Louisiana Tech University
Louisiana Tech Digital Commons

Mathematics Senior Capstone Papers | College of Engineering and Science

Spring 5-12-2020

# The Axiom of Choice and Related Topics

Bryan McCormick

# The Axiom of Choice and Related Topics

Bryan McCormick

May 12, 2020

Project Advisor: Dr. John Doyle

**Abstract**

In this paper I will be discussing the Axiom of Choice and its equivalent statements. The Axiom of Choice is an axiom of Zermelo-Fraenkel set theory that states that given a collection of non-empty sets, there exists a choice function which selects one element from each set to form a new set. The equivalents of the Axiom of Choice that I will be discussing include Zorn's Lemma, which states that a partially ordered set with every chain being bounded above contains a maximal element, and the Well-Ordering Theorem, which states that every set has a well ordering. In addition to proving the equivalence of these statements, I will be discussing the mathematics required to prove them individually, as well as each of their consequences across the field of mathematics.

*Keywords:* Axiom of Choice, Zorn's Lemma, Krull's Theorem, Well-Ordering Theorem

## 1 Introduction and Motivation

Prior to the 1870s, set theory had little formal structure, but instead was naïve and was defined by describing sets and their properties with real examples. Without any formal definitions or axioms, paradoxes were discovered. In the 1870s, mathematician Georg Cantor tried to formally define set theory by expanding on the discrete examples that were used to describe set theory at the time as described in [3]. In 1901, mathematician Bertrand Russell discovered Russell's Paradox as a result of Cantor's attempt at formalizing set theory. Cantor defined a set as any collection of objects, which turned out to be too weak of a definition. Russell's Paradox defined $R$ as the set of all sets that are not members of themselves. The paradox results when trying to decide if $R$ is in the set. If $R$ is in the set, then $R$ is included in itself, and therefore shouldn't be in the set. If $R$ isn't in the set, then by definition, it should be in the set. After Russell's Paradox was discovered, other mathematicians tried to fix Cantor's formalization of set theory so that there were not any paradoxes. Of those that did, Ernst Zermelo's and Abraham Fraenkel's formulation of set theory is the most used today. However, an additional axiom that Zermelo added to the axioms of Zermelo-Fraenkel set theory results in paradoxes, contradicting the goal of defining set theory free of paradoxes. One of these paradoxes, the Banach-Tarski Paradox, was the primary motivation for the research of the topics in this paper. In this paper, the additional axiom that Zermelo included in Zermelo-Fraenkel set theory that causes this paradox is discussed in detail, as well as the Banach-Tarski Paradox itself and various useful consequences of the axiom.

# 2 The Axiom of Choice

## 2.1 Zermelo-Fraenkel Axioms

The Zermelo-Fraenkel (ZF) Axioms are eight axioms that were created by Ernst Zermelo and Abraham Fraenkel that create the basis of modern set theory. The axioms are stated in [10] as follows:

1. Axiom of Extensionality: Let $X$ and $Y$ be sets. If $X$ and $Y$ have exactly the same elements, then $X = Y$.
2. Axiom of the Unordered Pair: For any $x$ and $y$ there exists a set $\{x, y\}$ that contains exactly $x$ and $y$.
3. Axiom of Subsets: If there exists some property $\varphi$ with parameter $p$, then for any set $X$ and $p$, there exists a set $Y = \{u \in X : \varphi(u, p)\}$ that contains all $u \in X$ with property $\varphi$.
4. Axiom of the Sum Set: For any set $X$ there exists a set $Y = \cup X$, which is the union of all elements of $X$.
5. Axiom of the Power Set: For any set $X$ there exists a set $Y = P(X)$, which is the set of all subsets of $X$.
6. Axiom of Infinity: There exists an infinite set.
7. Axiom of Replacement: If $F$ is a function, then for any set $X$ there exists a set $Y = F(X) = \{F(x) : x \in X\}$
8. Axiom of Foundation: Every nonempty set has a membership minimal element[1].

Zermelo and Fraenkel developed these axioms to create a notion of set theory that didn't result in any paradoxes. These axioms form what is called Zermelo-Fraenkel set theory, which is the basis for modern mathematics.

[1]Not to be confused with least element.

## 2.2 Statement of the Axiom of Choice

The eight Zermelo-Fraenkel axioms achieved their purpose of not resulting in any paradoxes. However, when trying to prove his idea of the Well-Ordering Theorem, Zermelo realized that he could not prove it with only the ZF axioms, so he stated a ninth axiom, the Axiom of Choice.

**Axiom 2.1** (Axiom of Choice)**.** Given a collection of non-empty sets, there exists a choice function which selects one element from each set to form a new set.

The Axiom of Choice allowed Zermelo to prove his Well-Ordering Theorem; however, other mathematicians of the time began to show that the inclusion of the Axiom of Choice in the ZF axioms allowed mathematical paradoxes to occur. This created a dilemma for mathematicians of the time. They could either utilize the Axiom of Choice, and with it accept its many uses as well as its paradoxes, or they could reject the axiom and the paradoxes that it created.

## 2.3 History of the Axiom of Choice

Zermelo first stated the Axiom of Choice in 1904 when he used it to prove his Well-Ordering Theorem. However, in doing so, he created a controversy among mathematicians of the time. The main criticism of the Axiom of Choice was that the choice function that it described was arbitrary. It stated no formal way to choose the elements from the sets, only that a choice was possible.

Mathematicians of the time, and to a smaller extent today, were split between two ideologies as described in [9]. The first was the belief that mathematics is a construct of the human mind, and as a result, mathematics must conform to what makes sense to the human mind. The other group believed that mathematics is natural. This group believed that if there was a consequence of mathematics that they discovered that didn't make sense but was logically sound, then it was to be believed, since everything does not necessarill make sense in nature when it is discovered. This sect of mathematicians believed that just because something can't be explained now, doesn't mean that it can't be explained ever.

Zermelo believed in the latter ideology, therefore to him, he didn't need to define his choice function. To him, it was merely something that existed in nature, and because it existed, he could use it. Other mathematicians of the time, such as Baire, Borel and Lebesgue [1], belonged to the former ideology, and therefore objected to Zermelo's choice function since it couldn't be defined for every set.

The axiom was debated throughout much of the early 20$^{\text{th}}$ century until the late 1930s when Kurt Gödel proved that it was consistent with the other axioms of set theory. In 1963, Paul Cohen expanded on Gödel's proof by proving that the axiom was also independent from the standard axioms of set theory. Due to the works of Gödel and Cohen, many mathematicians began to accept Zermelo's Axiom of Choice.

However, the problems that some mathematicians still have with the Axiom of Choice are not limited to differing ideologies. The Axiom of Choice also results in paradoxes; most notably, the Banach-Tarski Paradox.

## 2.4 The Banach-Tarski Paradox

The Banach-Tarski Paradox is a paradox that was discovered by Stefan Banach and Alfred Tarski in 1924. The Banach-Tarski Paradox is stated in [9] as follows:

**Theorem 2.2** (Banach-Tarski Paradox)**.** A solid ball may be seperated into a finite number of pieces and reassembled in such a way to create two solid balls, each identical in shape and volume to the original.

We will not prove the paradox in detail in this paper, for the full details of the proof are beyond the scope of this paper. However, a sketch of the proof will be given.

The first step in the proof is to define two different rotations in three-dimensional space. The first rotation, denoted $\tau$, is a rotation of $120°$ rotation about the $z$ axis. The second rotation, denoted $\sigma$, is a rotation of $180°$ about the line $z = x$. These two rotations can be combined and repeated to make an infinite amount of unique rotations. We denote the group of these infinite rotations as $G$. Then, $G$ is divided into three subsets, with rotations being placed in subsets based on their final combinations of $\tau$ and $\sigma$.

The next step of the proof defines a pole as a point on the sphere that remains fixed during a rotation and points out that each rotation in $G$ has two poles. Because there are infinite combinations of rotations, there are consequently infinite sets of poles. We define the set of all poles resulting from rotations in $G$ as $P$, and we defined the set $S - P$ as the set of every other point on the sphere. We then define an orbit as a set of points in $S - P$ that share a rotation in $G$. Using the Axiom of Choice, we can select one point from each orbit and form the set $C$. Now, the three subsets of rotations in $G$ are separately applied to $C$. Doing this results in three disjoint subsets of $S - P$ that together contain every point in $S - P$. Each of these subsets are congruent to each other, implying that they make up roughly one third of sphere. However, by the definitions of the rotations $\tau$ and $\sigma$, each subset is also equivalent to the union of the other two subsets, implying that each subset makes up one half of the sphere. This is called the Hausdorff Paradox. The Hausdorff Paradox allows us to separate $S - P$ into two copies of itself, resulting in six subsets of points plus the set of poles.

The final step of the proof takes the two copies of $S - P$ and reassembles them. This results in two nearly complete spheres that are each missing the set of poles. We can use the original set of poles to fill in the gaps in one sphere. After doing this, we have one complete sphere and a nearly complete sphere that is only missing poles. For each missing point in the sphere, there is a circle around the $y$ axis on the sphere, similar to the latitude lines we use to navigate earth's surface. Because there are infinitely many points on each of these circles and $\pi$ is irrational, we can "shift from infinity". This is a process where we move a point that is a distance $\pi$ from the missing point to the missing point. We then move the next point that is the same distance from the one we just moved into the now empty hole. Because we have an infinite number of points, we eventually fill the hole. This is similar to a popular example of "shifting from infinity" called Hilbert's Hotel. In Hilbert's Hotel, there are an infinite number of rooms and each is occupied. If one guest leaves, every other guest can move back one room and the hotel is still fully occupied. This is a property of infinite sets.

Now that we have "shifted from infinity" for every missing point in the second sphere, we have completed the second sphere and have two spheres that are identical to the original. This proof is dependent on the Axiom of Choice, for there is no other way to define the set $C$ without it. However, the subsets that we decompose the sphere into are not measurable sets, meaning they have no definable volume. So while the paradox is mathematically true, it cannot be used to duplicate items in the real world. The Banach-Tarski paradox is just a quirk of mathematics that has not been shown to have any real world applications or consequences.

# 3 Zorn's Lemma

## 3.1 Order Theory

Before discussing Zorn's Lemma, it is crucial to be familiar with order theory. Order theory is a subset of set theory that deals specifically with orders of sets. There are three types of orderings of sets as described in [8]: partial order, total order, and well-order.

**Definition 3.1.** A **partial order** is a binary relation that is reflexive, antisymmetric, and transitive.

Let $S$ be a set and $\prec$ be a binary relation. For $\prec$ to be reflexive means that for all $x \in S$, $x \preceq x$. Here, $\preceq$ means that either $x \prec x$ or $x = x$. To be antisymmetric, for all $x, y \in S$, if $x \preceq y$ and $y \preceq x$ then $x = y$. Finally, to be transitive implies that for all $x, y, z \in S$, if $x \prec y$ and $y \prec z$ then $x \prec z$.

For a set to be partially ordered, these conditions must be satisfied, however, since the set is partially ordered, not every element of the set may be comparable to every other element. A popular example is a set of people who are members of the same family. In this example, we define the partial order as "is a descendant of". In this example, a son is a descendent of a father and a grandson is a descendent of a son, but a son and his cousin cannot be compared, since neither is a descendant of the other. In this example, every member of the family is a descendant of someone else (except for the beginning of the genealogy).

The next type of order describes when every element of a set is comparable to every other element. We call this a total order.

**Definition 3.2.** A **total order** is a partial order that satisfies the trichotomy property.

Again, let $\prec$ be a binary relation on the set $S$. The trichotomy property states that for all $x, y \in S$, exactly one of the following is true: $x \prec y$, $y \prec x$, or $x = y$.

Consider the example of a partial order. Since the son is not a descendant of the cousin, the cousin is not a descendant of the son, and the son is not the same person as his cousin, the trichotomy property is not satisified, and the binary operation "is a descendant of" is not a total order, only a partial order. Now consider the set of real numbers $\mathbb{R}$. It is true that for any $x, y \in \mathbb{R}$, either $x < y$, $y < x$ or $x = y$. Therefore $<$ is a total order and $\mathbb{R}$ is a totally ordered set.

The final order to be discussed is a slightly stronger notion than a total order. We call it the well-order.

**Definition 3.3.** A **well-order** on a set $S$ is a total order that has the property that for any non-empty subset $T \subseteq S$, $T$ must have a least element.

Consider again the set of real numbers $\mathbb{R}$. It is true that $<$ is a total order on $\mathbb{R}$; however, considering the open interval $(0, 1)$, one can see that there is no least element of the subset of $\mathbb{R}$. Therefore $<$ is not a well-order on $\mathbb{R}$, and $\mathbb{R}$ is not a well-ordered set. Consider now the set of natural numbers $\mathbb{N}$. The binary relation $<$ satisifes the same total order as it does on the set of real numbers; however, in $\mathbb{N}$, one can see that every nonempty subset of $\mathbb{N}$ does in fact have a least element by the $<$ operator. Therefore $<$ is a well-order on $\mathbb{N}$ and the natural numbers are a well-ordered set.

There are a few more terms that are necessary to know in order to discuss Zorn's Lemma.

**Definition 3.4.** A **chain** is a totally ordered subset of a partially ordered set.

For example, consider once again the genealogy example of a partially ordered set. The subset containing a grandfather, his first son, and his first son's first son is a chain since it is a totally ordered subset of the partially ordered set.

**Definition 3.5.** Let $S$ be a partially ordered set with partial order $\prec$ and let $T$ be a subset of $S$. We say $x \in S$ is an **upper bound** for the subset $T$ if for every $a \in T$, $x \succeq a$.

Consider the subset of a family that contains a child and their siblings. The father of those children is an upper bound for the subset of siblings.

The final term to be defined before Zorn's Lemma is stated is a maximal element.

**Definition 3.6.** Let $S$ be a partially ordered set with partial order $\prec$. An element $x \in S$ is a **maximal element of *S*** if $x \not\prec y$ for all $y \in S$.

Note that this does not mean that $x \succ y$ for every $y \in S$, but only that there is no element of $S$ that is “larger” than $x$. For example, if a genealogy starts with a grandfather and contains his siblings and all of their descendants, the grandfather is a maximal element of the partially ordered set, since his ancestors are not members of the set. However, since his siblings are in the set and his siblings are not a descendant of him, he is not “larger” than every other member of the genealogy. It just means that no member of the family is “larger” than him.

Now that the necessary terms of order theory have been defined, Zorn’s Lemma can finally be stated.

## 3.2 Statement of Zorn’s Lemma

**Lemma 3.1** (Zorn’s Lemma)**.** Let $S$ be a partially ordered set with binary relation $\prec$, and suppose every chain of $S$ has an upper bound. Then $S$ has a maximal element.

Zorn’s Lemma will be proven later, but Zorn’s Lemma is a very useful theorem that is utilized in proofs of multitudes of other theorems across various fields of mathematics. For example, Zorn’s Lemma is used in linear algebra to prove that every vector space has a basis. In topology, Zorn’s Lemma is used to prove Tychonoff’s theorem that every product of compact spaces is compact. Another consequence of Zorn’s Lemma that is interesting is Krull’s Theorem in abstract algebra.

## 3.3 Krull’s Theorem

### 3.3.1 Rings, Ideals, and Maximal Ideals

Krull’s Theorem is a theorem in abstract algebra that deals with rings and ideals. A ring is a set with two binary operations that satisfy certain properties. Let $S$ be a set and let $+$ and $\cdot$ be binary operations called addition and multiplication respectively. We say $\langle S, +, \cdot\rangle$ is a **ring** if the following properties from [5] are true:

1. $S$ is closed under addition, that is if $x, y \in S$, then $x + y \in S$;
2. $S$ is closed under multiplication, that is if $x, y \in S$, then $x \cdot y \in S$;
3. $+$ is associative, meaning for any $x, y, z \in S$, $(x + y) + z = x + (y + z)$;
4. $\cdot$ is associative, meaning for any $x, y, z \in S$, $(x \cdot y) \cdot z = x \cdot (y \cdot z)$;
5. $+$ is commutative, meaning for any $x, y \in S$, $x + y = y + x$;
6. There is an additive identity 0 in $S$ such that for all $x \in S$, $x + 0 = x$;

7. For all $x \in S$, there is an additive inverse $-x$ such that $x + (-x) = 0$;
8. There is a multiplicative identity 1 in $S$, that is for all $x \in S$, $x \cdot 1 = x$;
9. Multiplication is left and right distributive with respect to addition, that is for all $x, y, z \in S$, $x \cdot (y + z) = (x \cdot y) + (x \cdot z)$ and $(x + y) \cdot z = (x \cdot z) + (y \cdot z)$.

For example, the set of integers $\mathbb{Z}$ satisfies these properties, meaning that the integers form a ring.

Let $\langle S, +, \cdot\rangle$ be a ring and let $N$ be an additive subgroup of that ring. An additive subgroup is a subset of the set $S$ with addition that satisfies properties 1, 6, and 7 from above. If for all $x \in S$, $xN \subseteq N$ and $Nx \subseteq N$, then $N$ is an ideal of $\langle S, +, \cdot\rangle$. This just means if you can multiply each element of $S$ by all of the elements of $N$ and get products that are also in $N$, then $N$ is an ideal. For example, the even numbers are an ideal in the set of integers $\mathbb{Z}$. It is true that the addition of two even numbers result in an even number, so property 1 is satisfied. We see that 0 is an even integer since $2 \cdot 0 = 0$, so 0 is in the subset of even integers. Therefore it is the additive identity for the even numbers and property 6 is satisfied. Finally, the additive inverses of all even numbers are also even, so property 7 is satisfied. Therefore, the even numbers are an additive subgroup of the integers. Now, since any integer multiplied by an even integer is also even, the even integers satisfy the properties of an ideal. Therefore the even integers are an ideal of the integers.

A maximal ideal of the ring $\langle S, +, \cdot\rangle$ is a proper ideal, meaning it is not just the ring itself, that is not contained by any other proper ideal. Similar to the notion of a maximal element of a partially ordered set, a maximal ring does not necessarily contain every other ideal. It just means that no other proper ideal contains it.

### 3.3.2 Statement and Proof of Krull's Theorem

This now leads to Krull's Theorem as stated in [7].

**Theorem 3.2** (Krull's Theorem)**.** Let $R$ be a nonzero ring. Then $R$ has a maximal ideal.

*Proof.* Let $S$ be the set of proper ideals of $R$ with the partial order of inclusion. The subset $\{0\}$ is a proper ideal since $r \cdot 0 = 0$ for all $r \in R$ and $R$ is a nonzero ring. Therefore $S$ is not empty. Now, let $C$ be a non-empty chain of proper ideals $I \subseteq S$ and let $U$ be the union of the ideals $I \in C$. Because every $I$ is a proper ideal, it does not contain 1, since $1 \cdot r = r$ for all $r \in R$. Therefore if $1 \in I$, then $I = R$. Since $U$ is a union of proper ideals, $U$ inherits the properties of an ideal from each $I$ and $U$ also does not contain 1 since no $I \in U$ contains 1. Therefore $U$ is a proper ideal of $R$. Since $U$ is a proper ideal of $R$, $U \in S$. Because $U$ is the union of all ideals $I$ in the chain $C$, $I \subseteq U$ for all $I \in C$. Therefore $U$ is an upper bound for $C$. Since $C$ is arbitrary, every chain in $S$ is bounded above. Because $S$ is nonempty and every chain in $S$ is bounded above, then by Zorn's Lemma, $S$ has a maximal element. Since $S$ is the set of proper ideals of $R$, this implies that $R$ has a maximal ideal. $\square$

Krull's Theorem is one example of the many useful consequences of Zorn's Lemma. Krull's Theorem is a theorem that is very useful in abstract algebra for proving theorems about ideals. Another unique result of Krull's Theorem is that it is equivalent to Zorn's Lemma. Mathematician Wilfrid Hodges proved in 1979 that by using the ZF Axioms of set theory and Krull's Theorem, Zorn's Lemma could be proven. Since we proved Krull's Theorem using Zorn's Lemma, they are equivalent statements.

# 4 The Well-Ordering Theorem

## 4.1 Cardinal and Ordinal Numbers

In order to discuss the Well-Ordering Theorem, it is critical to be familiar with ordinal numbers and how they differ from cardinal numbers. The cardinality of set $S$ is defined as the number of elements of the set $S$. For example, let $S = \{2, 3, 5\}$. In this example, the cardinality of $S$ is 3 and is denoted $|S| = 3$. Cardinality is a notion of size of sets; it merely describes how big a set is.

We define the cardinal numbers as the numbers used to measure cardinality of a set. Notice that in the example set $S$, $|S|$ is equal to the natural number 3. This is because for finite sets, cardinality can be easily described using the natural numbers (including zero) since there cannot be a negative amount of elements of a set or a fractional amount of elements of a set. Therefore, for finite sets, the cardinal numbers are the same as the natural numbers with zero. However, when dealing with infinite sets, the natural numbers no longer suffice. Consider the cardinality of $\mathbb{N}$. One cannot just find the greatest natural number and declare that number to be the cardinality of the natural numbers because the set $\mathbb{N}$ is infinite [2] and has no greatest element. Consequently, a new cardinal number must be defined to describe the cardinality of the set $\mathbb{N}$. We define $\aleph_0$ as this cardinal number.

Now that a way to measure the cardinality of an infinite set has been defined, the cardinality of other infinite sets can be measured. Consider the cardinality of the set of real numbers $\mathbb{R}$. Since the real numbers are infinite and the natural numbers are infinite, they may have the same cardinality. To check this, we can attempt to define a bijection between the naturals and the reals since sets with bijections between them have the same cardinality. In 1891, Georg Cantor proved that this was impossible. Cantor's diagonal argument showed that the set of real numbers $\mathbb{R}$ is uncountably infinite, implying that there are strictly more real numbers than natural numbers, despite both sets being infinite. Since there are more reals than naturals, the cardinality of $\mathbb{R}$ must be larger than the cardinality of $\mathbb{N}$. It has been shown to be impossible to prove whether the cardinality of the real numbers is the smallest notion of infinity that is greater than $\aleph_0$, a problem called the continuum hypothesis. Regardless of whether or not the continuum hypothesis is true, it is evident that not all infinite sets have the same cardinality. Therefore, additional cardinal numbers for larger infinite sets must be

[2] We call sets with the same cardinality as the natural numbers **countably infinite**.

defined. We define these numbers as $\aleph_1, \aleph_2, ..., \aleph_\omega, ...$ and so on. In conclusion, the cardinal numbers are the numbers used to describe the cardinalities of sets and are listed as follows:

$$0, 1, 2, ..., n, ..., \aleph_0, \aleph_1, \aleph_2, ..., \aleph_\omega, ...$$

where $n \in \mathbb{N}$.

Notice the last listed cardinal number $\aleph_\omega$. The number $\omega$ is what's called an ordinal number. Unlike the cardinal numbers, the ordinal numbers describe what place in a sequence an element of that sequence is at. More generally, ordinal numbers describe how elements of a well-ordered set are ordered. Consider again the set of natural numbers $\mathbb{N}$. One can see that 1 is the $1^{\text{st}}$ natural number by the usual ordering of $\mathbb{N}$, 2 is $2^{\text{nd}}$, 3 is the $3^{\text{rd}}$, and so on. To give a formal definition of the ordinal numbers, consider the empty set $\emptyset = \{\}$. We denote this as the ordinal number 0. Now consider the set containing 0, $\{0\}$. We denote this set as the ordinal number 1. Going forward, we'll define an ordinal number as the set containing all ordinal numbers prior to it. By this definition, $2 = \{0, 1\}$; $3 = \{0, 1, 2\}$; $4 = \{0, 1, 2, 3\}$ et cetera. This is called the Von Neumann definition of the ordinal numbers in [6]. Notice that for a finite ordinal number, the ordinal is equal to the cardinality of the set defining it, which is equal to the cardinal number of the same name. Now we have a sequence of ordinals that begins as

$$0, 1, 2, ..., \alpha, \alpha + 1, \alpha + 2, ...$$

continuing on infinitely. Note that $\alpha + 1$ is called a successor ordinal and it is defined as the smallest possible ordinal that is greater than $\alpha$.

Similar to the cardinal numbers, new ordinals have to be defined for infinite sequences. We define a limit ordinal as an ordinal that is not 0 and is not a successor ordinal. We say $\omega$ is the smallest limit ordinal and it is defined as being the first ordinal number greater than the natural numbers. The ordinal $\omega$ is obviously not 0, and since for every natural number less than $\omega$, you can add 1 and get another ordinal less than $\omega$, it is not a successor ordinal. Therefore $\omega$ is a limit ordinal. We then define the next successor ordinal after $\omega$ as $\omega+1$, then $\omega + 2$ after that, and so on until we get to $\omega + \omega = \omega \cdot 2$. Following this pattern, eventually we get to $\omega \cdot 3, ..., \omega \cdot 4, ...$ and so forth until we get to $\omega \cdot \omega = \omega^2$. Eventually we arrive at $\omega^\omega$, after which the ordinals continue on infinitely. It is important to note that there is no largest ordinal number. For any ordinal number, there is always an ordinal that succeeds it. Since we defined the ordinals as the set containing all ordinal numbers preceeding it, a set containing all ordinals cannot exist, since that set would be equivalent to the ordinal succeeding every ordinal. That would be an ordinal number by definition, and therefore that ordinal would be the set containing itself, which contradicts the ZF axioms of set theory.

The purpose of defining the ordinal numbers like this is so that we can apply the ordinal numbers to any well-ordered set up to order isomorphism. For example, say we have the ordered set $S = \{\alpha, \beta, \gamma\}$ where $\alpha < \beta < \gamma$. Then the subset $\{\emptyset\}$ is order isomorphic to the ordinal number 1. The subset $\{\emptyset, \alpha\}$ is order isomorphic to 2, and so on until every subset of $S$ is assigned an ordinal number. Now consider if $S$ is an infinite well-ordered set. We can still find order isomorphisms to $S$ since we defined limit ordinal numbers for infinite sequences. Because every well-ordered set can be "paired" with the ordinal numbers like this, it allows for the use of a proof technique called transfinite induction.

## 4.2 Transfinite Induction

Transfinite induction is similar to the typical mathematical induction that many people are familiar with. Mathematical induction is a proof technique that proves a theorem for a base case and then uses an induction step which assumes the proof is true for the first $n$ cases, and as a result shows that it is true for the $n+1$ case. If these steps can be shown, then the statement is true for every $n \in \mathbb{N}$. While this notion of induction is very useful for sequences of natural numbers, it doesn't always work for every sequence. This is where transfinite induction comes in. Transfinite induction is a proof technique with the following steps:

1. Prove the theorem is true for the 0 ordinal case.
2. Assume that the theorem is true for the ordinal number $\alpha$ case. Prove the theorem is true for the successor ordinal case $\alpha + 1$.
3. Assume that the theorem is true for the limit ordinal $\beta$. Prove that the theorem is true for any limit ordinal $\gamma$ where $\gamma > \beta$.

If all three of these conditions are met, then the theorem is true for every ordinal case. This allows induction to be applied to infinite sequences when the typical notion of mathematical induction could not.

## 4.3 Statement of the Well-Ordering Theorem

The Well-Ordering Theorem was the primary motivation for Ernst Zermelo to formulate the Axiom of Choice. Zermelo, along with other mathematicians at the time, like Georg Cantor, believed that the theorem logically must be true, and if it could be proven, it would have powerful consequences.

**Theorem 4.1** (Well-Ordering Theorem)**.** Every set can be well-ordered.

That is, every set can be given a relation, take $\preceq$ for example, such that every subset of the set has a "least" element. This relation however need not be the same $<$ relation that one uses with $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, or $\mathbb{R}$. In fact, $<$ only defines a well-ordering on $\mathbb{N}$. In $\mathbb{N}$, 1 is the least element for the set since $1 \leq n$ for all $n \in \mathbb{N}$. Since 1 is now defined as the smallest element of $\mathbb{N}$, we can now define the order $1 < 2 < 3 < ... < n < n+1 < n+2 < ...$ and so on for all $n \in \mathbb{N}$. Therefore the set $\mathbb{N}$ with the relation $<$ creates a well-ordered set.

Notice, however, that $<$ does not define a well-ordered set on $\mathbb{Z}$. Considering $1 \in \mathbb{Z}$, one can see that $0 < 1$. Therefore 1 is not the least element in $\mathbb{Z}$. Likewise, considering $0 \in \mathbb{Z}$, it can be seen that $-1 < 0$. In fact, for all $z \in \mathbb{Z}$ it is true that $z - 1 < z$. A well-ordering on $\mathbb{Z}$ still exists, however. Define $<_{\mathbb{Z}}$ as the relation where

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

Using this relation on $\mathbb{Z}$, we get $0 <_{\mathbb{Z}} -1 <_{\mathbb{Z}} 1 <_{\mathbb{Z}} -2 <_{\mathbb{Z}} 2 <_{\mathbb{Z}} ... <_{\mathbb{Z}} -z <_{\mathbb{Z}} z <_{\mathbb{Z}} ...$ and so on for all $z \in \mathbb{N}$. Therefore the set $\mathbb{Z}$ with the relation $<_{\mathbb{Z}}$ creates a well-ordered set.

A similar relation can be defined for $\mathbb{Q}$; however, no well-ordering has ever been defined for $\mathbb{R}$. Nonetheless, the Well-Ordering Theorem guarantees that a well-ordering of $\mathbb{R}$ does exist. In fact, the Well-Ordering Theorem guarantees that a well-ordering exists for every set, regardless of whether a well-order can be strictly defined.

### 4.4 Consequences of the Well-Ordering Theorem

There is only one primary consequence of the Well-Ordering Theorem to be discussed. That is that transfinite induction can be used on every set. Since every set has a well-order by the Well-Ordering Theorem, each element of an infinite set can be assigned a unique ordinal number, then transfinite induction can be utilized as described in Section 4.2.
Although it is not a direct consequence of the Well-Ordering Theorem, the Well-Ordering Principle is noteworthy because it defines the well-ordering of $\mathbb{N}$ that is guaranteed by the Well-Ordering Theorem. The Well-Ordering Principle states that every subset of $\mathbb{N}$ contains a least element. This is true by the definition of the well-order on $\mathbb{N}$ that was discussed in the previous section. The Well-Ordering Principle is useful when utilizing mathematical induction, since without a well-order on $\mathbb{N}$, induction could not be utilized. Although the amount of direct consequences of the Well-Ordering Theorem seems small, the ability to utilize induction and transfinite induction is a major proof strategy that has resulted in the proofs of many theorems across multiple fields of mathematics.

## 5 Equivalence of the Axiom of Choice, Zorn's Lemma, and the Well-Ordering Theorem

In previous sections, the Axiom of Choice, Zorn's Lemma, and the Well-Ordering Theorem were only stated and left unproven. Now we will prove that they are equivalent statements.

### 5.1 The Axiom of Choice Implies Zorn's Lemma

**Theorem 5.1.** Assume the Axiom of Choice is true. Then Zorn's Lemma is true.

*Proof.* We give a proof of Theorem 5.1 following [2]. Let $S$ be a partially ordered set. Define $F$ as the function that, when given an element $x$ of $S$, returns an element that succeeds $x$, denoted $x^+$, if $x$ is not maximal and returns $x$ if $x$ is maximal. By the axiom of choice, this function exists. Now define $G$ as the function that, when given a chain of $S$, returns an upper bound for that chain. By the axiom of choice, this function also exists. Now we index each element of $S$ with an ordinal number. Let $\alpha$ be an ordinal, and suppose $x_\beta$ has been defined for all $\beta < \alpha$. If $\alpha$ is a successor ordinal, then write $\alpha = \beta + 1$ for some $\beta$ and define $x_\alpha = x_\beta^+$. If $\alpha$ is limit ordinal, then we define $x_\alpha$ as an upper bound of the chain containing $x_\beta$ for all $\beta < \alpha$. Now suppose for a contradiction that the sequence of $x$'s is strictly increasing. Then this implies that there is a bijection between the ordinals and a subset of $S$. However, there cannot be a set that contains every ordinal number. Therefore the sequence of $x$'s is not strictly increasing. Therefore $S$ has a maximal element. Since $S$ is arbitrary, every partially ordered set with chains that are bounded above has a maximal element. $\square$

## 5.2 Zorn's Lemma Implies the Well-Ordering Theorem

**Theorem 5.2.** Assume Zorn's Lemma is true. Then the Well-Ordering Theorem is true.

*Proof.* The proof of Theorem 5.2 follows from [4]. Let $S$ be a nonempty set. Let $W = (T, \preceq)$ for all $T \subseteq S$ where $\preceq$ is a well-ordering on $T$. We define $\ll$ as follows:

$$(T_1, \preceq_1) \ll (T_2, \preceq_2) \text{ iff } T_1 \subseteq T_2 \text{ and for all } x, y \in T_1,\ x \preceq_1 y \text{ iff } x \preceq_2 y.$$

Then $W$ is partially ordered by $\ll$. Let $U$ be the union of all chains in $W$. Then $U$ is an upper bound for all sets in $W$ and is in $W$ by the definition of $\ll$. Therefore, by Zorn's Lemma, $W$ has a maximal element. Denote the maximal element of $W$ as $(M, \preceq_M)$. We claim that $M = S$. Suppose for a contradiction that this is not the case. Let $x_0 \in S \backslash M$. We define the ordering $\preceq'$ as follows:

$$\preceq' := \begin{cases} x \preceq_M y, & x, y \in M; \\ x_0 \preceq_M x, & x \in M. \end{cases}$$

This implies that $x_0 \in U$ and therefore $x_0 \in M$, negating the assumption that $x_0 \in S \backslash M$. Therefore $M = S$ and $S$ has a well-ordering $\preceq_M$. $\square$

## 5.3 The Well-Ordering Theorem Implies the Axiom of Choice

**Theorem 5.3.** Assume the Well-Ordering Theorem is true. Then the Axiom of Choice is true.

*Proof.* Let $C$ be a collection of arbitrary non-empty sets. Let $U$ be the union of sets in $C$. Since $U$ is a set, it has a well-ordering. Define $F$ as the function that, when given a set $S \in C$, returns the smallest element of $S$ by the well-ordering of $U$. Then $F$ is a choice function for the collection $C$. $\square$

# 6 Conclusion

Zorn's Lemma and the Well-Ordering Theorem are two very useful consequences of accepting the Axiom of Choice as an axiom of Zermelo-Fraenkel set theory. These two statements, as well as the Axiom of Choice itself, have many more uses and consequences across all of mathematics than just those that are discussed in this paper. Without the Axiom of Choice, modern mathematics would be radically different. While some paradoxes, such as the Banach-Tarski paradox, result from this axiom, their disadvantages are heavily outweighed by the advantages of accepting the Axiom of Choice as true. Because of this, many modern mathematicians choose to accept the axiom, despite these paradoxes.

# References

[1] John L. Bell. The axiom of choice. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2019 edition, 2019.

[2] Ken Brown. Zorn's lemma. `http://pi.math.cornell.edu/~kbrown/6310/zorn.pdf`. Accessed: 2019-12-09.

[3] José Ferreirós. The early development of set theory. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, summer 2019 edition, 2019.

[4] Gerald B. Folland. *Real Analysis: Modern Techniques and Their Applications 2nd Edition*. Wiley, 1999.

[5] John B. Fraleigh. *A First Course in Abstract Algebra, 7th Edition*. Pearson, New York City, New York, 2003.

[6] Jean Van Heijenoort. *From Frege to Gödel: A Source Book in Mathematical Logic*. Harvard University Press, 2002.

[7] Hideyuki Matsumura. *Commutative Ring Theory*. Cambridge University Press, 2012.

[8] Bernard Schröder. *Ordered Sets: An Introduction with Connections from Combinatorics to Topology*. Birkhäuser Basel, 2016.

[9] Leonard M. Wapner. *The Pea and the Sun: A Mathematical Paradox*. A K Peters, 888 Worcester Street, Suite 230, Wellesley, MA 02482, 2005.

[10] Eric W. Weisstein. Zermelo-fraenkel axioms. `http://mathworld.wolfram.com/Zermelo-FraenkelAxioms.html`. Accessed: 2020-02-27.