Set Theory

Labix

October 28, 2024

Abstract

These notes aim to develop basic notions of sets and logics following closely to ZFC set theory, as well as introducing a few examples and provide proofs for theorems that require careful inspection. Theorems that with proofs ommitted are expected to have readers be able to prove them. Although that there is a wide variety of axiomatic set theories that are accepted by different mathematicians, ZFC set theory is considered to be one of the most widely recognized theories among them all.

Beware that while set theory has no formal prequisites in terms of its content (except perhaps logic theory), the mathematical concepts are not at all easy to understand. In fact the more foundational the mathematics, the harder the proofs and the more abstract the content becomes. First year university students should aim to complete up to chapter 3 in order to develop the suitable language for further mathematical content in their degrees.

Famous mathematicians who contributed to this area in mathematics include Ernst Zermelo and Abraham Fraenkel, Kurt Gödel and Georg Cantor and many more.

References

- Naive Set Theory by Paul R. Halmos

Contents

1 The first 3 Axioms

1.1 Introduction

Sets are widely used in a variety of topics in Mathematics. It shows up naturally whenever we are trying to "group up" a bunch of related stuff. Examples would be perhaps the set of all natural numbers, the set of all functions with a fixed point 1, or even the set of all matrices such that its determinant, if defined, is 1.

However, the concept of sets was only introduced in the 1900's. This was because long ago, there was simply no need for vigorous mathematics. There was no need to "prove" a theorem vigorously with logic and axioms until Gödel came along and seek to unify separate branches of mathematics. Nowadays, set theory has become a universal language that every Mathematician knows by heart.

We start with the concept of belonging. There is no good definition of a set except by allowing elements to belong to a set. This is how we shall characterize sets, as well as identify sets.

Definition 1.1.1: The Concept of Belonging

If x belongs to a set A we write $x \in A$.

So we have the concept of belonging. However, we don't even know if sets even exists in our world! This is how obnoxious mathematicians are. If not given a proof, they would not rest on the notion that a set could even exists. Unfortunately, there is no real tool for us to even prove that a set, any set exists at all.

Axiomatic set theory sort of solves this problem by stating some universally agreed assumptions. These are statements that cannot be proved and is a compromise for all mathematicians that they will work on theorems and definitions under these assumptions.

Currently, the only way we can construct sets are by simply listing elements that belong to that set. For example, if I say that $1, 2, 3$ belongs to the set S and nothing more, then we can write S as

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

We use the open and close brackets at the begining and the end to indicate the start and the end of the contents of the set.

1.2 Axiom of Extensionality

Whenever presented an axiom, the reader should be prompt automatically to think about the reason for introducing this axiom. Could it be formulated in another way? Could it not be an axiom?

Axiom 1.2.1: Axiom of Extensionality

Two sets A, B are equal if and only if they have the same elements. We write $A = B$ in this case.

Later as we shall see, sets can be constructed using predicates or statements. The axiom of extensionality allows to different predicates to result in the same set, provided that they have the same elements. Moreover, the number of copies of the same element does not matter, as long as there is at least one. Therefore, we can shorten long sets as follows:

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

Finally, order in set also does not matter because once again, the requirement that two sets are equal are simply that they share the same elements. Therefore $\{\text{Tom}, \text{Mary}\}$ and $\{\text{Mary}, \text{Tom}\}$ are also equal.

We then give the definition of subsets.

Definition 1.2.2: Subsets

Let A, B be sets. A is a subset of B means that every elements in A is contained in B. We denote it by

$$A \subseteq B$$

if A is possibly equal to B and

$$A \subset B$$

if it is a proper subset, meaning A cannot be B.

We then give a crucial theorem that we will use throughout the entirety of the book.

Theorem 1.2.3

Let A, B be sets. Then

$$A = B \iff A \subseteq B \text{ and } B \subseteq A$$

Proof. Suppose that $A = B$. Then $x \in A \implies x \in B$ and $x \in B \implies x \in A$. Suppose that $A \subset B$ and $B \subset A$. Then $x \in A \implies x \in B$ and $x \in B \implies x \in A$. □

This theorem is particularly useful in proving sets are equal. Often we will use this theorem in its reverse direction. It is a characterization of equal sets. Whenever we are given to prove two sets are equal, the reader should immediately be able to refer to this theorem.

Readers who are already equipped with further concepts such as relations will realize that $\subseteq$ is a relation on sets. In fact, it is reflexive and transitive, as seen in the following theorem.

Theorem 1.2.4

Let, A, B, C be sets.

- $A \subseteq A$
- $A \subseteq B$ and $B \subseteq C \implies A \subseteq C$

Proof.

- $x \in A \implies x \in A$
- $x \in A \implies x \in B \implies x \in C$

□

1.3 Axiom of Regularity

The axiom of regularity, similar to the axiom schema of specification in the next chapter is meant to prevent paradoxes rather than construct new sets.

Axiom 1.3.1: Axiom of Regularity

Every non-empty set x contains a member y such that x and y are disjoint sets.

With this axiom, expressions such as $S \in S$ does not make sense anymore. Which is good, because it prevents self referencing as a potential paradox.

1.4 Axiom Schema of Specification

Undoubtedly one of the weirdest axioms to get a hold off is the axiom of specification. Essentially, the axiom means that we cannot create sets that are too big. Readers are free to look up more related information on this axiom, especially with regards to Russell's Paradox.

Axiom 1.4.1: Axiom Schema of Specification

To every set A and every condition set S(x) there corresponds a set B whose elements are exactly those elements of x of A for which S(x) holds. We write $B = \{x \in A \mid S(x)\}$

This axiom basically prevents us from creating arbitrarily large sets from thin air. Try and compare the statements $\{x \in A \mid S(x)\}$ and $\{x \mid S(x)\}$. For the first one we are essentially pulling things out from a set A, so inherently the new set is "smaller" than A, while the latter we are basically pulling things out of thin air. Here we assumed that there is some universal set that contains every single thing in the universe. And the latter statement is basically a condition towards this universal set. This universal set creates a lot of problems for us. By introducing the Axiom Schema of Specification, it allows us to route our way around Russell's Paradox:

If we define

$$R = \{x \mid x \notin x\}$$

which means R is the set of all elements that does not contain itself, then we reach a contradiction. Notably,

$$R \in R \iff R \notin R$$

We prove this as usual, assuming one side then reaching the other.

Assume that $R \in R$. Then since R contains itself, by definition of R, R should only contains elements that does not contain itself thus $R \notin R$. Now assume that $R \notin R$, then by definition of R, $R \in R$.

The reason that the Axiom Schema of Specification avoids this problem is that the predicate (condition) that we specify must be applied on a set (that is not the universal set). In other words, the underlying set for the predicate cannot be too vague since the new set created from the axiom of specification is necessarily a subset of the underlying set.

With this axiom in place, we can finally investigate our first concrete set!

Definition 1.4.2: Empty Set

Let A be a set. The set which contains no elements is the empty set, denoted by

$$\emptyset = \{x \in A : x \neq x\}$$

The axiom of extensionality guarantees that this empty set is unique. Although we do not have an all-encompassing universal set (the larger set possible), we still have the smallest set possible, characterized by the following theorem.

Theorem 1.4.3

$\emptyset \subseteq A$ for every set A.

Proof. Every element in $\emptyset$ is also in A. □

Its quite hard to wrap your head around this proof. Using the definition of a subset, we want to show that $x \in \emptyset$ implies $x \in A$. Do you think this is true for all $x \in \emptyset$? Well since there are no elements in the emptyset, this would be true! Can you see why?

2 Operations on Sets

2.1 Axiom of Unions

Axiom 2.1.1: Axiom of Unions

For every collection of sets A there exists a set that contains all the elements that belong to at least one set of the given collection. In other words, there exists a set B such that

$$B = \{x \in a \mid a \in A\}$$

One must wonder: why go the long away round and say that a unions are elements of a set living in a set of sets? Essentially this allows us to proceed with the Axiom of Pairing. The axiom of pairing allows new sets to be created from old sets. And in fact, the Axiom of Unions is simply recovering the elements hidden from the sets given by the axiom of pairing.

Definition 2.1.2: Union

Let A, B be sets. Define the union of A and B to be

$$A \cup B = \{x : x \in A \text{ or } x \in B\}$$

At this point the reader should not be confused between operations and statements. I assume the readers are crystal clear about it but I would elaborate on it as a reminder. Unions and the upcoming intersections are operations on sets that produce new sets. They are by no means able to be evaluated to be true or false. This property is preserved exclusively for statements and statements alone. And statements are given by symbols such as $=$ and $\subseteq$. If you think about it, one should be able to judge whether $A \subseteq B$ is true or false, depending on the contents of A and B while it does not make sense to judge the validity of $A \cup B$.

Proposition 2.1.3

Let A, B, C be sets.

- Identity: $A \cup \emptyset = A$
- Commutativity: $A \cup B = B \cup A$
- Associativity: $(A \cup B) \cup C = A \cup (B \cup C)$
- Idempotent: $A \cup A = A$
- $A \subset B \iff A \cup B = B$

Proof. We prove it in order.

- $A \cup \emptyset = \{x : x \in A \text{ or } x \in \emptyset\} = \{x : x \in A\} = A$
- $A \cup B = \{x : x \in A \text{ or } x \in B\} = \{x : x \in B \text{ or } x \in A\} = B \cup A$
- Proved similarly by expanding the definition of union and using the fact that the logic operator "or" is commutative.
- $A \cup A = \{x : x \in A \text{ or } x \in A\} = \{x : x \in A\} = A$
- Suppose that $A \subset B$. $x \in A \cup B \implies x \in A$ or $x \in B \implies x \in B$ or $x \in B \implies x \in B$. Thus $A \cup B \subset B$. $x \in B \implies x \in B$ or $x \in B \implies x \in A$ or $x \in B \implies x \in A \cup B$. Thus $B \subset A \cup B$. Now suppose that $A \cup B = B$. $x \in A \implies x \in B$ thus $A \subset B$.

□

There is a bunch of fancy names in front of the properties of the operations. They are simply jargons for the properties of any operations in general. Do not be frightened by it since I will mostly like not recall properties from their fancy names unless I want to shortent the length of a proof.

We are shown an operation which is some what equivalent to the "or" operation in logic theory. We shall present a similar notion for "and" as well.

Definition 2.1.4: Intersection

Let A, B be sets. Define the intersection of A and B to be

$$A \cap B = \{x \mid x \in A \text{ and } x \in B\}$$

The intersection is a another way to produce new sets from old, except that smaller sets are created, instead of larger sets. However this does not mean that the number of sets it procures is limited.

Proposition 2.1.5

Let A, B, C be sets.

- Identity: $A \cap \emptyset = \emptyset$
- Commutativity: $A \cap B = B \cap A$
- Associativity: $(A \cap B) \cap C = A \cap (B \cap C)$
- Idempotent: $A \cap A = A$
- $A \subseteq B \iff A \cap B = A$

Proof. The first four are proved in similarity to the counterpart with unions. We prove the last item. Suppose that $A \subset B$. $x \in A \cap B \implies x \in A$ and $x \in B$. Thus $x \in A$ and $A \cap B \subset A$. $x \in A \implies x \in A$ and $x \in A \implies x \in A$ and $x \in B$. Thus $x \in A \cap B$ and $A \subset A \cap B$. Now suppose that $A \cap B = A$. $x \in A \implies x \in A \cap B \implies x \in B$. Thus $A \subset B$. □

The following definition is a condition that can be satisfied by two sets.

Definition 2.1.6: Disjoint

Let A, B be sets. A and B are disjoint if and only if $A \cap B = \emptyset$.

Sometimes it is more convenient to simply say two sets are disjoint rather than stating that their intersection is empty.

Finally we have the distributive law that links the two operators between sets.

Theorem 2.1.7: Distributive Law

Let A, B, C be sets.

- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

Proof. The two are proved by supposing x in the left and the right and proving that the left and the right are subsets of each other by logic. □

2.2 Complements

Another important concept in set theory is complements of sets. However one shall see that there are exactly two notions of complements. One that involves complements from the universe and another complement from a relative set.

Definition 2.2.1: Relative Complements

Let A, B be sets. Define

$$A \setminus B = \{x \in A : x \notin B\}$$

to be the relative complement of B in A.

This complement depends on what the larger set is, therefore suggesting the name relative.

Definition 2.2.2: Absolute Complements

Let E be the set that contains all elements under study. Define the absolute complement of A in E to be

$$A^C = E \setminus A$$

From Russell's Paradox we already know that sets too large and vague is unacceptable, so we usually define what the universal set is. For example, if the universal set is natural numbers, them the absolute coplements of the even numbers is the odd numbers including zero (We have not formally talked about number systems, which we will in formal set theory courses). If the universal set is the real numbers then clearly the absolute complement is not only the odd numbers.

Notice that the concepts of relative complement and absolute complement coincide when $B \subseteq A$ in the definition of relative complements. The reason for two different notations is because relative complements does not necessarily require that $B \subseteq A$. It simply rules out elements of B the coincide with A, while in the case of absolute complements, the universal set is always a superset of A.

If we assume that there is a universal set that is a superset of both A and B, then we have the relation

$$B \cap A^C = B \setminus A$$

This can be proven once again using first order logic:

Proposition 2.2.3

Let A, B be sets. Then

$$A \setminus B = A \cap B^C$$

Proof. Let $x \in A \setminus B$. Then $x \in A$ and $x \notin B$. But $x \notin B$ implies $x \in B^C$. Thus $x \in A$ and $x \in B^C$ implies $x \in A \cap B^C$. This proves that $A \setminus B \subseteq A \cap B^C$.

The entire argument is reversible where implications are double sided. Thus also we have $A \cap B^C \subseteq A \setminus B$ and thus $A \setminus B = A \cap B^C$. □

Proposition 2.2.4

Let A, B be sets and subsets of E. The following four with respect to the absolute complement.

- $(A^C)^C = A$
- $\emptyset^C = E$ and $E^C = \emptyset$
- $A \cap A^C = \emptyset$ and $A \cup A^C = E$
- $A \subset B \iff B^C \subset A^C$

Proof. The four are proved by expanding on the definition of complement and proved by logic. □

Similar to the distributive law, we can associate the complement operator with the previous two operators, namely union and intersection.

Theorem 2.2.5: De Morgans Laws

Let A, B be sets,
- $(A \cup B)^C = A^C \cap B^C$
- $(A \cap B)^C = A^C \cup B^C$

Proof. Both are proved by expansion of the set language into logic. □

The following are a list of operations that are true with respect to the relative complements.

Proposition 2.2.6

Let A, B, C be sets.
- $A \subset B$ if and only if $A \setminus B = \emptyset$
- $A \setminus (A \setminus B) = A \cap B$
- $A \cap (B \setminus C) = (A \cap B) \setminus (A \cap C)$
- $A \cap B \subset (A \cap C) \cup (B \cap C^C)$
- $(A \cup C) \cap (B \cup C^C) \subset A \cup B$

Proof. Exercise. □

These obscure expressions are often explained clearly by drawing venn diagrams, which we will not discuss here because I do not know how to draw.

Finally, the rarely used symmetric difference will be defined here although there is no real usage for now.

Definition 2.2.7: Symmetric Difference

Let A, B be sets. Define the symmetric difference of A and B to be

$$A + B = (A \setminus B) \cup (B \setminus A)$$

2.3 Axiom of Powers

Finally, the axiom of power is simple: to assert the existence of power sets of a set.

Axiom 2.3.1: Axiom of Power Set

For each set there exists a collection of sets that contains among its elements all the subsets of the given set. Define that collection to be $\mathcal{P}(A)$, where A is any set.

The following is a proposition regarding integrating the operation of taking power sets and unions and intersections. It serves as a good exercise.

Proposition 2.3.2

Let E be a collection of sets. Then the following are true.
- $\bigcap_{X \in E} \mathcal{P}(X) = \mathcal{P}(\bigcap_{X \in E} X)$
- $\bigcup_{X \in E} \mathcal{P}(X) \subseteq \mathcal{P}(\bigcup_{X \in E} X)$

Proof. Let $A \in \bigcap_{X \in E} \mathcal{P}(X)$. Then $A \subseteq X$ for all $X \in E$. Thus $A \subseteq \bigcap_{X \in E} X$ and $A \in \mathcal{P}(\bigcap_{X \in E} X)$. For the reverse inclusion, the above implications are all double sided thus we are done.

Let $A \in \bigcup_{X \in E} \mathcal{P}(X)$. Then $A \subseteq \mathcal{P}(X)$ for some $X \in E$. Thus $A \subseteq \bigcup_{X \in E} X$ and $A \in \bigcup_{X \in E} \mathcal{P}(X)$ and we are done. Note that the reverse inclusion does not hold since we only have $A \subseteq \mathcal{P}(X)$ for some $X \in E$ as compared to the intersection of power sets. □

2.4 Cartesian Products

Definition 2.4.1: Ordered Pairs

Define the ordered pair of a and b to be

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

Note the cleverness here, the ordered pair is defined as a set of two elements, one of which is a singleton which defines which element in the ordere pair comes first. This allows order to be defined in a set which are a collection of unordered elements. Can you think of a way to extend this notion into triple, quadruples and even n-tuples?

Let us see an immediate consequence of this definition which proves that this notion indeed works out nicely.

Proposition 2.4.2

Let (a, b) and (c, d) be ordered pairs. Then $(a, b) = (c, d)$ if and only if $a = c$ and $b = d$.

Proof. Suppose that $(a, b) = (c, d)$. Then we have by definition,

$$\{\{a\}, \{a, b\}\} = \{\{c\}, \{c, d\}\}$$

There are two cases: $a = b$ and $a \neq b$. Suppose that $a = b$, then $\{\{a\}\} = \{\{c\}, \{c, d\}\}$. This forces $\{a\} = \{c\} = \{c, d\}$ and $a = c = d$. Suppose that $a \neq b$. Then $\{a\} = \{c\}$ and $\{a, b\} = \{c, d\}$. Thus $a = c$ and $b = d$. b cannot be c here since $a = b = c$ is a contradiction.

Now suppose that $a = c$ and $b = d$, then $\{a\} = \{c\}$ and $\{a, b\} = \{c, d\}$ thus $\{\{a\}, \{a, b\}\} = \{\{c\}, \{c, d\}\}$. □

We now define the cartesian product of two sets, whose elements are ordered pairs.

Definition 2.4.3: Cartesian Product

Let A, B be sets. Define the Cartesian Product of A and B to be

$$A \times B = \{(a, b) \mid a \in A, b \in B\}$$

Clearly by definition a cartesian product gives rise to a set of ordered pairs. But the converse is also true, where a set of ordered pairs can give rise to a cartesian product.

Proposition 2.4.4

Suppose that R is a set of ordered pairs. Then there exists A, B such that $R \subseteq A \times B$.

Proof. Simply define $A = \{a \mid (a, b) \in R\}$ and $B = \{b \mid (a, b) \in R\}$. These sets in fact have names as we will see in the next definition. □

Readers should make sure that the predicate $(a, b) \in R$ is a valid construct in the language of set theory, just as a sanity check.

As promised, we give names to the sets defined in the proof above.

Definition 2.4.5: The First and Second Projection

Let R be a set of ordered pairs. Define the first and second projections of R to be

$$A = \{a \mid (a, b) \in R\}$$

and

$$B = \{b \mid (a, b) \in R\}$$

Unfortunately these definition will be forgotten and rarely be used again.

We end the section with properties of the cartesian product when used in conjunction with other set operators.

Proposition 2.4.6

Let A, B, X, Y be sets.

- $(A \cup B) \times X = (A \times X) \cup (B \times X)$
- $(A \cap B) \times (X \cap Y) = (A \times X) \cap (B \times Y)$
- $(A \setminus B) \times X = (A \times X) \setminus (B \times X)$

Proof. Again just a practise of expansion into logic. □

3 Relations on a Set

3.1 Relations

Relations appear naturally in the course of studying mathematics. Some simple relations that are used commonly include $=$, $\leq$ and $\geq$ and $<, >$. In fact, $\subset$ and $\subseteq$ are also relations.

More generally, relations are used to associate two elements of a set. This could be done because that they share similar properties or used for comparison.

Definition 3.1.1: Relation

Define a relation R to be a set of ordered pairs. If $(a, b) \in R$, we write aRb instead.

The definition of relations with ordered pairs seems unnatural especially when considering the usual relations like $\leq$ and $\geq$. In practise we will rarely think about relations set theoreticly. This is only meant for formalization into the language of set theory, which in turn is the basis of modern mathematics.

Definition 3.1.2: Domain and Range

The first projection of a relation R is called $\mathrm{dom}(R)$ and the second projection is called $\mathrm{ran}(R)$. In other words,

$$\mathrm{dom}(R) = \{a \mid (a, b) \in R\} \text{ and } \mathrm{ran}(R) = \{b \mid (a, b) \in R\}$$

Bear with the notation for now. This is a natural extension of names from definitions of functions since functions are defined with relations.

We now classify relations based on some properties it exhibits. Note that this is not all the properties of relations but instead only the first few important ones.

Definition 3.1.3: Types of Relations

Let R be a relation. We say that

- R is reflexive if $(x, x) \in R$ for all $x \in R$
- R is symmetric if $(x, y) \in R \implies (y, x) \in R$ for all $x, y \in R$
- R is transitive if $(x, y) \in R$ and $(y, z) \in R$ implies $(x, z) \in R$ for all $x, y, z \in R$
- R is anti-symmetric if $(x, y) \in R$ and $(y, x) \in R$ implies $x = y$.

Readers can check that while $=$ is an equivalence relation, $<, >$ only satisfies transitivity while $\leq, \geq$ satisfies reflexivity in addition to transitivity.

Finally, we have inverse relations which proves itself useful when dealing with inverse functions.

Definition 3.1.4: Inverse Relations

Let R be a relation from A to B. Define the inverse relation to be

$$R^{-1} = \{(b, a) \mid (a, b) \in R\} \subseteq B \times A$$

Notice that here, every relation is guaranteed to have an inverse. But for function, additional requirements have to be satisfied in order for functions to be reversed.

3.2 Partitions and Equivalence Relations

The remainder of the section is mainly for exhibiting the relation between partitions and equivalence relations. This is often useful in the sense that whenever an equivalence relation appears, a partition will be possible from the set.

Firstly we define partitions.

Definition 3.2.1: Partition

Let X be a set. A partition of X is a disjoint collection E of subsets of X such that
- $A, B \in E$ and $A \neq B$ implies $A \cap B = \emptyset$
- $\bigcup_{A \in E} A = X$

Although we have yet to define the order of a set (number of elements in a set), it is important to note that partitions does not mean that the number of elements in each partition is the same.

Definition 3.2.2

Let X be a set. Let R be a relation on X. We say that R is an equivalence relation if the following are true.
- R is reflexive
- R is symmetric
- R is transitive

Definition 3.2.3: Equivalence Class

Let R be an equivalence relation on a set X. Denote $[x] = \{y \in X \mid (x, y) \in R\}$ the equivalence class of $x \in X$ and $X/R = \{[x] \mid x \in X\}$ the set of all equivalence classes.

Now comes the main theorem, it is split into two parts for readability. The proof is not particularly long but may take some time to digest and understand.

Theorem 3.2.4

An equivalence relation R on a set X induces a partition on X.

Proof. For every $x \in X$, $x \in x/R$ thus $\bigcup_{x \in X} x/R = X$. Now suppose that $z \in x/R \cap y/R$, then $(x, z) \in R$ and $(y, z) \in R$. By the symmetric property $(z, y) \in R$. By transitivity $(x, y) \in R$. Thus $x/R = y/R$. This proves that X/R is a partition. □

Theorem 3.2.5

A partition on X induces an equivalence relation R on X.

Proof. Suppose that X is partitioned. Define a relation R to be $(x, y) \in R$ if and only if x, y are in the same partition.

We now prove R is an equivalence relation. We have that $(x, x) \in R$ for every x since they necessarily appear in the same set and none others (by defition of partition) thus the reflexive property holds. If $(x, y) \in R$ then x, y are in the same partition thus naturally $(y, x) \in R$ holds thus the symmetric property holds. Finally if $(x, y) \in R$ and $(y, z) \in R$ then x, y, z are all in the same parition thus $(x, z) \in R$ holds which proves transitivity. We can now conclude that R is an equivalence relation and we are done. □

4 Functions between Sets

4.1 Functions

Functions play an integral role in all of mathematics. Therefore it is important to be able to express this concept in terms of set theoretic language.

Definition 4.1.1: Functions

Let X, Y be sets. A function from X to Y is a relation $f \subseteq X \times Y$ such that
- $\text{dom}(f) = X$
- $\exists y \in Y$ such that $(x, y) \in f$ for all $x \in X$ (existence of an output)
- $(x, y) \in f$ and $(x, z) \in f$ implies $y = z$ (uniqueness of an output)

In this case, we say that X is the domain of f and Y is the codomain of f. We often write f as $f : X \to Y$ to indicate the domain and codomain of f.

As one can see, functions are defined based on relations which is why the study of relations is also important. The additional rules the a relation has to satisfy in order to be a function is simply that all of X must be associated to exactly one thing in y, not zero, not two, meaning everything in the domain must be linked to something in Y.

An immediate consequence is that since relations are subsets of the cartesian product, all functions from X to Y must be encapsulated by the cartesian product in some way.

Proposition 4.1.2

The set of all functions from X to Y is a subset of $Y^X = \mathbb{P}(X \times Y)$.

Proof. Note that every function is a relation thus this induces a subset relation betweeen functions on a set and relations on a set. Then since we have shown that $R \subseteq X \times Y$, any function must be an element of $\mathbb{P}(X \times Y)$ and thus the set of all functions is a subset of $\mathbb{P}(X \times Y)$. □

Below we give three crucial functions which arises in most of the areas of mathematics.

Definition 4.1.3: Inclusion Map

Let $X \subset Y$ and $f : X \to Y$ where f is defined as $f(x) = x$. f is called the inclusion map of X into Y.

Notice the strict inclusion on $X \subset Y$. The inclusion map is meant to encorporate and identify X inside of Y. This is most often used when we want to extend the domain of a function. When the strict inclusion is relaxed and we have $X = Y$, it is called the identity map.

Definition 4.1.4: Identity Map

The inclusion map from X to X is called the identity map on X.

The name identity map will become clear once we reach inverse functions.

Definition 4.1.5: Restriction Map

Let $f : Y \to Z$ and $X \subset Y$. The restriction map of f is the function $g : X \to Z$ such that $g(x) = f(x)$ for all $x \in X$. Conversely, the extension map of $g \to Y$ is f. We write $g = f \mid X$.

Restriction maps are simply copies of the original map that is restricted to a certain subset of the original domain.

Finally, there are three important properties that a function can take.

Definition 4.1.6: Bijective Functions

Let $f : X \to Y$ be a function.
- f is injective if $f(x_1) = f(x_2) \implies x_1 = x_2$
- f is surjective if for all $y \in Y$ there exists $x \in X$ such that $f(x) = y$
- f is bijective if it is both injective and surjective

This properties of a function does not only depend on how the function/relation is defined, but the domain and codomain also plays a part.

4.2 Compositions

Functions are allowed to be composed. By doing this we are essentially creating a new function.

Definition 4.2.1: Composition of Functions

Let $f : A \to B$ and $g : B \to C$ be two functions. Define the composition of f and g to be

$$g \circ f : A \to C$$

If $a \in A$ then

$$(g \circ f)(a) = g(f(a))$$

Readers should verify that the new object $g \circ f$ is a function.

Lemma 4.2.2

Composition of functions result in a new function.

Proof. Easy exercise. □

While composition of functions is in general not commutative, it is fortunately associative.

Proposition 4.2.3: Associativity of Functions

Let $f : A \to B$, $g : B \to C$ and $h : C \to D$ be functions. Then the following is true.

$$(h \circ g) \circ f = h \circ (g \circ f)$$

Proof. Simple manipulation of definition of composition. □

In general, composition preserves injectivity and surjectivity, as seen by the folllowing proposition.

Proposition 4.2.4

Let $f : A \to B$ and $g : B \to C$ be functions.
- If f and g are injective then $g \circ f$ is injective
- If f and g are surjective then $g \circ f$ is surjective
- If f and g are bijective then $g \circ f$ is bijective

Proof. Easy exercise involving the use of definition of injectivity and surjectivity. □

4.3 Inverses

Inverse functions serve the important role of inverting functions so that we know what element in the domain is mapped to a fixed element in the codomain. However the condition that the inverse exists must be studied.

Definition 4.3.1: Inverse Functions

Let $f : X \to Y$ be a function. If the inverse relation $f^{-1} : Y \to X$ is also a function, then we say that f^{-1} is the inverse function of f.

The main criterion for inversability of a function is given by the below characterization.

Theorem 4.3.2

Let $f : X \to Y$ be a function. The inverse relation f^{-1} is a function from Y to X if and only if f is bijective.

Proof. We first suppose that the inverse of f is a function. We aim to prove injectivity and surjectivity.
Injectivity: Suppose that $f(x_1) = f(x_2) = y$. In terms of relations, this means that $(x_1, y) \in f$ and $(x_2, y) \in f$. By definition of inverse relations, $(y, x_1) \in f^{-1}$ and $(y, x_2) \in f^{-1}$. But since f^{-1} is a function, $x_1 = x_2$ and we are done.

Surjectivity: We want for every $y \in Y$ there exists $x \in X$ such that $f(x) = y$. So choose an arbitrary $y \in Y$. Since f^{-1} is a function from Y to X, $f^{-1}(y) = x$ lies in X. Then $(y, x) \in f^{-1}$ implies that $(x, y) \in f$, which we are done.

Finally, suppose now that f is bijective. We aim to show that f^{-1} is a function. The fact that $\mathrm{dom}(f^{-1}) = Y$ is trivially satisfied. There are two items to show: that for every $y \in Y$, there exists $x \in X$ such that $f^{-1}(y) = x$ and that $(y, x) \in f^{-1}$ and $(y, z) \in f^{-1}$ implies $x = z$. For the first item, we use the fact that f is surjective. Let $y \in Y$. By surjectivity, there exists $x \in X$ such that $(x, y) \in f$. Then by definition of inverse relation $(y, x) \in f^{-1}$ and we are done.

Now for the second item, suppose that $(y, x) \in f^{-1}$ and $(y, z) \in f^{-1}$. Then by definition of inverse relation $(x, y) \in f$ and $(z, y) \in f$. Then using injectivity, we see that this should imply $x = z$ and so we are done. □

As seen in the proof, for f^{-1} to be a function from Y to X, injectivity and surjectivity plays a crucial role.

From the theorem, inverses of a function exists if and only if f is bijective. We often call bijectivity a necessary and sufficient condition for inverses. Necessary here means that without this property, inverses would not exist. Sufficiency here means that with this property, inverses can exists. Both of which when used together, exactly means "if and only if".

Proposition 4.3.3

If $f : X \to Y$ is a bijective function then f^{-1} is bijective. Moreover, $f^{-1} \circ f$ is the identity function on X and $f \circ f^{-1}$ is the identity function on Y.

Proof. We know that since $f : X \to Y$ is bijective, $f^{-1} : Y \to X$ is a proper function. We first prove injectivity. Suppose that $(y, x) \in f^{-1}$ and $(z, x) \in f^{-1}$. By definition of inverse relation we have that $(x, y) \in f$ and $(x, z) \in f$. This means that $y = z$ by definition of a function. For surjectivity, suppose that $x \in X$. Since f is a function from X, there exists y such that $f(x) = y$. But this means that $(x, y) \in f$ and by definition of inverse relations, $(y, x) \in f^{-1}$ and so we are done.

Making sure that applying the inverse and the function itself is just the identity function is an easy exercise. □

Definition 4.3.4: Images and Preimages

Let $f : X \to Y$ be a function. Let $A \subseteq X$ and $B \subseteq Y$. Denote the image of A under f to be

$$f(A) = \{y \in Y \mid f(x) = y, \forall x \in A\} \subseteq Y$$

Define the preimage of B under f to be the set

$$f^{-1}(B) = \{x \in X \mid f(x) \in B\} \subseteq X$$

This is such an abuse of notation that I cannot stress enough. Try not to be confused with the usual inverse of f, which does not always exists as a function, while the preimage, always exists regardless of whether f has an inverse function.

Now that images and preimages have come into play, can you express surjectivity in terms of images and/or preimages? Can you also express the condition for bijecitivity and inverses in terms of images and/or preimages?

Next proposition is a crucial one because it tells us what happens by applying f and its inverse relation when inverses of the function may not exist.

Proposition 4.3.5

Let A, B, X, Y be sets. Let $f : X \to Y$.
- If $B \subseteq Y$ then $f(f^{-1}(B)) \subseteq B$
- If $B \subseteq Y$ and f is surjective then $f(f^{-1}(B)) = B$
- If $A \subseteq X$ then $A \subseteq f^{-1}(f(A))$
- If $A \subseteq X$ and f is injective then $f^{-1}(f(A)) = A$

Proof. Let $f : X \to Y$ be a function.
- Let $y \in f(f^{-1}(B))$. Then there exists $x \in f^{-1}(B)$ such that $y = f(x)$. But $x \in f^{-1}(B)$ implies $f(x) \in B$ by definition. Thus $y \in B$ and we are done.
- We just have to show that $B \subseteq f(f^{-1}(B))$. Let f be surjective and $y \in B$. Then by surjectivity there exists $x \in f^{-1}(B)$ such that $y = f(x)$. But $x \in f^{-1}(B)$ means that $f(x) \in f(f^{-1}(B))$. Thus $y \in f(f^{-1}(B))$ and we are done.
- Let $x \in A$. Then $f(x) \in f(A)$. But by definition of $f^{-1}(f(A))$, $x \in f^{-1}(f(A))$ thus we are done.
- We just have to show that $f^{-1}(f(A)) \subseteq A$. Let f be injective and $x \in f^{-1}(f(A))$. Then by definition, $f(x) \in f(A)$. By injectivity, there exists only one element in A that maps to $f(x)$, and that is precisely x. Thus $x \in A$ and we are done.

□

Interested readers may look up left and right inverses of a function. They are characterized by the fourth and second item respecrtively. By combining these two properties, left and right inverses combine to become an inverse of a function in the sense that it maps Y to X.

5 Number Systems

5.1 Axiom of Infinity

Before we formulate the natural numbers via set theoretic language, we need one more axiom to guarantee the existence of the set of natural numbers.

Definition 5.1.1: Successor

Define the successor of a set x to be

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

Notice that the successor of a set is also a set. With this notion we can define numbers as sets. For example, 0 is simply the empty set $\emptyset$, 1 would be $\{\emptyset\}$, 2 would be $\{\emptyset, \{\emptyset\}\}$ and so on. This gives a unique code for every natural number, as well as giving order to the set of natural numbers, as we will see soon.

Finally, we need to guarantee that such a set exists.

Axiom 5.1.2: Axiom of Infinity

There exists a set containing the empty set $\emptyset$ and containing the successor of each of its elements.

As with all the other axioms, to argue that an axiom is necessary and unprovable by other axioms is a very advanced topic. For now we will only state that it is needed.

5.2 The Peano Axioms

Definition 5.2.1: The Peano Axioms

We say that ω is a successor set if

1. $\emptyset \in \omega$
2. $n \in \omega \implies n^+ \in \omega$
3. $n^+ \neq \emptyset$ for all $n \in \omega$
4. $n, m \in \omega$ and $n^+ = m^+ \implies n = m$
5. $S \subset \omega$ and $\emptyset \in S$ and $n^+ \in S$ for all $n \in S$ implies $S = \omega$

Each of these axioms plays an important role in the formulation, without any one of them the natural numbers would not come into play.

The first item guarantees that the natural numbers is a non empty set. The second item guarantees that all successors, as required to define natural numbers, lie in the successor set. The third item guarantees that the natural numbers is not a huge loop of numbers that circulates back to 0. The fourth item guarantees that no two numbers has a common successor. The final axiom prevents extra loops to appear other than the natural numbers itself. For example, there cannot be loops such as $y = x^+$, $z = y^+$ and $x = z^+$. Notice that the third item does not guarantees this since it only prevents the natural numbers to be one big loop. The fifth item in turn guarantees this since the natural numbers has to satisfy that when elements near the number is a natural number, it will also be a natural number. Since x, y, z is a closed loop, they will never be natural numbers since the fifth item states the induction basis that starts from 0.

From now on, in a successor set, we now say 0 in place of $\emptyset$. If the readers wish to, we can now say that the successor of 0 is 1, the successor of 1 is 2 and vice versa.

The following theorem allows functions that work recursively to be defined.

Theorem 5.2.2: Recursion Theorem

Let X be a set and $a \in X$. Let $f : X \to X$. There exists a function $u : \omega \to X$ such that
- $u(0) = a$
- $u(n^+) = f(u(n))$ for all $n \in \omega$

Proof. Let

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

Since $\omega \times X \in E$, $E \neq \emptyset$. Define $u = \bigcap_{B \in E} B$. Then $u \in E$. We want to show that u is a function. We prove that the set of all n such that $(n, x) \in u$ and $(n, y) \in u \implies x = y$ is ω Let

$$S = \{n \in \omega \mid (n, x) \in u, (n, y) \in u \implies x = y\}$$

We prove it by the fifth item of Peano Axiom.

We first show that $0 \in S$. Suppose for a contradiction that $0 \notin S$ and $(0, b) \in u$ and $a \neq b$. Consider $v = u - \{(0, b)\}$. $(0, a) \in v$ and $(n, x) \in A \implies (n^+, f(x)) \in A$ since $n^+ \neq 0$ for all $n \in \omega$. Thus $v \in E$ contradicts the fact that $u = \bigcap_{B \in E} B$.

We now show that $n^+ \in S$ if $n \in S$. Let $n \in S$. Then there exists a unique $x \in X$ such that $(n, x) \in u$. Suppose that $n^+ \notin S$. Then there exists $y \neq f(x)$ such that $(n, y) \in u$. Consider $v = u - \{(n, y)\}$. $(0, a) \in v$ and $(n, x) \in v \implies (n^+, f(x)) \in v$. Thus $v \in E$ contradicts the fact that $u = \bigcap_{B \in E} B$. We thus have $S = \omega$, finishing the proof. □

Expanding things out, we see that $u(0) = a$, $u(1) = f(a)$, $u(2) = f(f(a))$ and so on. We have constructed a function u such that input the number in u means that you are compositing that number of f to the initial element. This is why through this theorem, we allowed the existence of recursive functions. The application of this immediate as we will use it to define addition and multplication through this theorem.

5.3 Arithmetic

We begin by defining the notion of addition in natural numbers.

Proposition 5.3.1

For every natural number m there exists a function $s_m : \mathbb{N} \to \mathbb{N}$ such that
- $s_m(0) = m$
- $s_m(n^+) = (s_m(n))^+$

$s_m(n)$ is by definition, the sum $m + n$.

Proof. Set X as $\mathbb{N}$, f as the successor function in recursion theorem and we are done. □

The recursion here used is the repeated use of successors, meaning we are applying $+1$ a certain amount of times.

We can now prove properties of addition in set theoretic language.

Proposition 5.3.2: Properties of Addition

Let $x, y, z \in \mathbb{N}$.
- (A1) $x + y \in \mathbb{N}$

- (A2) $(x+y)+z = x+(y+z)$
- (A3) $0+x = x = x+0$
- (A4) $x+y = y+x$

Proof. We prove associativity, identity and commutativity in order.

The closure under addition is direct from the definition of addition.

- For associativity, we induct on z. When $z = 0$, we have

$$\begin{aligned}(x+y)+0 &= x+y \\ &= x+(y+0)\end{aligned}$$

Suppose that $(x+y)+n = x+(y+n)$, we have

$$\begin{aligned}(x+y)+n^+ &= ((x+y)+n)^+ \\ &= (x+(y+n))^+ && \text{(Induction Hypothesis)} \\ &= x+(y+n)^+ \\ &= x+(y+n^+)\end{aligned}$$

Thus by the principle of induction, we have associativity.

- For identity, we induct on x. When $x = 0$, we have that

$$0+0 = 0 = 0+0$$

Suppose that $0+n = n = n+0$, we have

$$\begin{aligned}0+n^+ &= (0+n)^+ \\ &= n^+ && \text{(Induction Hypothesis)} \\ &= n^+ + 0\end{aligned}$$

Thus by the principle of induction, we have identity.

- For commutativity, we induct on y first to show that $x^+ + y = (x+y)^+$. When $y = 0$, we have $x^+ + 0 = x^+ = (x+0)^+$. Now suppose that $x^+ + n = (x+n)^+$

$$\begin{aligned}x^+ + n^+ &= (x^+ + n)^+ \\ &= ((x+n)^+)^+ && \text{(Induction Hypothesis)} \\ &= (x+n^+)^+\end{aligned}$$

Thus our first induction is complete. Now we prove commutativity by induction on x. When $x = 0$, we have $0+y = y+0$ from identity. Now suppose that $x+y = y+x$.

$$\begin{aligned}x^+ + y &= (x+y)^+ && \text{(First induction)} \\ &= (y+x)^+ && \text{(Induction Hypothesis)} \\ &= y+x^+\end{aligned}$$

Thus by the principle of induction, we have commutativity.

□

We then proceed to multiplication.

Proposition 5.3.3

For every natural number m there exists a function $p_m : \mathbb{N} \to \mathbb{N}$ such that

- $p_m(0) = 0$
- $p_m(n^+) = p_m(n) + m$

$p_m(n)$ is by definition, multiplication $m \times n$.

Proof. Take X as $\mathbb{N}$ and f as $f(x) = x + m$. Then using the recursion theorem we are done. □

We note here, that the successor function n^+ is in fact equivalent to $n + 1$ (using the definition of addition). This may be seen inherently when the successor notion is introduced, but it is only now that we can formulate it properly since addition is defined.

Similarly, the recursion here is that repeated addition of $+m$. And we can also prove remaining properties of multiplication in set theoretic language.

Proposition 5.3.4

Let $x, y, z \in \mathbb{N}$. Addition and multiplication follow the distributive law $x \cdot (y + z) = x \cdot y + x \cdot z$.

Proof. We prove the distributive law by induction on z. When $z = 0$, we have $x \cdot (y + 0) = x \cdot y$ and $x \cdot y + x \cdot 0 = x \cdot y$. Now suppose that $x \cdot (y + n) = x \cdot y + x \cdot n$.

$$\begin{aligned}
x(y + n^+) &= x(y + n)^+ && \text{(Definition of Addition)} \\
&= x(y + n) + x && \text{(Definition of Multiplication)} \\
&= (xy + xn) + x && \text{(Induction Hypothesis)} \\
&= xy + xn + x && \text{(Associativity of Addition)} \\
xy + xn^+ &= xy + (xn + x) && \text{(Definition of Multiplication)} \\
&= xy + xn + x && \text{(Associativity of Addition)}
\end{aligned}$$

Thus by the principle of induction, we have the distributive law. □

Proposition 5.3.5

Let $x, y, z \in \mathbb{N}$.
- (M1) $x \cdot y \in \mathbb{N}$
- (M2) $(x \cdot y) \cdot z = x \cdot (y \cdot z)$
- (M3) $1 \cdot x = x = x \cdot 1$
- (M4) $x \cdot y = y \cdot x$

Proof. We prove associativity, commutativity, identity in this order. Note we first prove commutativity then prove identity.
The closure under Multiplication is direct from the definition of multiplication.
- We prove associativity by induction on z. When $z = 0$, $(x \cdot y) \cdot 0 = 0$ and $x \cdot (y \cdot 0) = x \cdot 0 = 0$. Now suppose that $(x \cdot y) \cdot n = x \cdot (y \cdot n)$. We have

$$\begin{aligned}
(x \cdot y) \cdot n^+ &= (x \cdot y) \cdot n + x \cdot y && \text{(Definition of Multiplication)} \\
&= x \cdot y + (x \cdot y) \cdot n && \text{(Commutativity of Addition)} \\
&= x \cdot y + x \cdot (y \cdot n) && \text{(Induction Hypothesis)} \\
&= x \cdot (y + y \cdot n) && \text{(Distributivity)} \\
&= x \cdot (y \cdot n + y) && \text{(Commutativity of Addition)} \\
&= x(y \cdot n^+) && \text{(Definition of Multiplication)}
\end{aligned}$$

  Thus by the principle of induction, we have the associative law.
- We prove commutativity by induction on y. When $y = 0$, we have $x \cdot 0 = 0 = 0 \cdot x$. Now

suppose that $x \cdot y = y \cdot x$, we have

$$\begin{aligned} x \cdot y^+ &= (x \cdot y) + x && \text{(Definition of Multiplication)} \\ &= y \cdot x + x && \text{(Induction Hypothesis)} \\ &= x + \cdot(y + x) && \text{(Commutativity of Addition)} \\ &= y^+ \cdot x && \text{(Distributivity)} \end{aligned}$$

Thus by the principle of induction, we have the commutative law.

- The identity is a special case of the commutative law, taking $y = 1$.

□

Now that we have the natural numbers, we can also extend the numbers into integers, rationals, real numbers and even complex numbers. Most of them without the use of set theoretic language. However, this is another topic unrelated to set theory mostly.

6 Ordering and Cardinality

6.1 Total and Partial Orders

Definition 6.1.1: Partial Order

Let X be a set. Let R be a relation on X. We say that R is a partial order on X if the following are true.

- R is reflexive
- R is anti-symmetric
- R is transitive

In this case, we say that R is partially ordered, or that R is a poset.

Definition 6.1.2: Total Order

Let X be a set. Let R be a relation on X. We say that R is a total ordering on X if for any $x, y \in X$, either $(x, y) \in R$ or $(y, x) \in R$. In this case, we say that X is totally ordered.

Definition 6.1.3: Chains

Let X be a set. Let R be a partial order on X. A chain in X is a subset $S \subseteq X$ such that R restricted to S gives a total order on S.

Definition 6.1.4: Initial Segments

Let X be a partially ordered set and $a \in X$, the set

$$s(a) = \{x \in X : x < a\}$$

is the initial segment determined by a. The weak initial segment is denoted

$$\bar{s}(a) = \{x \in X : x \leq a\}$$

Definition 6.1.5: Least and Greatest Element

Let X be a partially ordered set and $a \in X$ such that $a \leq x$ for all $x \in X$. Then a is the least element of X. If $b \in X$ such that $x \leq b$ for all $x \in X$ then b is the greatest element of X.

Definition 6.1.6: Minimal and Maximal Elements

Let X be a partially ordered set and $a \in X$ such that $x \leq a$ implies $x = a$, then a is called the minimal element of X. If $b \in X$ such that $b \leq x$ implies $x = b$, then b is called the maximal element of X.

6.2 Ordering the Natural Numbers

Another important aspect of the natural numbers is that they are comparable. We have the notion of size in them. The Peano axioms for the natural numbers also encapusulates this nicely.

Definition 6.2.1: Comparable

Two natural numbers m, n are comparable if either $m \in n$ or $m = n$ or $n \in m$.

This is called a trichotomy, which means that two natural numbers are comparable when exactly one out of the three conditions are fulfilled. Since we do not have the concept of size yet, we use belonging symbols because if one remembers, natural numbers are defined recursively.

The following theorem states that the trichotomy holds for any two natural numbers, which leads us to being able to compare them.

Theorem 6.2.2

Any two natural numbers are comparable.

Proof. □

Now that we have seen that any two natural numbers are comparable, we can essentially order them according to whether one is a subset of another, or in terms of magnitude, whether one is larger than the other or not.

Definition 6.2.3: Order in $\mathbb{N}$

Define the relation $<$ for two natural numbers m, n such that if $m \in n$, then $m < n$.

As we will see later, even comparisons have different types, namely total orders and partial orders. Total order means that everything in the set can be rearranged into one long queue according to the way we want to order it. Partial order means that not everything in the set is comparable against each other, leading to a tree like appearance when we try and sort the elements.

Proposition 6.2.4

The relation $<$ on $\mathbb{N}$ is transitive.

It is worth noting here that $<$ is neither reflexive nor symmetric which the readers can check.

6.3 Cardinality of Sets

We begin not from the definition of cardinality, but rather by introducing a way to compare the size of two sets. Cardinality can be thought of as the number of elements in a set (for the finite case).

Definition 6.3.1: Cardinality of a Set

Two sets A and B have the same cardinality if there exists a bijection between A and B. In this case, we write

$$|A| = |B|$$

where $|A|$ stands for the cardinality of A.

We say that A has cardinality strictly less than B if there exists an injective function from A to B, but no bijective function exists. In this case, we write $|A| < |B|$.

Basically we compare the size of two sets simply by attempting to define a bijective function between of them. Naturally, we have the following proposition.

Proposition 6.3.2

Every natural number has cardinality strictly less than $|\mathbb{N}|$.

Definition 6.3.3: Finite Sets

A set A is called finite if it has the same cardinality as a natural number represented as a set. In that case, the cardinality of A is precisely the natural number.

For example, if $A = \{a, b, c, d\}$, then it has cardinality 4 and we write it as $|A| = 4$. From the finiteness of the sets, we can perform the usual addition and multiplication on cardinality, provided it makes sense.

Proposition 6.3.4

Let E, F be sets with finite cardinality.
- If E, F are disjoint, then $|E \cup F| = |E| + |F|$
- $|E \times F| = |E| \cdot |F|$

Definition 6.3.5: Countably Infinite Sets

We say that a set A is countably infinite if it has cardinality equal to $\mathbb{N}$. In other words, if

$$|A| = |\mathbb{N}|$$

In particular, mathematicians also say that a set is countable if it is either finite or countably infinite. This is due to the fact that there is also a notion of uncountability.

Definition 6.3.6: Uncountable Sets

We say that a set A is uncountable if it has cardinality strictly greater than $\mathbb{N}$. In other words, if

$$|A| > |\mathbb{N}|$$

7 Equivalent Formations of Zorn's Lemma

7.1 Axiom of Choice and Zorn's Lemma

Definition 7.1.1: Families

A family is a function $f : I \to X$ where I is an index set, usually ω or a natural number, such that every element of the range of f is of the form x_i, $i \in \omega$.

Axiom 7.1.2: Axiom of Choice

The Cartesian Product of a non-empty family of non-empty sets is non-empty.

Theorem 7.1.3

If a set is infinite, then it has a subset equivalent to ω.

Theorem 7.1.4: Zorn's Lemma

If X is a partially ordered set such that every chain in X has an upper bound, then X contains a maximal element.

7.2 Well Ordering

Definition 7.2.1: Well Ordered Set

A partially ordered set is well ordered if every non-empty subset of it has a smallest element.

Theorem 7.2.2

Every well ordered set is totally ordered.

Theorem 7.2.3: The Principle of Transfinite Induction

Suppose that S is a subset of a well ordered set X, and suppose that $x \in X$ such that $s(x) \subset S$, then $x \in S$.

Definition 7.2.4: Continuation

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
- B and A have the same ordering

Theorem 7.2.5

If E is an arbitrary collection of initial segment of a well ordered set, E is a chain with respect to continuation.

Theorem 7.2.6

If a collection E of well ordered sets is a chain with respect to continuation and if $U = \bigcup_{X \in E} X$, then there is a unique well ordering of U such that U is a continuation of each set.

Theorem 7.2.7: Well Ordering Theorem

Every set can be well ordered.

Proposition 7.2.8

The well ordering theorem implies the axiom of choice.

7.3 Transfinite Recursion

Definition 7.3.1: A sequence of type a in X

Let a be an element in a well ordered set W. Let X be an arbitrary set. The sequence of type a in X means a function from $s(a) \subset W$ into X.

Definition 7.3.2: A sequence of type W in X

A sequence of type W in X is a function f whose domain consists of all sequences of type a in X, for all elements a in W and range is included in X.

Theorem 7.3.3: Transfinite Recursion Theorem

If W is a well ordered set and if f is a sequence function of type W in a set X, then there exists a unique function U from W into X such that $U(a) = f(U^a)$ for each a in W, where U^a is the restriction of $U : W \to X$ to the initial segment $s(a)$.

Definition 7.3.4: Similarity

Two partially ordered sets are similar if there is a one to one correspondence that preserves order, or $f(a) \leq f(b) \implies a \leq b$.

Proposition 7.3.5

Let f be a similarity from X to Y.

- f^{-1} is a similarity from Y to X
- gf is a similarity from X to Z if g is a similarity.

Theorem 7.3.6

If f is a similarity of a well ordered set X to itself, then $a \leq f(a)$ for all $a \in X$.

Theorem 7.3.7

Let X, Y be well ordered sets. If X and Y are similar, then the correspondence function is unique.

Theorem 7.3.8

A well ordered set is never similar to one of its initial segments.

Theorem 7.3.9

[Compatibility Theorem] Let X and Y be well ordered sets. Either X and Y are similar or one of them is similar to an initial segment of the other.

8 Counting Beyond Infinity

8.1 Ordinal Numbers

Definition 8.1.1: ω Successor Function

Let f be a function from strict predecessors of some natural number n, and $f(0) = \omega$ and $f(m^+) = f(m)^+$ whenever $m^+ < n$. Then this function is an ω Successor Function.

Axiom 8.1.2: Axiom of Substitution

If $S(a, b)$ is a sentence such that for each a in a set A the set $\{b : S(a, b)\}$ can be formed, then there exists a function F with domain A such that $F(a) = \{b : S(a, b)\}$ for each a in A.

Definition 8.1.3

An ordinal number is defined as the well ordered set α such that $s(\xi) = \xi$ for all $\xi \in \alpha$, where $s(\xi) = \{\eta \in \alpha : \eta < \xi\}$

Theorem 8.1.4

If α is an ordinal number, α^+ is also an ordinal number.

Definition 8.1.5: Ordinal Numbers after ω

Define F from the axiom of substitution such that $F(0) = \omega$ and $F(n^+) = F(n)^+$ for each natural number n. We write $F(n) = \omega + n$ and ω^2 as the set consisting of all n and all $\omega + n$ with $n \in \omega$.

Theorem 8.1.6

If ξ is an element of an ordinal number α, ξ is also an ordinal number.

Theorem 8.1.7

If two ordinal numbers are similar, than they are equal.

Theorem 8.1.8

Every set of ordinal numbers is totally ordered. Every set of ordinal numbers is well ordered.

Theorem 8.1.9: Transfinite Ordinal Numbers

The natural numbers are finite ordinal numbers, the others are called transfinite. Each finite ordinal numbers other than 0 has an immediate predecessor. Transfinite ordinal numbers that do not have a predecessor is called limit numbers.

Theorem 8.1.10: Supremum of Ordinal Numbers

Every set of ordinal numbers have a supremum.

Theorem 8.1.11: Burali-Forti Paradox

No set contains all ordinal numbers.

Theorem 8.1.12: Counting Theorem

Each well ordered set is similar to a unique ordinal number.