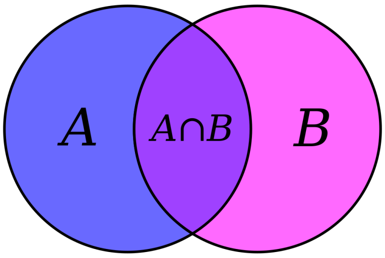
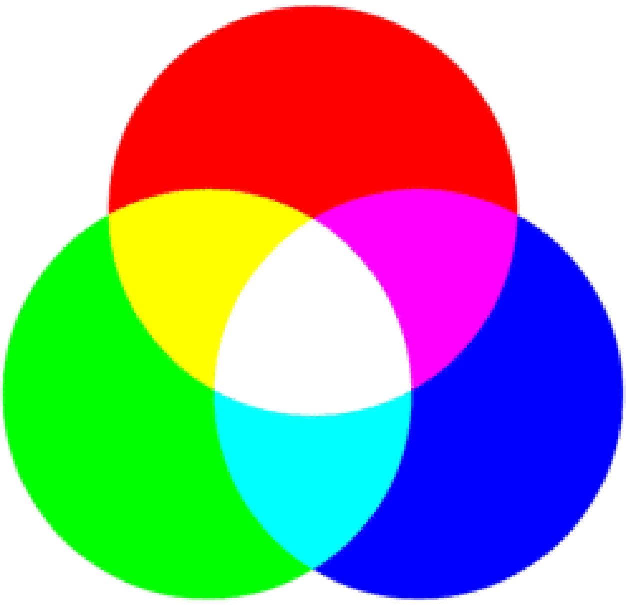
Set Theory and Logic

By: Shantanu Nene
Mentor: Shourya Pandey

July 2021





Contents

1 Introduction

Set theory is the fundamental building block of mathematics. It is the study of sets and classes, which are loosely aggregates or collections. This field originated due to the lack of mathematical rigor in the foundations of mathematics, in particular, definition of natural numbers, topology, abstract algebra and many such fields.

We start with naive set theory, setting up definitions and notations, then move on to a more rigorous axiomatic formulation of set theory. Then we construct the set of natural numbers using this, and generalize them to cardinal numbers. Then we study well-ordered sets and ordinal numbers. Then we introduce logic and build a theory of inference. Finally, we give a brief introduction to axiomatic theories.

2 Naive Set Theory

This was the first formulation of set theory, put forward by Georg Cantor. In it, a **set** is any collection of definite, distinguishable objects known as **elements**. For every set A and element x, either x **belongs** to A, denoted by $x \in A$, or x **does not belong** to A, denoted by $x \notin A$. Two sets A, B are said to be **equal** iff they have the same element. The set whose elements are $a, b, c, \dots$ is denoted as

$$\{a, b, c, \dots\}$$

2.1 Abstraction

We use the concept of **formula** to create sets. A **statement** is a declarative sentence capable of being classified as either true or false. A formula in some symbol x is a finite sequence of words and the symbol x such that when each occurrence of x is replaced by the same name of an object of an appropriate nature, a statement results. We will usually denote a formula in x as $P(x)$.

Naive Principle of Abstraction: For every formula $P(x)$, there exists a set A consisting of all objects a for which $P(a)$ is true.

Such a set is denoted by $\{x \mid P(x)\}$.

2.2 Subsets

We say that a set A is a **subset** of B if every element of A is also an element of B. We denote it by $A \subseteq B$. A is a **proper subset** of B if, in addition, we have $A \neq B$; we denote this by $A \subset B$. For every set S, we can define the **power set** of S, $\mathcal{P}(S)$, as the set whose elements are all the subsets of S.

2.3 Empty Set

The empty set can be defined by the formula $\{x \mid x \neq x\}$. This set will have no elements, by definition. We denote the empty set by $\varnothing$. Also, for any set A, $\varnothing$ is vacuously a subset of A.

2.4 Operations on Sets

For any sets A, B, we define their **intersection**, $A \cap B$, as the set of all elements that belong to *both* A and B, i.e.,

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

We also define their **union**, $A \cup B$, as the set of all elements that belong to *at least one of* A or B, i.e.,

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

For any two objects x, y, we define the **ordered pair** (x, y) to be the set $\{x, \{x, y\}\}$. Note that $(a, b) = (c, d)$ implies that $a = c$ and $b = d$. Now we can define the **Cartesian product**, $A \times B$, as

$$A \times B = \{(x, y) \mid x \in A \text{ and } y \in B\}$$

We also define the relative difference $A - B$ as:

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

2.5 Relations

A **relation** $\sim$ on a set A is a subset R of $A \times A$, and we write $x \sim y$ iff $(x, y) \in R$. A relation is **transitive** if $x \sim y$ and $y \sim z$ imply $x \sim z$. It is **symmetric** if $x \sim y$ implies $y \sim x$, and it is **reflexive** if $x \sim x$ for all $x \in A$. An **equivalence relation** is a transitive, symmetric and reflexive relation. For any $xinA$, the **equivalence class** containing x is the set of all $y \in A$ satisfying $x \sim y$. It is easy to see that any two distinct equivalence classes are disjoint, and the union of all possible equivalence classes is A.

An **order relation** is a transitive and reflexive relation that satisfies: $x \sim y$ and $y \sim x$ imply $x = y$ (sometimes this is stated as a **partial order**). We usually denote an order relation by $\leq$. We also use $x < y$ to denote $x \leq y$ and $x \neq y$. An order is a **total order** if $x \leq y$ or $y \leq x$ for all $x, y \in A$. A total order is a **well order** if for any $B \subseteq A$, there exists a $z \in B$ such that $z \leq x$ for all $x \in B$.

2.6 Functions

For any sets A, B, a **function** from A to B, denoted by $f : A \to B$, is a subset R of $A \times B$ such that for every $x \in A$, there exists a unique $y \in B$, denoted by $y = f(x)$, satisfying $(x, y) \in R$. A function is **injective** if for any distinct $x, z \in A$, we have $f(x) \neq f(z)$. A function is **surjective** if for any $y \in B$, there exists an $x \in A$ satisfying $f(x) = y$. A function is **bijective** if it is both injective and surjective. Also for any subset X of A, we define

$$f(X) = \{f(x) \mid x \in X\}$$

We say A is the **domain**, B is the **co-domain**, and $f(A)$ is the **range** of f.

2.7 Paradoxes of Naive Set Theory

The lack of rigor in naive set theory, in particular principle of abstraction, presents a bunch of paradoxes.

1. **Russel's Paradox**: Consider the set R defined by the formula $R = \{x \mid x \notin x\}$, i.e., the set of all sets that don't contain themselves. If $R \in R$, then by definition, $R \notin R$, and vice versa, so we have a contradiction.

2. **Cantor's Paradox**: Let S be the set of all sets, i.e., $\{x \mid x = x\}$. Then we must have $\mathcal{P}(S) \subseteq S$ since S contains all sets. That means we can define an injection from $\mathcal{P}(S)$ to S. But this is impossible, as we will prove in Theorem 10 (and as Cantor had proved).

Both of these paradoxes occur due to unrestricted usage of abstraction. This can be fixed in a more rigorous version of set theory.

3 Zermelo-Fraenkel Set Theory

This is an axiomatic approach to set theory developed by Ernst Zermelo and Abraham Fraenkel. We use the same concepts of sets, elements and inclusion developed in the first portion of naive set theory. We will also use the concept of formula. However these sets must follow the following axioms:

1. **Axiom of Extension**: For any sets A, B, if $x \in A$ iff $x \in B$, then $A = B$.
2. **Axiom Schema of Specification**: For any set A there exists a set B such that $x \in B$ iff $x \in A$ and $P(x)$; here $P(x)$ is a formula in x containing no occurrence of B. *We denote this in the same way as done in naive set theory, except as a subset of A.*
3. **Axiom of the Empty Set**: There exists a set, denoted by $\varnothing$, that contains no elements.
4. **Axiom of Pairing**: For any sets A, B, there exists a set C satisfying $A \in C$ and $B \in C$.
5. **Axiom of Union**: For any set A, there exists a set B such that $x \in B$ iff there exists a $C \in A$ satisfying $x \in C$. *This generalizes our previous definition of union.*
6. **Axiom of Power Set**: For any set A, there exists a set B such that $x \in B$ iff $x \subseteq A$.
7. **Axiom of Infinity**: There exists a set A such that $\varnothing \in A$, and $x \in A \implies x \cup \{x\} \in A$.
8. **Axiom Schema of Replacement**: If $P(x, y)$ is a formula such that for each x in a set A, $P(x, y)$ and $P(x, z)$ imply that $y = z$, then there exists a set B such that $y \in B$ iff there exists an $x \in A$ satisfying $P(x, y)$.
9. **Axiom of Regularity**: For any non-empty set A, there exists an element $B \in A$ such that for any $x \in A$, $x \notin B$.
10. **Axiom of Choice**: For any non-empty set A, there exists a function $f : \mathcal{P}(A) \to A$ such that $f(B) \in B$ for any non-empty $B \in \mathcal{P}(A)$. *Such a function is called a* ***choice function****.*

Theorem 1. For any sets x, y, the set $\{x, y\}$ exists.

Proof. By Pairing, we can find a set C containing both x and y. Then, using Specification, the subset $\{z \in C \mid z = x \text{ or } z = y\}$ exists. This is precisely the set $\{x, y\}$. □

This also implies that the ordered pair (x, y) exists. This guarantees existence of Cartesian product $A \times B$ as a subset of $\mathcal{P}(\mathcal{P}(A \cup B))$.

Theorem 2 (Existence of Intersection)**.** For any set A, there exists a set B such that $x \in B$ iff $x \in C$ for all $C \in A$.

Proof. Let D be any set in A. Then $\bigcap A = \{x \in D \mid x \in C \text{ for all } C \in A\}$, and we are done by Union and Specification. □

Exercise 3.1. There don't exist sets A, B satisfying $A \in B$ and $B \in A$.

Proof. Assume FTSOC that these sets do exist. By Theorem 1, the set $\{A, B\}$ exists. By Regularity, there exists an $C \in \{A, B\}$ such that $C \cap \{A, B\} = \varnothing$. But we can't have $C = A$ or $C = B$ by our initial assumption, contradiction! So such sets cannot exist. □

In particular, there is no set that contains itself.

Exercise 3.2. The only set A satisfying $A \subseteq A \times A$ is the empty set.

Proof. Assume FTSOC that such an A is non-empty. Then, by Regularity, we can find an element $z \in A$ such that $z \cap A = \varnothing$. Then note that $z \in A \times A$, so there exist $x, y \in A$ such that $z = (x, y)$. But then, $x \in z$ and $x \in A$, so $x \in z \cap A$, contradiction! So the only such A is the empty set. □

3.1 Consequences of the Axiom of Choice

The Axiom of Choice (or AOC) has many interesting and sometimes counter-intuitive applications.

For any ordered set S, we define a **chain** in S to be a subset $T \subseteq S$ such that for any $x, y, z \in S$ with $x \leq y \leq z$, if $x, z \in T$, then $y \in T$ as well. An **upper bound** of any subset $A \subseteq S$ is an element z such that $x \leq z$ for all $x \in A$; and z is a **least upper bound** if for any upper bound y of A, we have $z \leq y$.

Theorem 3 (Bourbaki-Witt Theorem)**.** Suppose S is a partially ordered set such that every chain in S has a least upper bound. If $f : S \to S$ satisfies $x \leq f(x)$ for all $x \in S$, then there exists a $y \in S$ satisfying $f(y) = y$.

Proof. Omitted due to length. □

Note that $\subseteq$ is a valid order on any set. Also, this set will satisfy that every chain has a least upper bound; indeed, for any chain C, the required lub is $\bigcup C$.

Theorem 4 (Hausdorff's Maximality Principle)**.** Every partially ordered set contains a maximal chain, i.e., a chain which is not a proper subset of any other chain.

Proof. Let S be the set, and let $C \subseteq \mathcal{P}(S)$ be the set of all chains of S. If the theorem is false, then for every $X \in C$, the set

$$C_X = \{Y \in C \mid X \subset Y\}$$

is non-empty. By AOC, there exists an $f : \mathcal{P}(C) \to C$ such that $f(C_X) \in C_X$ for all $X \in C$. Define $g : C \to C$ as $g(X) = g(C_X)$ for all $X \in C$. Then, if we consider C to be a partially ordered set with order $\subseteq$, g satisfies conditions of Bourbaki-Witt Theorem. But, $X \subset g(X)$ for all $X \in C$, so we can never have $g(Y) = Y$, contradiction! □

Theorem 5 (Zorn's Lemma)**.** Suppose S is a partially ordered set such that every chain in S has a least upper bound. Then S itself has a least upper bound.

Proof. By Hausdorff's Maximality Principle, there exists a maximal chain C in S. Let x be an lub of C. If x is not an upper bound of S, then there exists a $y \in S$ with $x < y$. But that means $C \cup \{y\}$ is also a chain, whcih contradicts maximality of C. Therefore x is an upper bound of S, and since $x \in S$, it must indeed be the required lub. □

Theorem 6 (Well-ordering Theorem)**.** Every set can be well ordered, i.e., for every set there exists a well order on it.

Proof. Omitted due to length. □

Exercise 3.3. Suppose A is partially ordered by some order $\leq$. Prove that there exists a total order $\leq'$ on A such that for any $x, y \in A$, $x \leq y \implies x \leq' y$.

The following puzzle also demonstrates the 'weirdness' of AOC:

Exercise 3.4. An infinite number of prisoners, each with an unknown and randomly assigned red or blue hat, line up in a single file. Each prisoner faces away from the beginning of the line, and each prisoner can see all the hats in front of him, and none of the hats behind. Starting from the beginning of the line, each prisoner must correctly identify the color of his hat or he is killed on the spot. The prisoners have a chance to meet beforehand to come up with a strategy, but once in line, no prisoner can hear what the other prisoners say. Prove that there is a way to ensure that only finitely many prisoners are killed.

4 Natural Numbers

The set of natural numbers can be derived using the Axiom of Infinity (in fact, without this axiom, we cannot guarantee the existence of infinite sets). Call a set a **successor set** if it satisfies the conditions of the Axiom of Infinity.

Theorem 7 (Existence of Natural Numbers)**.** There exists a unique successor set that is a subset of every successor set.

Proof. It is easy to see that intersection of a non-empty collection of successor sets is also successor. So if A is some successor set, the intersection of all successor sets contained in A is also a successor set, say ω. We claim that ω is the required set. Indeed, if B is any successor set, then $A \cap B$ is a successor set which is contained in A, so $\omega \subseteq A \cap B \subseteq B$. The uniqueness of ω is trivial. □

This is indeed the set of natural numbers, which we also denote by $\mathbb{N}$. We write $0 = \varnothing$, $1 = \{0\}$, $2 = \{0, 1\}$ and so on.

4.1 Peano's Axioms

Peano's Axioms describe the properties of the natural numbers, and are often called the 'building blocks' of much of mathematics. They are:

1. $\mathbb{N}$ is non-empty. In particular, there is an element which we will denote by 0 in $\mathbb{N}$.
2. There exists a function $S : \mathbb{N} \to \mathbb{N}$, which we will call the **successor function**.
3. S is injective.
4. There does not exist any $n \in \mathbb{N}$ satisfying $S(n) = 0$.
5. (**Principle of Induction**) For any $K \subseteq \mathbb{N}$, if K satisfies both of the following properties:
   - $0 \in K$; and
   - $\forall n \in \mathbb{N}, n \in K \implies S(n) \in K$;

   then $K = \mathbb{N}$.

These can be easily derived from our construction of $\mathbb{N}$. In particular, $S(m) = m \cup \{m\}$, and the Principle of Induction follows directly from Theorem 7.

Exercise 4.1. Prove that for any non-zero $n \in \mathbb{N}$, there exists an $m \in \mathbb{N}$ such that $n = S(m)$.

Such an m is called a **predecessor** of n.

4.2 Addition

Addition is a binary operation defined on $\mathbb{N}$, and denoted by $+ : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$. It satisfies the following properties:

- For all $n \in \mathbb{N}$, $n + 0 = n$
- For all $m, n \in \mathbb{N}$, $n + S(m) = S(n + m)$

Using the Principle of Induction, we can easily see that addition is well-defined.

This addition clearly aligns with the traditional concept of addition, and so must ideally follow the same laws. Indeed this is true, as we can prove in the following results, left as exercises:

Exercise 4.2 (Associativity of Addition)**.** For all $a, b, c \in \mathbb{N}$,

$$a + (b + c) = (a + b) + c$$

Proof. Consider $K = \{c \in \mathbb{N} \mid a + (b + c) = (a + b) + c \ \ \forall a, b \in \mathbb{N}\}$. First,

$$a + (b + 0) = a + b = (a + b) + 0$$

so $0 \in K$. Now, if $c \in K$, then

$$\begin{aligned} a + (b + S(c)) &= a + S(b + c) \\ &= S(a + (b + c)) \\ &= S((a + b) + c) \\ &= (a + b) + S(c) \end{aligned}$$

for all $a, b \in \mathbb{N}$, and so $S(c) \in K$. Therefore $K = \mathbb{N}$ by the principle of induction, and we are done $\square$

Exercise 4.3 (Commutativity of Addition)**.** For all $a, b \in \mathbb{N}$,

$$a + b = b + a$$

Exercise 4.4 (Cancellation Laws for Addition)**.** For all $a, b, c \in \mathbb{N}$,

- $a + c = b + c \implies a = b$
- $c + a = c + b \implies a = b$

4.3 Multiplication

Multiplication is another a binary operation defined on $\mathbb{N}$, and denoted by $\cdot : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$. It satisfies the following properties:

- For all $n \in \mathbb{N}$, $n \cdot 0 = 0$
- For all $m, n \in \mathbb{N}$, $n \cdot S(m) = n \cdot m + n$

Again, the well-definiteness of multiplication follows from the principle of induction. We sometimes omit the $\cdot$ sign, so xy is the same as $x \cdot y$.

Similar to addition, the traditional laws of multiplication also hold:

Exercise 4.5 (Distributive Property)**.** For all $a, b, c \in \mathbb{N}$,

- $a \cdot (b + c) = a \cdot b + a \cdot c$
- $(b + c) \cdot a = b \cdot a + c \cdot a$

Exercise 4.6 (Associativity of Multiplication)**.** For all $a, b, c \in \mathbb{N}$,

$$a \cdot (b \cdot c) = (a \cdot b) \cdot c$$

Exercise 4.7 (Commutativity of Multiplication)**.** For all $a, b \in \mathbb{N}$,

$$a \cdot b = b \cdot a$$

Exercise 4.8 (Cancellation Laws for Multiplication)**.** For all $a, b, c \in \mathbb{N}$, if $c \neq 0$,

- $a \cdot c = b \cdot c \implies a = b$
- $c \cdot a = c \cdot b \implies a = b$

We also briefly define **exponentiation** as follows:

- For all $n \in \mathbb{N}$, $n^0 = 1$.
- For all $m, n \in \mathbb{N}$, $n^{S(m)} = n^m \cdot n$

4.4 Order

For any $m \in \mathbb{N}$, using a proof similar to Theorem 7, we can prove the existence of a unique successor set containing m that is a subset of every successor set containing m. Call this set D_m, the set of **descendants** of m. Then we define a relation $\leq$ on $\mathbb{N}$ such that $m \leq n$ iff $n \in D_m$.

Theorem 8. $\leq$ well-orders $\mathbb{N}$.

Proof. Omitted due to length. □

Exercise 4.9. Prove that, for any $m, n \in \mathbb{N}$, $m \leq n$ iff $n = m + k$ for some $k \in \mathbb{N}$.

Exercise 4.10. Prove that, for any $m, n, k \in \mathbb{N}$, $m \leq n \implies m + k \leq n + k$.

Exercise 4.11. Prove that, for any $m, n, k \in \mathbb{N}$, where $k \neq 0$, $m < n \implies mk < nk$.

5 Cardinal Numbers

While we have defined natural numbers as sets, we can also think of them as a representative of 'size' of their sets. Naively, the size of a set is the number of elements in that set. More rigorously, two sets A, B are said to be **equipotent** iff there exists a bijection $f : A \to B$. To give a 'size' to any set, first we have to consider a universal set U (to avoid paradoxes with naive abstraction), whose subsets and elements are the only sets we consider. Then equipotency can be easily checked to be an equivalence relation on $\mathcal{P}(U)$, and its equivalence classes are called **cardinal numbers**. For any $A \subseteq U$, its equivalence class is denoted by $card(A)$ or $|A|$, and is called the **cardinality** of A. This definition does depend on U, however any property of cardinal numbers can be translated into a property about sets and bijections between them.

5.1 Comparison of Cardinals

We say that a set A **dominates** set B if B is equipotent with some subset of A, and we write this as $|B| \leq |A|$. Clearly $|A| = |B| \implies |A| \leq |B|$ and $|B| \leq |A|$. The converse is also true:

Theorem 9 (Schroeder-Bernstein Theorem)**.** For any sets A, B, $|A| \leq |B|$ and $|B| \leq |A|$ imply $|A| = |B|$.

Proof. This is just a sketch. The assumptions give an injective function $f : B \to A$ and an injective function $g : B \to A$. Note that it is sufficient to find a subset A_1 of A such that $g(B - f(A_1)) = A - A_1$. Consider the collection F of all subsets A_0 of A satisfying $A - g(B) \subseteq A_0$ and $g(f(A_0)) \subseteq A_0$. F is non-empty since $A \in F$. Then we can just take $A_1 = \bigcap F$, and this can be checked to satisfy the required condition. □

Therefore cardinal ordering is indeed a valid order. We also define $|A| < |B|$ when $|A| \leq |B|$ but $|A| \neq |B|$.

Theorem 10 (Cantor's Theorem)**.** $|S| < |\mathcal{P}(S)|$ for all non-empty sets S.

Proof. Assume the contrary; then there exists an injection f from $\mathcal{P}(S)$ to S. Define a new subset T of S as follows: $x \in T$ iff there exists an $A \subseteq S$ such that $f(A) = x$ and $x \notin A$ (note that A must be unique, if it exists). Now it is easy to see that $f(A)$ cannot exist, contradiction! □

The above theorem was first proved by Cantor using naive set theory, and thus forms the basis for Cantor's paradox.

5.2 Finite Cardinals

We call the cardinality of the empty set to be 0. Also, we define cardinality of any set $A \cup \{c\}$ where $c \notin A$ to be $|A| + 1$. It is easy to see that $|A| = |B| \implies |A| + 1 = |B| + 1$. It is also easy to see that $|A| \leq |B| \implies |A| + 1 \leq |B| + 1$. These two definitions allow us to assign natural numbers as cardinal numbers to some sets. They are called **finite cardinals** and the sets are called **finite sets**. A set which is not finite is called an **infinite set**. Now natural numbers can also be ordered using the ordering for cardinals; we will call this $\leq_C$ and reserve $\leq$ for the natural ordering.

Theorem 11. For any $n \in \mathbb{N}$, the finite cardinal n is the cardinality of the set of natural numbers which precede n in the natural ordering.

Proof. This is not hard to prove using induction. □

Theorem 12. For any $n \in \mathbb{N}$, if $|A| = n$ for some set A, then $|B| < |A|$ for all $B \subset A$.

Proof. This is just a sketch. We use induction on n. $n = 0$ is vacuously true. Assume the theorem is true for some $n \in \mathbb{N}$, and let $|A| = n + 1$. Then write $A = D \cup \{d\}$ where $|D| = n$. If $B \subseteq D$, we are done since $|B| \leq |D| < n + 1$. Else write $B = E \cup \{d\}$ where $E \subset D$, then $|B| = |E| + 1 < |D| + 1 = |A|$. □

Theorem 13. The natural ordering and cardinal ordering agree, i.e., $m \leq n$ iff $m \leq_C n$.

Proof. $m \leq n \implies m \leq_C n$ is trivial from Theorem 11. Also, if $n < m$, then by Theorem 11 and Theorem 12, we must have $n <_C m$ as well, so $m \leq_C n \implies m \leq n$, and we are done. □

Thus we can use the two orderings interchangeably.

5.3 Countable Sets

We introduce our first **infinite** or **transfinite cardinal**: the cardinality of $\mathbb{N}$, denoted by $\aleph_0$. We also define $\aleph_1 = |\mathcal{P}(\mathbb{N})| > \aleph_0$.

Theorem 14 ($\mathbb{N}$ is infinite)**.** $n < \aleph_0$ for all $n \in \mathbb{N}$.

Proof. $n + 1 \leq \aleph_0$ is trivial from Theorem 11. But note that $n < n + 1$, and we are done. □

A set is called **countable** if it is either finite or equipotent to $\mathbb{N}$. In the latter case it is called **countably infinite**.

Exercise 5.1. Prove that, for any $m \in \mathbb{N}$, the set of descendants of m is countably infinite.

Exercise 5.2. Prove that $\mathbb{N} \times \mathbb{N}$ is countably infinite.

Proof. We can find a direct bijection between $\mathbb{N} \times \mathbb{N}$ and $\mathbb{N} - \{0\}$ as follows:

$$f(a, b) = 2^a(2b + 1)$$

□

Theorem 15. Any subset of a countable set is countable.

Proof. WLOG the countable set is $\mathbb{N}$, and let A be any subset. If A is finite, we are done. Else inductively define $g(n)$ to be the smallest number in A which is not equal to $g(m)$ for any $m < n$. Note that $g(n) \geq n$ for all n, and from here it is easy to see that g is a bijection between $\mathbb{N}$ and A. □

Exercise 5.3. Prove that, if the domain of a function is countable, then so is its range.

Theorem 16. Let C be a countable collection of countable sets. Then $\bigcup C$ is countable.

Proof. It is not hard to find an injection from $\bigcup C$ to $\mathbb{N} \times \mathbb{N}$, by thinking of each set in C as a 'line' parallel to the 'Y axis'. □

5.4 Cardinal Arithmetic

We generalize addition, multiplication and exponentiation to arbitrary cardinals.

Sum of two cardinals u, v, corresponding to disjoint sets A, B, is the cardinality of $A \cup B$, and is denoted by $u + v$.

Theorem 17. For any cardinals u, v, w,

1. $u + v = v + u$
2. $(u + v) + w = u + (v + w)$
3. $u \leq v \implies u + w \leq v + w$

Proof. Omitted due to length. □

Product of two cardinals u, v, corresponding to sets A, B, is the cardinality of $A \times B$, and is denoted by $u \cdot v$ or simply uv.

Theorem 18. For any cardinals u, v, w,

1. $uv = vu$
2. $(uv)w = u(vw)$
3. $u \leq v \implies uw \leq vw$
4. $(u + v)w = uw + vw$

Proof. Omitted due to length. $\square$

Exercise 5.4. Prove that

$$n + \aleph_0 = n\aleph_0 = \aleph_0 \cdot \aleph_0 = \aleph_0$$

for any $n \in \mathbb{N}$.

For any cardinals u, v, corresponding to sets A, B, the v^{th} **power** of u is the cardinality of A^B, i.e., the set of functions with domain B and co-domain A, and is denoted by u^v.

Theorem 19. For any cardinals u, v, w,

1. $u^v \cdot u^w = u^{v+w}$
2. $u^w \cdot v^w = (uv)^w$
3. $(u^v)^w = u^{vw}$
4. $u \leq v \implies u^w \leq v^w$
5. $u \leq v \implies w^u \leq w^v$

Proof. Omitted due to length. $\square$

Exercise 5.5. Verify that cardinal operations agree with the usual operations on natural numbers.

Exercise 5.6. Prove that $|\mathcal{P}(S)| = 2^{|S|}$ for all sets S.

Exercise 5.7. Prove that

$$\aleph_0^{\aleph_0} = \aleph_1^2 = \aleph_1$$

5.5 Infinite Cardinals

$\aleph_0$ is the 'smallest' infinite cardinal in some sense, as shown below:

Theorem 20. If A is an infinite set, the $\aleph_0 \leq |A|$.

Proof. We use AOC for this result. Let f be a choice function on A. We define an injective function $g : \mathbb{N} \to A$ inductively as follows: $g(0) = f(A)$, and $g(n) = f(A - \{g(m) \mid m \in \mathbb{N} \text{ and } m < n\})$ for all $n > 0$. It is easy to see that g is indeed the required injection. $\square$

Exercise 5.8. For any infinite set A, prove that there exists a proper subset $B \subset A$ which is equipotent with A.

Theorem 21. If C is a set of cardinal numbers, then there exists a cardinal number greater than all cardinals in C.

Proof. Note that C is just a collection of equivalence classes of some universal set. Therefore, using AOC, we can choose a set A_u for each cardinality u in C. Let $F = \{A_u \mid u \in C\}$. Then $|\bigcup F| \geq u$, so $|2^{\bigcup F}| > u$ for all $u \in C$. $\square$

Finally, here are some properties of infinite cardinal arithmetic:

Theorem 22. For any infinite cardinal u, and any other cardinal v,

1. $v \leq u \implies u + v = u$
2. $v \leq u \implies uv = u$

Proof. Omitted due to length. $\square$

6 Ordinal Numbers

In the last section we studied sizes of sets, in this section we will study order relations. For two totally ordered sets X, Y, with order $\leq, \leq'$ respectively, we say that X and Y are **ordinally similar** iff there exists a bijection $f : X \to Y$ satisfying $a \leq b$ iff $f(a) \leq' f(b)$. Such an f is called an **order isomorphism** (or just isomorphism whenever implied) between the two sets. If f is not bijective, but still preserves order, it is called an **order embedding** (or just embedding whenever implied). We will only work with well orders, which exist for all sets due to the Well-Ordering Theorem.

6.1 Well-Ordered Sets

First we will prove some properties of well-ordered sets.

Theorem 23. If A is well ordered and $f : A \to A$ is an embedding, then $x \leq f(x)$ for all $x \in A$.

Proof. Assume that for some x we have $f(x) < x$. Let B be the set of all such x, and let y be the least element in B. Then, $f(y) < y$, so $f(f(y)) < f(y) \implies f(y) \in B$, which contradicts minimality of y. □

For any well-ordered set A and any $x \in A$, the **initial segment** determined by x is the set $\{y \in A \mid y < x\}$ and is denoted by A_x.

Theorem 24. A well ordered set is not ordinally similar to any of its initial segments.

Proof. Assume FTSOC that A is ordinally similar to some A_x, and let f be an isomorphism. Then f is an embedding from A to A, so by the previous theorem $x \leq f(x)$, contradiction since $f(x) \in A_x$. □

Exercise 6.1. Prove that if A is well ordered, then A_x and A_y are ordinally similar iff $x = y$.

Theorem 25. If A and B are well-ordered and ordinally similar, then there is a unique isomorphism from A to B.

Proof. Assume there are two distinct isomorphisms g, h from A to B. Then there exists a $y \in A$ with $g(y) \neq h(y)$; WLOG $g(y) > h(y)$ Then the function $f : A \to A$ satisfying $g(f(x)) = h(x)$ is an isomorphism. Therefore $x \leq f(x) \implies g(x) \leq g(f(x)) = h(x)$ for all $x \in A$. Contradiction! □

If A is well-ordered, and B is any non-empty set, then a **sequence of type** x is a function from A_x to B. Using this we can state a generalization of the principle of induction:

Theorem 26 (Principle of Transfinite Induction)**.** Let A be well-ordered having least element a_0. Let B be any set containing some element c. Let $\mathcal{G}$ be the set of all sequences of some type x in B over all $x \neq a_0$. Then for any $f : \mathcal{G} \to B$, there exists a unique function $g : A \to B$ satisfying $g(a_0) = c$ and $g(x) = f(g|_{A_x})$ for all $x \neq a_0$ in A, where $g|_{A_x}$ is the function g restricted to the domain A_x.

Proof. Omitted due to length. □

Exercise 6.2. Let A be well-ordered having least element a_0, and let $K \subseteq A$ satisfy:

- $a_0 \in K$; and
- $A_x \in K \implies x \in K$ for all $x \neq a_0$ in A.

Then prove that $K = A$.

Theorem 27. If A and B are well ordered, then exactly one of the following hold:

- A and B are ordinally similar.

- A is ordinally similar to some initial segment of B.
- B is ordinally similar to some initial segment of A.

Proof. Assume neither A nor B is ordinally similar to any initial segment of the other. Let a_0, b_0 be the least elements of A, B respectively., and choose any $x \neq a_0$ in A. For any sequence j of type x in B, let $f(j)$ be the least upper bound of j in B if it exists; else let $f(j) = b_0$. Therefore by using transfinite induction, it is easy to see that we can find a function $g : A \to B$ such that g maps an initial segment determined by x to an initial segment determined by $g(x)$. Now it is easy to verify that this is indeed the required isomorphism, and A and B are ordinally similar. □

Note that if A is similar to some initial segment of B, then $|A| \leq |B|$. Since every set can be well-ordered, the above theorem, along with Schroeder-Bernstein Theorem also proves:

Theorem 28 (Law of Trichotomy for Cardinals)**.** For any set A, B, exactly one of the following holds: $|A| = |B|$, $|B| < |A|$ or $|A| < |B|$.

6.2 Von Neumann's Construction

In this construction, we define ordinal numbers as specific well-ordered sets, rather than equivalence classes like in the case of cardinals. The natural numbers are themselves well-ordered sets, and for any $n \in \omega$ (we will use ω for natural numbers in this section) and any $m < n$, the initial segment determined by m (which we will just call $s(m)$) is m itself. That is, a natural number is a well ordered set such that the initial segment determined by each of its elements is the element itself. We generalize this to define **ordinal numbers**: well-ordered sets α that satisfy $s(x) = x$ for any $x \in \alpha$.

Exercise 6.3. Prove that, if α is an ordinal, then so is $\alpha^+ = \alpha \cup \{\alpha\}$, where we extend the well order on α by defining $x < \alpha$ for all $x \in \alpha$. Furthermore, prove that α and α^+ are not ordinally similar.

Note that ω itself is an ordinal, and hence so are $\omega^+, (\omega^+)^+, \ldots$ and so on. We will write $\alpha^+ = \alpha + 1$, $(\alpha^+)^+ = \alpha + 2$ and so on. Applying the Axiom Schema of Replacement with $A = \omega$ and $P(x, y)$ denoting $y = \omega + x$, we can see that $\{\omega, \omega + 1, \omega + 2, \ldots\}$ is a set. We will call the union of this set along with ω as $\omega 2$. Note that $\omega 2$ is also an ordinal (in the same way as ω is). In the same way we can define $\omega 2 + 1, \omega 2 + 2, \ldots$ up to $\omega 3$, then $\omega 4$, and so on. Using Replacement once again with $A = \omega$ and $P(x, y)$ as $y = \omega x$ gives us the ordinal ω^2, and we can continue doing this to define 'polynomials' in ω, even ω^ω and so on.

Exercise 6.4. Prove that each element of an ordinal number is itself an ordinal number.

The following property cements the uniqueness of ordinals:

Theorem 29. Two distinct ordinal numbers cannot be ordinally similar.

Proof. Let f be an order isomorphism from some ordinal α to some ordinal β. Then it is easy to prove using transfinite induction that $f(x) = x$ for all $x \in \alpha$; thus, $\alpha = \beta$. □

6.3 Ordering the Ordinals

Defining an order on the ordinals isn't very hard: $\alpha < \beta$ iff α is ordinally similar to some initial segment of β. Then Theorem 27 implies that this is indeed a total order.

Theorem 30. Any set of ordinals is well-ordered.

Proof. Let O be a set of ordinals, and let $\alpha \in O$. Assume α is not the least element, then $\alpha \cap O$ is a non-empty subset of α, and so is well-ordered. Let β be its smallest element. Note that $\beta < \alpha$. We claim that this is the required the least element of O. Indeed, $\gamma < \beta \implies \gamma < \alpha \implies \gamma \in \alpha \cap O$, which contradicts minimality of β. □

Exercise 6.5. Let O be a set of ordinals. Prove that there exists an ordinal greater than all ordinals in O.

Exercise 6.6. Prove that any non-empty set of ordinals has a least upper bound.

Proof. Let O be a set of ordinals. It is easy to see that $\bigcup O$ is also an ordinal, and is clearly an upper bound of O. It is also easy to see that it is the least upper bound, since $\alpha \leq \beta \implies \alpha \subseteq \beta$ for any ordinals α, β. $\square$

The next result talks about usage of ordinals to compare well-ordered sets:

Theorem 31. Each well-ordered set is ordinally similar to a unique ordinal number.

Proof. Omitted due to length. $\square$

Such an ordinal number associated with a well-ordered set A is called the **ordinality** of A, and write it as $ord(A)$.

Before we end this section, we derive the following result about cardinals from ordinals:

Theorem 32. Any set of cardinals is well-ordered.

Proof. First, note that using AOC we can assume that we have a set C of sets corresponding to the required cardinalities. For any set $A \in C$, let $Z(A)$ be the set of all ordinals associated with some well-ordering of A (this set indeed exists because the set of well-orders is a subset of $\mathcal{P}(A \times A)$, and so we can use Specification and Replacement). Let α_A denote the least member of A. Note that $\alpha_A = \alpha_B \implies |A| = |B| = |\alpha_A|$, which is impossible. Also $|A| < |B| \implies |\alpha_A| < |\alpha_B| \implies \alpha_A < \alpha_B$. So we get a corresponding set of ordinals that preserves order, and the later can be well-ordered by Theorem 30. $\square$

6.4 Ordinal Arithmetic

Ordinal arithmetic is a bit more complicated than its cardinal counterpart.

For any ordered set O of ordinals, we can define the sum of elements of O in the following way: to each ordinal $\alpha \in O$, we assign a set S_α with the same ordinality such that all the sets S_i are disjoint. Let T be the union of all the S_i. We define an ordering on T as follows: If x, y are in the same S_α, then we apply the ordering of that set; else, assume $x \in S_\alpha$ and $y \in S_\beta$ where α appears before β in the ordering of O, then define $x < y$. Then the sum of elements of O is the ordinality of T. Note that, if instead we took O to be a set of sets, and do a similar construction, then we can define addition of a collection of ordinals even if it contains repeats.

Products of ordinals can be defined in a similar way: given an ordered set of ordinals O, let T be the set of functions from O to $\bigcup O$ such that $f(\alpha) \in \alpha$ for all $\alpha \in O$. Given two such functions f, g, let α be the smallest element according to ordering of O such that $f(\alpha) \neq g(\alpha)$. Then define $f < g$ iff $f(\alpha) < g(\alpha)$. Then product of elements of O is defined as the ordinality of T. Again, we can accommodate for repeats here.

Exponentiation can just be defined as repeated multiplication. Indeed, for α^β, let O be a set of ordinality β containing sets of ordinality α. We can define T in the same way as above, and then define α^β as the ordinality of T.

Unfortunately, neither addition nor multiplication of ordinals is commutative. For example, $1 + \omega = \omega \neq \omega + 1$, and $2\omega = \omega \neq \omega 2$. However, a few other properties do hold:

Theorem 33. For any ordinals α, β, γ,

1. $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$
2. $(\alpha\beta)\gamma = \alpha(\beta\gamma)$
3. $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$
4. $\alpha < \beta \implies \gamma + \alpha < \gamma + \beta$ and $\alpha + \gamma \leq \beta + \gamma$

5. $\alpha < \beta$ and $\gamma > 0 \implies \gamma\alpha < \gamma\beta$ and $\alpha\gamma \leq \beta\gamma$
6. $\alpha^\beta \cdot \alpha^\gamma = \alpha^{\beta+\gamma}$
7. $(\alpha^\beta)^\gamma = \alpha^{\beta\gamma}$
8. $\alpha > 1$ and $\beta < \gamma \implies \alpha^\beta < \alpha^\gamma$
9. $\alpha < \beta \implies \alpha^\gamma \leq \beta^\gamma$

Proof. Omitted due to length. □

Exercise 6.7. Prove that there exists an ordinal ϵ satisfying the equation $\omega^\epsilon = \epsilon$.

7 Beyond the Natural Numbers

7.1 Integers

The **integers** are constructed from the natural number using equivalence classes and a concept of subtraction. In particular, consider a relation $\sim$ on $\mathbb{N}\times\mathbb{N}$ such that $(a,b)\sim(c,d)$ iff $a+d=b+c$. It is easy to check that this is indeed an equivalence relation. We identify the integers with the equivalence classes of this relation, and let $\mathbb{Z}$ denote the set of integers.

Addition and multiplication also carry over to the integers. Given $A, B \in \mathbb{Z}$, define

$$A + B = \{(a+c, b+d) \mid (a,b)\in A \text{ and } (c,d)\in B\}$$

$$A \cdot B = \{(ac+bd, bc+ad) \mid (a,b)\in A \text{ and } (c,d)\in B\}$$

Well-definiteness of these operations can be easily checked.

We can imbed the natural numbers in integers by defining n as the equivalence class containing $(n,0)$ for all natural numbers n. This also defines 0, the **additive identity**, in $\mathbb{Z}$. There's an additional concept of **additive inverse**, defined as

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

Well definiteness: $(a,b)\sim(a',b') \implies a+b'=b+a' \implies (b,a)\sim(b',a')$. We can check that $A+(-A)=0$. Thus we can define **subtraction** as follows: $A-B=A+(-B)$. We can also carry over order relations: $A>B \iff A-B$ is a non-zero natural number.

It is easy to check that addition and multiplication satisfy the same theorems as natural numbers, and that order satisfies transitivity and law of trichotomy. Further, Exercises 4.9 and 4.10 hold for all $m,n,k\in\mathbb{Z}$, and Exercise 4.11 also holds but with the additional condition that $k>0$. In addition:

Exercise 7.1. Prove that, for any $m,n,k\in\mathbb{Z}$ with $k<0$, $m<n \implies n\cdot k<m\cdot k$.

What about the cardinality of this set? Well, it is not hard to see that:

Exercise 7.2. Prove that $|\mathbb{Z}|=\aleph_0$.

7.2 Rational Numbers

The **rational numbers** are constructed from the integers using equivalence classes and a concept of division. In particular, consider a relation $\sim$ on $\mathbb{Z}\times\mathbb{Z}\setminus 0$ such that $(a,b)\sim(c,d)$ iff $a\cdot d=b\cdot c$. It is easy to check that this is indeed an equivalence relation. We identify the rational numbers with the equivalence classes of this relation, and let $\mathbb{Q}$ denote the set of rational numbers.

Addition, subtraction and multiplication also carry over to the rationals. Given $A,B\in\mathbb{Q}$, define

$$A + B = \{(ad+bc, bd) \mid (a,b)\in A \text{ and } (c,d)\in B\}$$

$$A - B = \{(ad-bc, bd) \mid (a,b)\in A \text{ and } (c,d)\in B\}$$

$$A \cdot B = \{(ac, bd) \mid (a,b)\in A \text{ and } (c,d)\in B\}$$

Well-definiteness of these operations can be easily checked.

We can imbed the integers in the rationals by defining n as the equivalence class containing $(n,1)$ for all integers n. This also defines 1, the multiplicative identity, in $\mathbb{Q}$. Note that $(0,a)\sim(0,b)$ for all integers $a,b\neq 0$ since $0\cdot b=0\cdot a=0$. Thus we can also define 0 in $\mathbb{Q}$ as the set $\{(0,n)\mid n\in\mathbb{Z} \text{ and } n\neq 0\}$. There's an additional concept of **multiplicative inverse** or **reciprocal**, for non-zero rationals A, defined as

$$\frac{1}{A} = \{(b,a) \mid (a,b)\in A\}$$

Well definiteness: $(a, b) \sim (a', b') \implies ab' = ba' \implies (b, a) \sim (b', a')$. We can check that $A \cdot \dfrac{1}{A} = 1$. Thus we can define **division** for rationals A, B with $B \neq 0$ as follows: $\dfrac{A}{B} = A \cdot \dfrac{1}{B}$. Definition of order is a bit non-trivial. We call a rational number A **positive** if $(m, n) \in A$ for some integers $m > 0$ and $n > 0$. Then define $A > B \iff A - B$ is positive.

It is easy to check that $(a, b) \in x \implies x = \dfrac{a}{b}$ for all $a, b \in \mathbb{Z}$. Thus every rational number can be written as a fraction of two integers.

It is easy to check that rationals satisfy the same properties of addition, subtraction, multiplication and order as integers. With the existence of multiplicative inverse as well, this means $\mathbb{Q}$ is, what is called, an **ordered field**.

We briefly define exponentiation: For any $a \in \mathbb{Q}$, define $a^0 = 1$, and $a^{n+1} = a^n \cdot a$ for all natural numbers n. Also define $a^{-1} = \dfrac{1}{a}$ and $a^{-n} = (a^{-1})^n$ for all integers $n > 0$. We also briefly define absolute value: $|0| = 0$, $|a| = a$ if $a > 0$ and $|a| = -a$ if $a < 0$. It is easy to check that $a^2 = |a|^2$ for all rationals a. Also check that $|a| < M \implies -M < a < M$. The following inequality is very important:

Exercise 7.3 (Triangle Inequality)**.** For all rational numbers x, y, prove that $|x + y| \leq |x| + |y|$.

The following result would also be useful later on:

Exercise 7.4. For any rationals $0 < a < 1$ and $\varepsilon > 0$, prove that there exists a positive integer N such that $a^n < \varepsilon$ for all integers $n > N$.

It is not very hard to prove the following about the cardinality of $\mathbb{Q}$:

Exercise 7.5. Prove that $|\mathbb{Q}| = \aleph_0$.

7.3 Cauchy Sequences

A **sequence of rational numbers** (shortened to just **sequence** for the next two subsections) is a function with domain $\mathbb{N}/0$ and co-domain $\mathbb{Q}$. In other words, we are given rational numbers x_n for all positive integers n. This is abbreviated as $\{x_n\}$. A sequence is a **Cauchy sequence** if for any rational $\varepsilon > 0$, there exists a positive integer N such that for all integers $m, n > N$, we have $|x_n - x_m| < \varepsilon$.

We can use Cauchy sequences to construct real numbers. Let S be the set of all Cauchy sequences. We define an equivalence relation $\sim$ on S as follows: $\{x_n\} \sim \{y_n\}$ iff for any rational $\varepsilon > 0$, there exists a positive integer N such that for all integers $m, n > N$, we have $|x_n - y_m| < \varepsilon$. This is indeed an equivalence relation: reflexivity by definition of Cauchy sequences, symmetry because $|x - y| = |y - x|$, and transitivity because of triangle inequality. Then the set of real numbers, $\mathbb{R}$, is just the set of all equivalence classes.

Given real numbers X, Y, we define addition and multiplication as follows:

$$X + Y = \{\{x_n + y_n\} \mid \{x_n\} \in X \text{ and } \{y_n\} \in Y\}$$

$$X \cdot Y = \{\{x_n y_n\} \mid \{x_n\} \in X \text{ and } \{y_n\} \in Y\}$$

Addition is indeed well defined since $\{x_n\} \sim \{x'_n\}$ and $\{y_n\} \sim \{y'_n\} \implies$ for any $\varepsilon > 0$, there exist positive integers N, M such that $|x_n - x'_m| < \dfrac{\varepsilon}{2}$ and $|y_n - y'_m| < \dfrac{\varepsilon}{2}$ for all integers $m, n > \max\{M, N\} \implies |x_n + y_n - x'_m - y'_m| \leq |x_n - x'_m| + |y_n - y'_m| < \varepsilon \implies \{x_n + y_n\} \sim \{x'_n + y'_n\}$. Before proving well-definiteness of multiplication, we need the following lemma:

Exercise 7.6. Prove that every Cauchy sequence is bounded, i.e., for every Cauchy sequence $\{x_n\}$, there exists a rational number Δ such that $|x_n| < \Delta$. This Δ is called the upper bound of $\{x_n\}$.

Now assume $\{x_n\} \sim \{x'_n\}$ and $\{y_n\} \sim \{y'_n\} \implies$ for any $\varepsilon > 0$, there exist positive integers N, M such that $|x_n - x'_m| < \dfrac{\varepsilon}{2\Delta}$ and $|y_n - y'_m| < \dfrac{\varepsilon}{2\Delta}$ for all integers $m, n > \max\{M, N\}$ (where Δ is the largest of the upper bounds for all four sequences). Therefore,

$$|x_n y_n - x'_m y'_m| = |y_n(x_n - x'_m) + x'_m(y_n - y'_m)| \leq |y_n| \cdot |x_n - x'_m| + |x'_m| \cdot |y_n - y'_m| < \varepsilon$$

$$\implies \{x_n y_n\} \sim \{x'_n y'_n\}$$

Again, we imbed $\mathbb{Q}$ in $\mathbb{R}$ as follows: a rational number a is just the equivalence class containing the constant sequence $x_n = a$ for all $n \in \mathbb{N}/0$. This also means the definition of 0 (additive identity) and 1 (multiplicative identity) is carried over to the reals. We can define additive inverse as

$$-X = \{\{-x_n\} \mid \{x_n\} \in X\}$$

and thus subtraction as $X - Y = X + (-Y)$. For reciprocal, we need the following result, which can be easily proved:

Exercise 7.7. Let $X \neq 0$ be a real number, and let $\{x_n\} \in X$. Then prove that there exists a positive integer N such that $x_n \neq 0$ for all integers $n > N$.

For any $\{x_n\} \in X$, let N satisfy the conditions of the lemma. Define a new sequence $\{y_n\}$ as $y_n = 0$ if $n \leq N$ and $y_n = \frac{1}{x_n}$ if $n > N$. Then $\dfrac{1}{X}$ is just the equivalence class containing $\{y_n\}$.

Finally, we define order. We say that $X > Y$ if for any $\{x_n\} \in X$ and $\{y_n\} \in Y$, there exists a positive rational number ε and a positive integer N such that $x_n - y_m > \varepsilon$ for all integers $m, n > N$. This can be easily checked to be well defined by definition of Cauchy sequences.

Now it is not hard to verify that all properties of addition, multiplication and order carry over from the rationals, i.e., $\mathbb{R}$ is also an ordered field.

7.4 Supremum Property

For a non-empty subset $S \subseteq \mathbb{R}$, we say that a real number M is an **upper bound** of S if $x \leq M$ for all $x \in S$. We say that M is the **least upper bound** (also called **supremum**, or just shorthanded as LUB) if $M \leq M'$ for any upper bound M' of S.

Theorem 34 (Supremum Property)**.** If a non-empty subset of $\mathbb{R}$ has an upper bound, then it has a least upper bound.

Before proving this, we need the following lemma:

Exercise 7.8. Prove that, for any real numbers $X > Y$, there exists a rational number q satisfying $X > q > Y$.

Now we return to the proof of the LUB property.

Proof. Let S be the non-empty set, and let M be an upper bound of S. WLOG we can assume that M is rational by using the above lemma on $M + 1$ and M if necessary. Choose any $s \in S$, and by applying the lemma on s and $s - 1$, choose a rational $N < s$. Let $q = M - N$, which is a positive rational. Construct a sequence $M_1, M_2, \ldots$ as follows: $M_1 = M$, and for every integer $n \geq 1$, $M_{n+1} = M_n - \dfrac{q}{2^n}$ if $M_n - \dfrac{q}{2^n}$ is an upper bound of S, and $M_{n+1} = M_n$ otherwise. Therefore $M_1, M_2, \ldots$ is a (non-strictly) decreasing sequence of upper bounds of S satisfying

$$|M_{n+1} - M_n| \leq \frac{q}{2^n}$$

$$\implies |M_n - M_m| \leq \frac{q}{2^n} \sum_{i=0}^{m-n-1} \frac{1}{2^i} = \frac{q}{2^n}\left(2 - \frac{1}{2^{m-n-1}}\right) < \frac{q}{2^{n-1}}$$

for all positive integers $m > n$. Now it is easy to prove that $\{M_n\}$ must be Cauchy. Further, by the definition of M_n, it is easy to see that for every positive integer n, there exists an $s_n \in S$ such that $M_n - s_n < \frac{q}{2^{n-1}}$. Let Δ be the real number associated with the Cauchy sequence $\{M_n\}$.

First we prove that Δ is an upper bound. Let $s \in S$, and let $\{u_n\} \in s$. Assume FTSOC that $\Delta < s$. Therefore, there exists a rational $\varepsilon > 0$ and a positive integer N such that $u_n - M_m > \varepsilon$ for all integers $m, n > N \implies u_n - M_{N+1} > \varepsilon$ for all $m > N \implies s > M_{N+1}$ since a Cauchy sequence of any rational is the constant sequence of the rational itself, contradiction since M_n are all upper bounds of S.

Finally we show that Δ is an LUB. Let x be any upper bound of S, with $\{y_n\} \in x$ and assume FTSOC that $x < \Delta$. Then, there exists a rational $\varepsilon > 0$ and a positive integer N_1 such that $M_n - y_m > \varepsilon$ for all integers $m, n > N_1$. Choose a positive integer N_2 such that $\frac{q}{2^{N_2-1}} < \frac{\varepsilon}{2}$. We can WLOG assume $N_2 > N_1$, since otherwise we can increase N_2 without changing the inequality. Then,

$$M_{N_2} - s_{N_2} < \frac{\varepsilon}{2} \implies M_{N_2} - \frac{\varepsilon}{2} < s_{N_2}$$

Let $\{v_n\} \in s_{N_2}$. Then by the above, and using the fact that a Cauchy sequence of any rational number is the constant sequence of the rational number itself, there must exist a rational $\varepsilon' > 0$ and a positive integer N_3 such that $v_n - M_{N_2} > \varepsilon' - \frac{\varepsilon}{2}$ for all integers $n > N_3$. Therefore, for all integers $m, n > \max\{N_1, N_3\}$,

$$v_n - y_m = (v_n - M_{N_2}) + (M_{N_2} - y_m) > \varepsilon' + \frac{\varepsilon}{2}$$

So $s_{N_2} > x$, contradiction since x is an upper bound of S. Therefore Δ is the required LUB of S. □

Therefore $\mathbb{R}$ is an ordered field with supremum property. Using this, the real numbers, in one way, can "plug" the "holes" in the rational numbers:

Exercise 7.9. Prove that there is a real number x satisfying $x^2 = 2$, and show that x is not rational.

This ultimately affects the cardinality of $\mathbb{R}$:

Theorem 35. $|\mathbb{R}| = 2^{\aleph_0}$.

Proof. The proof requires some foray into real analysis, and so will be skipped. □

Exercise 7.10. Prove that $|\{x \in \mathbb{R} \mid 0 < x < 1\}| = |\mathbb{R}|$

Proof. Let $(0, 1)$ denote the first set. We can construct an explicit bijection $f : (0, 1) \to \mathbb{R}$ as follows:

$$f(x) = \frac{1}{x} + \frac{1}{x-1} \quad \forall x \in (0, 1)$$

□

8 Logic

Logic is the analytic theory of reasoning which systematizes and makes it rigorous. It codifies stuff like statements and mathematical proofs. Mathematical theory can thus be an application of said logic. This was first developed by Russel and Whitehead in the book Principia Mathematica. However, the approach here will be quite informal.

8.1 Sentential Connectives

In mathematics, we use various connectors such as "and", "or", "implies", "if and only if", and also the negation "not". These are called **sentential connectives** and can be used to create **composite sentences** from **prime sentences** which do not contain such connectives.

"and", which is called **conjunction**, would be denoted by $\wedge$. "or" (always used in an inclusive sense), which is called **disjunction**, would be denoted by $\vee$. A sentence involving "implies" is called a **conditional sentence**; the first part is called "antecedent" and the second part is called "consequent". We will keep using the symbol $\implies$ for this. An "if and only if" sentence is called a **biconditional** and would be denoted by $\iff$. "not", which is called **negation**, would be denoted by $\neg$. If P and Q are sentences, then we can create the following new sentences:

$$\neg P \ , \ P \wedge Q \ , \ P \vee Q \ , \ P \implies Q \ , \ P \iff Q$$

We also use parentheses to denote denote order of precedence, e.g., $(\neg P) \wedge Q$ means the negation of P, and Q, while $\neg(P \wedge Q)$ means the negation of P and Q.

8.2 Truth Tables

All statements of logic can be given a unique **truth value**: "true", denoted by T, or "false", denoted by F. If we know the truth values of two statements P and Q, we can uniquely determine the truth values of the five basic sentences formed using each of the sentential connectives. The rules for finding these truth values can be summarised concisely using **truth tables**:

P	$\neg P$
T	F
F	T

Table 1: Negation

P	Q	$P \wedge Q$
T	T	T
T	F	F
F	T	F
F	F	F

Table 2: Conjunction

P	Q	$P \vee Q$
T	T	T
T	F	T
F	T	T
F	F	F

Table 3: Disjunction

P	Q	$P \implies Q$
T	T	T
T	F	F
F	T	T
F	F	T

Table 4: Conditional

P	Q	$P \iff Q$
T	T	T
T	F	F
F	T	F
F	F	T

Table 5: Biconditional

8.3 Statement Calculus

Statement calculus is the analysis of those logical relations among statements which depend solely on their composition from constituent statements using sentential connectives. The setting for this includes an initial collection of prime statements, and the following assumption: Each sentence under consideration is composed from prime statements using sentential connectives and, for a given assignment of truth values to the prime statements, receives a truth value in accordance to the truth tables mentioned in the previous part.

Now suppose we have a non-empty collection of initial distinct statements, and we extend this collection by adjoining precisely all those statements that can be formed using (repeatedly) the sentential connectives. Thus if A and B are in this collection, so are $\neg A$, $A \wedge B$, $A \vee B$, $A \implies B$ and $A \iff B$. We call members of this extended collection **formulas**. The initial statements are called **prime formulas**, and the others are called **composite formulas**.

A (composite) formula whose value is always T, regardless of the truth values assigned to the prime formulas, is called a **tautology**. We will write $\models A$ for "A is a tautology". Whether a statement is a tautology or not can be determined by studying its truth table.

Exercise 8.1. If A, B are statements, prove that $A \implies A$ and $A \vee (\neg A)$ are tautologies.

The following rule allows us to develop new tautologies from old:

Theorem 36. Let A be a formula, and let A^* be the formula obtained by replacing all occurences of some prime formula P in A by some other formula B. If $\models A$, then $\models A^*$.

Proof. Fairly easy, omitted. □

Given two composite statements A, B in the same collection of formulas, we say that A and B are equivalent, denoted by $A \equiv B$, if they have the same truth value regardless of the truth values assigned to the prime statements. Equivalence can be characterized by the following theorem:

Theorem 37. $\models A \iff B$ if and only if $A \equiv B$.

Proof. Follows fairly directly from the truth table of $\iff$. □

Theorem 38. If $\models A \implies B$ and $\models A$, then $\models B$.

Proof. Follows fairly directly from the truth table of $\implies$. □

Using these theorems, we can give the following list of tautologies and equivalences:

Exercise 8.2. For statements A, B, C, prove that the following are tautologies:

1. $(A \wedge (A \implies B)) \implies B$
2. $A \implies (B \implies (A \vee B))$
3. $(A \wedge B) \implies A$
4. $A \implies (A \vee B)$
5. $((A \implies B) \wedge (B \implies C)) \implies (A \implies C)$
6. $(A \implies (B \wedge \neg B)) \implies \neg A$

Exercise 8.3. For statements A, B, C, prove the following equivalences:

1. $A \equiv A$
2. $\neg(\neg A) \equiv A$
3. $A \iff B \equiv B \iff A$
4. $A \implies B \equiv (\neg B) \implies (\neg A)$
5. $A \wedge B \equiv B \wedge A$
6. $A \wedge (B \wedge C) \equiv (A \wedge B) \wedge C$
7. $A \wedge (B \vee C) \equiv (A \wedge B) \vee (A \wedge C)$
8. $A \wedge A \equiv A$
9. $\neg(A \wedge B) \equiv (\neg A) \vee (\neg B)$
10. $A \implies B \equiv (\neg A) \vee B$
11. $A \iff B \equiv (A \implies B) \wedge (B \implies A)$

The last few equivalences show that, in fact, only "and" and "not" are required to get relations that are equivalent to all other sentential connectives.

We give another method to construct tautologies. We consider formulas constructed from primes $P_1, P_2, \ldots$ and only $\wedge$, $\vee$ and $\neg$ (note that for any statement B, we can find an equivalent statement B' which only uses these connectives from the above remark). The **denial** A_d of such a formula A is formed by replacing each occurrence of $\wedge$ by $\vee$ and vice versa, and replacing each occurrence of P by $\neg P$ and vice versa for all primes P.

Theorem 39. $\neg A_d \equiv A$

Proof. Omitted due to length. □

8.4 Consequence

A statement B is the **consequence** of statements $A_1, A_2, \ldots$ if for every assignment of truth values of the involved prime formulas, B receives truth value T whenever every A_i receives the value T. This is denoted by

$$A_1, A_2, \ldots \models B$$

Theorem 40. $A \models B$ if and only if $\models A \implies B$.

Proof. Follows easily from definition. □

Theorem 41. $A_1, A_2, \ldots \models B$ if and only if $A_1 \wedge A_2 \wedge \ldots \models B$

Proof. Again follows easily from definition. □

Exercise 8.4. Prove that $A, B_1, B_2, \ldots \models A$ for all statements $A, B_1, B_2, \ldots$.

Exercise 8.5. Prove that, if $\models B$, then $A \models B$ for any statement A.

Theorem 42. If $A_1, A_2, \ldots \models B_i$ for $i = 1, 2, \ldots$, and $B_1, B_2, \ldots \models C$, then $A_1, A_2, \ldots \models C$.

Proof. Suppose each A_i has truth value T. Then each B_i has truth value T has well, and hence C has truth value T, and we are done. □

With this, we can formalise the concept of **proof**: It is just a finite string of formulas, starting with some formulas as **premise**, and satisfying the property that every next formula in the string is a consequence of all the previous ones. Generally, these premise formulas will be called **axioms**, so a proof will assert something of the form

$$A_1, A_2, \ldots \models B \implies C$$

where the A_i are the axioms. If we also have $A_1, A_2, \ldots \models B$, then we call C a **theorem**.

Call a collection of statements $\{A_1, A_2, \ldots\}$ **satisfiable** if there exists at least one assignment of truth values to the prime components such that each A_i is true. A **contradiction** is a statement that is always false, i.e., a statement whose negation is a tautology. For example, $A \wedge (\neg A)$ is a contradiction.

Theorem 43. A collection of statements is not satisfiable if a contradiction is a consequence of them.

Proof. Follows fairly directly. □

Theorem 44. $A_1, A_2, \ldots \models B$ if a contradiction can be derived from $\neg B, A_1, A_2, \ldots$

Proof. Assume some contradiction C can be derived from $\neg B, A_1, A_2, \ldots$. Then, $A_1, A_2, \ldots \models (\neg B) \implies C$. Therefore if $A_1, A_2, \ldots$ are true, then $(\neg B) \implies C$ is true, but C is false. Therefore $(\neg B)$ is false, and hence B is true. □

The above theorem forms the basis for the proof technique known as **proof by contradiction**.

8.5 Terms, Predicates and Quantifiers

While statement calculus makes general predictions regardless of which statements are used, it isn't to break down a statement into finer constituents and make deductions based on them.

For example, consider the statement "$\sin^2(x) + \cos^2(x) = 1$ for all real numbers x". Here we have used x as a placeholder for a bigger set; we call x an **individual variable**. On the other hand, the number 1 and the function sin represent specific, well-defined objects; we call them **individual constants**. Collectively, variables and constants are classified as **terms**.

Consider the sentence "x is an integer". If we substitute x for any specific object, we will get a statement. Such a sentence is called a **predicate** and we will normally denote it by $P(x)$. Predicates with multiple variables, e.g. "$x + y = z$", will be denoted by $P(x, y, z)$, and so on. A predicate with n variables will be called an **n-place predicate**, and we also define statements to be just 0-place predicates.

Another way to make statements from predicates is to attach a **quantifier** at the front, e.g., "For all x, x is an integer". The phrase "for all x" is called a **universal quantifier** and is denoted by $\forall x$. Another quantifier is the phrase "there exists an x"; it is called an **existential quantifier** and denoted by $\exists x$. For example, if we let $P(x)$ and $Q(x)$ stand for the predicates "x is an integer" and "x is a real number" respectively, then the statement "Every integer is a real number" can be written as

$$\forall x(P(x) \implies Q(x))$$

while "Some real numbers are not integers" can be written as

$$\exists x((\neg P(x)) \wedge Q(x))$$

Consider the following expression:

$$\exists x(P(x) \wedge \forall y(P(y) \implies Q(x,y)))$$

This says that any object y which satisfies $P(y)$ must also satisfy $Q(x,y)$, for some fixed object x. If Q is some equivalence relation, then this talks about all objects satisfying a certain property to be in the same equivalence class. In particular, if $Q(x,y)$ stands for $x = y$, then this talks about a unique x which satisfies $P(x)$. We will abbreviate this as $\exists! x(P(x))$, the **uniqueness quantifier**.

We will only allow quantification of individual variables; this will make our theory of inference a **first order theory**. In most cases, there will be an equality predicate $x = y$; in that case the theory would be a **first order theory with equality**.

8.6 Predicate Calculus

Predicate calculus is the theory of inference based on statement calculus, predicates and quantifiers.

Firstly, we assume that we are given a non-empty list of n-place predicates for each of $n = 0, 1, \dots$. These will be called **predicate letters**.

We redefine formulas in predicate calculus. A **prime formula** is an expression that we get by substituting any variables (not necessarily distinct, and not necessarily from the predicate letter) for those variables that appear in the predicate letter. For example, the prime formulas from a predicate letter $P(x,y,z)$ are $P(a,b,c)$, $P(u,u,v)$, $P(z,z,z)$ and so on. We extend this to a collection of all **formulas** by using connectives and quantifiers. The ones which are not prime are called **composite formulas**.

A **scope** of a quantifier is the part of the expression for which the quantifier is valid. An occurrence of a variable is **bound** if it is either in the scope of a quantifier employing that variable, or is an explicit occurrence in the quantifier. An occurrence of a variable is **free** if it is not bound. For example, in $\forall x(P(x,y))$, both occurrences of x are bound, while the occurrence of y is free.

A variable is called **free in a formula** or **bound in a formula** if at least one occurrence of that variable is free or bound respectively.

8.7 Truth Values in Predicate Calculus

Assignment of truth values in predicate calculus is a bit more complicated. First, we must assume that there is a non-empty set D, called the domain, such that each individual variable ranges over D. Further, it is assumed that for every n-place predicate letter, there is an associated **logical function**, that is, an assignment of T or F to every n-tuple of elements from D. For 0-place predicates, this function is assumed to be constant, one of T or F. Then, truth values can be assigned thus to prime formulas, relative to assignment of variables to elements of D: For a formula $P(y_1, y_2, \dots)$ (note that the y_i may not be all distinct), if the variable y_i is assigned to some $d_i \in D$, then the truth value of $P(y_1, y_2, \dots)$ is the truth value associated to the tuple $(d_1, d_2, \dots)$ in accordance with the predicate letter P.

For composite formulas, we first assume that every distinct free variable in the formula is assigned a value in D. We use the same truth tables as that of statement calculus for sentential connectives. Also, we use the following rules for quantifiers:

- The value of $\forall x(A)$ is T if the value of A is T for every assignment of x to some element in D, and the value is F otherwise.
- The value of $\exists x(A)$ is T if the value of A is T for at least one assignment of x to some element in D, and the value is F otherwise.

Again, these rules are fairly intuitive.

A formula is called **valid in a given domain** if it has truth value T for each assignment of its free variables and for each logical function associated with all the predicate letters. A formula is called **valid** if it is valid in every domain. Again, we use $\models A$ to denote "A is valid". We also define two formulas A and B to be **equivalent**, and write $A \equiv B$, if $\models A \iff B$.

Exercise 8.6. Prove that $\exists x(A) \equiv \neg(\forall x(\neg A))$.

So again, as in the case of connectives, having two quantifiers is redundant.

8.8 Substitution

For a formula A, we say we **substitute** a variable y for a variable x if we replace all free occurrences of x by y. Here it is convenient to use $A(x)$ for the original formula, and $A(y)$ for the substituted one. Note that this is different from a predicate notation, since x can have bound occurrences in A, and there could be other free variables. However, we cannot always do this substitution in the intended way, as there may be occurrences of y in A as well. We say that $A(x)$ is **free for** y if no free occurrence of x in $A(x)$ is in the scope of a quantifier $\forall y$ or $\exists y$.

Exercise 8.7. Prove that $\models A(y) \implies (\exists x(A(x)))$ whenever $A(x)$ is free for y.

Exercise 8.8. Prove that $\models (\forall x(A(x))) \implies A(y)$ whenever $A(x)$ is free for y.

Theorem 45. Let B be a formula without any free occurrences of x, and let $A(x)$ be any formula. Then,

1. If $\models B \implies A(x)$ then $\models B \implies (\forall x(A(x)))$.
2. If $\models A(x) \implies B$ then $\models (\exists x(A(x))) \implies B$

Proof. Follows fairly directly from definitions. □

Exercise 8.9. Prove that $\models A(x)$ if and only if $\models \forall x(A(x))$.

The above exercise basically shows a general proof technique - you take an arbitrary variable x, prove the result for that x, then argue that the result must be true for all x.

So far we have substituted only variables for other variables, but now we will talk about substituting formulas for other formulas. Suppose we have some formula B, not containing any of $x_1, x_2, \ldots x_k$, and containing a prime formula resulting from the predicate letter $P(x_1, x_2, \ldots, x_k)$, and we want to substitute it with a formula $A(x_1, x_2, \ldots, x_k)$ (here x_i may not be free, and there may be other free variables that occur in A). To do this we will substitute each prime of the form $P(y_1, y_2, \ldots, y_k)$ in B with $A(y_1, y_2, \ldots, y_k)$. This substitution is **admissible** if none of the variables in B occur bound in $A(x_1, x_2, \ldots x_k)$, and none of the free variables in $A(x_1, x_2, \ldots x_k)$ occur bound in B.

Theorem 46. Let B be a formula, not containing any of $x_1, x_2, \ldots x_k$, which contains a prime formula resulting from the predicate letter $P(x_1, x_2, \ldots, x_k)$. Let B^* be the formula resulting from B by an admissible substitution of $P(x_1, x_2, \ldots x_k)$ by some $A(x_1, x_2, \ldots x_k)$. If $\models B$, then $\models B^*$.

Proof. The statement and the proof are fairly intuitive, however the proof is omitted due to length. □

Using these results, we can prove many equivalences.

Exercise 8.10. Prove the following equivalences, where $A(x), B(x), A(x, y)$ are formulas and x, y are distinct variables:

1. $\exists x \exists y A(x, y) \equiv \exists y \exists x A(x, y)$

2. $\forall x \forall y A(x, y) \equiv \forall y \forall x A(x, y)$

3. $\exists x(A(x) \vee B(x)) \equiv (\exists x(A(x))) \vee (\exists x(B(x)))$

4. $\forall x(A(x) \wedge B(x)) \equiv (\forall x(A(x))) \wedge (\forall x(B(x)))$

Further, proofs and theorems in predicate calculus will work in the same way as statement calculus. We say $A_1, A_2, \ldots \models B$ if $\models A_1 \wedge A_2 \wedge \cdots \implies B$, and a proof will be a string of formulas such that every formula is a consequence of the previous ones.

9 Axiomatic Theory

Axiomatic theory is a systematic study of mathematical systems using various axioms - statements which are prime formulas - and deriving new results using logic. This approach was first applied by Greek mathematician Euclid to the study of plane geometry in his book *Elements*. Each mathematical theory has its base in axioms (mostly axioms of set theory). The formal axiomatization of geometry was first done by David Hilbert, the axiomatization of number theory and real analysis was done by Giuseppe Peano, and the axiomatization of set theory was done by Ernst Zermelo and Abraham Fraenkel. Subsequent axiomatic theories in which general set theory is assumed are called **informal theories**. For example, group theory bases its study on a set with an operation that satisfies certain axioms.

9.1 Set Theory in Predicate Language

Using predicate calculus, we can rewrite axioms of ZFC in a more formalized way. We will talk about two prime 2-place predicates: $\in(x, y)$, also written as $x \in y$, and $=(x, y)$, also written as $x = y$.

1. **Axiom of Extension**: $\forall x \forall y((x = y) \iff (\forall z((z \in x) \iff (z \in y))))$
2. **Axiom Schema of Specification**: $\forall x \forall y_1 \forall y_2 \ldots \exists z \forall w((w \in z) \iff ((w \in x) \wedge P))$
   Here P is a formula such that all of its free variables are among $w, x, y_1, y_2, \ldots$.
3. **Axiom of the Empty Set**: $\exists x(\forall y(\neg(y \in x)))$
   Again we will refer to the empty set as $\varnothing$, and add it to the language of ZFC instead of repeating this axiom everywhere.
4. **Axiom of Pairing**: $\forall x \forall y \exists z((x \in z) \wedge (y \in z))$
5. **Axiom of Union**: $\forall y \exists x \forall z \forall w(((w \in z) \wedge (z \in y)) \implies (w \in x))$
   While this does not directly give us the union, it gives us a set that contains the union, and then we can use specification to get the union.
6. **Axiom of Power Set**: $\forall x \exists y \forall z((z \in y) \iff \forall w((w \in z) \implies (w \in x)))$
7. **Axiom of Infinity**: $\exists x((\varnothing \in x) \wedge \forall y((y \in x) \implies (S(y) \in x)))$
   Here we have used $S(z)$ to abbreviate $z \cup \{z\}$.
8. **Axiom Schema of Replacement**:

   $$\forall x \forall y_1 \forall y_2 \ldots (\forall z((z \in x) \implies \exists! w(P)) \implies \exists t \forall u((u \in x) \implies \exists v((v \in t) \wedge P)))$$

   Here P is a formula such that all of its free variables are among $x, z, u, v, y_1, y_2, \ldots$.
9. **Axiom of Regularity**: $\forall x((\neg(x = \varnothing)) \implies \exists y((y \in x) \wedge \forall z((z \in A) \implies (\neg(z \in B)))))$
10. **Axiom of Choice**: $\forall x \exists y \forall z(((z \in x) \wedge (\neg(z = \varnothing))) \implies \exists! w((w \in z) \wedge (w \in y)))$

10 References

1. *Set Theory and Logic*, Robert R. Stoll