ZFC Set Theory

Muchang Bahng

Spring 2025

Contents

A few notes I have taken from reading Hrbacek and Jech's *Introduction to Set Theory*, plus some other supplementary resources.

Consider the set of all students at Duke University, or the set of all stars in our universe, or the set of all pink elephants. We can construct all kinds of sets, but they are not real objects of the world and are only created by our minds. When we say the number 2, can we *see* "2" with our own eyes, or *hold* "2" on the palm of our hand? No. They are merely an abstraction of our minds, and so care must be taken to treat them as abstractions. Often, we may rely on our intuition in the real world to talk about sets, but as we will find our shortly, this may not work. Therefore, we will go with Cantor's definition of a set, created in his naive attempt to construct set theory at the end of the 19th century.

Unlike axiomatic set theories, which are defined using formal logic, naive set theory was defined informally at the end of the 19th century by Cantor, in natural language (like English).

Definition 0.1 (Cantor)

A *set* is a collection into a whole of definite (i.e. well-defined), distinct objects of our intuition or our thought. These objects are called *elements.*

Okay, sounds simple enough, but there are a lot of holes in this definition. It all stems from the lack of specification of what "well-defined" means. According to Cantor, we can construct a set by specifying a defined property P and talking about the set of all elements x satisfying $P(x)$, denoted in *set-builder notation* as

$$\{x \mid P(x)\} \tag{1}$$

But the property P cannot by itself guarantee the consistency and unambiguity of what exactly constitutes and what does not constitute a set, and therefore this is not a formal definition. Since such a formalism is absent, we cannot have any sets to work with in the first place.

Perhaps this is all a bit too philosophical. If I talk about the set $\{1, 2, 3\}$ or the set of all prime numbers under 100, this is pretty concrete and well-defined, and there is really no problem at first glance. This is to be expected, since Cantor relied on natural language (like English) to construct his *naive set theory.* It describes the aspects of mathematical sets using words (e.g. *satisfying, such as, ...*) and sufficed for the everyday use of set theory in mathematics back then. Such issues did not arise until Russell formulated such a "well-defined" set that lead to a paradox.

Definition 0.2 (Russell Set)

Let $R = \{x \mid x \notin x\}$, i.e. the set of all sets that do not contain themselves as elements. This is called the **Russell set**.

While the property may sound a bit confusing, this is still well-defined since for any set, we can check this property to be either true or false.

Example 0.1 (Russell Property is Verifiable)

Let's consider some examples of sets.

1. $\{1\} \in R$ since it does not contain $\{1\}$ as an element.
2. The set of *everything*[a] is not in R since it contains itself (as an element of everything).

[a] everything is indeed well-defined... or is it?

However, this leads to a paradox that violates the law of the excluded middle.

Theorem 0.1 (Russell's Paradox)

The Russell set exists and does not exist.

Proof. We will determine if R is an element of itself.

1. If $R \in R$, then by it does contain itself, so it does not satisfy the property and $R \notin R$.
2. If $R \notin R$, then it satisfies the property, so $R \in R$.

Therefore, it is both the case that $x \in R$ and $x \notin R$, which contradicts the membership definition. Therefore, R is both a set from set-builder construction and not a set due to the membership definition.

There are other paradoxes out there, which are similar to Russell's paradox.

Corollary 0.2 (Universal Set)

The set of all sets E, called the **universal sets**, both exists and does not exist. Furthermore, the set of all singleton sets does not exist.

Proof. We can define $U' = \{x \mid \{\} = \{\}\}$, which defines a set. Then the property P that $\{\} = \{\}$ is always true, and U' would contain everything, and by the definition of equality $U = U'$. Now since the Russell set R is both a set and not a set from Russell's paradox, we have $R \in U$ and $R \notin U$, which means that U cannot exist. Therefore U does not exist.
Alternatively, we can prove it as such. E exists according to naive set theory due to set builder notation. Then by definition $E \in E$. Now we can construct Russell's paradox from this by defining

$$\{x \in E \mid x \notin x\} \tag{2}$$

To prove the second claim, note that if such a set S existed, then $\cup S = E$, which leads to the same paradox.

So the sufficiency a well-defined property to be able to construct a set is *too powerful* in that we can construct *any* set we want. This leads us to construct the Russell set, which opens up a lot of paradoxes. This ability to take a set of all objects satisfying such an arbitrary property is known as the **schema of general comphrension**, and is the major flaw of naive set theory. Therefore, we would like to restrict the notion of well-defined in a way, which leads to axiomatic set theories. Attempting to achieve this will be done in axiomatic set theory, like ZFC.

Exercise 0.1 (Math 631 Fall 2025, Final Exam Exercise 8)

Let $X_1, ..., X_n, ...$ be independent identically distributed random variables with expectation zero and variance 1. Define $S_n = X_1 + \cdots + X_n$. Use the central limit theorem and Kolmogorov's $0-1$ law to conclude that $\limsup \frac{S_n}{\sqrt{n}} = \infty$ almost surely.

Solution. By central limit theorem,

$$\lim_{n\to\infty} P(S_n/\sqrt{n} \geq B) = \frac{1}{\sqrt{2\pi}} \int_B^\infty e^{-x^2/2} dx \equiv \Phi(B). \tag{3}$$

Let $E_{n,B} = \{\omega | S_n(\omega)/\sqrt{n} \geq B\}$. Then $m(\cup_{j=n}^\infty E_{n,B}) \geq \Phi(B)$ for every n, and hence $m(\limsup E_{n,B}) \geq \Phi(B)$ by monotonicity and continuity of measure. It follows that for every B,

$$P(\limsup S_n/\sqrt{n} \geq B) \geq \Phi(B) > 0. \tag{4}$$

But the event $E_B = \{\omega | \limsup S_n/\sqrt{n} \geq B\}$ does not depend on any finite number of X_n and so belongs to the tail σ-algebra T. Hence by Kolmogorov $0-1$ law, $P(E_B) = 1$ for every B (since it has positive probability and so cannot be probability zero). Define $E = \cap_{B \in \mathbb{N}} E_B$. By continuity of measure, $P(E) = 1$. But E is exactly equal to $\{\omega | \limsup S_n / \sqrt{n} = \infty\}$.

1 Sets

So with these paradoxes in mind, we would like to construct an axiomatic formulation of sets. My take is to think that sets "exist" out there somewhere in the universe, and our job is to find them. Cantor with his naive set theory believed that for every meaningful property of things there is a set whose members are exactly all the things with that property. Russell shows this this cannot be the case. Nevertheless, *some* sets exist, and we have intuitive experience thinking about finite sets. Therefore, the axioms of set theory are a limited list of *assumptions* that we hope are true about that actually existing universe of sets. As long as they are true, then whatever we conclude from them by valid reasoning steps must also be true.[1] Hence we have the following definition, which first requires the familiar property of acting like a collection of something, and then obeys the axioms we set.

Definition 1.1 (Set)

A **set** X is anything

1. that has the innate property of containing elements, and
2. obeys the axioms of ZF(C).

Let's first talk about the language, where they are defined formally using the axioms in the next subsection. From first-order logic, note that we have the following symbols in our alphabet $\mathcal{L}_{\text{ZFC}}$.

1. The logical connectives $\neg$, $\vee$, $\wedge$.
2. The quantifier symbols $\exists, \forall$
3. Brackets ().

To represent sets, we also need symbols, and the membership property requires us to define a symbol for that too.

1. A countably infinite amount of variables used for representing sets.
2. The set membership symbol $\in$. In fact, when we say $x \in A$, this is a proposition formed from the predicate $P(x)$.

This is what we have to work with so far. We will construct the rest of the symbols $(=, \subset, \supset, \cup, \cap)$ from the axioms.

Now we state the axioms, which is the foundation of ZF set theory. A question one might ask is: how do we even know for sure that these axioms aren't contradictory? The answer is that we don't, and that is why we take them as axioms rather than provable theorems. Fortunately, from the formulation in the early 20th century up until now, no contradictions have been found, and if there is one, then it would be very bad news for us.

As obvious as the axioms may seem, none of them can be rigorously proven since we need to start from a set of assumptions and some principles of logic to use some deductive reasoning.

Axiom 1.1 (Axiom of Existence)

There exists a set which has no elements.[a]

[a] Technically, this does not need to be stated as an axiom, but it is often done for convenience. Indeed, by the axiom of infinity, there exists some set S. By the axiom of restricted comprehension, we can define $\{x \in S \mid x \neq x\}$, which is empty since the predicate always evaluates to false.

[1] This idea is called naive Platonism.

Definition 1.2 (Empty Set)

The **empty set** is denoted $\emptyset$.[a]

[a] Note that this is not determined to be unique... yet!

Why do we need to axiomize what seems to be such an obvious thing? The rest of the axioms that follow talk about what we can or cannot do with sets, but this whole theory would be useless if we didn't know if there exists *any* set that obeys the following axioms! Therefore, we would like to assert the existence of at least one set, namely the empty set. This asserts that our universe of discourse is not void, and so it gives us a starting set to work with, which we can build on to create more sets.

1.1 Axiom of Extensionality

The empty set is simply described as the set containing nothing, but it can be constructed in various ways. I can describe it as the set of humans that have negative age, or the set of married bachelors. All examples of this kind describe one and the same set $\emptyset$, but we cannot yet prove this. This is why we need our second axiom, which states that a set is uniquely characterized simply by its elements.

Axiom 1.2 (Axiom of Extensionality)

Two sets are equal (are the same set) if they have the same elements.

$$\forall A \forall B \big[\forall x (x \in A \iff x \in B) \iff A = B\big] \tag{5}$$

Definition 1.3 (Equality)

This axiom allows us to define the equality operator =, which we now add to our alphabet.

Theorem 1.1 (Uniqueness of Empty Set)

The empty set is unique.

Proof. Assume that A and B are sets with no elements. Then it is vacuously true that every element of A is an element of B, and vice versa. Therefore by the axiom of extensionality $A = B$.

This next theorem now shows that every set is uniquely characterized by its distinct elements, which aligns with our established notion that sets don't contain repeated elements.

Theorem 1.2 (Sets Don't Contain Repeated Elements)

Sets are unique up to distinct elements. That is, given 2 sets A, B,

$$\forall x (x \in A \iff x \in B) \implies A = B \tag{6}$$

As an example, we have the following:

$$\{1, 1, 2\} = \{1, 2\} = \{1, 1, 2, 2\} \tag{7}$$

Though note that we don't even know if any of the sets above even exist with our axioms so far!

1.2 Axiom of Restricted Comprehension

The axiom assists us in regulating which sets are viable and which are not, preventing Russell's paradox.

Axiom 1.3 (Axiom Schema of Restricted Comprehension)

Given P a formula with $P(x)$ specifying a property of x, for any set A, there exists a set B such that $x \in B$ if and only if $x \in A$ and $P(x)$. That is, if A exists, then the set, written in set-builder notation,

$$B = \{x \in A \mid P(x)\} \tag{8}$$

exists.[a]

[a] Note that this axiom does not allow the construction of entities of the more general form $\{x \mid P(x)\}$. This restriction is obviously needed to avoid Russell's paradox, hence the name *restricted* comprehension.

Lemma 1.3 (Uniqueness of Restricted Comprehension)

The set $B = \{x \in A \mid P(x)\}$ is unique, and therefore it makes sense to treat it as a unique object.

The axiom of specification allows us to denote subsets.

Definition 1.4 (Subset, Superset)

Notationally, if A is a subset of B, then we write $A \subset B$. Similarly, we say A is a **superset** of B, written $A \supset B$, if $B \subset A$.

We can extend this to the restriction of more sets.

Theorem 1.4 (Existence of Binary Intersection)

That is, if A, B are sets, then there is a set C such that $x \in C$ if and only if $x \in A$ and $x \in B$. This allows us to define intersection as

$$A \cap B \coloneqq \{x \in A \mid x \in B\} \tag{9}$$

Proof.

We can extend this proof to define the intersection of a set of sets.

Definition 1.5 (Intersection)

Given a nonempty[a] set of sets S, the **intersection** $\cap S$ is the unique set satisfying $x \in \cap S$ if and only if $x \in A$ for all $A \in S$. In set builder notation, we can write

$$\bigcap S \coloneqq \{x \in A \mid \forall B (B \in S \implies x \in B)\} \tag{10}$$

We can also expand our notation by defining the following.

1. Just another way of writing is
$$\cup S = \bigcap_{A \in S} A \tag{11}$$
2. By the axiom of extensionality, we can define
$$A_1 \cap \ldots \cap A_n = \cap \{A_1, \ldots, A_n\} \tag{12}$$

[a] $\bigcap \emptyset$ is not defined since every x belongs to all $A \in \emptyset$ vacuously, and so such a set would be the universal set, which does not exist.

Definition 1.6 (Disjoint Sets)

Two sets A and B are **disjoint** if $A \cap B = \emptyset$.

Unfortunately, the union cannot be expressed in this specification schema, and we need a separate axiom for this.

Definition 1.7 (Set Minus)

We can however define set minus. Given sets A, B

$$A \setminus B \coloneqq \{x \in A \mid x \notin B\} \tag{13}$$

Definition 1.8 (Set Complement)

Given B and a subset $A \subset B$, the **complement** of A with respect to B is

$$A^c \coloneqq \{x \in B \mid x \notin A\} = B \setminus A \tag{14}$$

Definition 1.9 (Symmetric Difference)

Given sets A, B, the **symmetric difference** between the two sets is defined

$$A \triangle B \coloneqq (A \setminus B) \cup (B \setminus A) \tag{15}$$

1.3 Axiom of Pairing

Okay this is all great, but the only set that we have claimed the existence of is $\emptyset$, and the schema of restricted comprehension is useless in constructing any new sets since

$$\{x \in \emptyset \mid P(x)\} = \emptyset \tag{16}$$

for any property P. Starting from now, we provide more helpful axioms to construct new sets.

Axiom 1.4 (Axiom of Pairing)

If A, B are sets, then there exists a set which contains A and B as elements.[a]

$$\forall A \forall B \exists C ((A \in C) \land (B \in C)) \tag{17}$$

This allows us to construct sets from old ones.

[a] For example, if $A = \{1, 2\}$ and $B = \{2, 3\}$,then $\{\{1, 2\}, \{2, 3\}\}$ exists.

Note that clearly, this set is not necessarily unique, since there can be other elements in C in addition to A and B.

Theorem 1.5 (Nested Sets)

By the axiom of pairing, if we have a set X, then $\{X\}$ is also a set, since we can set $A = B = X$ which asserts the existence of $\{X, X\} = \{X\}$.

Example 1.1 (More Sets)

Now we have unlocked our first sets that are not the empty set! Consider the following.

1. If $A = B = \emptyset$, then by the axiom of pairing we can get $C = \{\emptyset, \emptyset\}$, which is equal to $\{\emptyset\}$ by the axiom of extensionality.
2. Now we let $A = \emptyset, B = \{\emptyset\}$, and so $C = \{A, B\} = \{\emptyset, \{\emptyset\}\}$.

1.4 Axiom of Union

Axiom 1.5 (Axiom of Union)

For any set of sets S, there exists a set U such that $x \in U$ if and only if $x \in A$ for some $A \in S$.

$$\forall \mathcal{F} \exists U \forall X \forall x \big[(x \in X \land X \in \mathcal{F}) \implies x \in U\big] \tag{18}$$

Lemma 1.6 (Union is Unique)

The union U is unique.

Proof. Again use the axiom of extensionality.

Definition 1.10 (Union)

The set U, is called the **union** of sets $A \in S$, denoted $\cup S$. We can also expand our notation by defining the following.

1. Just another way of writing is
$$\cup S = \bigcup_{A \in S} A \tag{19}$$
2. By the axiom of extensionality, we can define
$$A_1 \cup \ldots \cup A_n = \cup\{A_1, \ldots, A_n\} \tag{20}$$

Sometimes, it is formulated alternatively as follows: For any set of sets S, there is a set U containing every element that is a member of S. This does not necessarily mean that U is unique, and so $\cup S$ is not defined yet. However, we can define it by using the axiom of restricted comprehension and defining

$$\cup S \coloneqq \{x \in A \mid \exists X(x \in X \land X \in S)\} \tag{21}$$

Example 1.2

$$\{\{\emptyset\}\} \cup \{\emptyset, \{\emptyset\}\} = \{\emptyset, \{\emptyset\}\} \tag{22}$$

1.5 Rules of Set Operations

Let's first talk about rules following the union, intersection, and set minus operators.

Theorem 1.7 (Commutativity)

Union and intersection are commutative.

$$A \cup B = B \cup A \tag{23}$$
$$A \cap B = B \cap A \tag{24}$$

Theorem 1.8 (Associativity)

Union and intersection are associative.

$$(A \cup B) \cup C = A \cup (B \cup C) \tag{25}$$
$$(A \cap B) \cap C = A \cap (B \cap C) \tag{26}$$

Theorem 1.9 (Distributivity)

Given sets A, B, C,

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \tag{27}$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \tag{28}$$

Proof. Listed.

1. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.
    (a) $A \cap (B \cup C) \subset (A \cap B) \cup (A \cap C)$. Assume $x \in A \cap (B \cup C)$. Then $x \in A$ and $x \in B \cup C$. If $x \in B$, then $x \in A \cap B$. If $x \in C$, then $x \in A \cap C$. Therefore, since $x \in B \cup C$, it must be the case that $x \in A \cap B$ or $x \in A \cap C$, which by definition implies $x \in (A \cap B) \cup (A \cap C)$.
    (b) $A \cap (B \cup C) \supset (A \cap B) \cup (A \cap C)$. Assume that $x \in (A \cap B) \cup (A \cap C)$. Then WLOG let $x \in A \cap B$. Then $x \in A$ and $x \in B \subset (B \cup C)$, so by definition $x \in A \cap (B \cup C)$.
2. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.
    (a) $A \cup (B \cap C) \subset (A \cup B) \cap (A \cup C)$. Assume $x \in A \cup (B \cap C)$. Then $x \in A$ or $x \in B \cap C$. If $x \in A$, then since $A \subset (A \cup B)$ and $A \subset (A \cup C)$, we have $x \in (A \cup B)$ and $x \in (A \cup C)$, which by definition means $x \in (A \cup B) \cap (A \cup C)$. If $x \notin A$, then $x \in B \cap C \implies x \in B \subset (A \cup B)$ and $x \in C \subset (A \cup C)$, and so $x \in (A \cup B) \cap (A \cup C)$.
    (b) $A \cup (B \cap C) \supset (A \cup B) \cap (A \cup C)$. Assume $x \in (A \cup B) \cap (A \cup C)$. Then $x \in A \cup B$. If $x \in A$, then since $A \subset A \cup (B \cap C)$, $x \in A \cup (B \cap C)$. If $x \notin A$, then $x \in B$. Since $x \in A \cup C$, $x \in C$ also. Therefore by definition $x \in (B \cap C) \subset A \cup (B \cap C) \implies x \in A \cup (B \cap C)$.

Theorem 1.10 (DeMorgan's Laws)

If X is a set and $A, B \subset X$, then

$$X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B) \tag{29}$$
$$X \setminus (A \cap B) = (X \setminus A) \cup (X \setminus B) \tag{30}$$

Proof. Listed.

1. $X \setminus (A \cup B) = (A \setminus A) \cap (X \setminus B)$.
    (a) $X \setminus (A \cup B) \subset (A \setminus A) \cap (X \setminus B)$. Assume $x \in X \setminus (A \cup B) \iff x \in X$ and $x \notin (A \cup B)$. Since $x \notin (A \cup B$, $x \notin A$ and $x \notin B$. However, $x \in X$, so $x \notin A \implies x \in X \setminus A$. Same goes for B, and so $x \in (X \setminus A) \cap (X \setminus B)$.

   (b) $X \setminus (A \cup B) \supset (A \setminus A) \cap (X \setminus B)$. Assume $x \in (X \setminus A) \cap (X \setminus B)$. Then $x \in X \setminus A \iff X \in X$ and $x \notin A$, and $x \in X \setminus B \iff x \in X$ and $x \notin B$. Since $x \notin A$ and $x \notin B$, $x \notin A \cup B$. Combined with the fact that $x \in X$, we have $x \in X \setminus (A \cup B)$.

2. $X \setminus (A \cap B) = (A \setminus A) \cup (X \setminus B)$.
   (a) $X \setminus (A \cap B) \subset (A \setminus A) \cup (X \setminus B)$. Let $x \in X \setminus (A \cap B)$. Then $x \in X$ and $x \notin A \cap B$. Since $x \notin A \cap B$, it must be the case that at least $x \notin A$ or $x \notin B$. WLOG let $x \notin A$. Then $x \in X$ and $x \notin A \implies x \in (X \setminus A) \subset (X \setminus) \cup (X \setminus B) \implies x \in (X \setminus A) \cup (X \setminus B)$.
   (b) $X \setminus (A \cap B) \supset (A \setminus A) \cup (X \setminus B)$. WLOG let $x \in (X \setminus A)$. Then $x \in X$ and $x \notin A$, and $x \notin A \implies x \notin (A \cap B) \subset A$ (contrapositive is trivial). Therefore, $x \in X$ and $x \notin (A \cap B) \iff x \in X \setminus (A \cap B)$.

Theorem 1.11 (Properties of Set Difference)

We have the following.
1. $A \cap (B \setminus C) = (A \cap B) \setminus C$
2. $A \setminus B = \emptyset$ if and only if $A \subset B$.

Proof.

Theorem 1.12 (Properties of Symmetric Difference)

We have the following.
1. $A \triangle A = \emptyset$
2. $A \triangle B = B \triangle A$
3. $(A \triangle B) \triangle C = A \triangle (B \triangle C)$
4. $A \triangle B = (A \cup B) \setminus (A \cap B)$

Proof.

1.6 Axiom of Regularity

Axiom 1.6 (Axiom of Regularity)

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

$$\forall A \big[A \neq \emptyset \implies \exists x (x \in A \land A \cap x = \emptyset) \big] \tag{31}$$

This, along with the axioms of pairing and union, implies that no set is an element of itself and that every set has an ordinal rank.

2 Correspondences

2.1 Axiom of Power Set

Now that we have constructed the Von Neumann ordinals, we are allowed to do *indexing*.

Axiom 2.1 (Axiom of Power Set)

The axiom of power set states that for any set A, there is a set B that contains every subset of A.

$$\forall A \exists B \forall S (S \subset A \implies S \in B) \tag{32}$$

The axiom of schema of restricted comprehension is then used to define the power set as the unique subset of such B containing the subset of A exactly.

$$\mathcal{P}(A) = 2^A = \{X \in B \mid X \subset A\} \tag{33}$$

The existence of the power set allows us to define subfamilies of subsets of a given set, namely things such as the topology or σ-algebra. But a perhaps more important consequence is the ability to construct Cartesian products of sets. So far, a set of say $\{a, b\}$ is an *unordered* pair since by the axiom of extensionality, $\{a, b\} = \{b, a\}$. We would like to consider some way to order the elements into (a, b), and this ordered pair (a, b) must be a set. There are many ways to do this, but the most established is to take (a, b) as an element of the power set of the power set of the union of sets.

Definition 2.1 (Cartesian Product)

The power set axiom allows for the definition of the **Cartesian product** of two sets A and B. Note that if $a \in A, b \in B$, then by the axiom of union $a, b \in A \cup B$ and by the axiom of power set $\{a\}, \{a, b\} \in \mathcal{P}(A \cup B)$. Therefore, using the axiom of power set again we can define

$$(a, b) \coloneqq \{\{a\}, \{a, b\}\} \in \mathcal{P}(\mathcal{P}(A \cup B)) \tag{34}$$

and the Cartesian product is defined

$$A \times B \coloneqq \{(a, b) \in \mathcal{P}(\mathcal{P}(A \cup B)) \mid a \in A \wedge b \in B\} \tag{35}$$

which is a valid set by the axiom schema of specification.

Theorem 2.1

$(a, b) = (a', b')$ if and only if $a = a'$ and $b = b'$.

Proof. The backwards implication is trivial. For the forward, let us assume that $\{\{a\}, \{a, b\}\} = \{\{a'\}, \{a', b'\}\}$. If $a \neq b$, then $\{a\} = \{a'\}$ and $\{a, b\} = \{a', b'\}$. So first $a = a'$ and then $\{a, b\} = \{a, b'\}$ implies $b = b'$. If $a = b$, then $\{\{a\}, \{a, a\}\} = \{\{a\}\}$. So $\{a\} = \{a'\}$ and $\{a\} = \{a', b'\}$, and we get $a = a' = b'$, and so $a = a'$ and $b = b'$.

Theorem 2.2

We have

$$A \times (B \cup C) = (A \times B) \cup (A \times C) \tag{36}$$

From this we can define the Cartesian product of any finite collection of sets recursively. It is indeed the case that $(X \times Y) \times Z$ is a different set from $X \times (Y \times Z)$, but as we will see in later functions, we can define a canonical bijection between them, treating them as equivalent. Furthermore, notice that we have not defined the Cartesian product of infinite sets yet. We can define them using functions actually.

The definition of Cartesian products allows us to formally define **correspondences**. The most notable correspondences are *functions*, *order relations*, and *equivalence relations*.

Definition 2.2 (Correspondence, Relation)

A **correspondence**, or a **binary relation**, R on a set A is a subset of $A \times A$. We write aRb if and only if $(a, b) \in R$.[a] But not all relations may be meaningful or interesting. Therefore we usually like to have certain properties on these relations, including but not limited to

1. *Reflexive*. For all $a \in A$, aRa
2. *Symmetric*. For all $a, b \in A$, if aRb then bRa
3. *Antisymmetric*. For all $a, b \in A$, if aRb and bRa then $a = b$
4. *Transitive*. For all $a, b, c \in A$, if aRb and bRc then aRc
5. *Total*. For all $a, b \in A$, either aRb or bRa

[a] It is a way of describing precisely which two elements are related to one another.

2.2 Functions

We explore our first—and most universally used—relation.

Definition 2.3 (Function)

Given two sets X, Y, a function $f : X \to Y$ is a subset $f \subset X \times Y$ satisfying the following

1. For all $x \in X$, there exists $y \in Y$ s.t. $(x, y) \in f$.[a]
2. For all $x \in X$ and $y, y' \in Y$, if $(x, y) \in f$ and $(x, y') \in f$, then $y = y'$.[b]

The set X is said to be the **domain** and Y the **codomain**.

$$\begin{tikzcd} X \arrow[r, "f"'] & Y \end{tikzcd}$$

Figure 1: A diagram representing the function $f : X \to Y$.

[a] This says that f must be defined for all inputs in X.

[b] In other words, f must map one element to exactly one element.

Lemma 2.3

Given functions f, g, $f = g$ if and only if the domains of f and g are equal and $f(x) = g(x)$ for all $x \in \text{domain}(f)$.

2.2.1 Sets Mapped Through Functions

Definition 2.4 (Image, Preimage)

Given some $f : X \to Y$ and $A \subset X$, the **image** of A under f is defined

$$f(A) := \{y \in Y \mid \exists x \in X (f(x) = y)\} \tag{37}$$

Given $B \subset Y$, the **preimage** of B under f is defined

$$f^{-1}(B) := \{x \in X \mid f(x) \in B\} \tag{38}$$

Now let's see how these operations behavior under functions.

Theorem 2.4 (Preservation Under Mapping Back and Forth)

Given $f : A \to B$, with $A_0, A_1 \subset A$ and $B_0, B_1 \subset B$, the following hold

1. $A_0 \subset f^{-1}(f(A_0))$, with equality holding if f is injective.
2. $f(f^{-1}(B_0)) \subset B_0$, with equality holding if f is surjective.

Proof. Listed.

1. Assume that $x \in A_0$. Then $f(x) \in f(A_0)$. The preimage is
$$f^{-1}(f(A_0)) \coloneqq \{y \in A \mid f(y) \in f(A_0)\} \tag{39}$$
and x certainly satisfies the condition that $f(x) \in f(A_0)$. Therefore $x \in f^{-1}(f(A_0))$ and so $A_0 \subset f^{-1}(f(A_0))$.
Now assume that f is injective. It suffices to prove that $f^{-1}(f(A_0)) \subset A_0$ since the other direction is proven for all functions. We prove this by proving the contrapositive, i.e. $x \notin A_0 \implies x \notin f^{-1}(f(A_0))$. Suppose $x \notin A_0 \implies f(x) \notin f(A_0) \implies f^{-1}(f(x)) \not\subset f^{-1}(f(A_0))$ by definition of the image and preimage. But note that since f is injective, $f^{-1}(f(x)) = x$.[a] and thus $x \notin f^{-1}(f(A_0))$.
2. We prove this using the contrapositive. Assume that $x \notin B_0$. Then, with abuse of notation, we have by definition of the preimage and the image $f^{-1}(x) \not\subset f^{-1}(B_0) \implies f(f^{-1}(x)) \not\subset f(f^{-1}(B_0))$. But $f(f^{-1}(x)) = \{x\}$, since we are just mapping the preimage of x back across to f. Therefore, $x \notin f(f^{-1}(B_0))$.
Now assume that f is surjective. It suffices to prove that $B_0 \subset f(f^{-1}(B_0))$. Assume $y \in B_0$. Since f is surjective, we know that $f^{-1}(y)$ is nonempty in A. We can state $f^{-1}(y) \subset f^{-1}(B_0)$[b] which then implies $f(f^{-1}(y)) \subset f(f^{-1}(B_0))$.[c] But $f(f^{-1}(y)) = y$ as mentioned previously, and so $y \in f(f^{-1}(B_0))$.

[a] More specifically, if we treat x as the singleton set, $f(x)$ is also a singleton set by definition of a function. Since f is injective, the preimage of a singleton set must be a singleton set. If it were not, then there exists x, y with $x \neq y$ that maps to the same z, which contradicts the definition of injectivity.

[b] The formal proof of this is given in Munkres 1.2.2.a.

[c] Again formal proof of this given in Munkres 1.2.2.e.

Example 2.1 (Counterexamples)

To see why equality does not hold in general for the two cases, look at the counterexamples below.

1. $A_0 \not\supset f^{-1}(f(A_0))$.
2. $f(f^{-1}(B_0)) \not\supset B_0$. Consider $X = Y = \{0, 1\}$ and $f : X \to Y$ defined $f(0) = f(1) = 0$. Consider $C = Y$. We have $f^{-1}(C) = f^{-1}(0) \cup f^{-1}(1) = X \cup \emptyset = X$. Then $f(f^{-1}(C)) = f(X) = \{0\} \neq C$.

Theorem 2.5 (Preservation Under Preimages)

Given $f : A \to B$, with $A_\alpha \subset A$ and $B_\alpha \subset B$, f preserves the inclusion, union, intersection, and set difference under the preimage.

1. *Inclusion.* $B_0 \subset B_1 \implies f^{-1}(B_0) \subset f^{-1}(B_1)$.
2. *Union.* $f^{-1}(\cup B_\alpha) = \cup_\alpha f^{-1}(B_\alpha)$.
3. *Intersection.* $f^{-1}(\cap B_\alpha) = \cap_\alpha f^{-1}(B_\alpha)$.
4. *Set Difference.* $f^{-1}(B_0 \setminus B_1) = f^{-1}(B_0) \setminus f^{-1}(B_1)$.

Proof. Listed.

1. *Inclusion.* If $x \in B_0$, then $f^{-1}(x) \subset A$ maps to x by definition. But since $x \in B_0$, $f^{-1}(x)$ maps to a point in B_0, and so $f^{-1}(x) \subset f^{-1}(B_0)$. Since $B_0 \subset B_1$ by assumption, $x \in B_1$, and by the previous logic but with B_0 replaced by B_1 we have $f^{-1}(x) \subset f^{-1}(B_1)$. We have just proved that $f^{-1}(x) \in f^{-1}(B_0) \implies f^{-1}(x) \in f^{-1}(B_1)$, and so $f^{-1}(B_0) \subset f^{-1}(B_1)$.
2. *Union.* We prove bidirectionally.
   (a) $f^{-1}(B_0 \cup B_1) \subset f^{-1}(B_0) \cup f^{-1}(B_1)$. Let $x \in f^{-1}(B_0 \cup B_1)$ which by definition of the preimage means $f(x) \in B_0 \cup B_1$. Therefore $f(x) \in B_0$ or B_1. Without loss of generality, let $f(x) \in B_0$. Then we have
   $$x \in f^{-1}(f(x)) \subset f^{-1}(B_0) \tag{40}$$
   where the first inclusion comes from [Munkres 1.2.1.a] when treating $A_0 = \{x\}$, and the second subset comes from [Munkres 1.2.2.a] when treating $B_0 = \{f(x)\}, B_1 = B_1$. Therefore $x \in f^{-1}(B_0) \subset f^{-1}(B_0) \cup f^{-1}(B_1)$.
   (b) $f^{-1}(B_0) \cup f^{-1}(B_1) \subset f^{-1}(B_0 \cup B_1)$. Let $x \in f^{-1}(B_0) \cup f^{-1}(B_1)$. Without loss of generality, let $x \in f^{-1}(B_0)$ which by definition of the preimage implies $f(x) \in B_0 \subset (B_0 \cup B_1) \implies f(x) \in (B_0 \cup B_1)$. Therefore, we have
   $$x \in f^{-1}(f(x)) \subset f^{-1}(B_0 \cup B_1) \tag{41}$$
   where the inclusion claim comes from [Munkres 1.2.1.a] when treating $A_0 = \{x\}$, and the subset claim comes from [Munkres 1.2.2.a] when treating $B_0 = \{f(x)\}, B_1 = B_0 \cup B_1$. Therefore $x \in f^{-1}(B_0 \cup B_1)$.

   Therefore, $f^{-1}(B_0) \cup f^{-1}(B_1) = f^{-1}(B_0 \cup B_1)$.
3. *Intersection.* We prove bidirectionally.
   (a) $f^{-1}(B_0 \cap B_1) \subset f^{-1}(B_0) \cap f^{-1}(B_1)$. Assume $x \in f^{-1}(B_0 \cap B_1)$, which by definition of the preimage means $f(x) \in B_0 \cap B_1$. So
   $$f(x) \in B_0 \implies x \in f^{-1}(f(x)) \subset f^{-1}(B_0) \tag{42}$$
   $$f(x) \in B_1 \implies x \in f^{-1}(f(x)) \subset f^{-1}(B_1) \tag{43}$$
   where the inclusion claim comes from [Munkres 1.2.1.a] when treating $A_0 = \{x\}$, and the subset claim comes from [Munkres 1.2.2.a] when treating $f(x)$ as a singleton set. Therefore x is in both of the preimages and so $x \in f^{-1}(B_0) \cap f^{-1}(B_1)$.
   (b) $f^{-1}(B_0) \cap f^{-1}(B_1) \subset f^{-1}(B_0 \cap B_1)$. Let $x \in f^{-1}(B_0) \cap f^{-1}(B_1)$. Then by definition of intersection and preimage,
   $$x \in f^{-1}(B_0) \implies f(x) \in B_0 \tag{44}$$
   $$x \in f^{-1}(B_1) \implies f(x) \in B_1 \tag{45}$$
   and so $f(x) \in B_0 \cap B_1$ by definition of intersection. This means by definition of the preimage that $x \in f^{-1}(B_0 \cap B_1)$.
4. *Set Difference.* We prove bidirectionally.
   (a) $f^{-1}(B_0 \setminus B_1) \subset f^{-1}(B_0) \setminus f^{-1}(B_1)$. Let $x \in f^{-1}(B_0 \setminus B_1)$ which by definition of preimage means $f(x) \in B|0 \setminus B_1$. This implies two things. First,
   $$f(x) \in B_0 \implies x \in f^{-1}(f(x)) \subset f^{-1}(B_0) \tag{46}$$
   where the inclusion comes from [Munkres 1.2.1.a] when treating $A_0 = \{x\}$ as the single set, and the subset claim comes from [Munkres 1.2.2.a] stating that inclusions are preserved under the preimage operator. Secondly, we claim that
   $$f(x) \notin B_1 \implies x \notin f^{-1}(B_1) \tag{47}$$
   since if $x \in f^{-1}(B_1)$, then $f(x) \in B_1$ by definition of the preimage.

(b) $f^{-1}(B_0) \setminus f^{-1}(B_1) \subset f^{-1}(B_0 \setminus B_1)$. Let $x \in f^{-1}(B_0) \setminus f^{-1}(B_1)$. Then the following holds

$$x \in f^{-1}(B_0) \implies f(x) \in B_0 \tag{48}$$

$$x \notin f^{-1}(B_1) \implies f(x) \notin B_1 \tag{49}$$

from the definition of the preimage and the contrapositive of its implication. Therefore $f(x) \in B_0 \setminus B_1$ which by definition of the preimage $x \in f^{-1}(B_0 \setminus B_1)$.

Theorem 2.6 (Preservation Under Images)

Given $f : A \to B$, with $A_\alpha \subset A$ and $B_\alpha \subset B$, f preserves the inclusion and union under the image, but inclusion properties for the intersection and set difference hold.

1. *Inclusion.* $A_0 \subset A_1 \implies f(A_0) \subset f(A_1)$.
2. *Union.* $f(\cup_\alpha A_\alpha) = \cup_\alpha f(A_\alpha)$.
3. *Intersection.* $f(\cap_\alpha A_\alpha) \subset \cap_\alpha f(A_\alpha)$, and equality holds if f is injective.
4. *Set Difference.* $f(A_0 \setminus A_1) \supset f(A_0) \setminus f(A_1)$, and equality holds if f is injective.

Proof. Listed.

1. *Inclusion.* Let $x \in A_0$. Then by definition of the image $f(x) \in f(A_0)$. Since $A_0 \subset A_1$, then $x \in A_1$ and it immediately follows that $f(x) \in f(A_1)$. Therefore $f(A_0) \subset f(A_1)$.
2. *Union.* We prove bidirectionally.
   (a) $f(A_0 \cup A_1) \subset f(A_0) \cup f(A_1)$. Let $y \in f(A_0 \cup A_1)$. Then by definition there exists some $x \in A_0 \cup A_1$ s.t. $f(x) = y$. WLOG let $x \in A_0$. Then by definition $y = f(x) \in f(A_0) \subset f(A_0) \cup f(A_1)$.
   (b) $f(A_0) \cup f(A_1) \subset f(A_0 \cup A_1)$. Let $y \in f(A_0) \cup f(A_1)$. WLOG $y \in f(A_0)$, and there exists some $x \in A_0$ s.t. $f(x) = y$. Since $x \in A_0$, $x \in A_0 \cup A_1$, and by definition $y = f(x) = f(A_0) \cup f(A_1)$.
3. *Intersection.* Assume that $y \in f(A_0 \cap A_1)$. Then by definition there exists some $x \in A_0 \cap A_1$ s.t. $f(x) = y$. So we have

$$x \in A_0 \implies f(x) \in f(A_0) \tag{50}$$

$$x \in A_1 \implies f(x) \in f(A_1) \tag{51}$$

   and therefore $y = f(x) \in f(A_0) \cap f(A_1)$.
   To prove equality, it suffices to show that $f(A_0) \cap f(A_1) \subset f(A_0 \cap A_1)$ if f is injective. Assume that $y \in f(A_0) \cap f(A_1)$. Then $y \in f(A_0)$, and so there exists an $x \in A_0$ s.t. $y = f(x) \in f(A_0)$. By the same logic there exists an $x' \in A_1$ s.t. $y = f(x') \in f(A_1)$. But since f is injective, this implies that $x = x'$. So $x \in A_0 \cap A_1$, and so $y = f(x) \in f(A_0 \cap A_1)$.
4. *Set Difference.* Assume that $y \in f(A_0) \setminus f(A_1)$. Since $y \in f(A_0)$, there exists some $x \in A_0$ s.t. $y = f(x)$. Since $y \notin f(A_1)$, there exists no $x' \in A_1$ s.t. $y = f(x')$. Therefore, $x \in A_0 \setminus A_1 \implies y = f(x) \in f(A_0 \setminus A_1)$.
   To prove equality, it suffices to show that $f(A_0 \setminus A_1) \subset f(A_0) \setminus f(A_1)$ if f is injective. Assume that $y \in f(A_0 \setminus A_1)$. Then there exists some $x \in A_0 \setminus A_1$ s.t. $f(x) = y$. We claim that x is unique since if there were two x, x', then $f(x) = f(x')$ with $x \neq x'$, which means f is not injective. We see that $x \in A_0 \implies y = f(x) \in f(A_0)$, and $x \notin A_1 \implies y = f(x) \notin f(A_1)$. Therefore, $x \in f(A_0) \setminus f(A_1)$.

Example 2.2 (Intersection Not Necessarily Preserved)

Note that intersection is not necessarily preserved. To see why, look at the counterexample. Consider $A = \{0, 1\}, B = \{1, 2\}$, and define

$$f(0) = f(2) = 0, f(1) = 1 \tag{52}$$

Then $f(A) = f(B) = \{0, 1\} \implies f(A) \cap f(B) = \{0, 1\}$. On the other hand, we have $A \cap B = \{1\} \implies f(A \cap B) = \{1\}$.

2.2.2 Composite Functions

Definition 2.5 (Composition)

Given functions $f : X \to Y$, $g : Y \to Z$, we define the **composite**, denoted $g \circ f$ or $g(f(\cdot))$, of f and g as the subset

$$g \circ f \coloneqq \{(x, z) \in X \times Z \mid \exists y \in Y (f(x) = y \land f(y) = z)\} \tag{53}$$

Theorem 2.7 (Compositions)

A composite is a function.

$$\begin{tikzcd}
X \arrow[r, "f"] \arrow[rd, "g\circ f"'] & Y \arrow[d, "g"] \\
 & Z
\end{tikzcd}$$

Figure 2: Commutative diagram representing a composition of functions.

Proof. Using the definition above, we prove the two properties.

1. For all $x \in X$, there exists $y \in Y$ s.t. $(x, y) \in f$. Similarly, for all $y \in Y$, there exists $z \in Z$ s.t. $(y, z) \in g$. Therefore, for all $x \in X$, there exists a $y \in Y$, which follows that there exists also a $z \in Z$. Therefore $g \circ f$ is defined for all inputs in X.
2. For all $x \in X$ and $z, z' \in Z$, say that $(x, z), (x, z') \in g \circ f$. Then by definition of composition there exists a $y, y' \in Y$ s.t. $f(x) = y, f(y) = z$ and $f(x) = y', f(y') = z'$. Since f is a function, $y = y'$. Since g is a function, $y = y' \implies z = z'$. Therefore $g \circ f$ is a function.

For the computer science students, note that a function behaves precisely like functional dependencies in a relational database. A composition represents a natural join.

Theorem 2.8 (Associativity)

Composition is associative. That is, consider $f : Y \to Z, g : X \to Y, h : W \to X$ functions. Then

$$(f \circ g) \circ h = f \circ (g \circ h) \tag{54}$$

Therefore, we write this as

$$f \circ g \circ h \tag{55}$$

$$\begin{tikzcd}
 & X \arrow[r, "g"] \arrow[rrd, "g\circ h"] & Y \arrow[rd, "h"] & \\
W \arrow[ru, "f"] \arrow[rru, "g\circ f"'] & & & Z
\end{tikzcd}$$

Figure 3

Proof. Consider any $w \in W$, and let us label $x = h(w)$, $y = g(x)$, $z = f(y)$, where x, y, z must be uniquely determined by w since it is a function. Then,

$$((f \circ g) \circ h)(w) = (f \circ g)(h(w)) = (f \circ g)(x) = z \tag{56}$$

$$(f \circ (g \circ h))(w) = f((g \circ h)(w)) = f(y) = z \tag{57}$$

and they coincide for all $w \in W$.

If we are familiar with algebra, this gives the set of functions $\{f : X \to X\}$ the structure of a *monoid* under composition. We can also talk about commutativity.

Definition 2.6 (Commutativity)

Two functions $f, g : X \to X$ are said to be **commute** if

$$f \circ g = g \circ f \tag{58}$$

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[d, "g"] & X \arrow[d, "g"] \\ X \arrow[r, "f"] & X \end{tikzcd}$$

Figure 4: Commutative diagram representing commuting functions f, g.

Theorem 2.9 (Composition)

Let $f : X \to Y$ and $g : Y \to Z$.

1. f injective and g injective $\implies$ $g \circ f$ injective.
2. f surjective and g surjective $\implies$ $g \circ f$ surjective.
3. f bijective and g bijective $\implies$ $g \circ f$ bijective.

2.2.3 Injective, Surjective, Bijective Functions

Definition 2.7 (Injectivity, Surjectivity, Bijectivity)

A function $f : X \to Y$ is said to be

1. **injective** if $\forall x \in X, \forall x' \in X\big(f(x) = f(x') \implies x = x'\big)$.
2. **surjective** if $\forall y \in Y \exists x \in X(y = f(x))$.
3. **bijective** if it is injective and surjective.

Definition 2.8 (Inverse Function)

If a function $f : X \to Y$ is bijective, then there exists an **inverse function** $f^{-1} : Y \to X$ satisfying

$$\forall x \in X\big[f(f^{-1}(x)) = f^{-1}(f(x)) = x\big] \tag{59}$$

Definition 2.9 (Restriction, Extension)

If $f : X \to Y$ and $X_0 \subset X$, we define the **restriction** of f to X_0 to be the function mapping to Y whose rule is

$$f|_{X_0} \coloneqq \{(x, f(x)) \in f x \in X_0\} \tag{60}$$

Letting $g : X_0 \to Y$, any function $f : X \supset X_0 \to Y$ satisfying $f(x) = g(x)$ for all $x \in X_0$ is said to be an **extension** of g to X.

Theorem 2.10 (Injectivity/Surjectivity)

Let $f : X \to Y$, $g : Y \to Z$, and $h = g \circ f$. The following hold:

1. h injective $\implies$ f injective.
2. h surjective $\implies$ g surjective.
3. h bijective $\implies$ f injective and g bijective.

Corollary 2.11 (Bijection Equals Existence of Inverse)

$f : X \to Y$ has a inverse function $f^{-1} : B \to A$ iff it is bijective.

Corollary 2.12 (Decomposition)

Any function $h : X \to Y$ can be decomposed to the form $h = g \circ f$, where f is injective and g is surjective.

Proof. Given X, let us define an equivalence class where for any $x, y \in X$, $x \sim y$ iff $f(x) = f(y)$. Call this quotient space $X/\sim$. Then we can define the mappings.

1. $\iota : X \to X/\sim$ which maps each element to its equivalence class. $\iota(x) = [x]$
2. $f' : X/\sim \to Y$ which maps each class to the element of Y that it maps to. $f'([x]) = f(x)$.

ι is surjective since for every $[x] \in X/\sim$, there exists at least one element $x \in X$ that maps to it. f' is injective since have squished all the points x that map to the same y into a single class $[x]$.

$$\begin{tikzcd}
X \arrow[r, "f"] \arrow[d, "\iota"] & Y \\
X/\sim \arrow[ur, "f'"'] &
\end{tikzcd}$$

Figure 5: Decomposition of f into surjective ι and injective f'.

Theorem 2.13 (Inverse of Inverses)

If f is bijective, then $f = (f^{-1})^{-1}$.

Theorem 2.14 (Inverse of Compositions)

If f, g are both bijective, then

$$(f \circ g)^{-1} = g^{-1} \circ f^{-1} \tag{61}$$

Theorem 2.15 (Finite Set Mappings)

Suppose X and Y are finite sets, each with n elements, and $f : X \to Y$. If f is injective or bijective, then f is bijective.

2.2.4 Cartesian Products as Functions

Remember that we have used the power set to construct the binary Cartesian product of sets. Now we extend this to arbitrary Cartesian products using functions.

Lemma 2.16

Let A and B be sets. Then the set of all functions $f : A \to B$ is denoted B^A. We claim that this set exists.

Proof.

Definition 2.10 (Cartesian Product)

Let I be some set (used for indexing) and $S = \{S_i\}_{i \in I}$ be an indexed set of sets. Then we define the **Cartesian product** of S_i's as

$$\prod S = \prod_{i \in I} S_i \coloneqq \{f \in (\cup S)^I \mid f(i) \in S_i \; \forall i \in I\} \tag{62}$$

Note that we need the property since we want $f(i)$ to map specifically into S_i.

With this more general construction, we can pretty much forget about the previous construction of binary Cartesian products.

2.3 Order Relations

Definition 2.11 (Partial Order)

An **order** is a relation R, usually denoted $\leq$, that satisfies the following properties.

1. *Reflexive.* For all $a \in A$, $a \leq a$
2. *Antisymmetric.* For all $a, b \in A$, if $a \leq b$ and $b \leq a$ then $a = b$
3. *Transitive.* For all $a, b, c \in A$, if $a \leq b$ and $b \leq c$ then $a \leq c$

Note that when we say $x \leq y$, this means "x is related to y" (but does not necessarily mean that y is related to x), or "x is less than or equal to y." A set X with a partial order is called a partially ordered set.

Example 2.3 (Partially Ordered Sets)

We list some examples of partially ordered sets.

1. The real numbers ordered by the standard "less-than-or-equal" relation $\leq$ (totally ordered set as well).
2. The set of subsets of a given set X ordered by inclusion. That is, the power set 2^X with the partial order $\subseteq$ is partially ordered.
3. The set of natural numbers equipped with the relation of divisibility.
4. The set of subspaces of a vector space ordered by inclusion.
5. For a partially ordered set P, the sequence space containing all sequences of elements from P,

where sequence a precedes sequence b if every item in a precedes the corresponding item in b.

Definition 2.12 (Comparable Elements)

Given elements a, b of partially ordered set A, if either $a \leq b$ or $b \leq a$, then a and b are **comparable**. Otherwise, they are **incomparable**.

While partial ordering is nice, we would often want a stricter structure so that the order "encompasses" the whole set, i.e. every element is comparable. This property is sometimes known as *totality*.

Definition 2.13 (Total Order)

A partial order in which every pair of elements is comparable is called a **total order**, or **linear order**. Note that from this $\leq$ relation, we can similarly define

1. $a < b \iff (a \leq b) \land (a \neq b)$.
2. $a \geq b \iff \neg(a < b)$.
3. $a > b \iff \neg(a \leq b)$.

Almost always when we talk about ordered sets, we mean totally ordered sets. So we will work with them by default and define following terms according to totally ordered sets. For convenience of notation, we also write $a < x < b \iff (a < x) \land (x < b)$.

Definition 2.14 (Interval)

Given a totally ordered set X, we denote the **intervals** as

1. $(a, b) \coloneqq \{x \in X \mid a < x < b\}$
2. $[a, b) \coloneqq \{x \in X \mid a \leq x < b\}$
3. $(a, b] \coloneqq \{x \in X \mid a < x \leq b\}$
4. $[a, b] \coloneqq \{x \in X \mid a \leq x \leq b\}$

Definition 2.15 (Extrema)

Given a totally ordered set X,

1. $x \in X$ is a **maximum** X if $y \leq x$ for all $y \in X$.
2. $x \in X$ is a **minimum** X if $x \leq y$ for all $y \in X$.

Definition 2.16 (Bounds)

Given a totally ordered set X and some subset $S \subset X$.

1. $x \in X$ is an **upper bound** of S if $x \geq y$ for all $y \in S$.
2. $x \in X$ is a **lower bound** of S if $x \leq y$ for all $y \in S$.
3. $x \in X$ is a **supremum**, or **least upper bound**, of S if x is the minimum of the set of all upper bounds of S.
4. $x \in X$ is a **infimum**, or **greatest lower bound**, of S if x is the maximum of the set of all lower bounds of S.

Note that we have defined max/min separately from the concept of bounds. You can define the maximum of a set with just knowing the set, but the bounds require *both* some subset S with respect to an enclosing set X.[2] Intuitively, the main difference between the supremum/infimum and maximum/minimum is that the supremum/infimum accounts for limit points of the subset S.

Definition 2.17 (Convexity on Ordered Sets)

Given a totally ordered set $(X, \leq)$, a subset S is said to be **convex** if for all $a, b \in S$,

$$a \leq c \leq b \implies c \in S \tag{63}$$

2.4 Equivalence Relations

Definition 2.18 (Equivalence Relation)

An **equivalence relation** on a set A is a relation, denoted $\sim$ satisfying

1. *Reflexive.* For all $a \in A$, $a \sim a$
2. *Symmetric.* For all $a, b \in A$, if $a \sim b$ then $b \sim a$
3. *Transitive.* For all $a, b, c \in A$, if $a \sim b$ and $b \sim c$ then $a \sim c$

Given an equivalence relation, we can define an **equivalence class** as

$$[y] \coloneqq \{x \in A \mid x \sim y\} \tag{64}$$

Definition 2.19 (Partition)

A **partition** of a set X is a collection of disjoint nonempty subset of X whose union is all of A.

Theorem 2.17 (Quotient Space, Map)

The set of equivalence classes of a set X with an equivalence relation $\sim$ is a partition of X, denoted as the **quotient set** $X/\sim$. Therefore, the map $\iota : X \to X/\sim$ is well-defined and is called a **quotient map**.

Proof. Assume the contrary. If X has one element, then its equivalence class is $[x]$ and this is trivially proven. If X has at least 2 elements, let us call them $x, y \in X$ with $x \neq y$. $[x], [y]$ are their equivalence classes. Clearly due to reflexivity, $x \in [x]$ and $y \in [y]$ and so they are nonempty. Since we assumed that this is not a partition, there exists some $z \in X$ in both $[x], [y]$. But $z \in [x] \implies z \sim x$ and $z \in [y] \implies y \sim z$. So by transitivity, $x \sim z$, meaning that $[x] = [y]$. Therefore they must be the same element of a partition.

Example 2.4 (Circles)

M is the set of circles in $\mathbb{R}^2$. Given $a, b \in M$, $a \sim b$ iff the radii are equal in length. We can denote each equivalence class by $\{r\}$, where r is the length of the radius. We can define addition as

$$\{a\} + \{b\} \equiv \{a + b\} \tag{65}$$

[2] For example, it makes sense to define the maximum of a set $S = [0, 1]$ by itself, but not an upper bound for it. If $X = \mathbb{Q}$, then the supremum is 1, but if X was the set of all irrationals, then this has no supremum.

3 Structures

3.1 Algebraic Structures

Now sets are very boring when studying by themselves. It is often the case that we *endow* a set with some additional information, which we call *structure*. Consider the familiar integers, which come with certain operations such as + and ×, along with the concepts identity elements—namely 0 and 1—which allows us to define (additive and multiplicative) inverses.

Definition 3.1 (Structure)

A **structure** on a set is some additional information on the set, such as relations, constants, and operations associated with the set.

A lot of these properties depend on what operations you can do with them.

Definition 3.2 (Operation)

A **p-ary operation**[a] $*$ on a set A is a map

$$* : A^p \longrightarrow A \tag{66}$$

where A^p is the p-fold Cartesian product of A. In specific cases,

1. If $p = 1$, then $*$ is said to be **unary**.
2. If $p = 2$, then $*$ is **binary**.

We can consider for $p > 2$ and even if p is infinite.

[a] or called an operation of arity p.

Definition 3.3 (Monoid)

A **monoid** is a set S with an operation $+$.

Definition 3.4 (Group)

Definition 3.5 (Ring)

Definition 3.6 (Field)

A **field** is an algebraic structure $(\mathbb{F}, +, \cdot)$ where

1. $\mathbb{F}$ is an abelian group under $+$, with 0 being the *additive identity.*
2. $\mathbb{F} \setminus \{0\}$ is an abelian group under $\cdot$, with 1 being the *multiplicative identity.*
3. It connects the two operations through the *distributive property.*

$$x \cdot (y + z) = x \cdot y + x \cdot z \tag{67}$$

Lemma 3.1 (Left = Right Distributivity)

Left and right distributivity are equivalent.

$$x \cdot (y + z) = (y + z) \cdot x \tag{68}$$

Proof.

$$\begin{aligned} x \cdot (y + z) &= x \cdot y + x \cdot z && \text{(Distributive)} \\ &= y \cdot x + z \cdot x && \text{(Commutative)} \\ &= (y + z) \cdot x && \text{(Distributive)} \end{aligned}$$

Lemma 3.2 (Properties of Addition)

The properties of addition hold in a field.

1. If $x + y = x + z$, then $y = z$.
2. If $x + y = x$, then $y = 0$.
3. If $x + y = 0$, then $y = -x$.
4. $(-(-x)) = x$.

Proof. For the first, we have

$$\begin{aligned} x + y = x + z &\implies -x + (x + y) = -x + (x + z) && \text{(addition is a function)} \\ &\implies (-x + x) + y = (-x + x) + z && \text{(+ is associative)} \\ &\implies 0 + y = 0 + z && \text{(definition of additive inverse)} \\ &\implies y = z && \text{(definition of identity)} \end{aligned}$$

For the second, we can set $z = 0$ and apply the first property. For the third, we have

$$\begin{aligned} x + y = 0 &\implies -x + (x + y) = -x + 0 && \text{(addition is a function)} \\ &\implies (-x + x) + y = -x + 0 && \text{(+ is associative)} \\ &\implies 0 + y = -x + 0 && \text{(definition of additive inverse)} \\ &\implies y = -x && \text{(definition of identity)} \end{aligned}$$

For the fourth, we simply follow that if y is an inverse of z, then z is an inverse of y. Therefore, $-x$ being an inverse of x implies that x is an inverse of $-x$. $-(-x)$ must also be an inverse of $-x$. Since inverses are unique[a], $x = -(-x)$.

[a] This is proved in algebra.

Lemma 3.3 (Properties of Multiplication)

The properties of multiplication hold in a field.

1. If $x \neq 0$ and $xy = xz$, then $y = z$.
2. If $x \neq 0$ and $xy = x$, then $y = 1$.
3. If $x \neq 0$ and $xy = 1$, then $y = x^{-1}$.
4. If $x \neq 0$, then $(x^{-1})^{-1} = x$.

Proof. The proof is almost identical to the first. Since $x \neq 0$, we can always assume that x^{-1} exists. For the first, we have

$$\begin{aligned} xy = xz &\implies x^{-1}(xy) = x^{-1}(xz) && \text{(multiplication is a function)} \\ &\implies (x^{-1}x)y = (x^{-1}x)z && \text{($\times$ is associative)} \\ &\implies 1y = 1z && \text{(definition of multiplicative inverse)} \\ &\implies y = z && \text{(definition of identity)} \end{aligned}$$

For the second, we can set $z = 1$ and apply the first property. For the third, we have

$$\begin{aligned} xy = 1 &\implies x^{-1}(xy) = x^{-1}1 && \text{(multiplication is a function)} \\ &\implies (x^{-1}x)y = x^{-1}1 && \text{($\times$ is associative)} \\ &\implies 1y = x^{-1}1 && \text{(definition of multiplicative inverse)} \\ &\implies y = x^{-1} && \text{(definition of identity)} \end{aligned}$$

For the fourth, we simply see that x^{-1} is a multiplicative inverse of both x and $(x^{-1})^{-1}$ in the group $(\mathbb{F} \setminus \{0\}, \times)$, and since inverses are unique, they must be equal.

Lemma 3.4 (Properties of Distribution)

For any $x, y, z \in \mathbb{F}$, the field axioms satisfy

1. $0 \cdot x = 0$.
2. If $x \neq 0$ and $y \neq 0$, then $xy \neq 0$.
3. $-1 \cdot x = -x$.
4. $(-x)y = -(xy) = x(-y)$.
5. $(-x)(-y) = xy$.

Proof. For the first, note that

$$0x = (0 + 0) \cdot x = 0x + 0x \tag{69}$$

and subtracting $0x$ from both sides gives $0 = 0x$. For the second, we can claim that $xy \neq 0$ equivalently claiming that it will have an identity. Since $x, y \neq 0$, their inverses exists, and we claim that $(xy)^{-1} = y^{-1}x^{-1}$ is an inverse. We can see that by associativity,

$$(y^{-1}x^{-1})(xy) = y^{-1}(x^{-1}x)y = y^{-1}y = 1 \tag{70}$$

For the third, we see that

$$0 = 0 \cdot x = (1 + (-1)) \cdot x = 1 \cdot x + (-1) \cdot x = x + (-1) \cdot x \tag{71}$$

which implies that $-1 \cdot x$ is the additive inverse. The fourth follows immediately from the third by the associative property. For the fifth we can see that

$$\begin{aligned} (-x)(-y) &= (-1)x(-1)y && \text{(property 3)} \\ &= (-1)(-1)xy && \text{($\times$ is commutative)} \\ &= -1 \cdot (-xy) && \text{(property 3)} \\ &= -(-xy) && \text{(property 3)} \\ &= xy && \text{(addition property 4)} \end{aligned}$$

Note that given a set, we can really put whatever order we want on it. However, consider the field with the following order.

$$\mathbb{F} = \{0, 1\},\ 0 < 1 \tag{72}$$

This does not behave well with respect to its operations because for example if we have $0 < 1$, then adding the same element to both sides should preserve the ordering. But this is not the case since $0+1 = 1 > 1+1 = 0$. While it may be easy to define an order, we would like it to be an ordered field.

Definition 3.7 (Ordered Ring/Field)

An **ordered ring** is a ring that has an order satisfying

1. $y < z \implies x + y < x + z$ for all $x \in \mathbb{F}$.
2. $x > 0, y > 0 \implies xy > 0$.

An **ordered field** has the same definition, and an ordered field is an ordered ring.

Theorem 3.5 (Properties)

In an totally ordered ring,

1. $x > 0 \implies -x < 0$.
2. $x \neq 0 \implies x^2 > 0$.
3. If $x > 0$, then $y < z \implies xy < xz$.

Proof. The first property is a single-liner

$$0 < x \implies 0 + -x < x + -x \implies -x < 0 \tag{73}$$

For the second property, it must be the case that $x > 0$ or $x < 0$. If $x > 0$, then by definition $x^2 > 0$. If $x < 0$, then

$$x^2 = 1 \cdot x^2 = (-1)^2 \cdot x^2 = (-1 \cdot x)^2 = (-x)^2 \tag{74}$$

and since $-x > 0$ from the first property, we have $x^2 = (-x)^2 > 0$. For the third, we use the distributive property.

$$\begin{aligned} y < z &\implies 0 < z - y && (75)\\ &\implies 0 = x0 < x(z - y) = xz - xy && (76)\\ &\implies xy < xz && (77) \end{aligned}$$

3.2 Topological and Metric Spaces

Definition 3.8 (Topology)

Note that an order can be used to generate an order topology, which we will define below.

Example 3.1 (Order Topology on $\mathbb{Q}$)

The order topology on $\mathbb{Q}$ is the topology generated by the set $\mathscr{B}$ of all open intervals

$$(a, b) \coloneqq \{x \in \mathbb{Q} \mid a < x < b\} \tag{78}$$

Definition 3.9 (Metric)

Given an arbitrary set X, a **metric** on X is a function

$$d : X \times X \to \mathbb{R} \tag{79}$$

satisfying

3.3 Vector Spaces

Definition 3.10 (Vector Space)

A vector space V over a field $\mathbb{F}$ is a.

Example 3.2 (Norm)

Given a vector space V over subfield $\mathbb{F} \subset \mathbb{C}$, the norm

$$|| \cdot || : V \to \mathbb{R} \tag{80}$$

Example 3.3 (Inner Product)

Given a vector space V over subfield $\mathbb{F} \subset \mathbb{C}$, the **inner product**

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C} \tag{81}$$

is a map satisfying

Definition 3.11 (Convex Sets)

A set S is convex if for every point $x, y \in S$, the point

$$z = tx + (1 - t)y \in S \tag{82}$$

where $0 \leq t \leq 1$.

3.4 Measure Spaces

4 The Naturals

4.1 Axiom of Infinity

Now we attempt to try and construct the natural numbers. We start with existence of the empty set $S_0 = \emptyset = \{\}$. Now given S_n, we can define inductively S_{n+1} through a *successor function* S that maps to the "next number."

$$S_{n+1} = S(W) \coloneqq S_n \cup \{S_n\} \tag{83}$$

where

1. S_n exists either by induction on n or from the base case where the axiom of empty set applies.
2. $\{S_n\}$ exists through the axiom of pairing.
3. $S_n \cup \{S_n\}$ exists through the axiom of union.

Note that even though we did not talk about what $n + 1$ means, we can just treat it as some elements of some indexing set. Great, so we can indeed prove that each element S_n exists. This motivates the following definition.

Definition 4.1 (Inductive Set)

A set I is called **inductive** if
1. $0 = \{\} \in I$.
2. If $n \in I$, then $n + 1 = S(n) = n \cup \{n\} \in I$

Therefore, an inductive set contains 0 and with each element, its sucessor. This makes us motivate the definition of the naturals as an inductive set which contains no other elements but the natural numbers, i.e. it is the smallest inductive set. We can express this minimal property by letting every natural number x be contained in *every* inductive set, leading us to the following definition.

$$\mathbb{N} = \{x \mid x \in I \text{ for every inductive set } I\} \tag{84}$$

But note that this does not follow the schema of restricted specification, so we must first take some *existing* inductive set A and define

$$\mathbb{N} = \{x \in A \mid x \in I \text{ for every inductive set } I\} \tag{85}$$

Now the question remains of whether there are any inductive sets at all. The intuitive answer is yes, but with our axioms so far, the existence of infinite sets cannot be proven actually. The reason is that we have started out with a finite set $\emptyset$, and the construction of new sets with our axioms only allows us to create more finite sets. Therefore, while we can construct finite sets with an unbounded number of elements, we can never reach an infinite set. This is why we must axiomatically claim that a finite inductive set exists.

Axiom 4.1 (Axiom of Infinity)

An inductive set exists.

This allows us to utilize the axiom of restricted specification to construct the primitive form of the naturals.

Definition 4.2 (Von Neumann Ordinals)

The **Von Neumann ordinals** is the minimal set X satisfying the axiom of infinity. It is the set containing

$$\begin{aligned} 0 &= \{\} = \emptyset \\ 1 &= \{0\} = \{\emptyset\} \\ 2 &= \{0,1\} = \{\emptyset, \{\emptyset\}\} \\ 3 &= \{0,1,2\} = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} \\ 4 &= \{0,1,2,3\} = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}\} \\ \ldots &= \ldots \end{aligned}$$

Now note that the von Neumann ordinals are not the only way to construct such a set. The **Zermelo ordinals** define the successor function to be $S(w) = \{w\}$ and state that the inductive set with this successor function exists. However, the von Neumann ordinals have the nice property that the cardinality n of the set is precisely the natural number that we would like to identify it with, and the fact that a natural number n contains all naturals $0, \ldots, n-1$ as elements of n.

4.2 Natural Numbers

From this, with a few more structures we can define the naturals as a commutative monoid with respect to both the addition and multiplication operations.

Definition 4.3 (Natural Numbers)

The **natural numbers** $\mathbb{N} = \{0, 1, 2, \ldots\}$ is the set of von Neumann ordinals with the following structure.

1. *Order.* The relation $\leq$ defined as
$$m \leq n \iff m \in n \tag{86}$$
is an order relation.
2. *Addition* can be defined recursively with the successor function as[a]
$$n + 0 \coloneqq n \tag{87}$$
$$n + 1 \coloneqq S(n) \tag{88}$$
$$n + S(m) \coloneqq S(n+m) \tag{89}$$
3. *Multiplication* is also defined recursively using the definition of addition.
$$n \times 0 \coloneqq 0 \tag{90}$$
$$n \times S(m) \coloneqq (n \times m) + n \tag{91}$$
which is familiar to the process of adding m to itself n times.
4. *Additive Identity* is 0, following directly from the definition of addition.
5. *Multiplicative Identity* is 1, following directly from the definition of multiplication.

[a] note that I am using $\coloneqq$ to denote that this is an *identity*, not an equation to be solved.

Example 4.1 (Order)

I assert that $3 \leq 4$ since
$$\{0,1,2\} = 3 \in \{0,1,2,3\} = 4 \tag{92}$$

Example 4.2 (Addition)

To define $5+2$, we can see that

$$\begin{aligned} 5+2 &= 5+S(1) && (93)\\ &\coloneqq S(5+1) && (94)\\ &= S(5+S(0)) && (95)\\ &\coloneqq S(S(5+0)) && (96)\\ &\coloneqq S(S(5)) && (97)\\ &= S(6) && (98)\\ &= 7 && (99) \end{aligned}$$

Example 4.3 (Multiplication)

To define 4×3, we apply the recursive definitions

$$\begin{aligned} 4\times 3 &= 4\times S(2) && (100)\\ &\coloneqq 4\times 2+4 && (101)\\ &= 4\times S(1)+4 && (102)\\ &\coloneqq 4\times 1+4+4 && (103)\\ &\coloneqq 4\times S(0)+4+4 && (104)\\ &= 4\times 0+4+4+4 && (105)\\ &= 0+4+4+4 && (106)\\ &= 12 && (107) \end{aligned}$$

Theorem 4.1 (Commutative)

$+$ and $\times$ are commutative in $\mathbb{N}$.

Theorem 4.2 (Associativity)

$+$ and $\times$ are associative in $\mathbb{N}$.

4.3 Induction

Lemma 4.3 (Well Ordering Principle)

Every nonempty subset of $\mathbb{N}$ has a minimal element.

Proof. Take a subset $A \subset \mathbb{N}$.

1. If $0 \in A$, the minimum is 0.
2. Else if $1 \in A$, the minimum is 1.
3. ...

We can use this inductive property of natural numbers to prove properties of them. Note that this can only be used to prove for finite (yet unbounded) numbers!

Lemma 4.4 (Induction Principle)

Given $P(n)$, a property depending on a natural number $n \in \mathbb{N}$,

1. if $P(n_0)$ is true for some $n_0 \in \mathbb{N}$, and
2. if for every $k \geq n_0$, $P(k)$ true implies $P(k+1)$ true,

then $P(n)$ is true for all $n \geq n_0$.

Lemma 4.5 (Strong Induction Principle)

Given $P(n)$, a property depending on a positive integer n,

1. if $P(n_0), P(n_0+1), \ldots, P(n_0+m)$ are true for some positive integer n_0, and nonnegative integer m, and
2. if for every $k > n_0 + m, P(j)$ is true for all $n_0 \leq j \leq k$ implies $P(k)$ is true,

then $P(n)$ is true for all $n \geq n_0$.

Theorem 4.6 (Equivalence of 3 Principles)

The well-ordering principle, induction principle, and the strong induction principle are all equivalent.

Proof. We prove the steps.

1. *Well Ordering $\implies$ Strong Induction.*
2. *Strong Induction $\implies$ Induction.*
3. *Induction $\implies$ Well-Ordering.*

The idea behind the strong induction principle leads to the proof using infinite descent. Infinite descent combines strong induction with the fact that every subset of the positive integers has a smallest element, i.e. there is no strictly decreasing infinite sequence of positive integers.

Theorem 4.7 (Infinite Descent)

Given $P(n)$, a property depending on positive integer, assume that $P(n)$ is false for a set of integers $\mathcal{S}$. Let the smallest element of $\mathcal{S}$ be n_0. If $P(n_0)$ false implies $P(k)$ false, where $k < n_0$, then by contradiction $P(n)$ is true for all n.

4.4 Sequences and The Recursion Theorem

Since we have defined the naturals, we can construct a function that takes in a natural number and outputs to an arbitrary set. This is called a *sequence.*

Definition 4.4 (Sequence)

Given a nonempty set X, a function $f : \mathbb{N} \to X$ is called a *sequence.* We usually denote it by

$$x_1, x_2, x_3, \ldots, \qquad x_n = f(n) \tag{108}$$

It is often shorthand written as (x_n).[a] There are two ways to define such a sequence:

1. *Explicitly.* We denote $x_n = f(n)$ in some closed form.
2. *Recursively.* We denote $x_{n+1} = g(x_n, n)$ for all $n \in \mathbb{N}$, with some base case x_0.

[a] Note that this is different from $\{x_n\}$, which is considered a *set* and is unordered.

The explicit definition gives us a well-defined sequence, but the recursive definition requires a bit more care. Given an initial condition x_0 and some recursive condition (similar to differential equations), such a definition is only well-formulated if there exists such a sequence (or function) that satisfies these two conditions. The recursion theorem proves both the existence and uniqueness of such an infinite sequence.

Theorem 4.8 (Recursion Theorem)

For any set A, any $a \in A$, and function $g : A \times \mathbb{N} \to A$, there exists a unique infinite sequence $f : \mathbb{N} \to A$ such that

1. $f(0) = a$.
2. $f(n+1) = g(f(n), n)$ for all $n \in \mathbb{N}$.

Sequences are particularly important in topology and analysis, where they are used as a tool to analyze properties of topological spaces or metric spaces. Since we are working in an arbitrary set, we cannot do much more than this.

5 Sequences of Sets

Let's talk about sequences of sets $(A_n)_n$.

Definition 5.1 (Monotone Sequence)

A sequence of sets $(A_n)_n$ is called

1. **(strictly) increasing** if $A_n \subsetneq A_{n+1}$.
2. **nondecreasing** if $A_n \subseteq A_{n+1}$.
3. **(strictly) decreasing** if $A_n \supsetneq A_{n+1}$.
4. **nonincreasing** if $A_n \supseteq A_{n+1}$.

Definition 5.2 (Limsup and Liminf of Sets)

Given a sequence of sets $(A_n)_n$, the **limsup** and **liminf** of them can be defined in the equivalent ways.

1. The **liminf** is the set of points that are missing in only a finite number of sets, and the **limsup** is the set of points that are in an infinite number of sets.

$$\liminf_{n\to\infty} A_n \coloneqq \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m \tag{109}$$

$$\limsup_{n\to\infty} A_n \coloneqq \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m \tag{110}$$

2. The **liminf** and **limsup** are the set of points x where the liminf and limsup of the indicator function function evaluated at x equals 1.

$$\liminf_{n\to\infty} A_n \coloneqq \{x \in X \mid \liminf_{n\to\infty} \mathbb{1}_{A_n}(x) = 1\} \tag{111}$$

$$\limsup_{n\to\infty} A_n \coloneqq \{x \in X \mid \limsup_{n\to\infty} \mathbb{1}_{A_n}(x) = 1\} \tag{112}$$

Both liminf and limsup always exist for any sequence of sets.

Proof. DeMorgan's law.

Lemma 5.1 (Monotonicity)

For any sequence of sets

$$\liminf_{n\to\infty} A_n \subseteq \limsup_{n\to\infty} A_n \tag{113}$$

Lemma 5.2 (Complements)

$$\liminf_{n\to\infty} A_n = \left(\limsup_{n\to\infty} A_n^c\right)^c \tag{114}$$

Proof.

Definition 5.3 (Limit of Sets)

6 The Integers and Rationals

6.1 Integers

We then construct the integers as an ordered commutative ring that embeds $\mathbb{N}$ through a canonical ordered monoid homomorphism.

Definition 6.1 (Integers)

The **integers**, denoted with $\mathbb{Z}$ is constructed by defining the equivalence relation on $\mathbb{N} \times \mathbb{N}$ given by

$$(a, b) \sim (c, d) \iff a + d = b + c \tag{115}$$

which we can think of as (a, b) being the solution to $a = x + b$. We call the quotient set $\mathbb{Z} \coloneqq (\mathbb{N} \times \mathbb{N})/\sim$, with the following structure.

1. *Order* defined by
$$[(a, b)] \leq_{\mathbb{Z}} [(c, d)] \iff a + d \leq_{\mathbb{N}} b + c \tag{116}$$
2. *Addition.*
$$[(a, b)] +_{\mathbb{Z}} [(c, d)] \coloneqq [(a + c, b + d)] \tag{117}$$
3. *Multiplication.*
$$[(a, b)] \times_{\mathbb{Z}} [(c, d)] \coloneqq [(ac + bd, ad + bc)] \tag{118}$$
4. *Additive Identity* is 0.
5. *Multiplicative Identity* is 1.
6. *Additive Inverse* of $z = (a, b) \in \mathbb{Z}$ is denoted $-z = (b, a)$.[a]

This makes $\mathbb{Z}$ an ordered ring.

[a] Since $(a, b) + (b, a) = (a + b, a + b) \sim (0, 0)$.

Note that we could have chosen other extensions, but the reason that we choose this is that it preserves the structure of $\mathbb{N}$. Even though set-theoretically, $\mathbb{N}$ and $\mathbb{Z}$ are disjoint, you can embed the naturals in the integers, which aligns with our intuition.

Theorem 6.1 (Embedding of $\mathbb{N}$ in $\mathbb{Z}$)

The map $f : \mathbb{N} \to \mathbb{Z}$ defined $f(n) = [(n, 0)]$ is an ordered monoid homomorphism. That is, it satisfies the following

$$f(n +_{\mathbb{N}} m) = f(n) +_{\mathbb{Z}} f(m), \tag{119}$$
$$f(n \times_{\mathbb{N}} m) = f(n) \times_{\mathbb{Z}} f(m) \tag{120}$$
$$n \leq m \iff f(n) \leq_{\mathbb{Z}} f(m) \tag{121}$$

Theorem 6.2 (Countability)

$\mathbb{Z}$ is countably infinite.

6.2 Rational Numbers

Just as we have done before, we construct the rationals $\mathbb{Q}$ as an ordered field that embeds $\mathbb{Q}$ through a canonical ordered ring homomorphism.

Definition 6.2 (Rationals)

Given the ordered ring of integers $(\mathbb{Z}, +_{\mathbb{Z}}, \times_{\mathbb{Z}}, \leq_{\mathbb{Z}})$ the **rational numbers** $(\mathbb{Q}, +_{\mathbb{Q}}, \times_{\mathbb{Q}})$ are defined as such.

1. $\mathbb{Q}$ is the quotient space on $\mathbb{Z} \times \mathbb{Z} \setminus \{0\}$ with the equivalence relation $\sim$
$$(a, b) \sim (c, d) \iff a \times_{\mathbb{Z}} d = b \times_{\mathbb{Z}} c \tag{122}$$
We denote this class as (a, b), where $b > 0$, since if $b < 0$, we know that $(-a, -b)$ are also in this order.
2. The additive and multiplicative identities are
$$0_{\mathbb{Q}} := (0_{\mathbb{Z}}, a), \quad 1_{\mathbb{Q}} := (a, a) \tag{123}$$
3. Addition on $\mathbb{Q}$ is defined
$$(a, b) +_{\mathbb{Q}} (c, d) := \big((a \times_{\mathbb{Z}} d) +_{\mathbb{Z}} (b \times_{\mathbb{Z}} c), b \times_{\mathbb{Z}} d\big) \tag{124}$$
4. The additive inverse is defined
$$-(a, b) := (-a, b) \tag{125}$$
5. Multiplication on $\mathbb{Q}$ is defined
$$(a, b) \times_{\mathbb{Q}} (c, d) := \big(a \times_{\mathbb{Z}} c, b \times_{\mathbb{Z}} d\big) \tag{126}$$
6. The multiplicative inverse is defined
$$(a, b)^{-1} := (b, a) \tag{127}$$
7. The order $\leq_{\mathbb{Q}}$ defined on the rationals as
$$(a, b) \leq_{\mathbb{Q}} (c, d) \iff ad \leq_{\mathbb{Z}} bc \tag{128}$$
is a total order.

Theorem 6.3 (Order on Rationals)

The order $\leq_{\mathbb{Q}}$ defined on the rationals as
$$(a, b) \leq_{\mathbb{Q}} (c, d) \iff ad \leq_{\mathbb{Z}} bc \tag{129}$$
is a total order. Remember that we have defined $b, d > 0$.

Proof. We prove the three properties.

1. Reflexive.
$$(a, b) \leq_{\mathbb{Q}} (a, b) \iff ab \leq_{\mathbb{Z}} ab \tag{130}$$
2. Antisymmetric.
$$(a, b) \leq_{\mathbb{Q}} (c, d) \implies ad \leq_{\mathbb{Z}} bc(c, d) \leq_{\mathbb{Q}} (a, b) \qquad \implies bc \leq_{\mathbb{Z}} ad \tag{131}$$
This implies that both $ad = bc$, which by definition means that they are in the same equivalence class.
3. Transitivity. Assume that $(a, b) \leq (c, d)$ and $(c, d) \leq (e, f)$. Then, we notice that $b, d, f > 0$ and

therefore by the ordered ring property[a] of $\mathbb{Z}$, we have

$$(a,b) \leq_{\mathbb{Q}} (c,d) \implies ad \leq_{\mathbb{Z}} bc \implies adf \leq_{\mathbb{Z}} bcf \tag{132}$$
$$(c,d) \leq_{\mathbb{Q}} (e,f) \implies cf \leq_{\mathbb{Z}} de \implies bcf \leq_{\mathbb{Z}} bde \tag{133}$$

Therefore from transitivity of the ordering on $\mathbb{Z}$ we have $adf \leq bde$. By the ordered ring property[b] we have $0 \leq bde - adf = d(be - af)$. But notice that $d > 0$ from our definition of rationals, and therefore it must be the case that $0 \leq be - af \implies af \leq_{\mathbb{Z}} be$, which by definition means $(a,b) \leq_{\mathbb{Q}} (e,f)$.

[a] If $a \leq b$ and $0 \leq c$, then $ac \leq bc$.
[b] If $a \leq b$, then $a + c \leq b + c$.

Theorem 6.4 (Rationals are a Field)

$\mathbb{Q}$ is a field.

Proof. We do a few things.

1. Verify the additive identity.

$$(a,b) + (0,c) = (ac + 0b, bc) = (ac, bc) \sim (a,b) \tag{134}$$

2. Verify the multiplicative identity.

$$(a,b) \times (c,c) = (ac, bc) \sim (a,b) \tag{135}$$

3. Additive inverse is actually an inverse.

$$(a,b) + (-a,b) = (ab + (-ba), bb) = (0, bb) \sim (0,1) \tag{136}$$

4. Multiplicative inverse is actually an inverse.

$$(a,b) \times (b,a) = (ab, ba) = (ab, ab) \sim (1,1) \tag{137}$$

5. Addition is commutative.

$$(a,b) + (c,d) = (ad + bc, bd) = (cb + ad, bd) = (c,d) + (a,b) \tag{138}$$

6. Addition is associative.

$$\begin{aligned} (a,b) + ((c,d) + (e,f)) &= (a,b) + (cf + de, df) && (139) \\ &= (adf + bcf + bde, bdf) && (140) \\ &= (ad + bc, bd) + (e,f) && (141) \\ &= ((a,b) + (c,d)) + (e,f) && (142) \end{aligned}$$

7. Multiplication is commutative.

$$(a,b) \times (c,d) = (ac, bd) = (ca, db) = (c,d) \times (a,b) \tag{143}$$

8. Multiplication is associative.

$$\begin{aligned} (a,b) \times ((c,d) \times (e,f)) &= (a,b) \times (ce, df) && (144) \\ &= (ace, bdf) && (145) \\ &= (ac, bd) \times (e,f) && (146) \\ &= ((a,b) \times (c,d)) \times (e,f) && (147) \end{aligned}$$

9. Multiplication distributes over addition.

$$\begin{aligned}(a, b) \times ((c, d) + (e, f)) &= (a, b) \times (c, d) + (a, b) \times (e, f) && (148)\\ &= (ac, bd) + (ae, bf) && (149)\\ &= (abcf + abde, b^2 df) && (150)\\ &= (acf + ade, bdf) \qquad = (a, b) \times (cf + de, df) && (151)\end{aligned}$$

Theorem 6.5 (Rationals are an Ordered Field)

$\mathbb{Q}$ is an ordered field.

Proof. We show that our defined order satisfies the definition.

1. Assume that $y = (a, b) \leq (c, d) = z$. Let $x = (e, f)$. Then $x+y = (af+be, bf)$, $x+z = (cf+de, df)$. Therefore

$$\begin{aligned}(af + be)df &= adf^2 + bedf && (152)\\ &\leq bcf^2 + bedf && (153)\\ &= (cf + de)bf && (154)\end{aligned}$$

But $(af+be)df = (cf+de)bf$ is equivalent to saying $(af+be, bf) \leq_{\mathbb{Q}} (cf+de, df)$, i.e. $x+y \leq x+z$!

2. Let $x = (a, b), y = (c, d)$. Since $0 < x, 0 < y$, by construction this means that $0 < a, 0 < c$ (since $b, d > 0$ in the canonical rational form). By the ordered ring property of the integers, $0 < ac$. So

$$0 < ac \iff 0 \cdot bd < ac \cdot 1 \iff (0, 1) < (ac, bd) \iff 0_{\mathbb{Q}} < (a, c) \times_{\mathbb{Q}} (b, d) = xy \qquad (155)$$

We have successfully defined the rationals, but now these are almost completely separate elements. We know that all integers are rational numbers, and so to show that the rationals are an extension of $\mathbb{Z}$ we want to identify a *canonical injection* $\iota : \mathbb{Z} \to \mathbb{Q}$. This can't just be any canonical injection; it must preserve the both the order and algebraic structure between the two sets and must therefore be a *ordered ring homomorphism.*

Theorem 6.6 (Canonical Injection of $\mathbb{Z}$ to $\mathbb{Q}$ is an Ordered Ring Homomorphism)

Let us define the canonical injection $\iota : \mathbb{Z} \to \mathbb{Q}$ to be $\iota(a) = (a, 1)$. This is a ring homomorphism.

Proof. We show a few things.

1. Preservation of addition.

$$\begin{aligned}\iota(a) +_{\mathbb{Q}} \iota(b) &= (a, 1) +_{\mathbb{Q}} (b, 1) && (156)\\ &= (1a +_{\mathbb{Z}} 1b, 1^2) && (157)\\ &= (a +_{\mathbb{Z}} b, 1) && (158)\\ &= \iota(a +_{\mathbb{Z}} b) && (159)\end{aligned}$$

2. Preservation of multiplication.

$$\begin{aligned}\iota(a) \times_{\mathbb{Q}} \iota(b) &= (a, 1) \times_{\mathbb{Q}} (b, 1) && (160)\\ &= (a \times_{\mathbb{Z}} b, 1^2) && (161)\\ &= (a \times_{\mathbb{Z}} b, 1) && (162)\\ &= \iota(a \times_{\mathbb{Z}} b, 1) && (163)\end{aligned}$$

3. Preservation of multiplicative identity.

$$\iota(1_{\mathbb{Z}}) = (1,1) = 1_{\mathbb{Q}} \tag{164}$$

4. Preservation of order.

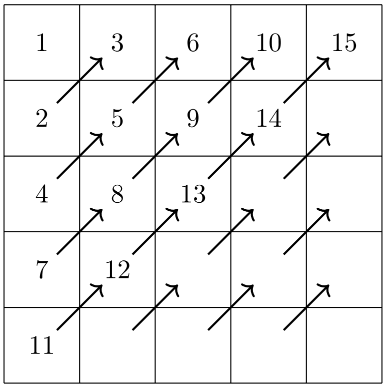
$$\begin{aligned} a \leq_{\mathbb{Z}} b &\iff a \cdot 1 \leq_{\mathbb{Z}} b \cdot 1 && (165)\\ &\iff (a,1) \leq_{\mathbb{Q}} (b,1) && (166)\\ &\iff \iota(a) \leq_{\mathbb{Q}} \iota(b) && (167) \end{aligned}$$

Theorem 6.7 (Rational Numbers)

$\mathbb{Q}$ is countable.

Proof. Since $\mathbb{N} \approx \mathbb{Z}$, it suffices to prove that $\mathbb{N} \times \mathbb{N}$ is countable. We wish to find the bijection $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$. We claim that

$$f(x,y) = \frac{1}{2}\big\{(x+y-1)^2 - (x+y-1) + 2\big\} + x - 1 \tag{168}$$



1	3	6	10	15
2	5	9	14	
4	8	13		
7	12			
11				

Figure 6: You can see that given (x,y) it is on the $(x+y+1)$th diagonal, which starts from the $\frac{1}{2}\big((x+y+1)^2 - (x+y+1) + 2\big)$th number and increments by $x-1$. Therefore, we have the formula above.

Theorem 6.8 (Finite Fields)

There are no finite ordered fields.

Proof. Assume $\mathbb{F}$ is such an ordered field. It must be the case that $0, 1 \in \mathbb{F}$, with $0 < 1$. Therefore, we also have $0 + 1 < 1 + 1 \implies 1 < 1 + 1$. Repeating this we get

$$0 < 1 < 1 + 1 < 1 + 1 + 1 < \ldots \tag{169}$$

where these elements must be distinct (since only one of $>, <, =$ must be true for a totally ordered set). Since this can be done for a countably infinite number of times, $\mathbb{F}$ cannot be finite.

6.2.1 Norm, Metric, and Topology on Rationals

Note that we can also define a norm on the rationals with just the order and algebraic properties.

Theorem 6.9 (Norm on $\mathbb{Q}$)

The following is indeed a norm on $\mathbb{Q}$.

$$|x| \coloneqq \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0 \end{cases} \tag{170}$$

It is well known that the metric induced by any norm is indeed a metric. Therefore we state the metric as a definition.

Definition 6.3 (Metric on $\mathbb{Q}$)

The Euclidean metric on $\mathbb{Q}$ is defined

$$d(x, y) \coloneqq |x - y| = \begin{cases} x - y & \text{if } x \geq y \\ y - x & \text{if } x < y \end{cases} \tag{171}$$

Thus we get to what we want: the induced topology of open balls. Again, since we know from point-set topology that metric topologies are indeed topologies, we will state this as a definition rather than a theorem.

Definition 6.4 (Open-Ball Topology on $\mathbb{Q}$)

The Euclidean topology on $\mathbb{Q}$ is the topology generated by the set $\mathscr{B}$ of all open balls

$$B(x, r) \coloneqq \{y \in \mathbb{Q} \mid |x - y| < r\} \tag{172}$$

Note that this is the same topology as the order topology. This should however be proved.

Theorem 6.10 (Metric and Order Topologies on $\mathbb{Q}$)

The metric and order topologies on $\mathbb{Q}$ are the same topologies.

Proof.

7 The Reals

8 Complex Numbers

9 Cardinal Numbers

Now we would like to rigorously construct the intuitive concept of the “size” of a set. Before we can even label any set with such a number, which we call the *cardinal number*, we can with our current tools compare the sizes, or *cardinalities*, of sets. This is a bit counterintuitive since we’re able to *compare* sizes but not know what the sizes actually are!

Definition 9.1 (Equipotence)

Two sets A and B are **equipotent**, written $A \approx B$, if there exists a bijective map $f : A \to B$. This implies that their cardinalities are the same: $|A| = |B|$. It has the following properties:

1. Reflexive: $A \approx A$
2. Symmetric: $A \approx B$ implies $B \approx A$
3. Transitive: $A \approx B$ and $B \approx C$ implies $A \approx C$

Now equipotence behaves like an equivalence relation, though we can’t define an equivalence class on the nonexistent set of all sets.

Theorem 9.1

The following hold for equipotence.

1. $A \approx A$.
2. $A \approx B \implies B \approx A$.
3. $A \approx B, B \approx C \implies A \approx C$.

Proof. Listed.

1. Take the identity map.
2. Take the inverse, which is well defined under bijection.
3. Take the composition of bijections which we proved is a bijection.

Definition 9.2 (Comparison of Cardinality)

The **cardinality** of A is said to be less than or equal to the cardinality of B, denoted $|A| \leq |B|$) if there is a one-to-one mapping of A into B.

Just like how equipotence behaves like an equivalence relation, comparison of cardinality behaves like an ordering on the collection of equivalence classes.

Theorem 9.2

The following hold.

1. $|A| \leq |A|$.
2. $|A| \leq |B|, |A| = |C| \implies |C| \leq |B|$
3. $|A| \leq |B|, |B| = |C| \implies |A| \leq |C|$.
4. $|A| \leq |B|, |B| \leq |C| \implies |A| \leq |C|$.

However, we still have to establish antisymmetry, which—unlike the other properties—is a major theorem.

Theorem 9.3 (Cantor-Bernstein)

If $|A| \leq |B|$ and $|B| \leq |A|$, then $|A| = |B|$.

Proof.

So far, we have proved many properties of cardinality without actually defining what cardinality is. We don't actually need to define such a thing, but for convenience and convention we do. The following is really a theorem, which can be proved by the axiom of choice, but we introduce it as a definition.

Definition 9.3 (Cardinal Numbers)

There exists sets called **cardinal numbers** with the property that for every set X, there is a unique cardinal $|X|$, called the *cardinality of* X, and sets X and Y are equipotent if and only if $|X|$ is equal to Y.[a]

[a] However, there does *not* exist a set containing all the cardinals!

We should intuitively see that the natural numbers can be treated as a subset of the cardinals, but it is not sufficient since by the axiom of infinity, there exists infinite sets X in which $|X| \not\in \mathbb{N}$. We will start off with finite set and continue onto finite sets.

9.1 Finite Sets

Definition 9.4 (Finite Set)

A set S is **finite** if it is equipotent to some natural number $n \in \mathbb{N}$.[a] We define $|S| = n$ and say S has n elements.

[a] Note that n is a set.

Definition 9.5 (Infinite Set)

A set S that is not finite is called **infinite**.

It follows that the cardinal numbers of finite sets are the natural numbers, and natural numbers are themselves finite sets, meaning that $|n| = n$ for all $n \in \mathbb{N}$. It remains to prove that $|S|$ for a finite set is unique.

Lemma 9.4 (No Proper Inclusion)

If $n \in \mathbb{N}$, then there is no bijective mapping of n onto a proper subset $X \subsetneq n$.

Proof. We do strong induction on n.

Corollary 9.5

The following is immediate.

1. If $n \neq m$, then there is no bijective mapping $f : n \to m$.
2. If $|S| = n$ and $|S| = m$, then $n = m$.
3. $\mathbb{N}$ is infinite.
4.

Theorem 9.6 (Preservation of Finite Cardinality)

Finiteness is preserved under many set operations and maps.

1. *Subset.* If X is finite and $Y \subset X$, then Y is finite and $|Y| \leq |X|$.
2. *Function.* If X is finite, and f is a function, then $f(X)$ is finite with $|f(X)| \leq |X|$.
3. *Binary Union.* If X and Y are finite, $X \cup Y$ is finite with $|X \cup Y| \leq |X| + |Y|$. If they are disjoint, then $|X \cup Y| = |X| + |Y|$.
4. *Arbitrary Union.* If a set of sets S is finite and every $X \in S$ is finite, then $\cup S$ is finite.
5. *Power Set.* If X is finite, then $\mathcal{P}(X)$ is finite.
6. *Cartesian Product.* If $X_1, \ldots, X_n$ is finite, then $\prod X_i$ is finite.

9.2 Countable Sets

Now we go into our first class of infinite sets, which we know are the naturals.

Definition 9.6 (Aleph Null)

The cardinal number of the naturals $\mathbb{N}$ is denoted $\aleph_0 = |\mathbb{N}|$.

Definition 9.7 (Countable Set)

A set S is **countable** if $|S| = \aleph_0 = |\mathbb{N}|$. A set is **at most countable** if $|S| \leq |\mathbb{N}|$.

At this point, we may already be familiar with the fact that $\mathbb{Q}$ is countable and $\mathbb{R}$ is uncountable. Let us formalize the statement that a countable infinity is the smallest type of infinity. We can show this by taking a countable set and showing that every infinite subset must be countable. If it was not countable (e.g. uncountable), then this would mean that a countable set contains some other class of infinite sets, which means that *that* infinite set would be "smaller" than $\aleph_0$.

Theorem 9.7

Every infinite subset of a countable set A is countable.

Now that we've established that $\aleph_0$ is the smallest infinity, let's try to see which set operations preserve this cardinality.

Theorem 9.8 (Countable Union is Countable)

An at most countable union of countable sets is countable.

Theorem 9.9 (Finite Product is Countable)

A finite Cartesian product of countable sets is countable.

Proof. By induction it suffices to prove that if X and Y are countable, then $X \times Y$ is countable. We can find an enumeration by going across the diagonals.

Theorem 9.10 (Further Properties)

1. The set of all finite sequences in countable X is countable.
2. The set of all finite subsets of a countable set is countable.
3. An equivalence relation on a countable set has at most countably many equivalence classes.

9.3 Cardinal Arithmetic

So far, the only sets we have to work with is the naturals, which are countable. To extend this to other infinities, we would want to use these cardinal numbers to hopefully create larger cardinals. Therefore, we need to define operations on cardinals. First, we need a lemma to prove that our definitions are consistent.

Lemma 9.11

If A, B, A', B' are sets such that $|A| = |A'|$ and $|B| = |B'|$, with $A \cap B = A' \cap B' = \emptyset$, then $|A \cup B| = |A' \cup B'|$.

Definition 9.8 (Addition of Cardinals)

Given cardinals κ, λ where $\kappa = |A|$ and $\lambda = |B|$ for some sets A, B satisfying $A \cap B = \emptyset$, we can define

$$\kappa + \lambda \coloneqq |A \cup B| \tag{173}$$

Now to define multiplication as the cardinality of Cartesian products, we also prove consistency.

Lemma 9.12

If A, B, A', B' are sets such that $|A| = |A'|$ and $|B| = |B'|$, then $|A \times B| = |A' \times B'|$.

Definition 9.9 (Multiplication of Cardinals)

Given cardinals κ, λ where $\kappa = |A|$ and $\lambda = |B|$ for some sets A, B, we can define

$$\kappa \cdot \lambda \coloneqq |A \times B| \tag{174}$$

These operations behave pretty nicely, as outlined below.

Theorem 9.13 (Commutativity, Associativity)

Given cardinals κ, λ, μ, we have

1. *Commutativity of Addition.* $\kappa + \lambda = \lambda + \kappa$.
2. *Associativity of Addition.* $\kappa + (\lambda + \mu) = (\kappa + \lambda) + \mu$.
3. *Commutativity of Multiplication.* $\kappa \cdot \lambda = \lambda \cdot \kappa$
4. *Associativity of Multiplication.* $\kappa \cdot (\lambda \cdot \mu) = (\kappa \cdot \lambda) \cdot \mu$.
5. *Distributivity.* $\kappa \cdot (\lambda + \mu) = \kappa \cdot \lambda + \kappa \cdot \mu$.

Theorem 9.14 (Inequalities)

We have

1. $\kappa \leq \kappa + \lambda$.

2. $\kappa_1 \leq \kappa_2, \lambda_1 \leq \lambda_2 \implies \kappa_1 + \lambda_1 \leq \kappa_2 + \lambda_2$.
3. $\kappa \leq \kappa \cdot \lambda$ if $\lambda > 0$.
4. $\kappa_1 \leq \kappa_2, \lambda_1 \leq \lambda_2 \implies \kappa_1 \cdot \lambda_1 \leq \kappa_2 \cdot \lambda_2$.
5. $\kappa + \kappa = 2 \cdot \kappa$.
6. $\kappa + \kappa \leq \kappa \cdot \kappa$ whenever $\kappa \geq 2$.

Now these don't actually help us create larger infinities, and to do this we define exponentiation of cardinal numbers.

Lemma 9.15

if $|A| = |A'|$ and $|B| = |B'|$, then $|A^B| = |(A')^{B'}|$.

Therefore the definition is consistent.

Definition 9.10 (Exponentiation of Cardinals)

Given cardinals κ, λ where $\kappa = |A|$ and $\lambda = |B|$ for some sets A, B, we can define

$$\kappa^\lambda = |A^B| \tag{175}$$

Theorem 9.16 (Properties of Exponentiation)

The following holds.

1. $\kappa \leq \kappa^\lambda$ if $\lambda > 0$.
2. $\lambda \leq \kappa^\lambda$ if $\kappa > 1$.
3. $\kappa_1 \leq \kappa_2, \lambda_1 \leq \lambda_2 \implies \kappa_1^{\lambda_1} \leq \kappa_2^{\lambda_2}$
4. $\kappa \cdot \kappa = \kappa^2$.
5. $\kappa^{\lambda+\mu} = \kappa^\lambda \cdot \kappa^\mu$.
6. $(\kappa^\lambda)^\mu = \kappa^{\lambda \cdot \mu}$.
7. $(\kappa \cdot \lambda)^\mu = \kappa^\mu \cdot \lambda^\mu$.

Now that we've established a useful collection of properties of cardinals, we are ready to extend beyond $\aleph_0$ and $2^{\aleph_0}$.

Theorem 9.17 (Cantor's Theorem)

Actually the first part is known as Cantor's theorem, but the second is also used often together.

1. For every set X, $|X| < |\mathcal{P}(X)|$.
2. For every set X, $|\mathcal{P}(X)| = 2^{|X|}$.[a]

[a] In terms of cardinal numbers, we have $\kappa < 2^\kappa$ for all cardinal κ.

Therefore, given any collection of sets, we can always find a set that has a strictly greater cardinality. Despite the existence of each cardinal number, you would be surprised to find that there exists *no* set containing all the cardinal numbers!

Theorem 9.18 (Set of Cardinal Numbers Does Not Exist)

The set of all cardinal numbers does not exist.

Proof. Assume that C was the set of all cardinals. Then $\cup C$ would be a cardinal exceeding all the cardinals in C, which is a contradiction.

9.4 Uncountable Sets

Definition 9.11 (Uncountable Set)

A set S is **uncountable** if it is infinite and not countable.

We know that $\mathbb{Q}$ is countable with cardinality $\aleph_0$ and $\mathbb{R}$ is uncountable with cardinality $2^{\aleph_0}$. The question of whether there exists an intermediate infinity κ such that $\aleph_0 < \kappa < 2^{\aleph_0}$ is still not fully resolved today.

Theorem 9.19 (Continuum Hypothesis)

There is no uncountable cardinal number κ such that $\kappa < 2^{\aleph_0}$.

Now, how do we prove that a set is uncountable? We can't really use the contrapositive of Theorem 9.7, since to prove that an arbitrary set A is uncountable, then we must find an infinite subset that is not countable. But now we must prove that this subset itself is not countable, too! Therefore, we can use this theorem.

Theorem 9.20

Given an arbitrary set A, if every countable subset B is a proper subset of A, then A is uncountable.

Proof. Assume that A is countable. Then A itself is a countable subset of A, but by the assumption, A should be a proper subset of A, which is absurd. Therefore, A is uncountable.

Theorem 9.21

The set of all functions $f : \mathbb{R} \to \mathbb{R}$ has cardinality $2^{2^{\aleph_0}} > 2^{\aleph_0}$.

10 Ordinal Numbers

We have introduced the natural numbers as an inductive set that contains the empty set and recursively adding each element through the sucessor function $S(x) = x \cup \{x\}$. This gives us the natural numbers $\mathbb{N}$, but there is no barrier to stopping, and so we can count beyond the natural numbers by imagining some infinite number ω and continuing the counting process into the transfinite.

$$\omega = \mathbb{N} = \{0, 1, 2, \ldots\} \tag{176}$$
$$S(\omega) = \omega \cup \{\omega\} = \{0, 1, 2 \ldots, \omega\} \tag{177}$$
$$S(S(\omega)) = S(\omega) \cup \{S(\omega)\} = \{0, 1, 2, \ldots, \omega, S(\omega)\} \ldots \qquad = \ldots \tag{178}$$

This is the motivation behind *ordinal numbers*, which are numbers that describe the order of some element in a set, analogous to how cardinal numbers were designed to describe the size of sets.

10.1 Axiom of Replacement

Axiom 10.1 (Axiom Schema of Replacement)

This axiom asserts that the image of a set under any definable function will fall inside a set.

10.2 Transfinite Induction and Recursion

The induction principle and recursion theorem are the main tools for proving theorems about the natural numbers and constructing functions with domain $\mathbb{N}$, respectively. We can generalize them to ordinal numbers.

10.3 Ordinal Arithmetic

11 Axiom of Choice

The axioms up to this point are pretty much undisputed and completes ZF set theory. Now that we've defined a function, let's quickly extend our previous definition of a Cartesian product into an arbitrary union of sets.

Definition 11.1 (Choice Function)

Given a set X of sets, a **choice function** of X is a function $f : X \to \cup X$, that assigns each $S \in X$ to one of its elements $f(S) \in S$.

Example 11.1

Given $X = \{\{1, 4, 7\}, \{9\}, \{2, 7\}\}$, one possible choice function is

$$f(\{1,4,7\}) = 4 \tag{179}$$
$$f(\{9\}) = 9 \tag{180}$$
$$f(\{2,7\}) = 2 \tag{181}$$

Definition 11.2 (Cartesian Product)

If $\{X_\alpha\}_{\alpha \in A}$ is an indexed family of sets, then their **Cartesian product** is defined as the set of choice functions of $\{X_\alpha\}_{\alpha \in A}$.

$$\prod_{\alpha \in A} X_\alpha \coloneqq \left\{ f : A \to \bigcup_{\alpha \in A} X_\alpha \;\middle|\; \forall \alpha \in A, f(\alpha) \in X_\alpha \right\} \tag{182}$$

Therefore, we have used the power set axiom to define a finite Cartesian product, to then define a function, to then define a general Cartesian product. But this detail is irrelevant later on. Note also that this definition of Cartesian product is not the same as that of the previous definition. The binary Cartesian product is defined as $(a, b) = \{\{a\}, \{a, b\}\}$ while this defines as a function $f : \{1, 2\} \to A, B$. But once we have overwritten the old definition (which is still necessary!) we can just forget about it and use this new definition of Cartesian product since there is a canonical bijection between them. It is a lot less annoying to think of ordered tuples as just tuples rather than as sets of sets.

However, in our definition, we just call this a "set of functions" and have never proved that it actually contains anything. But we can see obviously that if this Cartesian product is nonempty then there exists a choice function, and if there exists a choice function then the Cartesian product is nonempty. It would be ideal if we can prove one of the two conditions, but it turns out we can't, and therefore we introduce the final axiom, called the *axiom of choice*. Though controversial, it is required in the proofs of some notable theorems. If we include this axiom of choice, then we have ZFC set theory. The axiom of choice has many equivalent definitions.

Colloquially, the axiom of choice says that a Cartesian product of a collection[3] of non-empty sets is non-empty. That is, it is possible to construct a new set by choosing one element from each set, even if the collection is infinite.

[3] Note that this does not have to be finite

Axiom 11.1 (Axiom of Choice)

Let X be any set of nonempty sets. Then, AC states the equivalent things:

1. There exists a choice function on X.
2. The cartesian product $\prod X$ is nonempty.
3. There exists a set containing exactly one element from each set in the collection.

The controversy around AC is that is is nonconstructive by nature, and one cannot write down a specific formula or rule to define such a choice function. This is in contrast to—say, the axiom of infinity, since you can construct such a set inductively. For example, take a look at a choice function for the set of all subsets of the reals, which is considered as an "unruly" set.

Example 11.2 (Choice Function on Power Set of Reals)

Let I be the set of all nonempty subsets of $\mathbb{R}$, and $X_i = i \in I$. Then an element f in $\prod_{i \in I} X_i$ is a function which picks an element $f(T) \in T$ for every nonempty $T \subset \mathbb{R}$. How do you *define* such an f?

1. We might say, *pick the minimum element*, but subsets like $(0, 1)$ does not have a minimum.
2. We could write a rule that says *pick 1 if it is in the set, otherwise 2, otherwise 3, and so on*, but the choice function would not be defined for subsets that don't contain any natural number, such as $\{\pi, 2\pi, e\}$.

Therefore, there is no canonical choice of an element in a nonempty set of real numbers. But AC tells us that we don't have to worry about this. It gives us such a function, even if we cannot "write it down" (which means, construct it from the other ZF axioms). However, there are still sets which this is possible.

Example 11.3 (Choice Function on Power Set of Naturals)

Let I be the set of all nonempty subsets of $\mathbb{N}$. Then, we can define a choice function $f(T) = \min(T)$, which is always defined due to the well-ordering principle.

Example 11.4 (Choice Function on Open Sets of Reals)

If we let I be the set of all nonempty *open* subsets of $\mathbb{R}$, then there is a choice function. Choose any bijection $\tau : \mathbb{N} \to \mathbb{Q}$, and then assign to each nonempty open subset $U \subset \mathbb{R}$ the element $\tau(\min\{n \in \mathbb{N} \mid \tau(n) \in U\})$. This works since $U \cap \mathbb{Q} \neq \emptyset$, and by the well ordering principle, we are guaranteed a minimum element.

11.1 Well-Ordering and Zorn's Lemma

The examples indicate that we must try to find some representative element of every subset of a set. This motivates the definition, followed by two additional axioms that turn out to be equivalent to AC.

Definition 11.3 (Well-Ordered)

A set X is **well-ordered** by a strict total order $\leq$ if every nonempty subset of X has a least element under $\leq$.

Therefore, if the well-ordering theorem holds, then we can see that every subset of the reals has such a least element, and therefore we can construct a choice function, which supports AC. It turns out that the converse is true as well. If AC is true, we can see generally that we would like to use a choice function to select a representative element of each set in X. Then we can use these to construct an order.

Axiom 11.2 (Well-Ordering Theorem)

Every set can be well-ordered.

Despite the seeming equivalence between AC and the well-ordering theorem, the this result seems to be the most counterintuitive, since it claims the existence of such a total order on $\mathbb{R}$ such that *every* nonempty subset of $\mathbb{R}$ has a minimum! Nobody has been able to explicitly construct such an ordering for the reals, and at first glance, perhaps one may try to *prove* that such a well-ordering cannot exist. Let's move onto the second axiom.

Axiom 11.3 (Zorn's Lemma)

Let X be a partially ordered set that satisfies the two properties.

1. P is nonempty.
2. Every **chain** (a subset $A \subset P$ where A is totally ordered) has an upper bound in P.

Then P has at least one maximal element.

The validity of Zorn's lemma is a bit ambiguous, which motivates the following quote from Jerry Bona: *The axiom of choice is obviously true, the well-ordering principle obviously false, and who can tell about Zorn's lemma?* Ironically, all three results turn out to be equivalent.

Theorem 11.1 (Equivalence)

The following are equivalent.

1. Axiom of Choice.
2. Well-Ordering Theorem.
3. Zorn's Lemma.

Proof.

11.2 Banach-Tarski Paradox

Most of my notes on this paradox is from Wahlberg's set of notes.[4] It helped me to not worry as much about the axiom of choice, as well as worry more about the axiom of infinity. Let G_3 be the group of all 3-dimensional rigid transformations $x \mapsto Ax + b$, where $x \in \mathbb{R}^3, A \in \mathrm{SO}(3), b \in \mathbb{R}^3$.

Definition 11.4 (Equidecomposability)

Let G be a group acting on set X. We say that $A, B \subset X$ are **G-equidecomposable**, written $A \sim_G B$, if both sets have a decomposition

$$A = A_1 \cup \ldots \cup A_n, \quad B = B_1 \cup \ldots \cup B_n \tag{183}$$

and $A_i = g_i B_i$ for some $g_i \in G$. We claim that $\sim_G$ is an equivalence relation on 2^X.

Proof. Listed.

1. *Reflexive.* Clearly, $A = eA$ where $e \in G$ is the identity transformation.
2. *Symmetric.* If $A \sim_G B$, then we see that $A_i = g_i B_i$, but this means that $B_i = g_i^{-1} A_i$ for $g_i^{-1} \in G$, so $B \sim_G A$.

[4] https://arxiv.org/pdf/2206.13512.

3. *Transitive.* If $A \sim B$, $B \sim C$, then from $A \sim B$, we have

$$A = A_1 \cup \ldots \cup A_n, \quad B = B_1 \cup \ldots \cup B_n, \qquad A_i = f_i B_i \text{ for } f_i \in G \tag{184}$$

From $B \sim C$, we have

$$B = B'_1 \cup \ldots \cup B'_m, \quad C = C_1 \cup \ldots \cup C_m, \qquad B'_j = g_j C_j \text{ for } g_j \in G \tag{185}$$

Now we can take the common partition, which can have at most $n \cdot m$ partitions.

$$A = \bigcup_{i,j} \big(A_i \cap f_i B'_j\big), \quad C = \bigcup_{i,j} \big(g_j^{-1} B_i \cap C_j\big) \tag{186}$$

and see that

$$C \mapsto \bigcup_j g_j \bigg(\bigcup_i \big(g_j^{-1} B_i \cap C_j\big) \bigg) = \bigcup_j \bigcup_i \big(B_i \cap g_j C_j\big) = \bigcup_j \bigcup_i \big(B_i \cap B'_j\big) \tag{187}$$

$$\mapsto \bigcup_i f_i \bigg(\bigcup_j \big(B_i \cap B'_j\big) \bigg) = \bigcup_i \bigcup_j \big(f_i B_i \cap f_i B'_j\big) = \bigcup_{ij} \big(A_i \cap f_i B'_j\big) = A \tag{188}$$

So equidecomposablity is stronger than a bijection, but not as strong as an isometry. It's in between, like a piecewise rigid transformation.

Definition 11.5 (Paradoxical Sets)

Let G be a group acting on set X, and let $E \subset X$ be nonempty. Then, E is G**-paradoxical** if

$$E = A \sqcup B, \qquad E \sim_G A, \quad E \sim_G B \tag{189}$$

Note that this essentially means that E can be duplicated since

$$E \sim A \cup B \sim g_1 A \cup g_2 B \sim g_1 E \cup g_2 E \tag{190}$$

Lemma 11.2 (Equidecomposable Sets Share Paradoxicality)

Let $A \sim_G B$. Then, A is G-paradoxical iff B is G-paradoxical.

Proof.

Theorem 11.3 (The Banach-Schröder-Bernstein Theorem)

Let G be a group acting on set X and $A, B \subset X$. If A is G-equidecomposable with a subset of B and B is G-equidecomposable with a subset of A, then $A \sim_G B$.

Corollary 11.4 (Conditions for Paradoxical)

Let G be a group acting on set X. Then, $A \subset X$ is G-paradoxical if it contains disjoint subsets $A_1, A_2 \subset A$ both equidecomposable with A.

Example 11.5 (Vitali Paradox)

Let SO(2) be the group of rotations in $\mathbb{R}^2$ and S^1 be the unit circle. For $p_1, p_2 \in S^1$, let $p_1 \sim p_2$ if the angle of rotation between them is a rational multiple of 2π, which is an equivalence relation. Let us invoke the axiom of choice to define the choice set C where each element contains a representative element of each equivalence class. Then, each point in S^1 can be expressed as an element of C, rotated by some rational $q \in [0, 1)$. By enumerating the rationals in the unit interval (q_n), we get

$$S^1 = q_1 C \sqcup q_2 C \sqcup q_3 C \sqcup \ldots = C_1 \sqcup C_2 \sqcup C_3 \sqcup \ldots \tag{191}$$

We can end up recreating S^1 be using only the sets of even or odd indices by applying a suitable rotation to them.

$$S^1 = C_1 \sqcup \underbrace{C_3 + (q_2 - q_3)}_{C_2} \sqcup \underbrace{C_5 + (q_3 - q_5)}_{C_3} \sqcup \ldots \tag{192}$$

$$S^1 = \underbrace{C_2 + (q_1 - q_2)}_{C_1} \sqcup \underbrace{C_3 + (q_2 - q_3)}_{C_2} \sqcup \underbrace{C_5 + (q_3 - q_5)}_{C_3} \sqcup \ldots \tag{193}$$

Therefore, S^1 has a decomposition into two subsets such that each of them is "*countably* SO(2)-equidecomposable" with S^1, indicating that S^1 is "countably SO(2)-paradoxical."

We say that a set S can generate a group G.

Definition 11.6 (Free Group)

Let G be the group generated by S. Then, G is **free** if it satisfies the following equivalent definitions.

1. No nonempty reduced word in S represents the identity element in G.
2. Every element in G can be represented by exactly one reduced word of S.

The number of elements of S—called the generators—is the **rank** of the free group.

Example 11.6 (Free Group of Rank 1)

Let $S = \{1\}$. Then, it generates the group $(\mathbb{Z}, +)$. Similarly, we can think of an irrational rotation in S^1, which will also give us a free generator.

Let's extend this by one more dimension.

Lemma 11.5 (Rank 2 Free Group is Paradoxical)

If a free group G is of rank 2, then it is G-paradoxical, where we view G as acting on itself.

Proof. Let G be freely generated by $S = \{\rho, \tau\}$, and for each $g \in \{\rho, \tau, \rho^{-1}, \tau^{-1}\}$, define G_g as the set of all elements from G represented by reduced words in S having the leftmost letter as g. Since G is a free group, we can partition G as

$$G = \{e\} \sqcup G_\rho \sqcup G_\tau \sqcup G_{\rho^{-1}} \sqcup G_{\tau^{-1}} \tag{194}$$

Note that by separating out the first letter, we can decompose $G_\rho = \{\rho\} \sqcup \rho G_\rho \sqcup \rho G_\tau \sqcup \rho G_{\tau^{-1}}$. By transforming all elements by ρ^{-1}, we can "remove" the element as $\rho^{-1} G_\rho = \{e\} \sqcup G_\rho \sqcup G_\tau \sqcup G_{\tau^{-1}}$. Therefore, $G = \rho^{-1} G_\rho \sqcup G_{\rho^{-1}}$, and so

$$G \sim_G G_\rho \sqcup G_{\rho^{-1}} \tag{195}$$

Similarly, we have $G \sim_G G_\tau \sqcup G_{\tau^{-1}}$. Since G is G-equidecomposable with its two disjoint subsets $G_\rho \sqcup G_{\rho^{-1}}$ and $G_\tau \sqcup G_{\tau^{-1}}$, it follows that G is G-paradoxical.

Theorem 11.6 (Free Subgroup in SO(3))

SO(3) has a subgroup that is free on two generators.

Proof. The general idea is to take motivation from the irrational angles as free generators. We pick the following rotations around the x and y axes as our free generators. For convenience, we also list their inverses.

$$\sigma = \frac{1}{5}\begin{bmatrix} 5 & 0 & 0 \\ 0 & 4 & -3 \\ 0 & 3 & 4 \end{bmatrix}, \quad \tau = \frac{1}{5}\begin{bmatrix} 4 & 0 & 3 \\ 0 & 5 & 0 \\ -3 & 0 & 4 \end{bmatrix}, \quad \sigma^{-1} = \frac{1}{5}\begin{bmatrix} 5 & 0 & 0 \\ 0 & 4 & 3 \\ 0 & -3 & 4 \end{bmatrix}, \quad \tau^{-1} = \frac{1}{5}\begin{bmatrix} 4 & 0 & -3 \\ 0 & 5 & 0 \\ 3 & 0 & 4 \end{bmatrix} \tag{196}$$

We claim that the subgroup $G = \langle \sigma, \tau \rangle$ is free. Since all 4 matrices can be treated as $\mathbb{Q}$-linear maps, we define the integer matrices $A_+ = 5\sigma$ and $B_+ = 5\tau$, along with their corresponding matrices for the inverses $A_- = 5\sigma^{-1}$ and $B_- = 5\tau^{-1}$:

$$A_+ = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 4 & -3 \\ 0 & 3 & 4 \end{bmatrix}, A_- = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 4 & 3 \\ 0 & -3 & 4 \end{bmatrix}, \quad B_+ = \begin{bmatrix} 4 & 0 & 3 \\ 0 & 5 & 0 \\ -3 & 0 & 4 \end{bmatrix}, B_- = \begin{bmatrix} 4 & 0 & -3 \\ 0 & 5 & 0 \\ 3 & 0 & 4 \end{bmatrix} \tag{197}$$

Let $\rho = s_k \cdots s_1$ be any non-trivial reduced word of length $k \geq 1$, and we wish to show that ρ cannot be the identity map. Let $M_i = 5s_i$ be the integer matrix corresponding to each generator. If $\rho = I$, then the product $M = M_k \cdots M_1$ must equal $5^k I$. We now consider the images of these matrices under the ring homomorphism $\pi : \mathbb{Z} \to \mathbb{Z}_5$. We define $\bar{A}_+, \bar{A}_-, \bar{B}_+, \bar{B}_- \in M_3(\mathbb{Z}_5)$ as:

$$\bar{A}_+ = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 4 & 2 \\ 0 & 3 & 4 \end{bmatrix}, \bar{A}_- = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 4 & 3 \\ 0 & 2 & 4 \end{bmatrix}, \quad \bar{B}_+ = \begin{bmatrix} 4 & 0 & 3 \\ 0 & 0 & 0 \\ 2 & 0 & 4 \end{bmatrix}, \bar{B}_- = \begin{bmatrix} 4 & 0 & 2 \\ 0 & 0 & 0 \\ 3 & 0 & 4 \end{bmatrix} \tag{198}$$

If $\rho = I$, then $\pi(M) = \pi(M_k) \cdots \pi(M_1) \equiv \mathbf{0} \pmod 5$. To show this is impossible, we will investigate the range and kernel of each map. With some calculations[a], we find that

$$\text{range}(\bar{A}_+) = \{(0, 3m, m) \in \mathbb{Z}_5^3\}, \quad \ker(\bar{A}_+) = \{(m, n, 3n) \in \mathbb{Z}_5^3\}, \tag{199}$$

$$\text{range}(\bar{A}_-) = \{(0, m, 3m) \in \mathbb{Z}_5^3\}, \quad \ker(\bar{A}_-) = \{(m, 3n, n) \in \mathbb{Z}_5^3\}, \tag{200}$$

$$\text{range}(\bar{B}_+) = \{(m, 0, 3m) \in \mathbb{Z}_5^3\}, \quad \ker(\bar{B}_+) = \{(3m, n, m) \in \mathbb{Z}_5^3\}, \tag{201}$$

$$\text{range}(\bar{B}_-) = \{(3m, 0, m) \in \mathbb{Z}_5^3\}, \quad \ker(\bar{B}_-) = \{(m, n, 3m) \in \mathbb{Z}_5^3\}. \tag{202}$$

Note that $\ker(\bar{A}_\pm) \cap \text{range}(\bar{B}_\pm) = \{0\}$ and $\ker(\bar{B}_\pm) \cap \text{range}(\bar{A}_\pm) = \{0\}$. Therefore, there are no overlappings of kernels and ranges for pairs of mappings that are not inverses of each other. Since M is given by a composition of mappings where none of $\bar{A}_+\bar{A}_-$, $\bar{A}_-\bar{A}_+$, $\bar{B}_+\bar{B}_-$, or $\bar{B}_-\bar{B}_+$ can occur, the only common element between the kernel and range of each mapping in the composition is the zero vector. Therefore, if a vector is not in the kernel of the first mapping, it gets mapped to a vector that is not in the kernel of the next, and continuing for all mappings gives a nonzero vector. One such vector is $(0, 0, 1)$, which can be verified.

[a] For example, $\bar{A}_+$ maps (a, b, c) to $(0, 4b + 2c, 3b + 4c)$. But in $\mathbb{Z}_5$, we have that $4b + 2c = 3(3b + 4c)$.

Theorem 11.7 (Lift of Paradoxical Decomposition of Group Action to Set)

Let G be a G-paradoxical group acting on set X. If only the identity in G has any fixed points[a] in X, then X is also G-paradoxical.

[a] An element $g \in G$ has a fixed in X if there exists some $x \in X$ s.t. $gx = x$.

Proof. Consider the orbits (with each orbit being a set $\{gx \in X \mid g \in G\}$) of the elements in X. By invoking the axiom of choice, we obtain a set C, containing one element from each unique orbit. Therefore, by construction applying all group actions to all elements of C results in the whole set.

$$GC \coloneqq \{gc \mid g \in G, c \in C\} = \{Gc \mid c \in C\} = X \tag{203}$$

Since G is G-paradoxical, there exists a paradoxical decomposition $G = A \cup B$ with $G \sim_G A$, $G \sim_G B$. Just as $X = GC$, we can apply the group actions of A and B on C to obtain a partition of X by $Y \coloneqq AC$ and $Z \coloneqq BC$. We claim that Y and Z are disjoint. Assume that there is a common element, i.e. there exists $a \in A, b \in B$, and $c_1, c_2 \in C$ such that $ac_1 = bc_2$. But then, $c_1 = a^{-1}bc_2$, so both c_1, c_2 belong to the same orbit. Since A, B are disjoint, $a \neq b$ and so $a^{-1}b$ cannot be the identity, and hence cannot have any fixed points. Thus, $c_1 \neq c_2$, and C contains more than one element from an orbit, which is a contradiction on our construction of C.

Now we wish to show that Y, Z form a paradoxical decomposition of X. It only remains to check that $X \sim_G Y$ and $X \sim_G Z$. Let's do $X \sim_G Y$ first. Since $G \sim_G A$, by definition, there exists decompositions

$$G = \bigcup_{i=1}^n G_i, \quad A = \bigcup_{i=1}^n A_i \tag{204}$$

with $g_1, \ldots, g_n \in G$, such that $A_i = g_i G_i$. Now, from our previous construction, we can define the decompositions

$$X = GC = \bigg(\bigcup_{i=1}^n G_i\bigg) C = \bigcup_{i=1}^n (G_i C) = \bigcup_{i=1}^n X_i \tag{205}$$

$$Y = AC = \bigg(\bigcup_{i=1}^n A_i\bigg) C = \bigcup_{i=1}^n (A_i C) = \bigcup_{i=1}^n Y_i \tag{206}$$

and now, we can easily verify the paradoxicality by showing that $Y_i = g_i X_i$ for all $i = 1, 2, \ldots, n$. For each i, since $A_i = g_i G_i$, we have

$$\begin{aligned} Y_i = A_i C = (g_i G_i) C &= \{hc \mid h \in g_i G_i, c \in C\} && (207) \\ &= \{g_i g' c \mid g' \in G_i, c \in C\} && (208) \\ &= \{g_i x \mid x \in G_i C\} = g_i (G_i C) = g_i X_i && (209) \end{aligned}$$

So $X \sim_G Y$, and similarly, we can verify that $X \sim_G Z$, establishing paradoxicality of X.

Note that the full strength of AC is needed since the number of orbits is uncountably infinite.

Theorem 11.8 (Simplified Hausdorff Paradox)

Let S^2 be the unit sphere in $\mathbb{R}^3$. There is a countable subset $D \subset S^2$ such that $S^2 \setminus D$ is SO(3)-paradoxical.

Proof. Let's extract a free group G of rank 2 from SO(3), which we can view as acting on S^2. We know that G is G-paradoxical, and our goal is use the theorem above to lift the paradoxicality of G onto the G-set S^2.

The problem is that S^2 contains fixed points for elements of G that are not the identity. We wish to construct D such that $S^2 \setminus D$ doesn't contain these fixed points *and* still remains closed under G. This is pretty straightforward. Since every rotation in SO(3) is a rotation around some line through the origin[a], every element of G has two fixed points. But G is countable, so the set of all fixed points—which we denote as D—must also be countable.

Now it remains to show that $S^2 \setminus D$ is closed under G. Assume that it wasn't, i.e. there exists a $p \in S^2 \setminus D$ and $\rho \in G$ such that $\rho p \in D$. But by definition of D, ρp must be a fixed point for some other nontrivial rotation $\gamma \in G$. That is,

$$\gamma \rho p = \rho p \tag{210}$$

Now applying the inverse rotation ρ^{-1} to both sides gives $\rho^{-1}\gamma\rho p = p$, which indicates that $\rho^{-1}\gamma\rho$ has p as a fixed point. But p is not in D, and so p cannot be a fixed point for any rotation except the identity, indicating that $\rho^{-1}\gamma\rho = e$, which implies that $\gamma = \rho\rho^{-1} = e$. This contradicts that γ is nontrivial. Therefore, $S^2 \setminus D$ is G-paradoxical, and since G is a subset of SO(3), $S^2 \setminus D$ is also SO(3)-paradoxical.

[a] See Euler's rotation theorem, or consider the fact that such a matrix must have 1 real eigenvalue (which must be 1) and 2 complex conjugate eigenvalues.

The next result formalize the phenomena on how you can add or remove points simply by rotating a set. That is, if we have an irrational rotation relative to a countable set D, then the rotated sets $\rho^n D$ will never overlap.

Lemma 11.9 (Irrational Rotation Relative to Points)

Let $D \subset \mathbb{R}^3$ be at most countably infinite. If there is a line ℓ not passing through any point in D, then there exists a rotation $\rho \in G_3$ with ℓ as its axis, such that for $A \coloneqq \cup_{n=0}^{\infty} \rho^n D$, we have

$$\rho A = A \setminus D \tag{211}$$

If ℓ passes through the origin, then $\rho \in \mathrm{SO}(3)$.

Proof. We focus on the case when ℓ passes through the origin, since its extension to G_3 is straightforward. Denote the rotation of angle $\theta \in (0, 2\pi]$ by ρ_θ around ℓ. Since D is at most countable, it follows that only a countable number of angles can give $\rho_\theta D \cap D \neq \emptyset$. Similarly, for each $n \in \mathbb{N}$, the set of angles

$$\{\theta \in (0, 2\pi] \colon \rho_\theta^n D \cap D \neq \emptyset\} \tag{212}$$

is at most countable. Since we can choose θ from an uncountable set, it turns out that we can choose such a $\rho = \rho_\theta$ such that $\{\rho^n D\}_{n=1}^{\infty}$ are disjoint. Let $A \coloneqq \cup_{n=0}^{\infty} \rho^n D$. By disjointness, we have

$$\rho A = \rho\bigg(\bigcup_{n=0}^{\infty} \rho^n\bigg) = \bigcup_{n=0}^{\infty} \rho^{n+1} D = \bigcup_{n=1}^{\infty} \rho^n D = \bigg(\bigcup_{n=0}^{\infty} \rho^n D\bigg) \setminus D = A \setminus D \tag{213}$$

Then the following theorem effectively allows us to add certain points to a paradoxical set to obtain a larger set that remains paradoxical, sort of like a Hilbert's Hotel logic to geometry. It essentially says that if you have a shape X and you remove a small (countable) set of "troublesome" points D, the remaining set $X \setminus D$ is still "the same size" (equidecomposable) as the original X.

Theorem 11.10 (Equidecomposability by Irrational Rotation)

Let $X \subset \mathbb{R}^3$ with $D \subset X$ at most countably infinite. If there is a line ℓ not passing through any point in D, and D remains in X after any rotation around ℓ, then $X \setminus D$ is G_3-equidecomposable with X. If ℓ passes through the origin, then $X \setminus D$ is also SO(3)-equidecomposable with X.

Proof. By the lemma, there exists a rotation ρ around ℓ such that for all rotated points $A \coloneqq \cup_{n=0}^{\infty} \rho^n D$, we have $\rho A = A \setminus D$. Since $\rho^n D \subset X$ for all $n \in \mathbb{N}$, it follows that $A \subset X$.
Now let $B \coloneqq X \setminus A$ be the rest of the points we are not shuffling. Now we can see

$$B \cup \rho A = (X \setminus A) \cup (A \setminus D) = X \setminus D \tag{214}$$

But $X = (B \cup A) \sim (B \cup \rho A)$, so by transitivity, it follows that $X \sim X \setminus D$.

Theorem 11.11 (Paradoxical Sphere)

S^2 is $SO(3)$-paradoxical.

Proof. According to the Hausdorff paradox, we can find a countable subset $D \subset S^2$ such that $S^2 \setminus D$ is SO(3)-paradoxical. Now, select a line ℓ passing through the origin but not intersecting any point in D.[a] Since rotations ρ of D around ℓ keeps ρD inside S^2. Therefore, by the previous theorem, we know that

$$S^2 \sim S^2 \setminus D \tag{215}$$

Since $S^2 \setminus D$ is SO(3)-paradoxical, it follows that S^2 must also be SO(3)-paradoxical.

[a] This is always possible since D is a countable subset of uncountable set S^2.

Now, we want to turn the paradoxicality of the sphere into paradoxicality of the entire ball. This is pretty straightforward since we just consider all the points "under" the set on the surface.

Theorem 11.12 (Paradoxical Punctured Ball)

Let B^3 be the closed unit ball in $\mathbb{R}^3$. Then, $B^3 \setminus \{0\}$ is SO(3)-paradoxical.

Proof. For each $E \subset S^2$, define

$$c(E) \coloneqq \{tx \mid x \in E, t \in (0, 1]\} \tag{216}$$

Note that $c(S^2) = B^3 \setminus \{0\}$, and that $A, B \subset E$ disjoint implies that $c(A) \cap c(B)$ are also disjoint. Since S^2 is paradoxical, let $S^2 = C \sqcup D$ be its paradoxical decomposition. Note that this makes $c(S^2) = s(C \sqcup D) = s(C) \sqcup s(D)$ a decomposition of $B^3 \setminus \{0\}$.
Since $S^2 \sim C$, there exists a decomposition $S^2 = E_1 \sqcup \ldots \sqcup E_n$ and $C = F_1 \sqcup \ldots \sqcup F_n$ with $g_1, \ldots, g_n \in$ SO(3) such that $F_i = g_i E_i$. This defines a bijective function $f : S^2 \to C$ that is piecewise rotational. We can extend f to a mapping $g : c(S^2) \to c(C)$ by defining

$$g(x) \coloneqq \|x\| \cdot f(x/\|x\|) \tag{217}$$

which can be checked to be indeed an extension . It is surjective since

$$g(c(S^2)) = \{tf(x) \mid x \in S^2, 0 < t \leq 1\} = \{ty \mid y \in C, 0 < t \leq 1\} = c(C) \tag{218}$$

and it is injective since if $g(x) = g(y)$, then $\|x\| = \|y\|$ and $f(x/\|x\|) = f(y/\|x\|)$. This implies that $x/\|x\| = y/\|x\| \implies x = y$. Therefore,

$$B^3 \setminus \{0\} = c(S^2) \sim c(C) \tag{219}$$

Similarly, we can show that $(B^3 \setminus \{0\}) \sim c(D)$. Therefore, $B^3 \setminus \{0\}$ is equidecomposable with both $c(C)$ and $c(D)$, which partitions it, and so $B^3 \setminus \{0\}$ is SO(3)-paradoxical.

Using the same steps we did in proving equidecomposablity by irrational rotations, the origin point can also be absorbed through equidecomposablity.

Corollary 11.13 (Banach-Tarski Paradox)

B^3 is G_3-paradoxical.

Proof. We basically want to define some at most countable set $D \subset B^3$ such that we can invoke our previous theorem on equidecomposability. We just proved that $B^3 \setminus \{0\}$ is SO(3)-paradoxical. Let ℓ be a line that passes within distance 1/2 from 0 but not intersecting it (e.g. intersecting at $(1/3, 0, 0)$). Let $D = \{0\}$. Then, for any rotation in G_3[a] around ℓ, 0 will stay within B^3. Therefore, now we can invoke our theorem to give us our result that

$$B^3 \sim_{G_3} B^3 \setminus \{0\} \tag{220}$$

[a] Note that these are rotations in G_3, not SO(3), since ℓ does not pass through the origin.

Theorem 11.14 (Strong Banach-Tarski Paradox)

Let $A, B \subset \mathbb{R}^3$ be two bounded sets with nonempty interiors. Then, A and B are G_3-equidecomposable.

11.3 Predicting Random Real Numbers

Another weird result is the following riddle.

There is a house with 100 rooms, and each room contains countably many boxes indexed with the natural numbers. Each box contains a random real number, which is the same over all the rooms (that is, box n contains the same real number in every room). 100 set theorists play a game. Each person will go into a unique room and open as many boxes as they like (perhaps countably many) as long as they leave at least one box in their room unopened. Then, each of them need to pick an unopened box in their room, and guess what real number is inside of it.

In order to win, 99 of them need to guess correctly. The mathematicians can discuss a strategy beforehand, but after they go into their respective rooms, no more communication is allowed. What is a winning strategy for this seemingly impossible task?

At first glance, a solution seems to be impossible, but here is one that uses AC.

1. Let S be the set of all sequences of real numbers $(x_i \in \mathbb{R}) \in S$. Let $\sim$ be an equivalence relation on S where $(x_i) \sim (y_i)$ if the sequences differ on finitely many terms. Using axiom of choice, the mathematicians agree on a representative sequence from each equivalence class of S.

2. The mathematicians go into their respective rooms. For $1 \leq n \leq 100$, let $(s_i^{(n)})_i$ denote the sequence of reals in the contents of the boxes as

$$s^{(n)} = n, 100 + n, 200 + n, 300 + n, \ldots \tag{221}$$

   Now, let player n open every box except for those in $s^{(n)}$. In this way, player p opens up 99 sequences of boxes $s^{(n)}$ for $n \neq p$.

3. Each player looks at the 99 sequences they can see, identifies which equivalence class each of them belongs to, and recalls the chosen representative sequence for every one of them.

4. For each player, they compare each of the 99 sequences $s^{(n)}$ with their respective representative sequence, and writes down the greatest index at which each observed sequence does not agree with its corresponding representative (which exists by the definition of $\sim$). Therefore, player p will have written down 99 integers $\{x_n\}_{n \neq p, 1 \leq n \leq 100}$. Note that between any two players, the 98 of the integers will overlap since they are looking at the same sequences.

5. Now, each player p takes the maximum $X_p \coloneqq \max\{x_n\}_{n \neq p, 1 \leq n \leq 100}$. Out of the remaining boxes in player p's room (which are precisely the indices that are congruent to p mod 100), they leave the first $X_p + 1$ of them closed but opens every box beyond that. Since they opened up all but a finite number of boxes, each player p can determine the equivalence class of $s^{(p)}$ and furthermore recall the representative sequence $[s^{(p)}]$.

6. For each player, they choose the highest index unopened box in their room—which is the $(X_p + 1)$th box in $s^{(p)}$ and guess it according to the corresponding element in $[s^{(p)}]$.

We claim that X_p agrees for at least 99 of $p = 1, \ldots 100$. Let the maximum of all 100 x_n's be denoted x_{n^*}. If n^* is unique, for $p \neq n^*$, $X_p = n^*$, and for $p = n^*$, X_p will be the next greatest number. If it is not unique, then $X_p = n^*$ for all 100 p's. Therefore, for at least 99 of the players p,

$$X_p \coloneqq \max\{x_n\}_{n \neq p, 1 \leq n \leq 100} = \max\{x_n\}_{1 \leq n \leq 100} \tag{222}$$

Remember that each x_n represents the greatest index at which $s^{(n)}$ doesn't agree with $[s^{(n)}]$, and X_p is the true maximum for at least 99 of them. So for 99 players, when they pick the $(X_p + 1)$th index, by definition, $s^{(p)}$ and $[s^{(p)}]$ must be guaranteed to match, since $X_p + 1 > X_p \geq x_p$.

11.4 Nonmeasurable Sets

11.5 Countable Choice

11.6 Dependent Choice

12 Exercises

Exercise 12.1 (Math 531 Spring 2025, PS2.6)

Assume that S is a set with exactly n elements. Assume that $T : S \to S$. Prove that there exists some $x \in S$ so that

$$T^j(x) = x, \tag{223}$$

for some $j \in \{1, 2, ..., n\}$. Here T^j means the composition of T with itself j-times.

Solution. Assume that the statement is false, and there exists no such $x \in S$. Let's choose any $x \in S$ and write out

$$x = T^0(x), T^1(x), T^2(x), \ldots, T^n(x) \tag{224}$$

These are $n + 1$ elements living in a space S of size n, so by pigeonhole principle there exists a repeat. Let us choose any of these repeats and label them $0 \leq i < j \leq n$ s.t. $T^i(x) = T^j(x)$. It cannot be the case that $i = 0$ since we assumed that it was false. Therefore, it must be the case that $1 \leq i < j \leq n \implies j - i \leq n - 1$. Consider the sequence

$$y = T^0(y) = T^i(x), T^1(y) = T^{i+1}(x), \ldots, T^n(y) \tag{225}$$

Starting from $y = T^i(x) \in S$. Since $0 < j - i \leq n - 1$, we know that $T^{j-i}(y) = T^j(x)$ lies in this sequence. Since both $y = T^0(y) = T^i(x)$ and $T^{j-i}(y) = T^j(x)$ are equal and present, we have shown an instance of when this claim is true, and the statement is true.

Exercise 12.2 (Shifrin Appendix A.3.1)

Consider the following relations on the set $\mathbb{Z}$:

(i) $(a, b) \in R_1$ if $ab \geq 0$
(ii) $(a, b) \in R_2$ if $ab > 0$
(iii) $(a, b) \in R_3$ if $ab > 0$ or $a = b = 0$

Decide whether each is an equivalence relation. (If not, which requirements fail?)

Solution. Listed.

1. Not a relation since transitivity fails. $4 \sim 0$ since $4 \cdot 0 = 0 \geq 0$. $0 \sim -2$ since $0 \cdot -2 = 0 \geq 0$. But $4 \not\sim -2$ since $4 \cdot -2 = -8 < 0$.
2. Not a relation since 0 is not related to itself. $0 \not\sim 0$ since $0 \cdot 0 = 0 \not> 0$.
3. It is a relation.

Exercise 12.3 (Shifrin Appendix A.3.2)

Find the flaw in the so-called proof in the Remark on p. 380.

Exercise 12.4 (Shifrin Appendix A.3.3)

Define a relation on $\mathbb{R}$ as follows: $x \sim y$ if and only if $x - y$ is an integer. Prove that $\sim$ is an equivalence relation and describe the set of equivalence classes.

Solution. We prove the three properties.

1. *Identity.* For $x \in \mathbb{R}$, $x - x = 0 \in \mathbb{Z}$ so $x \sim x$.
2. *Symmetricity.* For $x, y \in \mathbb{R}$, if $x - y \in \mathbb{Z}$, then by the ring properties $-(x - y) = -1 \cdot (x - y) =$

$-1 \cdot x + -1 \cdot -y) = -x + y = y - x \in \mathbb{Z}$ and so $y \sim x$.

3. *Transitivity.* For $x, y, z \in \mathbb{R}$, if $x \sim y$ and $y \sim z$. Then $x - y, y - z \in \mathbb{Z}$. Since $\mathbb{Z}$ is closed under addition and addition is associative, we have

$$(x - y) + (y - z) = x + (-y + y) - z = x + 0 + -z = x - z \in \mathbb{Z} \tag{226}$$

and so $x \sim z$.

The set of all equivalence classes can be represented by the interval $[0, 1)$, where each $x \in \mathbb{R}$ gets mapped to $x \pmod 1$.

Exercise 12.5 (Shifrin Appendix A.3.4)

Define a relation on $\mathbb{N}$ as follows: $x \sim y$ if and only if x and y have the same last digit in their base-ten representation. Prove that $\sim$ is an equivalence relation, and describe the set of equivalence classes.

Exercise 12.6 (Shifrin Appendix A.3.5)

Define a relation on $\mathbb{R}^2$ as follows: $(x_1, x_2) \sim (y_1, y_2)$ if and only if $x_1^2 + x_2^2 = y_1^2 + y_2^2$. Prove that $\sim$ is an equivalence relation, and describe the set of equivalence classes.

Exercise 12.7 (Shifrin Appendix A.3.6)

(a) Define a relation on $\mathbb{R}$ as follows: x and y are related if $|x - y| < 1$. Decide whether this is an equivalence relation.

(b) Define an equivalence relation on $\mathbb{R}$ whose equivalence classes are intervals of length 1.

Exercise 12.8 (Shifrin Appendix A.3.7)

Which of the following functions $f : \mathbb{Q} \to \mathbb{Q}$ are well-defined?

(a) $f(\frac{a}{b}) = \frac{a+1}{b+1}$

(b) $f(\frac{a}{b}) = \frac{a+b}{b}$

(c) $f(\frac{a}{b}) = \frac{2a^2}{3b^2}$

(d) $f(\frac{a}{b}) = \frac{b}{a}$

(e) $f(\frac{a}{b}) = \frac{a^2+ab+b^2}{a^2+b^2}$

Exercise 12.9 (Shifrin Appendix A.3.8)

Define an equivalence relation on $X = \mathbb{N} \times \mathbb{N}$ as follows:

$$(a, b) \sim (c, d) \Leftrightarrow a + d = c + b.$$

(a) Prove that $\sim$ is indeed an equivalence relation.

(b) Identify the set of equivalence classes.

Exercise 12.10 (Shifrin Appendix A.3.9)

Define a relation on $\mathbb{N}$ as follows: $x \sim y$ if and only if there are integers j and k so that $x|y^j$ and $y|x^k$.

(a) Show that $\sim$ is an equivalence relation.

(b) Determine the equivalence classes [1], [2], [9], [10], and [20].

(c) Describe explicitly the equivalence classes $[x]$ in general.

Solution. We first prove the three properties.

1. *Identity.* Since we can set $j = k = 1$ to get $x|x$, trivially $x \sim x$.
2. *Symmetricity.* Assume that $x \sim y$. Then there exists some j, k such that $x|y^j$ and $y|x^k$. We can just swap j, k to identify the integers that hold for $y \sim x$.
3. *Transitivity.* Assume that $x \sim y, y \sim z$. Then there exists some integers i, j, k, l s.t. $x|y^i, y|x^j, y|z^k, z|y^l$. We would like to find integers a, b such that $x|z^a$ and $z|x^b$. We can set $a = ik$ and $b = jl$. By replacing unimportant integer constants with $*$, we write

$$z^{ik} = (z^k)^i = (*y)^i = *y^i = *x \tag{227}$$

$$x^{jl} = (x^j)^l = (*y)^l = *y^l = *z \tag{228}$$

and therefore $x \sim z$.

The equivalence classes are as shown, with brief explanations.

1. $[1] = \{1\}$. Since we are looking for naturals where $y|1^k$, i.e. $y|1$, $y = 1$ is the only natural.
2. $[2] = \{2^n\}_{n=1}^{\infty}$, i.e. all naturals with powers of 2. $2|y^j$ tells me that y must at least be even, but $y|2^k$ tells me that y must be a power of 2.
3. $[9] = \{3^n\}_{n=1}^{\infty}$. $9|y^j$ tells me that as long as y is a multiple of 3, we can set $j \geq 2$. $y|9^k$ tells me that y must not have any other prime divisors.
4. $[10] = \{2^i \cdot 5^j\}_{i,j \in \mathbb{N}}$. $10|y^j$ tells me that y must have at least 2 and 5 as prime divisors. $y|10^k$ tells me that it cannot have any other prime divisors.
5. $[20] = [10]$. $20|y^j$ tells me that y must have at least 2 and 5, and then by setting $j = 2$, y^2 is guaranteed to be divisible by 20. $y|20^k$ tells me that it cannot have any other divisors. This is the same definition as that of [10].

The equivalence class $[x]$ is described as such. Given a natural $x \in \mathbb{N}$, let $P = \{p_1, \ldots, p_n\} \subset \mathbb{N}$ be the set of prime divisors of x, which is guaranteed to be finite and unique by the fundamental theorem of arithmetic. Then,

$$[x] \coloneqq \{p_1^{i_1} \cdot p_2^{i_2} \cdot \ldots \cdot p_n^{i_n} \mid i_1, \ldots, i_n \geq 1\} \tag{229}$$

Exercise 12.11 (Shifrin Appendix A.3.10)

Given a function $f : S \to T$, consider the following relation on S:

$$x \sim y \Leftrightarrow f(x) = f(y).$$

(a) Prove that $\sim$ is an equivalence relation.

(b) Prove that if f maps onto T, then there is a one-to-one correspondence between the set of equivalence classes and T.

Exercise 12.12 (Math 531 Spring 2025 PS1.2)

Prove that $\mathbb{Q}$ is countable.

Solution. Proved in theorem above.

Exercise 12.13 (Math 531 Spring 2025, PS2.5)

Prove that if X is a set and $A, B, C \subset X$, we have that

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C), \tag{230}$$

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C), \tag{231}$$

$$X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B), \tag{232}$$

$$X \setminus (A \cap B) = (X \setminus A) \cup (X \setminus B). \tag{233}$$

Solution. Proved in deMorgan's laws.

Exercise 12.14 (Shifrin Abstract Algebra Appendix 2.3)

Let $f : X \to Y$. Let $A, B \subset X$ and $C, D \subset Y$. Prove or give a counterexample (if possible, provide sufficient hypotheses for each statement to be valid):

1. $f(A) \cup f(B) = f(A \cup B)$
2. $f(A) \cap f(B) = f(A \cap B)$
3. $f(A - B) = f(A) - f(B)$
4. $f^{-1}(C) \cup f^{-1}(D) = f^{-1}(C \cup D)$
5. $f^{-1}(C) \cap f^{-1}(D) = f^{-1}(C \cap D)$
6. $f^{-1}(C - D) = f^{-1}(C) - f^{-1}(D)$
7. $f(f^{-1}(C)) = C$
8. $f^{-1}(f(A)) = A$

Solution. Listed.

1.

Exercise 12.15 (Munkres 1.1)

Check the distributive laws for $\cup$ and $\cap$ and DeMorgan's laws.

Solution.

Exercise 12.16 (Munkres 1.2)

Determine which of the following statements are true for all sets A, B, C, and D. If a double implication fails, determine whether one or the other of the possible implications holds. If an equality fails, determine whether the statement becomes true if the "equals" symbol is replaced by one or the other of the inclusion symbols $\subset$ or $\supset$.

1. $A \subset B$ and $A \subset C \iff A \subset (B \cup C)$.
2. $A \subset B$ or $A \subset C \iff A \subset (B \cup C)$.
3. $A \subset B$ and $A \subset C \iff A \subset (B \cap C)$.
4. $A \subset B$ or $A \subset C \iff A \subset (B \cap C)$.
5. $A - (A - B) = B$.
6. $A - (B - A) = A - B$.
7. $A \cap (B - C) = (A \cap B) - (A \cap C)$.
8. $A \cup (B - C) = (A \cup B) - (A \cup C)$.
9. $(A \cap B) \cup (A - B) = A$.

10. $A \subset C$ and $B \subset D \Rightarrow (A \times B) \subset (C \times D)$.
11. The converse of (j).
12. The converse of (j), assuming that A and B are nonempty.
13. $(A \times B) \cup (C \times D) = (A \cup C) \times (B \cup D)$.
14. $(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$.
15. $A \times (B - C) = (A \times B) - (A \times C)$.
16. $(A - B) \times (C - D) = (A \times C - B \times C) - A \times D$.
17. $(A \times B) - (C \times D) = (A - C) \times (B - D)$.

Solution.

Exercise 12.17 (Munkres 1.3)

1. Write the contrapositive and converse of the following statement: "If $x < 0$, then $x^2 - x > 0$," and determine which (if any) of the three statements are true.
2. Do the same for the statement "If $x > 0$, then $x^2 - x > 0$."

Solution.

Exercise 12.18 (Munkres 1.4)

Let A and B be sets of real numbers. Write the negation of each of the following statements:

1. For every $a \in A$, it is true that $a^2 \in B$.
2. For at least one $a \in A$, it is true that $a^2 \in B$.
3. For every $a \in A$, it is true that $a^2 \notin B$.
4. For at least one $a \notin A$, it is true that $a^2 \in B$.

Solution.

Exercise 12.19 (Munkres 1.5)

Let $\mathcal{A}$ be a nonempty collection of sets. Determine the truth of each of the following statements and of their converses:

1. $x \in \bigcup_{A \in \mathcal{A}} A \Rightarrow x \in A$ for at least one $A \in \mathcal{A}$.
2. $x \in \bigcup_{A \in \mathcal{A}} A \Rightarrow x \in A$ for every $A \in \mathcal{A}$.
3. $x \in \bigcap_{A \in \mathcal{A}} A \Rightarrow x \in A$ for at least one $A \in \mathcal{A}$.
4. $x \in \bigcap_{A \in \mathcal{A}} A \Rightarrow x \in A$ for every $A \in \mathcal{A}$.

Solution.

Exercise 12.20 (Munkres 1.7)

Given sets A, B, and C, express each of the following sets in terms of A, B, and C, using the symbols $\cup$, $\cap$, and $-$.
$D = \{x \mid x \in A \text{ and } (x \in B \text{ or } x \in C)\}$,
$E = \{x \mid (x \in A \text{ and } x \in B) \text{ or } x \in C\}$,
$F = \{x \mid x \in A \text{ and } (x \in B \Rightarrow x \in C)\}$.

Solution.

Exercise 12.21 (Munkres 1.8)

If a set A has two elements, show that $\mathcal{P}(A)$ has four elements. How many elements does $\mathcal{P}(A)$ have if A has one element? Three elements? No elements? Why is $\mathcal{P}(A)$ called the power set of A?

Solution.

Exercise 12.22 (Munkres 1.9)

Formulate and prove DeMorgan's laws for arbitrary unions and intersections.

Solution.

Exercise 12.23 (Munkres 1.10)

Let $\mathbb{R}$ denote the set of real numbers. For each of the following subsets of $\mathbb{R} \times \mathbb{R}$, determine whether it is equal to the cartesian product of two subsets of $\mathbb{R}$.

1. $\{(x, y) \mid x \text{ is an integer}\}$.
2. $\{(x, y) \mid 0 < y \leq 1\}$.
3. $\{(x, y) \mid y > x\}$.
4. $\{(x, y) \mid x \text{ is not an integer and } y \text{ is an integer}\}$.
5. $\{(x, y) \mid x^2 + y^2 < 1\}$.

Solution.

Exercise 12.24 (Munkres 2.1)

Let $f : A \to B$. Let $A_0 \subset A$ and $B_0 \subset B$.

1. Show that $A_0 \subset f^{-1}(f(A_0))$ and that equality holds if f is injective.
2. Show that $f(f^{-1}(B_0)) \subset B_0$ and that equality holds if f is surjective.

Solution. Solved in 2.4.

Exercise 12.25 (Munkres 2.2)

Let $f : A \to B$ and let $A_i \subset A$ and $B_i \subset B$ for $i = 0$ and $i = 1$. Show that f^{-1} preserves inclusions, unions, intersections, and differences of sets:

a) $B_0 \subset B_1 \Rightarrow f^{-1}(B_0) \subset f^{-1}(B_1)$.
b) $f^{-1}(B_0 \cup B_1) = f^{-1}(B_0) \cup f^{-1}(B_1)$.
c) $f^{-1}(B_0 \cap B_1) = f^{-1}(B_0) \cap f^{-1}(B_1)$.
d) $f^{-1}(B_0 - B_1) = f^{-1}(B_0) - f^{-1}(B_1)$.

Show that f preserves inclusions and unions only:

e) $A_0 \subset A_1 \Rightarrow f(A_0) \subset f(A_1)$.
f) $f(A_0 \cup A_1) = f(A_0) \cup f(A_1)$.
g) $f(A_0 \cap A_1) \subset f(A_0) \cap f(A_1)$; show that equality holds if f is injective.
h) $f(A_0 - A_1) \supset f(A_0) - f(A_1)$; show that equality holds if f is injective.

Solution. (a)-(d) proved in 2.5. (e)-(h) proved in 2.6.

Exercise 12.26 (Munkres 2.3)

Show that (b), (c), (f), and (g) of Exercise 2 hold for arbitrary unions and intersections.

Solution.

Exercise 12.27 (Munkres 2.4)

Let $f : A \to B$ and $g : B \to C$.

1. If $C_0 \subset C$, show that $(g \circ f)^{-1}(C_0) = f^{-1}(g^{-1}(C_0))$.
2. If f and g are injective, show that $g \circ f$ is injective.
3. If $g \circ f$ is injective, what can you say about injectivity of f and g?
4. If f and g are surjective, show that $g \circ f$ is surjective.
5. If $g \circ f$ is surjective, what can you say about surjectivity of f and g?
6. Summarize your answers to (b)–(e) in the form of a theorem.

Solution.

Exercise 12.28 (Munkres 2.5)

In general, let us denote the identity function for a set C by i_C. That is, define $i_C : C \to C$ to be the function given by the rule $i_C(x) = x$ for all $x \in C$. Given $f : A \to B$, we say that a function $g : B \to A$ is a left inverse for f if $g \circ f = i_A$; and we say that $h : B \to A$ is a right inverse for f if $f \circ h = i_B$.

1. Show that if f has a left inverse, f is injective; and if f has a right inverse, f is surjective.
2. Give an example of a function that has a left inverse but no right inverse.
3. Give an example of a function that has a right inverse but no left inverse.
4. Can a function have more than one left inverse? More than one right inverse?
5. Show that if f has both a left inverse g and a right inverse h, then f is bijective and $g = h = f^{-1}$.

Solution.

Exercise 12.29 (Munkres 2.6)

Let $f : \mathbb{R} \to \mathbb{R}$ be the function $f(x) = x^3 - x$. By restricting the domain and range of f appropriately, obtain from f a bijective function g. Draw the graphs of g and g^{-1}. (There are several possible choices for g.)

Solution.

Exercise 12.30 (Munkres 3.1)

Define two points (x_0, y_0) and (x_1, y_1) of the plane to be equivalent if $y_0 - x_0^2 = y_1 - x_1^2$. Check that this is an equivalence relation and describe the equivalence classes.

Solution.

Exercise 12.31 (Munkres 3.2)

Let C be a relation on a set A. If $A_0 \subset A$, define the restriction of C to A_0 to be the relation $C \cap (A_0 \times A_0)$. Show that the restriction of an equivalence relation is an equivalence relation.

Solution.

Exercise 12.32 (Munkres 3.3)

Here is a "proof" that every relation C that is both symmetric and transitive is also reflexive: "Since C is symmetric, aCb implies bCa. Since C is transitive, aCb and bCa together imply aCa, as desired." Find the flaw in this argument.

Solution.

Exercise 12.33 (Munkres 3.4)

Let $f : A \to B$ be a surjective function. Let us define a relation on A by setting $a_0 \sim a_1$ if $f(a_0) = f(a_1)$.

1. Show that this is an equivalence relation.
2. Let A^* be the set of equivalence classes. Show there is a bijective correspondence of A^* with B.

Solution.

Exercise 12.34 (Munkres 3.5)

Let S and S' be the following subsets of the plane:
$S = \{(x, y) \mid y = x + 1 \text{ and } 0 < x < 2\}$,
$S' = \{(x, y) \mid y - x \text{ is an integer}\}$.

1. Show that S' is an equivalence relation on the real line and $S' \supset S$. Describe the equivalence classes of S'.
2. Show that given any collection of equivalence relations on a set A, their intersection is an equivalence relation on A.
3. Describe the equivalence relation T on the real line that is the intersection of all equivalence relations on the real line that contain S. Describe the equivalence classes of T.

Solution.

Exercise 12.35 (Munkres 3.6)

Define a relation on the plane by setting $(x_0, y_0) < (x_1, y_1)$ if either $y_0 - x_0^2 < y_1 - x_1^2$, or $y_0 - x_0^2 = y_1 - x_1^2$ and $x_0 < x_1$. Show that this is an order relation on the plane, and describe it geometrically.

Solution.

Exercise 12.36 (Munkres 3.7)

Show that the restriction of an order relation is an order relation.

Solution.

Exercise 12.37 (Munkres 3.8)

Check that the relation defined in Example 7 is an order relation.

Solution.

Exercise 12.38 (Munkres 3.9)

Check that the dictionary order is an order relation.

Solution.

Exercise 12.39 (Munkres 3.10)

1. Show that the map $f : (-1, 1) \to \mathbb{R}$ of Example 9 is order-preserving.
2. Show that the equation $g(y) = 2y/[1 + (1 + 4y^2)^{1/2}]$ defines a function $g : \mathbb{R} \to (-1, 1)$ that is both a left and a right inverse of f.

Solution.

Exercise 12.40 (Munkres 3.11)

Show that an element in an ordered set has at most one immediate successor and at most one immediate predecessor. Show that a subset of an ordered set has at most one smallest element and at most one largest element.

Solution.

Exercise 12.41 (Munkres 3.12)

Let $\mathbb{Z}_+$ denote the set of positive integers. Consider the following order relations on $\mathbb{Z}_+ \times \mathbb{Z}_+$:

1. The dictionary order.
2. $(x_0, y_0) < (x_1, y_1)$ if either $x_0 - y_0 < x_1 - y_1$, or $x_0 - y_0 = x_1 - y_1$ and $y_0 < y_1$.
3. $(x_0, y_0) < (x_1, y_1)$ if either $x_0 + y_0 < x_1 + y_1$, or $x_0 + y_0 = x_1 + y_1$ and $y_0 < y_1$.

In these order relations, which elements have immediate predecessors? Does the set have a smallest element? Show that all three order types are different.

Solution.

Exercise 12.42 (Munkres 3.13)

Prove the following theorem. If an ordered set A has the least upper bound property, then it has the greatest lower bound property.

Solution.

Exercise 12.43 (Munkres 3.14)

If C is a relation on a set A, define a new relation D on A by letting $(b, a) \in D$ if $(a, b) \in C$.

1. Show that C is symmetric if and only if $C = D$.
2. Show that if C is an order relation, D is also an order relation.
3. Prove the converse of the theorem in Exercise 13.

Solution.

Exercise 12.44 (Munkres 3.15)

Assume that the real line has the least upper bound property.

1. Show that the sets
$[0, 1] = \{x \mid 0 \leq x \leq 1\}$,
$[0, 1) = \{x \mid 0 \leq x < 1\}$
have the least upper bound property.
2. Does $[0, 1] \times [0, 1]$ in the dictionary order have the least upper bound property? What about $[0, 1) \times [0, 1]$?

Solution.

Exercise 12.45 (Munkres Topology 5.5)

Which of the following subset of $\mathbb{R}^\omega$ can be expressed as the Cartesian product of subsets of $\mathbb{R}$?[a]

[a] Note that the existence of these sets depend on the axiom of choice.

Solution. Listed. We will denote the sets in question as A.

1. We claim that
$$A = \mathbb{Z} \times \mathbb{Z} \times \dots \tag{234}$$
2. Let us denote $\mathbb{R}_{\geq i}$ be the set of reals greater than or equal to i. This is clearly a subset of $\mathbb{R}$. Then
$$A = \prod_{i=1}^{\infty} \mathbb{R}_{\geq i} \tag{235}$$
3. We claim
$$A = \left(\prod_{i=1}^{100} \mathbb{R} \right) \times \left(\prod_{j=1}^{\infty} \mathbb{Z} \right) \tag{236}$$
4. This is not possible

Exercise 12.46 (Math 531 Spring 2025, PS1.1)

Find a formula for the sum of the first n odd numbers and prove that it is correct.

Solution. I claim that $f(n) = n^2$. I prove using induction. For $n = 1$, $f(1) = n^2 = 1^2 = 1$. Now assume f holds for some $k \in \mathbb{N}$. Then, the kth off number is $2k - 1$. Therefore

$$f(k + 1) = f(k) + (2k + 1) = k^2 + 2k + 1 = (k + 1)^2 \tag{237}$$

and the formula holds for $k = 1$. By the principle of induction, $f(n) = n^2$ is true for all $n \in \mathbb{N}$.

Exercise 12.47 (Shifrin Abstract Algebra 1.1.4.C)

We check for $n = 1$ denoting our formula as f. Indeed, we have

$$f(1) = \frac{1 \cdot 2 \cdot 3}{6} = 1 = 1^2 \tag{238}$$

For the induction step, assume that $f(k)$ is true for some $k \in \mathbb{N}$. Then,

$$\begin{aligned} f(k+1) &= f(k) + (k+1)^2 && (239) \\ &= \frac{k(k+1)(2k+1)}{6} + \frac{6(k+1)^2}{6} && (240) \\ &= \frac{(k+1)\{k(2k+1) + 6(k+1)\}}{6} && (241) \\ &= \frac{(k+1)(2k^2 + 7k + 6)}{6} && (242) \\ &= \frac{(k+1)(k+2)(2(k+1)+1)}{6} && (243) \\ &= f(k+1) && (244) \end{aligned}$$

Therefore f holds for all $n \in \mathbb{N}$.

Exercise 12.48 (Shifrin Abstract Algebra 1.1.4.G)

We prove the base cases for $n = 1, 2, 3$.

1. $n = 1$. $n + 2 = 3$ is divisible by 3.
2. $n = 2$. $n + 4 = 6$ is divisible by 3.
3. $n = 3$. $n + 2 = 3$ is divisible by 3.

For our inductive step, assume that for some $n = k \in \mathbb{N}$, one of the elements in $S_k = \{k, k+2, k+4\}$ is divisible by 3. Let us denote this element a. We wish to show that this claim is true for $n = k+3$ on the set $S_{k+3} = \{k+3, k+5, k+7\}$. Since $a \in S_k$, this means that $a + 3 \in S_{k+3}$, and $3 | a \implies 3 | (a+3)$. So we can always identify the element $a + 3$. Since we proved the base cases for $n = 1, 2, 3$, and proved the recursive step, we have essentially proved the claim for all naturals of the form $3k+1, 3k+3, 3k+3$ ($k \in \mathbb{N}_0$), which is precisely the natural numbers.

Exercise 12.49 (Shifrin Abstract Algebra 1.1.4.J)

Let $n = 1$. Then $1 + x \geq 1 + x$ trivially. For the induction step, assume that this inequality holds for some $n \in \mathbb{N}$. Then, we have

$$\begin{aligned} 1 + (n+1)x &= 1 + nx + x && (245) \\ &\leq (1+x)^n + x && (246) \\ &\leq (1+x)^n + x(1+x)^n && (247) \\ &= (1+x)^{n+1} && (248) \end{aligned}$$

where the prove the penultimate step by applying the ordered field axioms to the 2 cases:

1. If $x \geq 0$, then addition preserves order so $1 + x \geq 0 + 1 = 1$. Since $1 + x, 1 > 0$, order is preserved under multiplication by a positive element, so $(1+x)^2 \geq 1 + x \geq 1$. Using induction, we can show that for all $n \in \mathbb{N}$, $(1+x)^n \geq 1$, and again by preservation of order under multiplication by a positive element, this implies $x(1+x)^n \geq x$ for all $n \in \mathbb{N}$.
2. If $0 > x > -1$, we have $0 < 1 + x < 1$ and by the same induction proof, we can bound $0 < (1+x)^n < 1$ for all n. Finally by reversal of order under multiplication by a negative

element, we have $x(1+x)^n > x$.

Therefore, we take the less restrictive of the 2 bounds: $x(1+x)^n \geq x$.

Exercise 12.50 (Shifrin Abstract Algebra 1.1.7)

Let us denote

$$x = \frac{1+\sqrt{5}}{2}, \quad y = \frac{1-\sqrt{5}}{2} \tag{249}$$

Note the identities

$$x^2 = \left(\frac{1+\sqrt{5}}{2}\right)^2 = \frac{3+2\sqrt{5}}{2} = 1+x \tag{250}$$

$$y^2 = \left(\frac{1-\sqrt{5}}{2}\right)^2 = \frac{3-2\sqrt{5}}{2} = 1+y \tag{251}$$

We check the base case for $n = 1$

$$a_1 = \frac{1}{\sqrt{5}}(x+y) = \frac{1}{\sqrt{5}}\frac{2\sqrt{5}}{2} = 1 \tag{252}$$

and for $n = 2$

$$a_2 = \frac{1}{\sqrt{5}}(x^2 - y^2) = \frac{1}{\sqrt{5}}((1+x) - (1+y)) = \frac{1}{\sqrt{5}}(x+y) = a_1 = 1 \tag{253}$$

For the inductive step, assume that this formula holds for some $k-1, k \in \mathbb{N}$. Then, we have

$$\begin{aligned} a_{k+1} &= a_k + a_{k-1} && (254) \\ &= \frac{1}{\sqrt{5}}(x^{k-1} - y^{k-1}) + \frac{1}{\sqrt{5}}(x^k - y^k) && (255) \\ &= \frac{1}{\sqrt{5}}\big\{x^{k-1}(1+x) - y^{k-1}(1+y)\big\} && (256) \\ &= \frac{1}{\sqrt{5}}(x^{k+1} - y^{k+1}) && (257) \end{aligned}$$

and we are done.