal.*

$$\aleph_\alpha^+ = \aleph_{\alpha+1} = 2^{\aleph_\alpha}$$

This means we would have

$$\aleph_0 < \aleph_1 = 2^{\aleph_0} < \aleph_2 = 2^{\aleph_1} < \aleph_3 = 2^{\aleph_2} < \dots$$

$$\aleph_0 < \aleph_1 = 2^{\aleph_0} < \aleph_2 = 2^{2^{\aleph_0}} < \aleph_3 = 2^{2^{2^{\aleph_0}}} < \dots$$

These three theorems introduce and state the continuum hypothesis, and give us the background needed to discuss its more significant results.

5.2 PROVABILITY

One of the main reasons why the continuum hypothesis became well-known in the world of mathematics is due to its unprovability. It challenged the prior notion that every statement in mathematics can either be proven or disproven. Mathematician Kurt Gödel proved that the statement cannot be proven false. However, another mathematician, Paul Cohen, later proved that it cannot be proven true. We will only discuss some background and strategies of Gödel's proof as giving his formal proof is beyond the scope of this project. Gödel's proof consists of proving three theorems. Before we state those theorems, we need some more background on the concepts he uses beginning with the **universe of sets** V discovered by von Neumann.

We call V the universe of all sets which may seem like a fact that we should just accept, but let's try to understand exactly what that means and define it more explicitly. Within V there exists a hierarchy such that each rank in the hierarchy is based on ordinal numbers. For this reason, V can also be referred to as the cumulative hierarchy of sets or the von Neumann Hierarchy. This hierarchy is created by iterations of power sets [11]. The first level of the hierarchy V_0 consists

of the empty set.

$$V_0 = \emptyset$$

Then for each ordinal number α , there is a corresponding rank V_α in the hierarchy. The next rank succeeding V_0 would be V_1 . To get this we use the power set $\mathcal{P}(\emptyset)$ [5]. We can see how these iterations might look by finding the next few succeeding ranks.

$$V_1 = \mathcal{P}(\emptyset) = \{\emptyset\}$$

$$V_2 = \mathcal{P}(\{\emptyset\}) = \{\emptyset, \{\emptyset\}\}$$

$$V_3 = \mathcal{P}(\{\emptyset, \{\emptyset\}\}) = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\}$$

Then the process continues forever forming the hierarchy of sets.

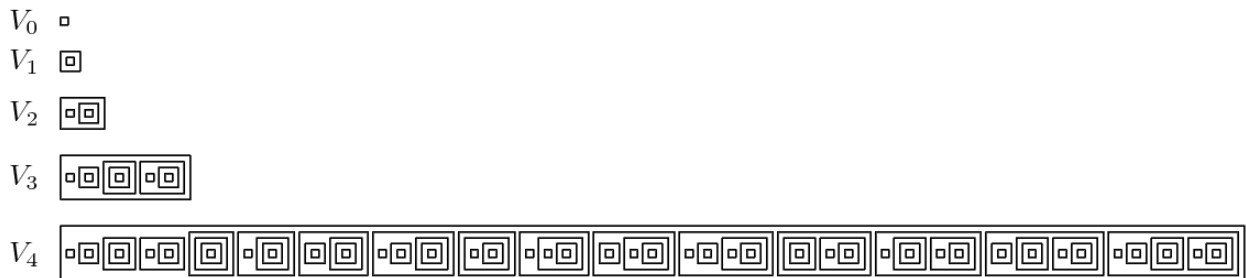


Figure 5.1: Visual representation of the cumulative hierarchy of sets [3].

In fig. 5.1, we can visually see how this hierarchy works. In this image, the singular square in V_0 represents the empty set. Then, V_1 is the set of the empty set since there is another square around our initial square. Every time there is another square, we can read that as "the set of" whatever is inside it. For every rank $V_{\alpha+1}$, $V_{\alpha+1}$ will contain all subsets of V_α that is,

$$V_{\alpha+1} = \mathcal{P}(V_\alpha)$$

Additionally, $V_\alpha \subseteq V_{\alpha+1}$. From what we have discussed we are only dealing with finite ordinals. Let's look at what will happen if we have a limit ordinal. Let's suppose we have a limit ordinal γ . Then,

$$V_\gamma = \bigcup_{\alpha < \gamma} V_\alpha$$

Note, with the limit ordinals we are not taking any power set operation since we form limit ordinals by collecting all the previous ordinals, so we are not introducing anything new. Up to this point, we still have not established exactly why V is the universe of all sets. As an example, consider this: we are able to write all the natural numbers in terms of empty sets. This means that every set of natural numbers is also a collection of elements from some rank. And since we are including limit ordinals, the power set of the natural numbers $\mathcal{P}(\mathbb{N})$ is also contained in one of the ranks. And since $|\mathcal{P}(\mathbb{N})| = |\mathbb{R}|$, we can also represent any collection of real numbers in terms of sets of empty sets.

Now that we have discussed the universe of sets, let's discuss the **constructible universe of sets** L . This universe is an inner model of Zermelo-Fraenkel set theory. By **inner model** we mean that it is a substructure of Zermelo-Fraenkel set theory, so both its domain and relations are restricted to that in ZF.

The constructible universe denoted L is similar to V but follows a few different principles. The hierarchy is still based on ordinal numbers, but instead, each level contains the sets in the previous level that satisfy a well-formed formula. Similar to V , we start with $L_0 = \emptyset$. Then $L_{\alpha+1}$ is defined as follows.

$$L_{\alpha+1} = \text{all } x \in L_\alpha \text{ that can be described by means of a well-formed formula}$$

If we have a limit ordinal γ then the following is true.

$$L_\gamma = \bigcup_{\alpha < \gamma} L_\alpha$$

The constructible universe L is the union of these sets in L_α over all ordinals.

$$L = \bigcup_{\alpha} L_\alpha$$

The sets that belong to L are called constructible sets. More specifically, the way constructing these hierarchies works is let's start with a constructed set L_α . Let's suppose that $\psi(v_n)$ is a well-formed formula as defined in the axiom of separation with one free variable v_n . If the sets a_1, a_2, \dots, a_m , which denote ψ , are in L_α then $X \in L_{\alpha+1}$ where X is all $x \in L_\alpha$ such that $\psi(x)$ holds. In other words, if an open sentence $a_1, a_2, \dots, a_m \in L_\alpha$, then the solutions $x \in L_\alpha$ for that open sentence, where $\psi(x)$ is true for $x \in L_\alpha$, make up the collection X . And $X \in L_{\alpha+1}$ [5]. Understanding the exact rules that this model follows is not as crucial for our sake. The importance here is that it is similar to our universe V , and with some rule exceptions we create a different model called the constructible universe. Noticing this difference gives us enough knowledge to state our next axiom, the axiom of constructibility.

Axiom 5.2.1 (Axiom of Constructibility). $V = L$, so $V = \bigcup_{\alpha} L_\alpha$.

This axiom, along with the rest of the Zermelo-Fraenkel axioms are the two axioms required for *constructible* set theory. The axiom of constructibility is consistent with ZFC and independent of ZFC, so we must add it as an additional axiom. A system is considered **consistent** if there are no contradicting statements within the system. When we say that an axiom is **independent** of ZFC, we mean that the axiom is not provable with just the axioms in ZFC [5]. This model of ZFC + ($V = L$) is an

extension to our model of ZFC, and the fact that this axiom is independent of ZFC will be a significant point to keep in mind as we begin discussing Gödel's proof.

We now have enough information to discuss Gödel's proof that the continuum hypothesis cannot be proven false. As I mentioned earlier, Gödel proves this result using three theorems which can be found in [2]. We will now state those theorems and try to understand exactly what they are saying and why they are useful. We will not formally prove any of the theorems as they are lengthy and require much more additional background.

Theorem 5.2.2. *Let A be any axiom of the Zermelo-Fraenkel axioms. Then, A_L is provable in ZF where A_L denotes the fact that the axiom is restricted to only the constructible sets $x \in L$.*

Essentially, what this first theorem states is that every axiom in Zermelo-Fraenkel set theory (we are not including the axiom of choice) is provable in our inner model L . To formally prove this, we would go through each axiom to show that they all are consistent in L ; however, we will omit this proof.

Theorem 5.2.3. *The axiom of constructibility restricted to the model L is provable in ZF. So $(V = L)_L$ is provable.*

This theorem provides us with a smaller point, yet it is still important. Here we are saying that a constructible set is constructible when it is relative to the constructible universe L .

Theorem 5.2.4. *The axiom of constructibility $(V = L)$ implies that the axiom of choice combined with the continuum hypothesis is provable in ZF.*

This final theorem shows the implication of the continuum hypothesis. By proving these three theorems, Gödel has shown that by constructing this extended model of ZFC using the axiom of constructibility, where the axiom itself is independent of

ZFC, the continuum hypothesis holds. Similarly, Paul Cohen constructed a model extending on ZFC using a method called *forcing*. In this model, the continuum hypothesis fails. Both mathematicians constructed models that were consistent and independent of ZFC. In one of these systems, the continuum hypothesis holds, and in the other, it fails. This leads us to the conclusion that the continuum hypothesis is not provable in ZFC.

5.3 SUBSEQUENT INQUIRIES AND IMPLICATIONS

The idea that a mathematical statement is unprovable may seem counter-intuitive, and many mathematicians would agree. However, I will leave the reader with one last theorem which allows for these kinds of results: Gödel's incompleteness theorem [15].

Theorem 5.3.1. *If S is a consistent formal system, then there is a statement of the language S which is true, but not provable in S .*

This result has been proven extensively by Gödel in 1931. If we take ZFC to be our consistent formal system, then this theorem allows for statements that are not provable within ZFC such as the continuum hypothesis. This theorem was influential as it had challenged prior ideas about mathematical truth and logic. There have been other statements that have been proven to be not provable in ZFC such as the diamond principle in the field of order theory. However, among these lists of statements we can use a statement to imply another, similar to how we did with the axiom of constructibility to imply the continuum hypothesis in theorem 5.2.4 and the axiom of choice to imply its equivalences in theorem 4.2.4.

We have stated that the continuum hypothesis is independent of ZFC which is even more convincing after stating Gödel's incompleteness theorem. However, we can take this one step further. Similar to how we created a model using ZFC

and the axiom of constructibility, we could create a model using ZFC and the continuum hypothesis. We could actually construct two models: one that takes the continuum hypothesis as an axiom and one that takes its negation as an axiom, that is $ZFC + CH$ and $ZFC + \sim CH$ respectively. We could utilize these models to prove other statements to be independent of ZFC. Mathematicians have already done so in the same way that Gödel used the axiom of constructibility to help prove the independence of the continuum hypothesis. An example of one of these statements is the Whitehead Problem in the field of abstract algebra. There have also been statements that are consistent in ZFC only if the continuum hypothesis is taken to be an axiom or its negation is taken to be an axiom. For example, an axiom called Martin's axiom in the field of set theory is consistent in ZFC only when the negation of the continuum hypothesis is included in the system.

We could take this one more step further and ask whether or not there are statements that are independent of $ZFC + CH$ or $ZFC + \sim CH$, or independent of both systems. We would find that there are in fact statements that are independent of one or both of these systems. Going beyond content covered in this independent study requires further investigation into this topic and exploring other statements independent of ZFC. However, we end with having gained an understanding of the important role that the continuum hypothesis plays in these inquiries of logic and consistency of mathematical systems.

AFTERWORD

With the completion of the project, I have gained a better understanding of different objects used in set theory, set theoretic systems in general, and the significance of the result of the continuum hypothesis. I was able to draw connections between concepts that at first I thought were unrelated. I think being able to make these connections demonstrates my improved expertise on the subject. There were many topics I struggled to understand throughout this process such as proving axiom of choice implications and the universe of constructible sets. However, through these struggles, I learned new approaches to different problems and determined which strategies worked best for me. Being able to accomplish this will be useful for any future endeavor I wish to pursue.

Moving forward with this project, I would like to continue exploring the ideas mentioned in the concluding paragraph. Specifically, I would like to look at how the continuum hypothesis may relate to statements in other fields besides set theory. I am curious to know more details about the implications that we would see if we treated the continuum hypothesis as an axiom. In addition, I would like to go into more detail with Gödel's proof and look further into Cohen's proof. I have an understanding of how the strategies of the proofs work, but I would like to have a better understanding of the specifics.

REFERENCES

1. Joan Bagaria. Set Theory. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2020 edition, 2020.
2. Paul J. Cohen. *Set Theorey and the Continuum Hypothesis*. W.A. Benjamin Inc., New York, NY, 1966.
3. Wikimedia Commons. File:von neumann universe 4.png — wikimedia commons, the free media repository, 2020. URL https://commons.wikimedia.org/w/index.php?title=File:Von_Neumann_universe_4.png&oldid=450829378. [Online; accessed 29-January-2022].
4. Wikimedia Commons. File:diagonal argument.svg — wikimedia commons, the free media repository, 2020. URL https://commons.wikimedia.org/w/index.php?title=File:Diagonal_argument.svg&oldid=454129138. [Online; accessed 17-January-2022].
5. Keith Devlin. *The Joy of Sets: Fundamentals of Contemporary Set Theory*. Springer-Verlag, New York, NY, 1993.
6. Herbert B. Enderton. *Elements of Set Theory*. Elsevier, San Diego, CA, 1977.
7. Theodore G. Faticoni. *The Mathematics of Infinity: A Guide to Great Ideas*. John Wiley & Sons, Hoboken, NJ, 2006.
8. Abraham A. Fraenkel. *Abstract Set Theory*. North Holland Publishing Company Amsterdam, Amsterdam, Netherlands, 1961.
9. A.G. Hamilton. *Numbers, Sets, and Axioms: the Apparatus of Mathematics*. Cambridge University Press, Cambridge, UK, 1982.
10. Richard Hammack. *Book of Proof*. Creative Commons, 3 edition, 2018.
11. Juliette Kennedy. Kurt Gödel. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Winter 2020 edition, 2020.

12. Alex Lopez-Ortiz. Relevance of the axiom of choice, 1998. URL <https://cs.uwaterloo.ca/~alopez-o/math-faq/node69.html>. [Online; accessed 17-January-2022].
13. GREGORY H. MOORE. Early History of the Generalized Continuum Hypothesis: 1878-1938. *The Bulletin of Symbolic Logic*, 17(4):489–532, 2011. ISSN 10798986. URL <http://www.jstor.org/stable/41302100>.
14. Dag Normann. <https://www.mn.uio.no/math/tjenester/kunnskap/kompendier/acwozl.pdf>, January 2012.
15. Janice Padula. The Logical Heart of a Classic Proof Revisited: A Guide to Godel's "Incompleteness" Theorems. *Australian Senior Mathematics Journal*, 25(1):32 – 44, 2011. ISSN 0819-4564. URL <http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ936613&site=ehost-live>.
16. Robert W. Vallin. *The Elements of Cantor Sets– With Applications*. John Wiley & Sons, Hoboken, NJ, 2013.
17. Wikipedia contributors. Cantor set — Wikipedia, the free encyclopedia, 2022. URL https://en.wikipedia.org/w/index.php?title=Cantor_set&oldid=1065545696. [Online; accessed 17-January-2022].

