A BRIEF INTRODUCTION TO ZFC

CHRISTOPHER WILSON

Abstract. We present a basic axiomatic development of Zermelo-Fraenkel and Choice set theory, commonly abbreviated ZFC. This paper is aimed in particular at students of mathematics who are familiar with set theory from a "naive" perspective, and are interested in the underlying axiomatic development. We will quickly review some basic concepts of set theory, before focusing on the set theoretic definition of the natural numbers, and the equivalence of the axiom of choice, Zorn's lemma, and well ordering principle.

Contents

1. Motivation and Russel's Paradox

Before beginning with the Axioms of Zermelo-Fraenkel Set Theory (ZF), it is worthwhile to engage with the reader's intuitive notion of a set, and justify the axiomatic approach to set theory.

Informally, a set is often thought of as a collection of objects. The primitive, undefined concept of sets is that of belonging. If x and A are sets, either x is in A (in which case we write $x \in A$), or x is not in A (in which case we write $x \notin A$).

Note that this excludes duplication. Either an object is in a set or it is not, it doesn't make sense to think of an object as being in a set multiple times.

At this point, the intrepid reader may think sets could be created in a straightforward manner. Namely, we can define a set as being the collection of all objects x which satisfy a particular property $P(x)$. In such a system, if A were this set, we would write

$$A = \{x \mid P(x)\}.$$

Such a theory is said to allow *unrestricted comprehension.*

Theories of unrestricted comprehension were common in the early days of attempting to formalize mathematics via logic. A particularly famous example is the theory developed by Gottlob Frege in his work Grundgesetze der Arithmetik.

However, theories of unrestricted comprehension are fundamentally unsound. *Russel's Paradox* is a well known example of an inconsistency which arises in such a theory. Namely, if we assume unrestricted comprehension, we could consider the following set

$$A = \{x \mid x \notin x\}.$$

That is, A is the set of all sets which do not contain themselves. We proceed with the following question: is $A \in A$? If $A \in A$, then by the definition of A, we must conclude $A \notin A$. However, if $A \notin A$, then again by the definition of A, we conclude $A \in A$. Ultimately we are left with the highly problematic statement that

$$A \in A \iff A \notin A.$$

Thus, we will proceed with an axiomatic development which more carefully characterizes sets, and allows us to avoid such paradoxes.

2. The First Axioms of ZF Set Theory

2.1. **Exetension, Specification, and the Empty Set.** We will now state some of the basic axioms of ZF set theory with minimal discussion. It is assumed the reader has worked with sets before in some capacity, and this section will merely be formalizing what is already known. We begin with a basic statement about set equality.

Axiom of Extension. For every set A and every set B, $A = B$ if and only if for every set x, $x \in A$ if and only if $x \in B$.

In other words, sets are said to be equal if they contain precisely the same objects. This axiom tells us that when considering a set, the only thing which matters are what elements the set has.

Axiom of Specification. If A is a set and $P(x)$ is a formula of first order logic, then there exists a set B containing precisely each $x \in A$ such that $P(x)$ is true.

For instance,

$$P(n) := \exists k \in \mathbb{N} \text{ such that } n = 2k,$$

is a valid formula. If we applied the axiom of specification to the natural numbers using this sentence, the result would be precisely the even numbers.[1] In general, the set B given by the axiom of specification is usually written in the following way

$$B = \{x \in A \mid P(x)\}.$$

[1]The natural numbers are not yet formally defined, but hopefully this example is still illustrative.

This helps illustrate that B is dependent on A and $P(x)$. It immediate that B must be the unique set containing every $x \in A$ such that $P(x)$ holds by the axiom of extension.

At this point, we haven't actually asserted any sets exist; we may well be working in a vacuum. Let us suppose for the sake of argument (temporarily) a particular set A exists. Even with this we reach difficulties, for we don't know anything about A, and thus we will have trouble using it to build further sets. There is however, one set which can easily be constructed out of A. This is the set $\{x \in A \mid x \neq x\}$. Clearly this set can have no elements. This empty set seems like a reasonable starting point, so we discard the temporary notion of a given set, and instead make the existence of the empty set an axiom.

Axiom of Empty Set. There exists a set $\emptyset$ such that for every set x, $x \notin \emptyset$.

This set is also sometimes denoted $\{\}$.

2.2. **Subsets.** We now introduce a basic set theoretic concept.

Definition 2.1. A set A is said to be a *subset* of a set B if for every set x, $x \in A$ implies $x \in B$. When this is the case, we write $A \subset B$.

If $A \subset B$, and $A \neq B$, we write $A \subsetneq B$, and say A is a *proper subset* of B.

We conclude with the statement of two basic results.

Theorem 2.2. *$A = B$ if and only if $A \subset B$ and $B \subset A$.*

Theorem 2.3. *For every set A, $\emptyset \subset A$.*

2.3. **Pairing, Union, Intersection, and Difference.** At this point, we have a way to make "smaller" sets out of given sets. The following axioms allow us to "combine" given sets.

Axiom of Pairing. If a and b are sets, then there exists a set containing precisely a and b.

It is common to denote this set $\{a, b\}$. In the case where $a = b$, we would simply write $\{a\}$ (i.e, $\{a, a\} = \{a\}$) At this point, we may pair $\emptyset$ with itself to obtain the set $\{\emptyset\}$. We may repeat this process to obtain the series $\{\{\emptyset\}\}, \{\{\{\emptyset\}\}\}$, etc. We may also pair $\emptyset$ with $\{\emptyset\}$ to obtain $\{\emptyset, \{\emptyset\}\}$. It's easy to see there are many more combinations which we might make.

Rather than combining a collection of sets into one set, we may wish to combine the elements of a collection of sets into one set. The next axiom guarantees this is possible.

Axiom of Unions. Let $\mathcal{A}$ be a collection of sets. Then there exists a set $\bigcup_{A \in \mathcal{A}} A$ such that $x \in \bigcup_{A \in \mathcal{A}} A$ if and only if there exists an $A \in \mathcal{A}$ such that $x \in A$.

The uniqueness of the union over a collection is guaranteed by the axiom of extension. Usually we denote the union over a collection of two sets by $A \cup B$. That is,

$$A \cup B := \bigcup_{X \in \{A, B\}} X.$$

Note that a set x is an element of a union if *there exists* a set in the collection containing x. There is a symmetric concept, the intersection of a collection, that has the defining property that x is a member of the intersection if it is a member of *every* set in the collection. We restate this more formally.

Definition 2.4. Let $\mathcal{A}$ be a nonempty collection of sets. Let $X \in \mathcal{A}$. Then the *intersection* of $\mathcal{A}$ is the set

$$\bigcap_{A \in \mathcal{A}} A = \{x \in X \mid x \in A \text{ for every } A \in \mathcal{A}\}.$$

Note that we didn't need a new axiom to define this concept, it relied on the axiom of specification. In fact, the particular choice of $X \in \mathcal{A}$ does not matter. This is because $x \in \bigcap_{A \in \mathcal{A}} A$ if and only if $x \in A$ for every $A \in \mathcal{A}$, regardless of which $X \in \mathcal{A}$ is chosen, so the axiom of extension guarantees all of these sets are the same.

Also, unlike unions, we were unable to define the intersection over an empty collection.[2] As before, we can define the intersection between two sets as

$$A \cap B := \bigcap_{X \in \{A,B\}} X.$$

We proceed to define one more operation between two sets.

Definition 2.5. Let A and B be sets. The *difference* of A and B is the set

$$A \setminus B = \{x \in A \mid x \notin B\}.$$

These three operations have a variety of interesting and useful properties, which we assume the reader is already familiar with, and thus omit (see [1], pg. 17).

2.4. **The Power Set, Ordered Pairs, Relations, and Functions.** We have been considering a variety of special subsets. The next axiom allows us to collect every subset of a given set into a set.

Axiom of powers. For every set X, there exists a set $\mathcal{P}(X)$ such that $A \in \mathcal{P}(X)$ if and only if $A \subset X$.

This axiom will prove to be powerful tool in our construction of relations and functions. To begin, we will need to formalize the concept of an ordered pair.

Definition 2.6. Let a and b be sets. The *ordered pair* (a, b) is the set $\{\{a\}, \{a, b\}\}$.

Theorem 2.7. *$(a, b) = (c, d)$ if and only if $a = c$ and $b = d$.*

This theorem allows us to think of ordered pairs in the usual way; given an ordered pair (a, b), we can think of a as the *first coordinate* of (a, b), and b the *second coordinate*.

Suppose A and B are given sets, $a \in A$, and $b \in B$. Note then that the ordered pair (a, b) is an element of $\mathcal{P}(\mathcal{P}(A \cup B))$. This allows for the following definition

Definition 2.8. Let A and B be given sets. The *cartesian product* of A and B is

$$A \times B = \{z \in \mathcal{P}(\mathcal{P}(A \cup B)) \mid z = (a, b) \text{ for some } a \in A, b \in B\}.$$

[2]If we tried to ignore this, $x \in \bigcap_{A \in \emptyset} A$ if and only if $x \in A$ for every $A \in \emptyset$. Clearly, every set x satisfies this criteria, leading to the intersection being the set of all sets, which is problematic.

Then in particular, $A \times B$ is the set of all ordered pairs (a, b) with $a \in A$ and $b \in B$. With this in hand, we can now define some more complex structures.

Definition 2.9. Let A and B be sets. A *relation* R from A to B is a subset of $A \times B$. If $(a, b) \in R$, we write aRb.

The *domain* and *range* of R are respectively the sets

$$\mathrm{dom}(R) = \{a \in A \mid \exists b \in B \text{ s.t. } aRb\}, \qquad \mathrm{ran}(R) = \{b \in B \mid \exists a \in A \text{ s.t. } aRb\}.$$

If R is a relation from A to A, we say R is a relation in A.

Definition 2.10. A *function* f from A to B is a relation from A to B with the following property. For every $a \in A$, there exists a unique $b \in B$ such that $(a, b) \in f$. Given $a \in A$, we write $f(a)$ for the element of B such that $(a, f(a)) \in f$.

If f is a function from A to B, we write $f : A \to B$. A is said to be the *domain* of f, and B is the *codomain*.

We conclude this section with some useful concepts relating to functions.

Definition 2.11. Let $f : A \to B$. f is said to be *injective* if $f(x) = f(y)$ implies $x = y$ for every $x, y \in A$. f is said to be *surjective* if for every $b \in B$, there exists an $a \in A$ such that $f(a) = b$. f is said to be *bijective* if f is injective and surjective.

Definition 2.12. Let $f : A \to B$, and $A' \subset A$. The *restriction* of f to A' is the function $f|_{A'}$, which maps from A' to B, and is defined by setting $f|_{A'}(x) = f(x)$ for all $x \in A'$.

Definition 2.13. A *family* of sets is a function A with domain I. When A is a family over the set I, we write $\{A_i\}_{i \in I}$, and A_i for $A(i)$.

3. The Natural Numbers

3.1. **Peano Axioms, The Axiom of Infinity.** At this point, we have discussed most of the axioms of ZF set theory. There are two other axioms traditionally included, the axiom of foundation, and the axiom of schema replacement. While these axioms are interesting and have useful consequences when developing set theory in general, they are not used for any future results in this paper, and thus omitted.

This section is dedicated to discussing the final axiom of ZF set theory, the axiom of infinity, and a basic development of its consequences. For the purpose of this paper, the main consequence axiom of infinity is to allow for us to describe the set of natural numbers in terms of sets.[3] However, this is not the only way we might go about describing the natural numbers. One may be familiar with the with the Peano Axioms for the natural numbers, which are as follows.

Peano Axioms. The natural numbers are the set ω associated with a successor function $S : \omega \to \omega$, with the following five properties:

(I) There exists an element 0 in ω.
(II) If $n \in \omega$, then $S(n) \in \omega$.
(III) For all $n \in \omega$, $S(n) \neq 0$.

[3]As the name suggests, this axiom will also allow for the construction of infinite sets. Beyond this, the axiom of infinity also enables the construction of infinte ordinals and cardinals, which is discussed further in [1]. Much of the work to follow for the natural numbers can be extended to ordinals and cardinals in general.

(IV) If $S(n) = S(m)$, then $n = m$ (S is injective).
(V) Suppose $A \subset \omega$ with the following two properties:
  (i) $0 \in A$;
  (ii) $n \in A$ implies $S(n) \in A$.
  Then $A = \omega$.

Axiom V is commonly referred to as the principle of mathematical induction.

We could, at this point, take all of the Peano Axioms as axioms, and proceed with whatever mathematics we wished. However, in this case, each natural number would not be a set, but an atomic object given by the axioms. The axiom of infinity will allow us to create a particular set which satisfies the Peano Axioms, and whose elements are indeed sets which are compatible with our previous axioms.

With these as axioms as a guide, we wish to construct a set which satisfies the above properties, as well as a suitable function S. Our first guiding principle is that a given natural number should contain the number of elements it is describing. In other words, if n is a natural number, there should be n elements in the set n. This immediately forces us to define $0 = \emptyset$. From here, we should ask supposing n is defined as we wish, how should we define $S(n)$. We want $S(n)$ to have $n + 1$ elements, and n has n elements, so a reasonable candidate is $S(n) = n \cup \{n\}$. If we use the construction, we obtain the following results:

$$0 = \emptyset$$
$$1 = S(0) = \{\emptyset\} = \{0\}.$$
$$2 = S(1) = \{\emptyset, \{\emptyset\}\} = \{0, 1\}.$$
$$3 = S(2) = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} = \{0, 1, 2\}.$$

And so on. In general,

$$S(n) = \{0, 1, \ldots, n\}.$$

While it's clear the process could be repeated an arbitrary finite number of times, there is no way to collect all of our numbers into a set via a finite number of operations under our current axioms. A further problem is that the S described is not actually a function. This will not prove to be a problem, but to highlight this difference, we switch notation.

Definition 3.1. If x is any set, the *successor* of x is the set $x^+ = x \cup \{x\}$.

The $+$ operator is nothing deeper than a notational convenience, we could write $x \cup \{x\}$ in every place we will write x^+.

We are now nearly ready to formalize the above process with a new axiom, though to simplify its statement, we precede it with a short definition.

Definition 3.2. A set A is said to be a *inductive* (or a *successor set*) if:
(i) $0 \in A$ (where $0 = \emptyset$ as before).
(ii) $n \in A$ implies $n^+ \in A$.

Axiom of Infinity. There exists an inductive set.

Note the inductive set given by this axiom may not be the set of natural numbers. However, this axiom does guarantee the existence of the natural numbers, in the following sense.

Theorem 3.3. *There exists a minimal inductive set. That is, there is an inductive set ω such that for every inductive set B, $\omega \subset B$.*

Proof. Suppose $\mathcal{A}$ is a non-empty collection of inductive sets. Then $\bigcap_{A\in\mathcal{A}} A$ is inductive, as:

(i) Each set in $\mathcal{A}$ is inductive. Thus, 0 is in every set in $\mathcal{A}$, which implies 0 is in their intersection.
(ii) If $n \in \bigcap_{A\in\mathcal{A}} A$, then $n \in A$ for each $A \in \mathcal{A}$. This implies $S(n)$ in A for each $A \in \mathcal{A}$, and thus $S(n) \in \bigcap_{A\in\mathcal{A}} A$.

Let C be the inductive set given by the axiom of infinity. We consider the collection

$$I = \{A \subset C \mid A \text{ is inductive}\}.$$

Note this collection is non-empty, as $C \in I$. Let

$$\omega = \bigcap_{A\in I} A.$$

Then ω is inductive. Suppose B is any inductive set. Then $C \cap B \subset C$ is inductive, so $C \cap B \in I$. Thus, $\omega \subset C \cap B$. But $C \cap B \subset B$, so $\omega \subset B$. This completes the proof. □

Definition 3.4. The set of natural numbers is the minimal inductive set produced by 3.3. This set is denoted by ω. A natural number is an element of ω.

Theorem 3.5. *ω satisfies the Peano Axioms (I), (II), (III), and (V)*

Proof. (I) and (II) follow from the fact that ω is inductive. (III) is the case as for all $n \in \omega$, $n \in S(n)$, so $S(n) \neq \emptyset = 0$. (V) follows by noting A is an inductive set, and ω is minimally inductive, so $\omega \subset A$. □

The proof that ω satisfies (IV) is the most complex, and will require some preliminary results.

Definition 3.6. A set X is said to be *transitive* if each element is a subset, i.e., $a \in X$ implies $a \subset X$.

Lemma 3.7. *Let n be a natural number ($n \in \omega$). Then*

(a) If $x \in n$, then $n \not\subset x$.
(b) n is transitive.

Proof. (a) By induction. Let $S = \{n \in \omega \mid \text{for all } x \in n, n \not\subset x\}$. It's easy to see $0 \in S$. Assume $n \in S$, and consider n^+. Suppose there is some set x such that $n^+ \subset x$. $n \subset n^+$, so $n \subset x$, and thus $x \notin n$. Further, $n \subset n$, and thus $n \notin n$, which shows that $x \neq n$. Therefore, $x \notin n^+$, completing the induction.

(b) Also induction. Let $S = \{n \in \omega \mid n \text{ is transitive}\}$. $0 \in S$. Suppose $n \in S$, and let $x \in n^+$. Then either $x \in n$ or $x = n$. In the former case, we can conclude $x \subset n$, and thus $x \subset n^+$. In the latter case, $n \subset n^+$ by the definition of successor. From this, we conclude $n^+ \in S$, completing the induction.

□

Theorem 3.8. *ω satisfies Peano Axiom (IV). That is, for all $n, m \in \omega$, if $n^+ = m^+$, then $n = m$.*

Proof. If $n^+ = m^+$, then $n \in m^+$, so either $n = m$ or $n \in m$. We can similarly conclude either $m = n$ or $m \in n$. If $n \neq m$, then by 3.7 (b), $n \in m$ implies $n \subset m$. However, by 3.7 (a), $m \in n$ implies $n \not\subset m$, a contradiction. □

3.2. **The Recursion Theorem.** The previous section demonstrated how induction can be used to prove statements about ω. Induction is also useful for defining functions on ω, as the next theorem demonstrates.

Theorem 3.9 (Recursion Theorem)**.** *Let X be a set, $a \in X$, and $f : X \to X$. Then there exists a function $u : \omega \to X$ such that $u(0) = a$ and $u(n^+) = f(u(n))$ for all $n \in \omega$.*

Proof. Let

$$\mathscr{C} = \{A \subset \omega \times X \mid (0, a) \in A \text{ and } (\forall (n, x) \in A)\big((n, x) \in A \implies (n^+, f(x)) \in A\big)\}.$$

Note $\omega \times X \in \mathscr{C}$, so this collection is not empty. Let

$$u = \bigcap_{A \in \mathscr{C}} A.$$

It's clear $u \in \mathscr{C}$, and in fact u is the smallest set in $\mathscr{C}$ (that is, if $A \in \mathscr{C}$, then $u \subset A$). We wish to show u is a function. Let

$$S = \{n \in \omega \mid \text{there exists a unique } x \in X \text{ such that } (n, x) \in u\}.$$

We now wish to show $0 \in S$. Suppose not. Then there exists $b \in X$ such that $b \neq a$ and $(0, b) \in u$. But then $u \setminus \{(0, b)\}$ would still be a member of $\mathscr{C}$, as:

(i) $a \neq b$, and $(0, a) \in u$, so $(0, a) \in u \setminus \{(0, b)\}$.
(ii) If $(n, x) \in u \setminus \{(0, b)\}$, then $(n, x) \in u$. This implies $(n^+, f(x)) \in u$. But $n^+ \neq 0$, so $(n^+, f(x)) \in u \setminus \{(0, b))\}$.

Then $u \setminus \{(0, b)\} \in \mathscr{C}$. This implies $u \subset u \setminus \{(0, b)\}$, which is a contradiction. Therefore, $0 \in S$.

Suppose now $n \in S$. Then there exists a unique $x \in X$ such that $(n, x) \in u$. This implies $(n^+, f(x)) \in u$. Assume there is a $y \in X$ such that $y \neq f(x)$ and $(n^+, y) \in u$. Then $u \setminus \{(n^+, y)\} \in \mathscr{C}$, as:

(i) $n^+ \neq 0$, so $(0, a) \in u \setminus \{(n^+, y)\}$
(ii) Suppose $(m, t) \in u \setminus \{(n^+, y)\}$. If $n = m$, then $t = x$, and thus
$$(m^+, f(t)) = (n^+, f(x)) \in u \setminus \{(n^+, y)\}.$$
If $n \neq m$, then $n^+ \neq m^+$, so $(m^+, f(t)) \in u \setminus \{(n^+, y)\}$.

Then $u \setminus \{(n^+, y)\} \in \mathscr{C}$. This implies $u \subset u \setminus \{(n^+, y)\}$, which is a contradiction. Therefore, $n^+ \in S$. This completes the induction, so $S = \omega$, and thus we can conclude u is a function. Further, $u \in \mathscr{C}$, so u satisfies the necessary requirements. □

3.3. **Arithmetic of the Natural Numbers.** With Recursion Theorem in hand, we are now prepared to define arithmetic on the natural numbers.

Definition 3.10. Fix $m \in \omega$. Define $s_m : \omega \to \omega$ by setting $s_m(0) = m$ and $s_m(n^+) = (s_m(n))^+$.

We define the *sum* of m and n to be

$$m + n = s_m(n).$$

Addition is the binary operation given by $+$.

We now prove the expected properties are true.

Theorem 3.11. *Addition is associative and commutative. That is, for all $k, m, n \in \omega$,*

(a) $(k+m)+n = k+(m+n)$,
(b) $m+n = n+m$.

Proof. (a) We induct on n. First note,

$$(k+m)+0 = k+m = k+(m+0).$$

This shows the base case. Then if associativity holds at n,

$$(k+m)+n^+ = ((k+m)+n)^+ = (k+(m+n))^+ = k+(m+n)^+ = k+(m+n^+).$$

This shows that if associativity holds at n, then it holds at n^+, completing the inductive step.

(b) We begin by showing $0+n = n$ for all $n \in \omega$ by inducting on n. We have

$$0+0 = 0,$$

and if $0+n = n$ then

$$0+n^+ = (0+n)^+ = n^+.$$

Next, we wish to show $m^+ + n = (m+n)^+$ for all $n, m \in \omega$. We induct on n. We have,

$$m^+ + 0 = m^+ = (m+0)^+.$$

And if $m^+ + n = (m+n)^+$, then

$$m^+ + n^+ = (m^+ + n)^+ = ((m+n)^+)^+ = (m+n^+)^+.$$

We are now prepared to show $n+m = m+n$, again by induction on n. For the base case,

$$0+m = m = m+0.$$

And for the induction step,

$$n^+ + m = (n+m)^+ = (m+n)^+ = m+n^+.$$

□

We can think of addition as having been defined inductively by adding 1 to the sum of the previous sum, i.e., $m+0 := m$ and $m+(n+1) := (m+n)+1$. In a similar way, we can define a product $m \cdot n$ inductively by setting $m \cdot 0 := 0$ and $m \cdot (n+1) := m \cdot n + m$. A more formal formulation follows.

Definition 3.12. Fix $m \in \omega$. Define $p_m : \omega \to \omega$ inductively by setting $p_m(0) = 0$ and $p_m(n^+) = p_m(n) + m$.

The *product* of m and n is defined to be

$$m \cdot n = p_m(n).$$

Multiplication is the binary operation given by $\cdot$.

Theorem 3.13. *Multiplication distributes over addition. That is, for all* $k, m, n \in \omega$,

$$k(m+n) = km + kn.$$

and

$$(k+m)n = kn + mn.$$

Proof. Induct on n. The equality

$$k(m+0) = km = km + k0$$

satisfies the base case. For the inductive step, suppose $k(m+n) = km + kn$. Then

$$k(m+n^+) = k(m+n)^+ = k(m+n) + k = km + kn + k = km + kn^+.$$

The second equality is an equally simple induction. □

Corollary 3.14. *Multiplication is associative and commutative. That is, for all* $k, m, n \in \omega$*,*

(a) $(k \cdot m) \cdot n = k \cdot (m \cdot n)$*,*
(b) $m \cdot n = n \cdot m$*.*

Proof. (a) Induct on n. For the base case,

$$(k \cdot m) \cdot 0 = 0 = k \cdot (m \cdot 0).$$

Now suppose $(km)n = k(mn)$. Then

$$(km)n^+ = (km)n + km = k(mn) + km = k(mn + m) = k(mn^+).$$

(b) We need three preliminary results. The first is that for all $n \in \omega$, $n + 1 = n^+$. This follows directly from our definition of addition. The second is that for all $n \in \omega$, $n = 1 \cdot n$. The proof of this is a straightforward induction. The final result is that for all $n \in \omega$, $0 \cdot n = 0$. This is also a straightforward induction. With these three results in hand, we can show commutativity by induction on n. For the base case, we have

$$m \cdot 0 = 0 = 0 \cdot m.$$

And for the induction step,

$$m \cdot n^+ = m(n+1) = mn + m = nm + 1m = (n+1)m = n^+ m.$$

□

The last concept of arithmetic we will discuss is exponentiation.

Definition 3.15. Fix $m \in \omega$. Define $e_m : \omega \to \omega$ inductively by

$$e_m(0) = 1, \qquad e_m(n^+) = e_m(n) \cdot m.$$

Then m to the *power* n, written m^n, is defined to be $e_m(n)$.

Theorem 3.16. *For all* $k, m, n \in \omega$*:*

(a) $k^{m+n} = k^m k^n$*.*
(b) $k^n m^n = (km)^n$
(c) $(k^m)^n = k^{mn}$

Proof. (a) Induct on n. For the base case

$$k^{m+0} = k^m = k^m \cdot k^0.$$

And for the induction step,

$$k^{m+n^+} = k^{(m+n)^+} = k^{m+n} \cdot k = k^m k^n k = k^m k^{n^+}.$$

(b) Induct on n. For the base case,

$$(km)^0 = 1 = k^0 m^0.$$

And for the induction step,

$$(km)^{n^+} = (km)^n \cdot (km) = k^n m^n km = k^{n^+} m^{n^+}.$$

(c) Induct on n. For the base case,

$$(k^m)^0 = 1 = k^{m \cdot 0}.$$

And for the induction step,

$$(k^m)^{n^+} = (k^m)^n \cdot k^m = k^{mn} \cdot k^m.$$

Then by (a),

$$k^{mn} \cdot k^m = k^{mn+m} = k^{mn^+}.$$

□

3.4. **Order on ω.**

Definition 3.17. Let $m, n \in \omega$. m and n are *comparable* if $m \in n$, $n \in m$, or $m = n$.

Lemma 3.18. *All natural numbers are comparable.*

Proof. Let $n \in \omega$. Define

$$S_n = \{m \in \omega \mid m \text{ and } n \text{ are comparable } \}.$$

and then define

$$S = \{n \in \omega \mid S_n = \omega\}.$$

Then we wish to show $S = \omega$, which we will do by induction.

First, we must show $S(0) = \omega$, which we will also show by induction. $0 \in S(0)$ by definition $(0 = 0)$. Suppose $m \in S(0)$. Then either $0 \in m$ or $m = 0$ (clearly $m \in 0$ cannot be the case). In either case, this implies $0 \in m^+$, so $m^+ \in S(0)$.

Now suppose $S_n = \omega$. We wish to show $S(n^+) = \omega$ by induction on m. Because $S(0) = \omega$, $n^+ \in S(0)$ in particular, which implies n^+ and 0 are comparable, and thus $0 \in S(n^+)$.

Now suppose $m \in S(n^+)$. There are three cases:

(i) $n^+ \in m$. Then $n^+ \in m^+$, so $m^+ \in S(n^+)$.
(ii) $n^+ = m$. Then $n^+ \in m^+$, so $m^+ \in S(n^+)$.
(iii) $m \in n^+$. Then either $m \in n$ or $m = n$. In the later case, $m^+ = n^+$, so $m^+ \in S(m^+)$. In the former case, by 3.7 (a), $n \not\subset m$. Recall $S_n = \omega$, so $m^+ \in S_n$. Then we have three cases
- $m^+ = n$. Then $m^+ \in n^+$, so $m^+ \in S(n^+)$.
- $m^+ \in n$. Then $m^+ \in n^+$, so $m^+ \in S(n^+)$.
- $n \in m^+$. Then either $n \in m$ or $n = m$. In either case, we can conclude $n \subset m$ (by 3.7 (b)), which is a contradiction.

Therefore, in all cases, $m^+ \in S(n^+)$, so we can conclude $S(n^+) = \omega$, completing the induction. □

Corollary 3.19. *For any $n, m \in \omega$, exactly one of the following hold:*

- $n = m$
- $n \in m$

- $m \in n$

Proof. 3.18 shows that for any two natural numbers, one of the above hold. We can then show by 3.7 that if any one of the above holds, the other two must not hold. □

Corollary 3.20. *For all $n, m \in \omega$ such that $n \neq m$, $n \in m$ if and only if $n \subset m$.*

Proof. The forward direction is given by 3.7 (b). If $n \notin m$ and $n \neq m$, by 3.19, $m \in n$, which implies by 3.7 (a) that $n \not\subset m$. □

Definition 3.21. We let $<$ be a relation on ω such that $n < m$ when $n \in m$.

3.20 shows $n < m$ if and only if $n \subsetneq m$.

Definition 3.22. For all $n, m \in \omega$, $n \leq m$ if $n < m$ or $n = m$.

It follows again from 3.20 that $n \leq m$ if and only if $n \subset m$. An immediate corollary of the subset formulation is that this order is transitive. That is, if $a < b$ and $b < c$, then $a < c$ (and the same for the non-strict case).

We conclude this section by showing how arithmetic and order interact.

Lemma 3.23. *(a) If $m, n \in \omega$ such that $m < n$, then $m^+ \leq n$.*
(b) Let $m \in \omega$. Then for all $n \in \omega$, $m \leq m + n$.
(c) Let $q \in \omega$ such that $q > 0$. Then for all $n \in \omega$ such that $n > 0$, $qn > 0$.

Proof. (a) If $n < m^+$, then $n \in m$ or $m = n$, and thus $n \leq m$.
(b) Induct on n. For the base case, $m = m + 0 \geq m + 0$. For the induction step, $m + n \leq (m + n)^+ = m + n^+$ by the order definition, and $m \leq m + n$ by the induction hypothesis, so $m \leq m + n^+$.
(c) Let $S = \{n \in \omega \mid n = 0 \text{ or } (n > 0 \text{ and } qn > 0\}$ and induct. $0 \in S$ by definition. If $n \in \omega$, either $n = 0$, in which case $q0^+ = q > 0$, or $n > 0$, in which case $qn^+ = qn + q \geq qn > 0$.

□

Theorem 3.24. *Let $k, m \in \omega$. $k < m$ if and only if there exists $n \in \omega$ such that $n > 0$ and $k + n = m$.*

Proof. First we show if there exists $n > 0$ such that $k + n = m$, then $k < m$. We do this by contrapositive. Suppose $m \leq k$. Then we wish to show for all $n \in \omega$ such that $n > 0$, $k + n \neq m$. We consider the set

$$S = \{n \in \omega \mid n = 0 \text{ or } (n > 0 \text{ and } m < k + n\}.$$

$0 \in S$ by definition. Suppose $n \in S$. There are two cases:

(i) $n = 0$. Then either $m = k$ or $m < k$. It's clear that in either case $m < k^+ = k + 0^+$, so the induction proceeds.
(ii) $n \neq 0$. Then $m < k + n < k + n^+$.

This completes the induction, so we can conclude that if $n \neq 0$, then $m \neq k + n$.

Suppose now $k < m$, and assume for contradiction $k + n \neq m$ for all $n > 0$. This implies $k + n \neq m$ for all $n \in \omega$. We will show this implies $k + n < m$ for all $n \in \omega$ by induction. The base case is clear from our assumption. Suppose now $k + n < m$. Then by 3.23 (a), $k + n^+ \leq m$. However, $k + n^+ \neq m$ by our assumption, so we are left with $k + n^+ < m$, completing the induction.

Then in particular, $m \leq k + m < m$ by 3.23 (b), which is a contradiction. □

Corollary 3.25. *Let $k, m, n \in \omega$ such that $k < m$. Then*

(a) $k + n < m + n$.
(b) $kn < mn$ when $n > 0$.

Proof. (a) If $k < m$, by 3.24, there exists $q \in \omega$ such that $q > 0$ and $k + q = m$. Then $(k + n) + q = m + n$, so by 3.24, $k + n < m + n$.
(b) Again, by 3.24, there exists $q > 0$ such that $k + q = m$. Then $kn + qn = mn$. Further, if $n > 0$, then $qn > 0$ by 3.23 (c), so we can conclude by 3.24 $kn < mn$. □

4. Finite and Infinite Sets, Partial Order, and Choice

At this point, we have successfully completed much (though certainly not all) of the development ZF set theory necessary for further mathematics outside set theory. However, most mathematicians today use Zermelo-Fraenkel as well as an additional axiom, the axiom of choice, as their underlying axiom system (abbreviated ZFC). The goal of this section is to state this axiom, and show it is equivalent to two other statements: Zorn's lemma, and well ordering principle.[4]

4.1. **Finite and Infinite Sets.** Before discussing axiom of choice, we digress briefly to discuss finite and infinite sets.

Definition 4.1. Let R be a relation on a set X.
We say R is *reflexive* if for all $x \in R$,

$$xRx.$$

We say R is *symmetric* if for all $x, y \in R$,

$$xRy \quad \text{implies} \quad yRx.$$

We say R is *asymmetric* if for all $x, y \in R$,

$$xRy \text{ and } yRx \quad \text{implies} \quad x = y.$$

We say R is *transitive* if for all $x, y, z \in R$,

$$xRy \text{ and } yRz \quad \text{implies} \quad xRz.$$

Definition 4.2. A relation R on a set X is an *equivalence relation* if R is reflexive, symmetric, and transitive.

Equality is perhaps the most obvious equivalence relation. The following definition is in fact another example of an equivalence relation (the proof of this is left as an exercise).

Definition 4.3. Two sets E and F are said to be equivalent (written $E \sim F$) if there exists a bijection between E and F.

Definition 4.4. Let E be a set. If E is equivalent to some natural number, then E is *finite*. If E is equivalent to no natural number, E is infinite.

We now set out to show finite sets are equivalent to a unique natural number.

[4]One important fact which we will not prove in this paper is that the axiom of choice is independent from the axioms of ZF. That is, the axioms of ZF do not imply the axiom of choice, nor do they imply not axiom of choice.

Lemma 4.5 (Pigeonhole Principle)**.** *If $n \in \omega$, then n is not equivalent to any of its proper subsets.*

Proof. We will show by induction that if $f : n \to n$ is injective, then f is surjective. Let

$$S = \{n \in \omega \mid \text{Every injective funciton } f : n \to n \text{ is surjective}\}.$$

The empty function is the only function from 0 to 0, and this function is bijective, so $0 \in S$. Suppose $n \in S$, and $f : n^+ \to n^+$ is injective.

(i) Suppose $\text{ran}(f|_n) \subset n$. Because $f|_n$ is injective, by our inductive hypothesis, $\text{ran}(f_n) = n$. Then because f is injective, we can conclude $f(n) = n$, and thus that f is surjective.

(ii) Now suppose $\text{ran}(f|_n) \not\subset n$. Then there exists $k \in n$ such that $f(k) = n$. We define $g : n^+ \to n^+$ to be $(f \setminus \{(k, f(k)), (n, f(n)\}) \cup \{(k, f(n)), (n, f(k))\}$. It's easy to see g is injective, and has the same range as f. In fact, $g|_n \subset n$, so by (i), we can conclude g is surjective, and thus f is surjective.

Therefore, $n^+ \in S$, completing the induction. □

Corollary 4.6. *If E is a set, and $n \in \omega$ such that $E \sim n$, then n is the unique natural number equivalent to E.*

Proof. If $m \neq n$, either $m \subsetneq n$ or $n \subsetneq m$. 4.5 shows that $n \not\sim m$, so from this we can conclude $E \not\sim m$. □

This allows for the following definition.

Definition 4.7. If E is finite, the *cardinality* of E is the natural number which is equivalent to E, denoted $|E|$.

We include one more interesting result.

Corollary 4.8. *ω is infinite.*

Proof. By 4.5, we can conclude if a set is finite, then it is not equivalent to any of its proper subsets. Thus, if a set is equivalent to one of its proper subsets, it is infinite. In particular, $\omega \sim \omega \setminus \{0\}$ under the mapping $n \mapsto n^+$. □

4.2. **Axiom of Choice.** A first observation is if we are given a non-empty set, then we may choose an element of that set. Specifically, there exists $x \in X$, and if we needed a specific element of X for any reason, we can denote that element by x and proceed. The next natural question to ask is what if, rather than one set, we had a collection of sets? The forthcoming theorem considers the finite case, but first some terminology.

Definition 4.9. Let X be a collection of sets. A *choice function* on X is a function $f : X \to \bigcup_{A \in X} A$ such that $f(A) \in A$ for all $A \in X$.

Proposition 4.10. *Let X be a collection of sets. If $\emptyset \in X$, then there does not exist a choice function on X.*

Proof. Suppose not, and let f be a choice function. Then $f(\emptyset) \in \emptyset$. □

Theorem 4.11. *Let X be a finite collection of sets. Then there exists a choice function on X if and only if no set in X is empty.*

Proof. 4.10 gives the forward direction. Suppose now X is a finite collection of non-empty sets. We induct on the cardinality of X. If $|X| = 0$, then the empty function is a choice function on X. Suppose a choice function exists for any set with cardinality n, and $|X| = n + 1$. Let $A \in X$. Then $|X \setminus \{A\}| = n$, so there is some choice function f on $X \setminus \{A\}$. Let $a \in A$. Then $f \cup \{(A, a)\}$ is a choice function on X. □

We now consider the infinite case. 4.10 shows that if a choice function exists on *any* collection of sets (even an infinite one), each set in the collection must not be empty. The other direction of this statement (if a collection of sets does not contain empty set, then there is a choice function on that set), seems obvious.

In some cases it's easy to name an example of a choice function.

Example 4.12. For $a, b \in \mathbb{R}$, let $(a, b) = \{x \in \mathbb{R} \mid a < x < b\}$ be the *interval* from a to b. Let X be the collection of all non-empty intervals. Note then an interval is non-empty if and only if $a < b$.

To define a choice function on X, let f map $(a, b) \mapsto \frac{a+b}{2}$.

In other cases however, there is no obvious choice function.

Example 4.13. Let X be the collection of non-empty subsets of $\mathbb{R}$. In order to define a choice function, we need a "rule" to assign elements of X to an element of $\bigcup_{A \in X} A$. Specifically, in 4.12, we applied the axiom of specification to $X \times \bigcup_{A \in X} A$. Here, we have no way to specify what element to assign to each subset. Further, we can not define X inductively like in 4.11 because X is infinite.

In fact, if we wish for a choice function to exist in general in the infinite case, we must take it as an axiom.

Axiom of Choice. Let X be an infinite collection of non-empty sets. Then there exists a choice function on X.

4.3. **Partial Order.** A useful application of Axiom of Choice is to the theory of partial orders. We begin with a variety of basic definitions.

Definition 4.14. Let $\preceq$ be a relation on X. If $\preceq$ is reflexive, asymmetric, and transitive, $\preceq$ is said to be a *partial order* on X.

If $x, y \in X$ such that $x \preceq y$ or $y \preceq x$, then x and y are *comparable*.

Of particular note is that in a partially ordered set, there may exist a pair of elements which is not comparable. The following definition introduces terminology for the case where all elements are comparable.

Definition 4.15. Let $\preceq$ be a partial order on X. $\preceq$ is a *total order* if every $x, y \in X$ are comparable.

Example 4.16. • $\leq$ on the natural numbers is a total order.
- Let X be any collection of sets. Then $\subseteq$ is a partial order on X.
- Let $\mathscr{F}$ be a collection of functions such that for all $f \in \mathscr{F}$, $\text{dom}(f) \subset X$ and $\text{ran}(f) \subset Y$. For $f, g \in \mathscr{F}$, say $f \preceq g$ if $\text{dom}(f) \subset \text{dom}(g)$ and for all $x \in \text{dom}(f)$, $f(x) = g(x)$. Then $\preceq$ is a partial order on $\mathscr{F}$.

Definition 4.17. Let $\preceq$ be a partial order on X. For all $x, y \in X$, we write $x \prec y$ when $x \preceq y$ and $x \neq y$.

$\prec$ is a relation in X, which is transitive, and the property $x \prec y$ implies $y \not\preceq x$.

Definition 4.18. Let X be a partially ordered set, and $a \in X$. The (strict) *initial segment* of a is the set

$$s(a) = \{x \in X \mid x \prec a\}.$$

The *weak initial segment* is the set

$$\overline{s}(a) = \{x \in X \mid x \preceq a\}.$$

Definition 4.19. Let X be partially ordered, and $a \in X$.
a is *least, smallest,* or *first* if for all $x \in X$, $a \preceq x$.
a is *greatest, largest*, or *last* if for all $x \in X$, $x \preceq a$.
a is *minimal* if for all $x \in X$, $x \preceq a$ implies $a = x$.
a is *maximal* if for all $x \in X$, $a \preceq x$ implies $a = x$.

An easy equivalence is $a \in X$ is minimal if $x \not\prec a$ for all $x \in X$. Similarly, $a \in X$ is maximal if $a \not\prec x$ for all $x \in X$.

4.4. **Zorn's Lemma.** The first major consequence (in fact, equivalent statement to) axiom of choice is Zorn's lemma. The goal of this section is to prove that axiom of choice implies Zorn's Lemma.

From this point forward, when a set is partially ordered, we will notate the ordering as $\leq$ (which is not necessarily the ordering for the natural numbers). We begin with some relevant definitions.

Definition 4.20. Let X be partially ordered. A *chain* in X is a totally ordered subset of X.

That is, $A \subset X$ is a chain if for all $a, b \in A$, $a \leq b$ or $b \leq a$.

Definition 4.21. Let A be a chain in X. $u \in X$ is an upper bound of A if $a \leq u$ for all $a \in A$.

Of particular note is that a chain's upper bound may not be a member of the chain.

We are now prepared to state Zorn's lemma.

Theorem 4.22 (Zorn's lemma)**.** *Suppose X is a non-empty partially ordered set such that every chain in X has an upper bound. Then X contains a maximal element.*

We now fix X to be the set given in the hypothesis of Zorn's lemma. In order to begin with this proof, we will want to transform X into a specific partially ordered set for which we know the ordering. We let

$$\mathcal{X} = \{A \subset X \mid A \text{ is a chain in X}\}.$$

We let $\mathcal{X}$ be partially ordered by inclusion. That is, for $x, y \in \mathcal{X}$, $x \leq y$ when $x \subset y$.

Lemma 4.23. *(a) $\mathcal{X} \neq \emptyset$.*
(b) If $\mathcal{C}$ is a chain in $\mathcal{X}$, then $\bigcup_{C \in \mathcal{C}} C$ is an upper bound of $\mathcal{C}$.
(c) If $A \in \mathcal{X}$ is maximal, then X has a maximal element.

Proof. (a) $\emptyset$ is a chain in X.

(b) Let $\mathcal{C}$ be a chain in $\mathcal{X}$. We first need to show $\bigcup_{C\in\mathcal{C}} C \in \mathcal{X}$. Suppose $x, y \in \bigcup_{C\in\mathcal{C}} C$. Then there exist $C_x, C_y \in \mathcal{C}$ containing x and y respectively. $\mathcal{C}$ is a chain, so $C_x \subset C_y$ or $C_y \subset C_x$. If $C_x \subset C_y$, then $x, y \in C_y$, and C_y is a chain in X, so x and y are comparable. The argument is the same if $C_y \subset C_x$. Therefore, $\bigcup_{C\in\mathcal{C}} C$ is indeed a chain in X, and thus an element of $\mathcal{X}$.

From this, it immediately follows $\bigcup_{C\in\mathcal{C}} C$ is an upper bound of $\mathcal{C}$, because $\mathcal{X}$ is ordered by inclusion.

(c) Let $A \in \mathcal{X}$ be maximal. A is a chain in X, so let $u \in X$ be an upper bound of A. We wish to show u is maximal in X.

Suppose $z \in X$ such that $u \leq z$. For all $x \in A$, $x \leq u$, and thus $x \leq z$. Therefore, $A \cup \{z\}$ is a chain, and thus an element of $\mathcal{X}$. Further, $A \subset A \cup \{z\}$, and A is maximal, so $A = A \cup \{z\}$. This implies $z \in A$, and thus $z \leq u$. We can therefore conclude $z = u$, proving that u is maximal.

□

With this lemma, we now need to show $\mathcal{X}$ has a maximal element to prove Zorn's lemma. We are now prepared to begin proving the main result.

Let f be a choice function on $\mathcal{P}(X) \setminus \{\emptyset\}$ (this is where we invoke axiom of choice). For each $A \in \mathcal{X}$, let $\hat{A} = \{x \in X \mid A \cup \{x\} \in \mathcal{X}\}$. We define $g : \mathcal{X} \to \mathcal{X}$ as follows:

$$g(A) = \begin{cases} A \cup f(\hat{A} \setminus A) & \hat{A} \setminus A \neq \emptyset \\ A & \hat{A} \setminus A = \emptyset \end{cases}.$$

In other words, $\hat{A}$ contains every $x \in X$ such that if x is added to A, then the new set is still a chain. Thus, $A \cup f(\hat{A} \setminus A)$ is always in $\mathcal{X}$, so our definition makes sense. It also follows A is maximal if and only if $\hat{A} \setminus A = \emptyset$, so we wish to find an $A \in \mathcal{X}$ such that $g(A) = A$.

We now will introduce some temporary terminology.

Definition 4.24. $\tau \subset \mathcal{X}$ is a *tower* if:

- $\emptyset \in \tau$.
- If $A \in \tau$, then $g(A) \in \tau$.
- If $\mathcal{C}$ is a chain in τ, then $\bigcup_{C\in\mathcal{C}} C \in \tau$.

Note in particular $\mathcal{X}$ is a tower, so we may consider the set

$$\tau_0 = \bigcap_{\tau \text{ is a tower}} \tau.$$

It's easy to see τ_0 is itself a tower.

Definition 4.25. $C \in \tau_0$ is *comparable* if for every $A \in \tau_0$, either $A \subset C$ or $C \subset A$.

Lemma 4.26. *Let $C \in \tau_0$ be comparable. If $A \in \tau_0$ such that $A \subsetneq C$, then $g(A) \subset C$.*

Proof. $g(A) \in \tau_0$, so either $g(A) \subset C$ or $C \subsetneq g(A)$. In the latter case, we have

$$A \subsetneq C \subsetneq g(A),$$

which contradicts the definition of g. □

Lemma 4.27. *Let $C \in \tau_0$ be comparable, and $\mathcal{U} = \{A \in \tau_0 \mid A \subset C \text{ or } g(C) \subset A\}$. Then $\mathcal{U}$ is a tower.*

Proof. $\emptyset \subset C$, so $\emptyset \in \mathcal{U}$.

Suppose $A \in \mathcal{U}$. If $g(C) \subset A$, then $g(C) \subset g(A)$ as well, so $g(A) \in \mathcal{U}$. If $A \subset C$, then $A \subsetneq C$ implies $g(A) \subset C$ by 4.26, whereas $A = C$ implies $g(A) = g(C)$, so in either case $g(A) \in \mathcal{U}$.

Suppose $\mathcal{C}$ is a chain in $\mathcal{U}$. We consider two cases:

(i) For all $A \in \mathcal{C}$, $A \subset C$. Then $\bigcup_{A\in\mathcal{C}} A \subset C$, and is thus an element of $\mathcal{U}$.
(ii) There exists an $A \in \mathcal{C}$ such that $A \not\subset C$. Then $g(C) \subset A$, so $g(C) \subset \bigcup_{A\in\mathcal{C}} A$, and is thus an element of $\mathcal{U}$.

□

Corollary 4.28. *If $C \in \tau_0$ is comparable, then $g(C)$ is comparable.*

Proof. $\mathcal{U}$ is a tower, so $\tau_0 \subset \mathcal{U}$, and $\mathcal{U} \subset \tau_0$ by definition. Then each $A \in \tau_0$ is in $\mathcal{U}$, so either $A \subset C$, which implies $A \subset g(C)$, or $g(C) \subset A$. □

Lemma 4.29. *Let $\mathcal{V} = \{A \in \tau_0 \mid A \text{ is comparable }\}$. $\mathcal{V}$ is a tower.*

Proof. $\emptyset \in \mathcal{V}$.

By 4.28 if $A \in \mathcal{V}$, then $g(A) \in \mathcal{V}$.

Suppose $\mathcal{C}$ is a chain in $\mathcal{V}$, and $A \in \tau_0$. There are two cases:

(i) There exists $C \in \mathcal{C}$ such that $A \subset C$. Then $A \subset \bigcup_{C\in\mathcal{C}} C$.
(ii) For all $C \in \mathcal{C}$, $C \subset A$. Then $\bigcup_{C\in\mathcal{C}} C \subset A$.

Therefore, $\bigcup_{C\in\mathcal{C}} C$ is comparable, and thus an element of $\mathcal{V}$. □

From this, we can conclude $\tau_0 = \mathcal{V}$, and thus τ_0 is a chain. Let $A = \bigcup_{T\in\tau_0} T$. Then $A \in \tau_0$, and A is an upper bound of τ_0 (by 4.23 (b)). Then $g(A) \in \tau_0$, so $g(A) \subset A$. However, $A \subset g(A)$ by definition, so we can conclude $A = g(A)$. This suffices to prove Zorn's Lemma.

4.5. **Well Ordering.** The next major equivalent statement to axiom of choice and Zorn's lemma is the well ordering principle. The goal of this section is to show that Zorn's lemma implies the well ordering principle and then that the well ordering principle implies axiom of choice. We begin with a definition.

Definition 4.30. Let X be partially ordered. X is *well ordered* if each nonempty subset of X has a least element.

We can immediately strengthen this without losing generality.

Theorem 4.31. *Every well ordered set is totally ordered.*

Proof. Suppose X is well ordered, and $x, y \in X$. The $\{x, y\}$ is a subset of X, so it must have a least element. If x is the least element, $x \leq y$, and if y is the least element, $y \leq x$. Therefore, in either case, x and y are comparable. □

We even have a concrete example of a well ordered set.

Theorem 4.32. *The natural numbers are well ordered.*

Proof. Suppose $A \subset \omega$ has no least element. Let

$$S = \{n \in \omega \mid m < n \text{ implies } m \notin A\}.$$

We have $0 \in S$, as there are no numbers less than 0 in ω. Suppose now $n \in S$, and $m < n^+$. There are two cases:

(i) $m < n$. Then $m \notin A$ directly by the inductive hypothesis.
(ii) $m = n$. If $n \in A$, then the inductive hypothesis would imply n is the least element of A, which cannot be. Therefore, $m \notin A$.

From this, we can conclude $n^+ \in S$, completing the induction. This shows that $A = \emptyset$, which suffices to prove the theorem. □

In fact, well ordered sets share a property similar to Peano Axiom V.

Theorem 4.33 (Transfinite Induction)**.** *Suppose X is well ordered, and $S \subset X$ such that for all $x \in X$, if $s(x) \subset S$, then $x \in S$. Then $S = X$.*

Proof. Suppose not. Then $X \setminus S$ is a nonempty subset of X, so let x be its least element. Then each element of $s(x)$ must be an element of S, which implies $x \in S$, a contradiction. □

Definition 4.34. A well ordered set A is a *continuation* of a well ordered set B if:
- $B \subset A$.
- B is an initial segment of A.
- The ordering in B is preserved in A.

Note then that we can partially order any collection of initial segments of a well ordered set by continuation. We may say a collection is a chain with respect to continuation if for any pair of distinct sets in the collection, one is a continuation of the other.

There is a converse of this statement, that given a chain with respect to continuation, we can use it to build a well ordered set.

Lemma 4.35. *Let $\mathcal{C}$ be a chain with respect to continuation, and*

$$U = \bigcup_{C \in \mathcal{C}} C.$$

Then there exists a unique well ordering of U such that U is a continuation of each element of $\mathcal{C}$ which is distinct from U.

Proof. Let $a, b \in U$. Then there are initial segments A and B which contain a and b respectively. Then either $A = B$, or one is a continuation of the other. Therefore, we can conclude there is a set in $\mathcal{C}$ which contains both a and b. We can thus order a and b by copying the order from any set which contains both a and b. $\mathcal{C}$ is a chain, so this order must be the same for each set which contains a and b, so our definition is unambiguous.

This relation inherits reflexivity, anti-symmetry, and transitivity from the sets in $\mathcal{C}$, so this relation is indeed a partial order. Further, the choice of ordering is forced on us if we wish to ensure each set in $\mathcal{C}$ distinct from U is continued by U. Thus, this ordering is uniquely determined. We now need to show this ordering is a well ordering.

Suppose $A \subset U$ and $A \neq \emptyset$. Then there is some $C \in \mathcal{C}$ such that $A \cap C \neq \emptyset$. Further, $A \cap C \subset C$, and C is well ordered, so $A \cap C$ has a least element u. We will show u is in fact a least element of A. Suppose not. Then there exists $x \in A$ such that $x < u$. $A \subset U$, so let $x \in B \in \mathcal{C}$. Then $B \neq C$, as in particular $x \notin C$. However, $\mathcal{C}$ is a chain with respect to continuation, and $B \not\subset C$, so B must be a continuation of C. But then C is an initial segment in B, and $u \in C$, and $x < u$, so $x \in C$, which is a contradiction. □

With this, we are now prepared to prove the main result of this section.

Theorem 4.36 (Well Ordering Principle)**.** *If we assume the axioms of ZF and Zorn's lemma, then every set can be well ordered.*

Proof. Let X be a given set. Consider the collection

$$\mathcal{W} = \{Y \subset X \mid Y \text{ can be well ordered}\}.$$

Partially order $\mathcal{W}$ by continuation. $\emptyset \in \mathcal{W}$, so $\mathcal{W} \neq \emptyset$. By 4.35, each chain in $\mathcal{W}$ is bounded above by its union. Therefore, by Zorn's lemma, there exists $M \in \mathcal{W}$ which is maximal.

We now wish to show $M = X$. Suppose not. Let $x \in X \setminus M$. Then let $M' = M \cup \{x\}$. Order M' by copying the order from M, and adding $y < x$ for all $y \in M$. Then M' is a continuation of M, contradicting the maximality of M. □

In summary, assuming the axioms of ZF, we have now proven axiom of choice implies Zorn's lemma, and Zorn's lemma implies well ordering principle. The following theorem completes the equivalence of these three statements.

Theorem 4.37. *If we assume the axioms of ZF and well ordering principle, then axiom of choice holds.*

Proof. Let X be an infinite collection of non-empty sets. For each $A \in X$, let $\min(A)$ be the smallest element of A in its well ordering. Define a choice function f by mapping $A \mapsto \min(A)$. □

Acknowledgments. I would like to thank my mentor Tung Nguyen for his guidance and feedback. I would like to thank Professor May for his comments and organizing the REU. I would also like to thank Professor Malliaris for her comments. Lastly, I would like to acknowledge the extreme influence of Naive Set Theory by Halmos [1] on the structure and content of this paper. Most of proofs in this paper are the same or highly similar to those found in Halmos's text.

References

[1] Paul Halmos. *Naive Set Theory.* Van Nostrand. 1960.