An Introduction to Proofs

Hassium, Fung San Gaan, Tingyu Wu

Introduction

In higher-level mathematics, students need a certain level of "mathematical maturity" to understand and apply abstract ideas. However, there is no clear way to measure this maturity, nor a definitive method to teach someone how to write a proof. This note is intended to be a transition from high school math to proof-based formal mathematics.

The first four sections are the solid prerequisites for any kind of mathematics. Sections 6 to 10 are selected topics that help students to prepare for advaneced math courses. There are mainly four parts: linear algebra, abstract algebra, real analysis, and general topology. For linear algebra, please read section 5. For abstract algebra, we recommend the readers to check section 6 and 7. For real analysis, we will use some basic ring theory to construct $\mathbb{R}$, so please check section 8 and 9. For general topology, please read section 6 and 10. Those parts are designed to be independent of each other.

You may notice that we use different names for the same object—a common practice in mathematics. For example, the set of all real numbers $\mathbb{R}$ can be viewed as a set, a group, a ring, a field, a topological space, a metric space ... (hopefully you will learn every term in this note) Each term highlights a particular aspect of the same structure. As Poincaré famously said, "Mathematics is the art of giving the same name to different things."

Many people have contributed to this note. Special thanks to my collaborators Fung San Gaan and Tingyu Wu. Fung San Gaan wrote the real analysis part of this note and collaborate with me on section 10. Tingyu Wu wrote the linear algebra part and collaborate with me on section 4. I shall also thank Bryce for his suggestions on the contents.

1 Basic Logic

Logic is the formal framework and rules of inference that ensure the validity and coherence of arguments in math.

Remark. We shall accept that sentences can be either true or false. Moreover, we assume that every English sentence can be stated in symbolic logic form.

A *proposition* is a sentence that is either true or false in a mathematical system. The label "true" or "false" assigned to a proposition is called its *truth value*. We use the letters T and F to represent "true" and "false", respectively. An *axiom* is a proposition that is assumed to be true within a mathematical system without requiring proof. Axioms serve as the foundational building blocks of a mathematical theory, from which other propositions can be derived. A *theorem* is a proposition that has been proven to be true using logical reasoning and the accepted axioms and previously established theorems of the mathematical system. The proof demonstrates why the theorem must hold based on these foundations.

Consider the proposition "π is not a rational number", which is trivially true. However, we could always find some false companion of this proposition, such as "π is a rational number". Similarly, we can find a true companion of a false proposition. Let P be a proposition, such a companion of P is called the *negation* of P, denoted $\neg P$.

Let P and Q be propositions. Those sentences can be combined using the word "and", denoted $P \wedge Q$, and called the *conjunction* of P and Q. The proposition $P \wedge Q$ is true if both P and Q are true. We can combine the propositions by the word "or", denoted $P \vee Q$, and called the *disjunction* of P and Q. The proposition $P \vee Q$ is true if at least one of P or Q is true. A *truth table* is shown below.

P	Q	$\neg P$	$P \wedge Q$	$P \vee Q$
T	T	F	T	T
T	F	F	F	T
F	T	T	F	T
F	F	T	F	F

Two propositions P and Q are *logically equivalent* if they have the same truth value in every possible combination of truth values for the variables in the statements, denoted $P \equiv Q$.

Example. Let P be a proposition, then $P \equiv \neg(\neg P)$ is logically equivalent. To prove this statement, consider $\neg P$ as a proposition Q, then we obtain the following truth table.

P	$Q \equiv \neg P$	$\neg Q \equiv \neg(\neg P)$
T	F	T
F	T	F

Here P and $\neg Q$ has the same truth value in each case, so $P \equiv \neg(\neg P)$.

Problem 1.1. Let P, Q, and R be propositions. Consider the following statements:

1. $P \vee (Q \vee R) \equiv (P \vee Q) \vee R$;
2. $P \wedge (Q \wedge R) \equiv (P \wedge Q) \wedge R$.

Try to prove or disprove the statements.

Problem 1.2. Let P, Q, and R be propositions. Consider the following statements:

1. $\neg(P \vee Q) \equiv (\neg P) \wedge (\neg Q)$;
2. $\neg(P \wedge Q) \equiv (\neg P) \vee (\neg Q)$;
3. $P \vee (Q \wedge R) \equiv (P \vee Q) \wedge (P \vee R)$.

Try to prove or disprove the statements. Based on your results, can you find more properties?

Let P and Q be propositions. Consider the proposition "if n is a natural number, then $2n$ is an even number". Let P denotes "n is a natural number" and let Q denotes "$2n$ is an even number", then the sentence becomes "if P, then Q", denoted $P \implies Q$. This implication called a *conditional proposition*, P is called the *antecedent* and Q is called the *consequent*. The proposition $P \implies Q$ is true if P is true and Q is true. What if P is false? The answer arises from one's intuition.

Imagine your high school teacher say "if you didn't submit your homework, then you haven't completed it". How would you argue against this sentence? The most likely response would be, "I did the homework but I didn't submit it". Whether or not you submitted your homework does not affect the truth value of the implication.

You should be convinced by your own intuition. This case is called a *vacuous truth*. In the proposition $P \implies Q$, when P is false, $P \implies Q$ is true. The truth table of $P \implies Q$ is shown below.

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

Let P and Q be propositions, $(P \implies Q) \wedge (Q \implies P)$ is called a *biconditional proposition*, denoted $P \iff Q$. We will write this by "P is true if and only if Q is true".

Problem 1.3. Let P and Q be propositions, show $(\neg P \equiv \neg Q) \iff (P \equiv Q)$.

Example. Let P and Q be propositions. Consider the conditional proposition $P \implies Q$. It is false only if P is true and Q is false, that is, $\neg(P \implies Q) \equiv P \wedge (\neg Q)$. Now we take the negation of the right side, $\neg(P \wedge (\neg Q)) \equiv (\neg P) \vee (\neg(\neg Q)) \equiv (\neg P) \vee Q$.

Problem 1.4. Write down the truth table of a biconditional proposition. Based on your truth table and the previous example, try to find a proposition R by "$\vee$", "$\wedge$", and "$\neg$" such that $R \equiv (P \iff Q)$. If $P \iff Q$ is true, does $P \equiv Q$?

Problem 1.5. Let P, Q, R, and S be propositions. Rewrite $P \implies (Q \implies (R \implies S))$ by '$\vee$", "$\wedge$", and "$\neg$". What is the negation of this sentence?

Problem 1.6. Let P, Q, and R be propositions. Try to prove or disprove $P \implies (Q \vee R) \equiv (\neg P) \vee Q \vee R$. What about $P \implies (Q \wedge R)$?

Given a proposition $P \implies Q$, the *converse* is defined as $Q \implies P$ and the *contrapositive* is defined as $(\neg Q) \implies (\neg P)$. The truth table is shown below, and it suffices to conclude that $(P \implies Q) \equiv (\neg Q \implies \neg P)$.

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

Problem 1.7. Let P and Q be propositions, when does $(P \implies Q) \equiv (Q \implies P)$?

Let P be the proposition "x is a natural number". Here x is a *variable*, and the truth value of this proposition depends on x. For instance, if $x = 1$, then P is true; if $x = 0.86$, then P is false. A *propositional function* is a family of propositions depending on one or more variables. The collection of permitted variables is the *domain*. Now we write $P(x)$ instead of P, so $P(1)$ is true and $P(0.86)$ is false.

Problem 1.8. Let x be a variable and let x be a natural number. Give a proposition $P(x)$ such that $P(x)$ is true when $x \leq 2024$ and false when $x \geq 2025$.

Propositional functions are often quantified. The *universal quantifier* is denoted by "$\forall$", and the proposition $\forall x(P(x))$ is true if and only if $P(x)$ is true for every x in its domain. The *existential quantifier* is denoted by "$\exists$", and the proposition $\exists x(P(x))$ is true if and only if $P(x)$ is true for at least one x in its domain. Consider the proposition $\forall x(P(x))$, this means all x make $P(x)$ true, so there does not exist some x such that $P(x)$ is false, which is $\neg(\exists x(\neg P(x)))$.

Example. Let $P(x)$ be a proposition, then $\neg(\forall x(P(x))) \iff \neg(\neg(\exists x(\neg P(x)))) \iff \exists x(\neg P(x))$.

Problem 1.9. Let $P(x)$ be a proposition, show that $\neg(\exists x(P(x))) \iff \forall x(\neg P(x))$.

The order of quantifiers does matter the meaning of a proposition. Consider the proposition "for all natural number x, there exists a natural number y such that $y > x$". Pick some x, let $y = x+1$, then $y > x$ and y is a natural number, so the proposition is true. However, switching the order of quantifiers gives "there exists a natural number y, for all natural number x, $y > x$". Suppose there exists such y, then $y + 1$ is a natural number, so let $x = y + 1$, it is trivial that $y < x$, hence the proposition is false.

Example. Let $P(x)$ and $Q(y)$ be propositions. Consider the proposition $\forall x(\exists y(P(x)\vee Q(y)))$. To find its negation, let $R(x) \equiv \exists y(P(x)\vee Q(y))$, now the negation becomes $\exists x(\neg R(x))$. Since P only depends on x, let $S(y) \equiv (P(x)\vee Q(y))$, then we have $\exists x(\neg(\exists y(S(y)))) \equiv \exists x(\forall y(\neg S(y))) \equiv \exists x(\forall y(\neg(P(x) \vee Q(y)))) \equiv \exists x(\forall y((\neg P(x)) \wedge (\neg Q(y))))$.

Problem 1.10. Let $P(x, y, z)$ be a proposition, consider the following propositions.

1. $Q(x, y, z) \equiv \exists x(\forall y(\forall z(P(x, y, z))))$
2. $R(x, y, z) \equiv \forall x(\exists y(\forall z(P(x, y, z))))$
3. $S(x, y, z) \equiv \forall x(\forall y(\exists z(P(x, y, z))))$

What are the negations of those propositions? What is the negation of $Q \vee (R \wedge S)$?

Example. Let $P(x)$ and $Q(x)$ be propositions. Consider the negation of $P(x) \implies Q(x)$, $\neg(P(x) \implies Q(x)) \equiv \neg((\neg P(x)) \vee Q(x)) \equiv P(x) \wedge (\neg Q(x)) \equiv \forall x(P(x) \wedge (\exists x(\neg Q(x)))) \equiv \exists x(P(x) \wedge (\neg Q(x)))$. Notice that taking the negation brings an existential quantifier.

In the following sections, we shall assume readers are familiar with basic logic and use it as a tool to understand or prove propositions. Several expressions and their "translations" are shown below.

$P \implies Q$	$P \iff Q$
P implies Q; if P, then Q	P if and only if Q
P is sufficient for Q; Q is necessary for P	P is necessary and sufficient for Q

Problem 1.11. Given the following propositions, analyze their structures.

1. The number $\sqrt{2}$ is not a rational number.
2. If x is a natural number, then x is an integer.
3. For all natural number x, for all rational number y with $x < y < x + 1$, there exists a real number z such that $y < z < y + 1$ and z is irrational.
4. Given a sequence (x_n) of real numbers, we say (x_n) converges to a real number L if, for all real number $\epsilon > 0$, there exists a real number N such that, for all natural number n, $n > N$ implies $|x_n - L| < \epsilon$.

Find the negation of each proposition.

2 Some Axioms of Sets

In this section, we begin investigating sets, the most basic entities in mathematics. It is natural to ask: What is a set? There is no precise definition of sets. Intuitively, a *set* is a collection of objects that satisfy some property, and the objects are called *elements*.

Remark. This note is based on the ZFC set theory. In this system, every object is a set, so there are sets of sets. From now on, assume that there exists a set.

If S is a set and x is an element in S, then we say x belongs to S, denoted $x \in S$. If x does not belong to S, then we write $x \notin S$. If S has no element, then we call it an *empty set*, denoted $\varnothing$.

Axiom of empty set. There exists an empty set.

Axiom of extensionality. Two sets A and B are equal if and only if they have the same elements.

Axiom schema of separation. If P is a property, then for any set X there exists a set $Y = \{x \in X \mid P(x)\}$.

Elements determine a set. One way to describe a set is to explicitly list the elements. For example, we can write a set $S = \{6, 7, 8\}$. Another way is to express the elements by the properties they satisfy.

Example. Here are several examples of sets.

1. the set $S = \{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}\}$ has three elements;

2. the set $\{2n \mid n \in \mathbb{N}\}$ is the set of all even numbers, where $\mathbb{N}$ is the set of natural numbers;
3. the set $\mathbb{Q} = \{a/b \mid a, b \in \mathbb{Z} \text{ and } b \neq 0\}$ is the set of rational numbers, where $\mathbb{Z}$ is the set of integers.

We shall provide constructions for $\mathbb{N}$ and $\mathbb{Z}$ later.

Problem 2.1. Write out the set of all positive integers and the set of all prime numbers.

Definition 2.1. Let S be a set. A set R is a *subset* of S, denoted $R \subset S$, if for all $x \in R$, $x \in S$. If there exists some $x \in S$ such that $x \notin R$, then R is called a *proper subset* of S, denoted $R \subsetneq S$.

It suffices to check that axiom schema of separation guarantees that subsets are sets.

Remark. Some textbooks use "$\subseteq$" for subsets and "$\subset$" for proper subsets.

Remark. From now on, try to vertify whether those constructions are actually sets.

Proposition. Let A be a set, then $A \subset A$.

Proof. For all $x \in A$, $x \in A$, so $A \subset A$. □

Proposition. Let X and Y be sets, then $X = Y$ if and only $X \subset Y$ and $Y \subset X$.

Remark. For a biconditional proposition $P \iff Q$, we use the notation "$(\Rightarrow)$" in the proof to show $P \implies Q$ and "$(\Leftarrow)$" for $Q \implies P$.

Proof. Let X and Y be sets. $(\Rightarrow)$ For all $x \in X$, since $X = Y$, $x \in Y$, so $X \subset Y$. For all $y \in Y$, since $X = Y$, $y \in X$, so $Y \subset X$. $(\Leftarrow)$ Suppose $X \neq Y$. If $X \subset Y$, then there exists $a \in Y$ such that $a \notin X$, so $X \not\subset Y$, a contradiction. □

Proposition. Let A be any set, then $\varnothing \subset A$.

Proof. Suppose $\varnothing \not\subset A$, then there exists $x \in \varnothing$ such that $x \notin A$, since $x \in \varnothing$ is false, a contradiction. □

Problem 2.2. Prove that a set is independent of the order of its elements. For example, $\{1, 2, 3\} = \{3, 2, 1\}$.

Problem 2.3. If X, Y, and Z are sets such that $X \subset Y$ and $Y \subset Z$, prove that $X \subset Z$.

Axiom of pairing. For two objects a and b, there exists a set $\{a, b\}$ containing exactly a and b.

Definition 2.2. Let a and b be some objects. An *ordered pair* (a, b) is defined as the set $\{\{a\}, \{a, b\}\}$.

Problem 2.4. Show that an ordered pair is indeed a set.

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

Proof. We have $(a, b) = \{\{a\}, \{a, b\}\}$ and $(c, d) = \{\{c\}, \{c, d\}\}$. $(\Rightarrow)$ Suppose $a \neq c$, then $\{a\} \neq \{c\}$. If $\{a\} = \{c, d\}$, then $c = d = a$, a contradiction. Suppose $b \neq d$. If $a = c$, then $\{a\} = \{c\}$ and $\{a, b\} \neq \{c, d\}$, a contradiction. $(\Leftarrow)$ If $a = c$ and $b = d$, then $\{a, b\} = \{c, d\}$ and $\{a\} = \{c\}$, hence $(a, b) = (c, d)$. □

The definition of ordered pairs can be extended to multiple elements. We call $(a_1, \ldots, a_n)$ a *n-tuple*.

Problem 2.5. Prove that $\{a\} = \{a, a\}$. A set with one element is called a *singleton*.

Problem 2.6. Write out the definition to n-tuples, where n is a positive integer.

Axiom of union. For all set X, there exists a set $Y = \bigcup X$, the union of all elements of X.

Definition 2.3. Let A and B be sets. The *union* of A and B is the set $\{x \mid x \in A \text{ or } x \in B\}$, denoted $A \cup B$. The *intersection* of A and B is the set $\{x \mid x \in A \text{ and } x \in B\}$. We say A and B are *disjoint* if $A \cap B = \varnothing$. The *complement* of A in B is the set $\{x \mid x \in B \text{ and } x \notin A\}$, denoted $B \setminus A$.

Problem 2.7. Let A and B be sets. Prove that $A \cup B$, $A \cap B$, and $A \setminus B$ are sets based on the axioms.

Proposition. Let A and B be sets, then $A \cup B = B \cup A$.

Proof. For all $x \in A \cup B$, if $x \in A$, then $x \in B \cup A$; if $x \in B$, then $x \in B$, hence $A \cup B = B \cup A$. □

Problem 2.8. Let A, B, and C be sets. Prove the following propositions.

1. $A \cap B = B \cap A$;
2. $A \cup (B \cup C) = (A \cup B) \cup C$;
3. $A \cap (B \cap C) = (A \cap B) \cap C$;
4. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$;
5. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

Theorem 2.1 (De Morgan's law)**.** Let A, B, and C be sets, then $C \setminus (A \cap B) = (C \setminus A) \cup (C \setminus B)$ and $C \setminus (A \cup B) = (C \setminus A) \cap (C \setminus B)$.

Proof. Let $x \in C \setminus (A \cap B)$, then $x \in C$ and $x \notin A \cap B$, that is, $x \notin A$ and $x \notin B$. If $x \notin C \setminus A$, then $x \notin A$, so $x \notin B$ and $x \in C \setminus B$. Hence $C \setminus (A \cap B) \subset (C \setminus A) \cup (C \setminus B)$. Now let $x \in (C \setminus A) \cup (C \setminus B)$, then $x \in C$ and $x \notin A$ or $x \notin B$, so $x \notin A \cap B$, that is, $x \in C \setminus (A \cap B)$, hence $(C \setminus A) \cup (C \setminus B) \subset C \setminus (A \cap B)$. The proof of the second part is left as an exercise. □

Some texts assume the existence of an "universal set", denoted U, which has all objects as elements including itself, so we can define complements of any set S as the set $U \setminus S$. However, this assumption leads to a paradox. Consider the set S, defined as the set of all sets that are not members of themselves, that is, $S = \{X \mid X \notin X\}$. Does S belong to S? This is known as Russell's Paradox. Assume $S \in S$, then by the definition of S, $S \in S$ implies $S \notin S$, a contradiction. Assume $S \notin S$, then $S \in S$. This is also a contradiction.

Axiom of regularity. For all nonempty sets, there is an element of the set that shares no element with the set.

Proposition. A set is not an element of itself.

Proof. Suppose X is a set such that $X \in X$ and $X \in \{X\}$. By contradiction, $X \neq \varnothing$. We have $X \cap \{X\} = \{x \mid x \in X,\ x \in \{X\}\} = X$. By the axiom of regularity, $X \cap \{X\} = \varnothing$, then $X = \varnothing$, a contradiction. □

Definition 2.4. Let X be a set. A *partition* of X is a set $\{X_i\}$ of non-empty subsets of X such that every elements $x \in X$ lies in exactly one of there subsets.

Problem 2.9. Prove that a partition is a set.

Definition 2.5. Let X be a set, and let the *successor* of X be $X^+ = X \cup \{X\}$. A set S is called an *inductive set* if $\varnothing \in S$ and for all $X \in S$, $X^+ \in S$.

Axiom of infinity. There exists an inductive set.

Proposition. The intersection of two inductive sets is an inductive set.

Proof. Let A and B be inductive sets, then $\varnothing \in A \cap B$. For all $S \in A \cap B$, $S \in A$ and $S \in B$. Since A and B are inductive, $S^+ \in A$ and $S^+ \in B$, hence $A \cap B$ is inductive. □

Definition 2.6. The set of all *natural numbers*, denoted $\mathbb{N}$, is the intersection of all inductive sets.

We denote $0 = \varnothing$, $1 = 0^+$, $2 = 1^+$, ...

Problem 2.10. Prove that the set of all natural numbers is a subset of any inductive set.

Axiom of power set. For any X there exists a set consisting of all subsets of X.

Definition 2.7. Given a set X, the set of all subsets of X is called its *power set*, denoted $\mathscr{P}(X)$.

Example. Let $X = \{a, b\}$, the power set $\mathscr{P}(X) = \{\varnothing, \{a\}, \{b\}, \{a, b\}\}$.

Definition 2.8. Let X and Y be sets. The *Cartesian product* $X \times Y$ is the set of all ordered pairs (a, b), where $a \in X$ and $b \in Y$.

Problem 2.11. Let X and Y be sets. Write out the set $\mathscr{P}(\mathscr{P}(X \cup Y))$. Prove that $X \times Y = \{z \in \mathscr{P}(\mathscr{P}(X \cup Y)) \mid$ there exists $x \in X$ and $y \in Y$ such that $z = (x, y)\} \subset \mathscr{P}(\mathscr{P}(X \cup Y))$, hence $X \times Y$ is a set.

Problem 2.12. Let S be a set, prove that $S \subsetneq \mathscr{P}(S)$.

Problem 2.13. Let A, B, and C be sets. Prove the following propositions.

1. $A \times B = B \times A$ if and only if $A = B$;
2. $A \times (B \times C) = (A \times B) \times C$;
3. $(A \cup B) \times C = (A \times C) \cup (B \times C)$;
4. $(A \cap B) \times C = (A \times C) \cap (B \times C)$;
5. $(A \setminus B) \times C = (A \times C) \setminus (B \times C)$.

Definition 2.9. The *disjoint union* of two sets A and B, denoted $A \amalg B$, is the set $A \amalg B = (A \times \{0\}) \cup (B \times \{1\})$.

Problem 2.14. Let A and B be sets, prove that $A \amalg B$ is a set.

Problem 2.15. Let X and Y be sets. The *symmetric difference* $X \triangle Y$ is defined to be $(X \setminus Y) \cup (Y \setminus X)$. Prove that $X \triangle Y$ is a set.

Definition 2.10. A *binary operation* R is a set of ordered pairs. If $(x, y) \in R$, we write xRy. The *domain* of R is the set $\mathrm{dom}(R) = \{u \mid$ there exists v such that $(u, v) \in R\}$. The *range* of R is the set $\mathrm{ran}(R) = \{v \mid$ there exists u such that $(u, v) \in R\}$.

It suffices to show a binary operation is indeed a set. Let R be a binary operation, then $R \subset X \times Y$ for some sets X and Y. By the axiom schema of separation, R is a set.

Problem 2.16. Let R be a binary operation. Prove that $\mathrm{dom}(R), \mathrm{ran}(R) \subset \bigcup(\bigcup R)$, hence, by axiom of union, $\mathrm{dom}(R)$ and $\mathrm{ran}(R)$ are sets.

Definition 2.11. Let R be a binary operation on a set S, that is, $R \subset S \times S$. We say R is an *equivalence relation* if the following properties hold.

1. For all $a \in X$, aRa. (reflexive)
2. If aRb, then bRa. (symmetric)
3. If aRb and bRc, then aRc. (transitive)

For all $a \in A$, the set $S_a = \{b \mid aRb\}$ is the *equivalence class* of a.

Problem 2.17. Prove that an equivalence class is a set.

Problem 2.18. Prove that "=" is an equivalence relation in $\mathbb{N}$.

Problem 2.19. Let R be a binary operation on a set X. For all $a, b \in A$, prove that $S_a \cap S_b$ is either $\varnothing$ or S_a. Prove that the collection of S_a forms a partition of X.

Definition 2.12. Let $\leq$ be a binary relation on a set X. We say $\leq$ is a *partial ordering* if the following conditions hold.

1. For all $x \in X$, $x \leq x$.
2. For all $x, y \in X$, $x \leq y$ and $y \leq x$ implies $x = y$.
3. For all $a, b, c \in X$, if $a \leq b$ and $b \leq c$, then $a \leq c$.

The set with a partial ordering is called a *partially ordered set*.

Definition 2.13. A partially ordered set $(X, \leq)$ is *linearly ordered* if for all $p, q \in X$, either $p \leq q$ or $q \leq p$.

Example. The set of natural numbers $\mathbb{N}$ forms a linearly ordered set in set inclusions.

Proposition. Let $(X, \leq)$ be a partially ordered set and let $Y \subset X$, then Y is partially ordered.

Proof. For all elements $a, b, c \in Y$, $a, b, c \in X$, so Y inherits the partial ordering of X. □

Problem 2.20. Let $(X, \leq)$ be a linearly ordered set and let $Y \subset X$, prove that Y is linearly ordered.

Problem 2.21. Let X be a set. If $(\mathscr{P}(X), \subset)$ is a linearly ordered set, prove that X is either a singleton or the empty set.

Definition 2.14. Let $(X, \leq)$ be a partially ordered set and let $Y \subset X$ be an nonempty subset. An element a is the *upper bound* of X if for all $x \in X$, $x \leq a$. An element b is the *lower bound* of X if for all $x \in X$, $b \leq x$. The least upper bound of X is called the *supremum* and the greatest lower bound of X is called the *infimum*.

Definition 2.15. Let $(X, \leq)$ be a partially ordered set. The set is *well-ordered* if for every nonempty subset S of X, there exists $a \in S$ such that for all $s \in S$, $a \leq s$.

Theorem 2.2 (well-ordering theorem)**.** Every set is well-orderable.

The well-ordering theorem is equivalent to the axiom of choice, which will be discussed later. You may assume it is correct for now.

Theorem 2.3 (finite induction)**.** Given a subset $S \subset \mathbb{N}$ of the natural numbers with $0 \in S$ and $n \in S$ implies $n + 1 \in S$, then $S = \mathbb{N}$.

Proof. Suppose $S \neq \mathbb{N}$, then $X = \mathbb{N} \setminus S$ is a nonempty set. By the well-ordering principle, X has a smallest element. Since $0 \in S$, $0 \notin X$, so the minimal element of X can be written in the form $k + 1$, where $k \in \mathbb{N}$. Recall that $k + 1 = k^+ = k \cup \{k\}$ is the successor of $k \in \mathbb{N}$, since $k + 1$ is the smallest element, $k \notin X$, so $k \in S$. Now we have $k \in S$ and $k + 1 \notin S$, a contradiction. □

Example. Consider the statement: let $n \in \mathbb{N}$, show that $\sum_{i=0}^{n} = (n(n+1))/2$. If $n = 0$, then the equation trivially holds. Assume $\sum_{i=0}^{k} = (k(k+1))/2$ holds for some $k \in \mathbb{N}$, then $\sum_{i=0}^{k+1} = (k(k+1))/2 + (k+1) = ((k+1)(k+2))/2$. Hence, by induction, $\sum_{i=0}^{n} = (n(n+1))/2$ for all $n \in \mathbb{N}$.

Problem 2.22. Prove that given a subset $S \subset \mathbb{N}$ of the natural numbers with $0 \in S$ and $\{0, 1, \ldots, n\} \subset S$ implies $n + 1 \in S$, then $S = \mathbb{N}$. This is known as the *complete finite induction*. Prove that the complete finite induction implies the finite induction, conclude that they are equivalent.

3 Functions

Definition 3.1. Let X and Y be sets. A *function* f is a binary operation $f \subset X \times Y$ such that for all $x \in X$, there exists a unique $y \in Y$ such that $(x, y) \in f$. We say f is a function from X to Y, denoted $f : X \to Y$. The set Y is called the *codomain* of f, denoted $\mathrm{cod}(f)$.

Problem 3.1. Given two sets A and B, prove that the collection of functions from A to B is a subset of $\mathscr{P}(A \times B)$, hence it is a set.

Definition 3.2. Let $f : X \to Y$ be a function, the *image* of X under f, denoted $\mathrm{im}(f)$, is the range of f. For all $(x, y) \in f$, we write $f(x) = y$. The *preimage* of Y under f, denoted $f^{-1}(Y)$, is the set $\{x \mid x \in X \text{ and } f(x) \in Y\}$. Two functions f and g are the same if $\mathrm{dom}(f) = \mathrm{dom}(g)$, $\mathrm{cod}(f) = \mathrm{cod}(g)$, and for all $x \in \mathrm{dom}(f)$, $f(x) = g(x)$.

Example. The operation $f : \mathbb{N} \to \mathbb{N}$ defined by $f(n) = n^+$, where $n \in \mathbb{N}$, is a function.

Example. The operation $f : \mathbb{R} \to \mathbb{R} \times \mathbb{R}$ defined by $f(x) = (x, x)$, where $x \in \mathbb{R}$, is a function.

Problem 3.2. Let f be a function, prove that $\text{ran}(f) \subset \text{cod}(f)$.

Problem 3.3. Consider a function $f : \varnothing \to A$, prove that $f = \varnothing$. Prove that such f is unique for every set A.

Problem 3.4. Verify whether the following binary operations are functions.

1. $f : \mathbb{R} \to \mathbb{R}$ with $f(x) = \sqrt{x}$ for all $x \in \mathbb{R}$.
2. $f : \{1, 2, 3\} \to \{2, 3\}$ with $f = \{\{1, 2\}, \{2, 2\}, \{3, 2\}\}$.
3. $f : \mathbb{N} \to \mathbb{Q}$ with $f(x) = x^4 - x^2$ for all $x \in \mathbb{N}$.
4. $f : \mathbb{Q} \to \mathbb{Z}$ with $f(x) = |x|$ for all $x \in \mathbb{Q}$.

Definition 3.3. Let $f : X \to Y$ and $g : Y \to Z$ be functions. The *composition* of f and g, denoted $g \circ f$, is defined as $g \circ f : X \to Z$ with $(g \circ f)(x) = g(f(x))$ for all $x \in X$.

Problem 3.5. Let $f : X \to Y$ and $g : Y \to Z$ be functions. Prove that $g \circ f$ is a well-defined function.

Proposition. The composition of functions is associative, that is, for all $f : X \to Y$, $g : Y \to Z$, and $h : Z \to S$, $h \circ (g \circ f) = (h \circ g) \circ f$.

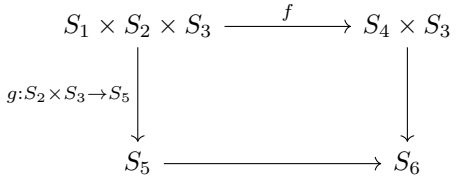
Proof. For all $x \in X$, $(h \circ g \circ f)(x) = h(g(f(x))) = (h \circ g)(f(x)) = ((h \circ g) \circ f)(x)$. □

Consider a diagram, where every vertex is an object and every arrow preserves the structure of those objects. Such a diagram is said to be *commutative* if all paths between two vertices are equivalent. In this section, every vertex is a set and every arrow is a function.

Example. Consider the following diagrams.

The left diagram is commutative if $g \circ f = h$. The right diagram is commutative if $g \circ f = h$ and $\psi \circ \varphi = h$.

Problem 3.6. Let $f : S_1 \times S_2 \to S_3$ be a function. We say f is commutative as a composition if $f(a, b) = f(b, a)$. Similarly, f is associative as a composition if $f(f(a, b), c) = f(a, f(b, c))$. Prove that if f is commutative, then $S_1 = S_2$. Prove that if f is associative, then the following diagram commutes.



Definition 3.4. A function $f : X \to Y$, where $X, Y \subset \mathbb{R}$, is said to be an *odd function* if for all $x \in X$, $f(x) = -f(-x)$. The function f is said to be an *even function* if for all $x \in X$, $f(x) = f(-x)$.

Example. Here are some examples of odd and even functions.

1. The function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$ for all $x \in \mathbb{R}$ is even.
2. The function $g : \mathbb{R} \to \mathbb{R}$ defined by $g(x) = x$ for all $x \in \mathbb{R}$ is odd.
3. The function $0 : \mathbb{R} \to \mathbb{R}$ defined by $0(x) = 0$ for all $x \in \mathbb{R}$ is both odd and even.

Proposition. Any function $f : X \to Y$, where $X, Y \subset \mathbb{R}$, can be written as $f = \varphi + \psi$, where φ is an odd function and ψ is an even function.

Proof. For all $x \in X$, define $\varphi = (f(x) - f(-x))/2$ and $\psi = (f(x) + f(-x))/2$, then $\varphi(-x) = (f(-x) - f(x))/2 = -\varphi(x)$ and $\psi(-x) = (f(-x)+f(x))/2 = \psi(x)$. Moreover, $\varphi(x)+\psi(x) = (f(x)-f(-x)+f(x)+f(-x))/2 = f(x)$. □

Problem 3.7. Write a function $f : \mathbb{R} \to \mathbb{R}$ that is neither odd nor even. Decompose f as a sum of an even and an odd function.

Problem 3.8. Prove that such decomposition for each $f : X \to Y$, where $X, Y \subset \mathbb{R}$, is unique.

Definition 3.5. Let $f : A \to B$ be a function. We say f is *injective* if for all $x, y \in B$ and $x = y$, then $f^{-1}(x) = f^{-1}(y)$. We say f is *surjective* if $\text{ran}(f) = \text{cod}(f)$. The function f is said to be *bijective* if it is both injective and surjective.

Problem 3.9. State an example if such a function exists.

1. $f : \{1, 2, 3\} \to \{4, 5, 6\}$ that is injective but not surjective.
2. $f : \{1, 2, 3\} \to \{4, 5, 6\}$ that is surjective but not injective.
3. $f : \mathbb{R} \to \mathbb{R}$ that is injective but not surjective.
4. $f : \mathbb{R} \to \mathbb{R}$ that is surjective but not injective.
5. $f : \mathbb{R} \to \{1, 2, 3\}$ that is injective but not surjective.
6. $f : \mathbb{R} \to \{1, 2, 3\}$ that is surjective but not injective.

Proposition. The composition of two surjective functions is surjective.

Proof. Let $f : A \to B$ and $g : B \to C$ be surjective functions, then $\text{ran}(f) = \text{cod}(f) = B$ and $\text{ran}(g) = \text{cod}(g) = C$. For all $x \in C$, $g^{-1}(x) \in B$ and $f^{-1}(g^{-1}(x)) \in A$, so $C \subset \text{ran}(g \circ f)$, hence $g \circ f$ is surjective. □

Problem 3.10. Let $f : X \to Y$ be a surjective function. Prove that there exists an injective function $g : Y \to X$.

Problem 3.11. Let f and g be functions. Prove or disprove the following statements.

1. If f and g are injective, then $f \circ g$ is injective.
2. If f and g are bijective, then $f \circ g$ is bijective.
3. If f is surjective and g is injective, then $f \circ g$ is injective.

Definition 3.6. A set S is said to be *finite* if there exists a bijective function $f : S \to n$, where $n \in \mathbb{N}$. The *cardinality* of S, denoted $|S|$, is the number n. If S is not finite, we say S is *infinite*.

Problem 3.12. Let S and X be finite sets with $|S| = |X|$. Let $f : S \to X$ be a function. Prove that if f is injective, then f is surjective. Does the converse hold?

Problem 3.13. Let S be a finite set. Prove that there does not exist a surjective function $f : S \to \mathscr{P}(S)$.

Definition 3.7. Let $f : X \to Y$ be a function. Let $Z \subset X$, then the *restriction* of f onto Z is the map $f|_Z : Z \to Y$ defined by $f|_Z(z) = f(z)$ for all $z \in Z$.

Definition 3.8. Let A be a set. The *identity function*, denoted id_A, is the function $\text{id}_A(x) = x$ for all $x \in A$.

Problem 3.14. Let $f : A \to B$ be a function, prove that $f \circ \text{id}_A = f = \text{id}_B \circ f$.

Definition 3.9. Let $f : A \to B$ be a function. The function $g : B \to A$ is a *left inverse* of f if $g \circ f = \text{id}_A$. The function $h : B \to A$ is a *right inverse* of f if $f \circ h = \text{id}_B$. A function φ is called an *inverse* of f if it is both a left inverse and a right inverse of f.

Proposition. Let f be a function. If g is an inverse of f, then g is unique.

Proof. Suppose g and h are inverses of f, then $h = h \circ f \circ g = (h \circ f) \circ g = g$. □

Problem 3.15. Let f be a function with a left inverse g. Prove that if f has a right inverse, then the right inverse is g, conclude that g is the inverse of f.

Problem 3.16. Prove that a function has a left inverse if and only if it is injective. Prove that a function has a right inverse if and only if it is surjective, conclude that a function is bijective if and only if it has an inverse.

Definition 3.10. Let $\sim$ be an equivalence relation on a set X. The *quotient set*, denoted $X_\sim$, is the set $\{[x] \mid x \in X\}$.

Example. Let $S = \{a, b, c, d\}$. Let $\sim$ be an equivalence relation on S such that $a \sim b$ and $c \sim d$. The quotient set $S_\sim$ is $\{a, c\}$.

Problem 3.17. Consider the set of all integers $\mathbb{Z}$. Fix $n \in \mathbb{Z}$, Let $\sim_n$ be an equivalence relation on $\mathbb{Z}$ such that $a \sim_n b$ if and only if $a - b = kn$ for some $k \in \mathbb{Z}$. Prove that $\sim_n$ is indeed an equivalence relation. Find the quotient set $\mathbb{Z}_{\sim_2}$. Find the quotient set $\mathbb{Z}_{\sim_n}$ for all n. We also denote it

Definition 3.11. Let X and Y be sets. Let $\sim$ be an equivalence relation on X. A function $f : X \to Y$ is *invariant* under the equivalence relation such that, $x \sim y$ if and only if $f(x) = f(y)$.

Problem 3.18. Let X be a set and let $\sim$ be an equivalence relation on X. Prove that $\pi : X \to X_\sim$ defined by $x \mapsto [x]$ is a well-defined surjective function. Prove it is invariant under $\sim$.

Proposition. Let $f : X \to Y$ be a function, then f is an invariant fuction under some equivalence relation on X.

Proof. Define $\sim$ on X by $x \sim y$, where $x, y \in X$, if and only if $f(x) \sim f(y)$. For all $x = y$, $f(x) = f(y)$. For all $x = y = z$, $f(x) = f(y) = f(z)$. Hence $\sim$ is a well-defined equivalence relation on X, and f is trivially invariant under $\sim$. □

Definition 3.12. If $\varphi(x, p_1, p_2, \dots)$ is some formula, then we say $C = \{x \mid \varphi\}$ is a *class*, that is, x is a member of C if and only if x satisfies φ.

Remark. Although a class might not be a set, we extend the notion $\in$ to classes, that is, $x \in C$, where C is a class, means x satisfies the formula φ corresponding to C.

Example. The collection of all sets is a class.

Problem 3.19. Prove that every set is a class.

Axiom schema of replacement. If a class F is a funciton, then for all X, there exists a set $Y = F(X) = \{F(x) \mid x \in X\}$.

Problem 3.20. Prove that axiom schema of replacement implies axiom schema of separation.

Notice that the axiom schema of separation guarantees the existence of functions, so we need both axioms.

Definition 3.13. A set S is said to be *transitive* if for all $s \in S$, $s \subset S$. A set is an *ordinal* if it is transitive and well-ordered by $\in$. The *successor ordinal* of an ordinal α is defined to be $\alpha \cup \{\alpha\}$, and we denote it by $\alpha + 1$.

Problem 3.21. Prove that the element of an ordinal is an ordinal.

Proposition. The successor ordinal of an ordinal is an ordinal.

Proof. Let α be an ordinal. For all $s \in \alpha + 1$, if $s \in \alpha$, since α is an ordinal, $s \subset \alpha \subset \alpha + 1$; if $s \in \{\alpha\}$, then $s = \alpha \subset \alpha + 1$. Since α is well-ordered by $\in$, α is the $\in$-maximal element in $\alpha + 1$, hence $\alpha + 1$ is also well-ordered. □

Let α and β be ordinals. We define $\alpha < \beta$ if and only if $\alpha \in \beta$.

Example. The set $0 \in \mathbb{N}$ is an ordinal.

Problem 3.22. The intersection of ordinals is an ordinal.

Proposition. Every well-ordered set is isomorphic to an ordinal.

Proof. □

Definition 3.14. An ordinal λ is said to be a *limit ordinal* if $\lambda \neq \varnothing$ and for all $\alpha < \lambda$, $\alpha + 1 < \lambda$.

Problem 3.23. Prove that every limit ordinal is an ordinal.

Suppose there exists a set of all ordinals, denoted On. Since every element of an ordinal is an ordinal, On is transitive. Let "$<$" be the ordering on On, then it is trivially a partial ordering. Let S be a nonempty subset of On. Consider $\bigcap s_i \subset s_i$, where $s_i \in S$, if $T \notin S$, then $T \subset s_i$, so $T \subset s$. Now we have $T \in \bigcap s_i = T$, a contradiction. Hence $T \in S$ and S is T is trivially the desired infimum. This proves On is an ordinal, so On $\in$ On. Recall that On is a set, so On $\notin$ On, a contradiction. This is known as the *Burali-Forti paradox*.

Proposition. Every non-empty ordinal is either a successor ordinal or a limit ordinal.

Proof. Let α be an ordinal and $\alpha \neq \varnothing$. Suppose α is not a successor ordinal, then $\alpha \neq \beta + 1$ for all ordinal β. □

Theorem 3.1 (transfinite induction)**.** Let C be a class of ordinals that satisfies:

1. $0 \in C$;
2. if $\alpha \in C$, then $\alpha + 1 \in C$;
3. if α is a limit ordinal and $\beta \in C$ for all $\beta < \alpha$, then $\alpha \in C$.

Then C is the class of all ordinals.

Proof. □

Axiom of choice. Let X be a set and $X \neq \varnothing$. Then there exists a map $f : \mathscr{P}(X) \to X$ such that for every $x \subsetneq X$, where $x \neq \varnothing$, $f(x) \in x$.

We call such a function f a *choice function*. Therefore, we can reformulate axiom of choice as: for every nonempty set X, there exists a choice function.

Theorem 3.2 (well-ordering theorem)**.** Every set is well-orderable.

Proof. Let S be a set, by the axiom of choice, there exists a choice function $f : \mathscr{P}(S) \setminus \{\varnothing\} \to S$. Let an ordinal $\alpha = 0$ and let $s_0 = f(S)$, this is the basis for the transfinite induction. Assume s_β has been defined, where $\beta < \alpha$. If $S \setminus \{s_\beta\} = \varnothing$, then the map $f : \{s_\beta\} \to \alpha$ defined by $f(\beta) = s_\beta$ is a bijection, so S is well-ordered. If $S \setminus \{s_\beta\} \neq \varnothing$, define $s_\alpha = f(S \setminus \{s_\beta\})$. Suppose the process does not stop, then we have an injection from the class of all ordinal to the set S, by Burali-Forti paradox, this is a contradiction. Hence we have a bijection between α and S, so S is well-ordered. □

By the well-ordering theorem, $\mathbb{N}$ is well-ordered, this is called the *well-ordering principle*.

Remark. Well-ordering principle is based on ZF system, the system without axiom of choice. The well-ordering theorem is an equivalent form of axiom of choice. For more information, we recommend you to check [3].

4 Integers and Cardinality

Recall that we have constructed $\mathbb{N}$ by axioms. By the convention of quotient sets, we could further construct the integers and rational numbers.

Theorem 4.1 (recursion)**.** Given a set X and $x \in X$. Let $f : X \to X$ be a function, then there exists a unique function $F : \mathbb{N} \to X$ such that $F(0) = x$ and $F(n^+) = f(F(n))$ for all $n \in \mathbb{N}$.

Proof. Construct such F inductively. Suppose F is not well-defined and $F(n^+) = f(F(n)) = f(F(m))$ for some $n, m \in \mathbb{N}$, since f is well-defined, a contradiction. Suppose G is another function satisfies the property, then $F(0) = x = G(0)$. Assume $F(n) = G(n)$ for some $n \in \mathbb{N}$, then $F(n^+) = f(F(n)) = f(G(n)) = G(n^+)$. Hence, by induction, $F = G$. □

Definition 4.1. The *addition* on $\mathbb{N}$ is defined to be $\{+_n\}_{n\in\mathbb{N}}$ such that $+_n(0) = n$ and $+_n(m^+) = (+_n(m))^+$ for a fixed n and every $m \in \mathbb{N}$. We denote $+_n(m)$ by $n + m$.

Proposition. For all $n \in \mathbb{N}$, $n + 0 = n = 0 + n$.

Proof. It is trivial that $n + 0 = n$. Let $n = 0$, then $0 + 0 = 0$. Suppose $0 + k = k$ for some $k \in \mathbb{N}$, then $0 + (k^+) = (0 + k)^+ = k^+$. Hence, by induction, $n + 0 = n = 0 + n$. □

Problem 4.1. Prove that the following properties hold.

1. $n + 1 = n^+$ for all $n \in \mathbb{N}$.
2. $m + n = n + m$ for all $m, n \in \mathbb{N}$.
3. $a + (b + c) = (a + b) + c$ for all $a, b, c \in \mathbb{N}$.

Problem 4.2. Consider a relation $\sim_{\mathbb{N}+} \subset \mathbb{N} \times \mathbb{N}$ defined by $(a, b) \sim (c, d)$ if and only if $a + d = b + c$. Prove that $\sim_{\mathbb{N}+}$ is an equivalence relation.

Definition 4.2. The set of all *integers*, denoted $\mathbb{Z}$, is defined to be $(\mathbb{N} \times \mathbb{N})_{\sim_{\mathbb{N}+}}$. Let $[(a, b)] \in \mathbb{Z}$, the *inverse* of $[(a, b)]$ is defined to be $[(b, a)]$.

Problem 4.3. Prove that there exists a bijection between $\{[(a, b)] \mid a, b \in \mathbb{N} \text{ and } a \subsetneq b\}$ and $\{[(b, a)] \mid a, b \in \mathbb{N} \text{ and } b \subsetneq a\}$, conclude that the inverse of an integer is unique.

Let $n \in \mathbb{Z}$, the inverse of n is denoted by $-n$.

Definition 4.3. Let $[(a, b)], [(c, d)] \in \mathbb{Z}$. The *addition* on $\mathbb{Z}$ is defined to be $[(a, b)] + [(c, d)] = [(a + c, b + d)]$.

Proposition. The addition on $\mathbb{Z}$ is well-defined.

Proof. Suppose $(a, b) \sim (a', b')$ and $(c, d) \sim (c', d')$, then $a+b' = b+a'$ and $c+d' = d+c'$. We have $(a+c)+(b'+d') = a + b' + c + d' = b + a' + d + c' = (b + d) + (a' + c')$, hence $(a + c, b + d) \sim (a' + c', b' + d')$. □

Definition 4.4. Let $[(a, b)], [(c, d)] \in \mathbb{Z}$, the *multiplication* on $\mathbb{Z}$ is defined to be $[(a, b)] \cdot [(c, d)] = [(ac + bd, ad + bc)]$.

Problem 4.4. Prove that the multiplication on $\mathbb{Z}$ is well-defined.

Definition 4.5. Let $p \in \mathbb{Z}$. We say p is a prime if

Theorem 4.2 (fundamental theorem of arithmetic)**.** For all $n \in \mathbb{Z}$, n can be factored as a product of prime numbers and the factorization is unique.

Proof. □

Problem 4.5. Consider the relation $\sim_{\mathbb{Z}\times} \subset \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ defined by $(a, b) \sim (b, c)$ if and only if $ad = bc$. Prove that $\sim_{\mathbb{Z}\times}$ is an equivalence relation.

Definition 4.6. The *rational numbers*, denoted $\mathbb{Q}$, is defined to be the set $(\mathbb{Z} \times (\mathbb{Z} \setminus \{0\}))_{\sim_{\mathbb{Z}\times}}$. Let $[(a, b)], [(c, d)] \in \mathbb{Q}$, the *addition* is defined to be $[(a, b)] + [(c, d)] = [(ad + bc, bd)]$ and the *multiplication* on $\mathbb{Q}$ is defined to be $[(a, b)] \cdot [(c, d)] = [(ac, bd)]$.

Problem 4.6. Prove that addition and multiplication on $\mathbb{Q}$ are well-defined.

Problem 4.7. Let 0 be the element $[(0, x)] \in \mathbb{Q}$ and let 1 be the element $[(x, x)] \in \mathbb{Q}$ for any x. Verify the following algebraic properties of $\mathbb{Q}$.

1. $a + b = b + a$ and $ab = ba$ for all $a, b \in \mathbb{Q}$. (*commutativity*)
2. $a + b + c = a + (b + c)$ and $abc = a(bc)$ for all $a, b, c \in \mathbb{Q}$. (*associativity*)
3. $a + 0 = a$ for all $a \in \mathbb{Q}$. (*additive identity*)
4. $a \cdot 1 = a$ for all $a \in \mathbb{Q}$. (*multiplicative identity*)
5. For all $a \in \mathbb{Q}$, there exists a unique $-a \in \mathbb{Q}$ such that $a + (-a) = 0$. (*additive inverse*)
6. For all $a \in \mathbb{Q}$ and $a \neq 0$, there exists a unique $a^{-1} \in \mathbb{Q}$ such that $aa^{-1} = 1$. (*multiplicative inverse*)
7. $a(b + c) = ab + ac$. (*distributivity*)

Definition 4.7. Let $a, b \in \mathbb{Q}$ and $b \neq 0$. The *subtraction* on $\mathbb{Q}$ is defined to be $a - b = a + (-b)$. The *division* on $\mathbb{Q}$ is defined to be $a/b = a \cdot b^{-1}$.

Recall that we have defined $\leq$ on $\mathbb{N}$ by the natural inclusion, that is, for all $a, b \in \mathbb{N}$, $a \leq b$ if and only if $a \subset b$. Let $[(a, b)], [(c, d)] \in \mathbb{Z}$, we say $[(a, b)] \leq [(c, d)]$ if and only if $a + d < b + c$. Let $[(a, b)], [(c, d)] \in \mathbb{Q}$ with $b, d \geq 0$ and $b, d \neq 0$, then $[(a, b)] \leq [(c, d)]$ if and only if $ad \leq bc$.

Proposition. For all $a, b, c \in \mathbb{Q}$, the following properties hold.

1. $a + c = b + c$ implies $a = b$.
2. $a \cdot 0 = 0$.
3. $(-a)b = -ab$.
4. $(-a)(-b) = ab$.
5. $ac = bc$ and $c \neq 0$ imply $a = b$.
6. $ab = 0$ implies either $a = 0$ or $b = 0$.

Proof. (i) We have $a = a + (c + (-c)) = a + c + (-c) = b + c + (-c) = b + (c + (-c)) = b$. (ii) We have $a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0$, then $a \cdot 0 + 0 = a \cdot 0 + a \cdot 0$, so $a \cdot 0 = 0$. (iii) We have $ab + (-a)b = (a + (-a))b = 0b = 0$. Hence $(-a)b = -ab$. (iv) We have $(-a)(-b) + (-ab) = (-a)(-b) + (-a)b = (-a)(-b + b) = 0$. Hence $(-a)(-b) = ab$. (v) We have $a = a(cc^{-1}) = acc^{-1} = bcc^{-1} = b(cc^{-1}) = b$. (vi) Let $b \neq 0$, then $ab = 0 = 0b$, so $a = 0$. □

Problem 4.8. Prove that for all $a, b, c \in \mathbb{Q}$, the following properties hold.

1. If $a \leq b$, then $-b \leq -a$.
2. If $a \leq b$ and $c \leq 0$, then $bc \leq ac$.
3. If $0 \leq a$ and $0 \leq b$, then $0 \leq ab$.
4. $0 \leq a^2$.
5. If $0 < a$, then $0 < a^{-1}$.
6. If $0 < a < b$, then $0 < b^{-1} < a^{-1}$.

Now we consider the cardinality of infinite sets. We have constructed $\mathbb{Z}$ and $\mathbb{Q}$ based on $\mathbb{N}$, and all of those sets are considered to have the same cardinality.

Definition 4.8. A set S is said to be *countable* if S is finite or there exists a bijective function $f : S \to \mathbb{N}$. If a set is not countable, then we say the set is *uncountable*.

The function $\mathrm{id}_{\mathbb{N}} : \mathbb{N} \to \mathbb{N}$ is trivially bijective, so $\mathbb{N}$ is countable.

Definition 4.9. Let X and Y be sets. We say X has smaller cardinality than Y, denoted $|X| \leq |Y|$, if there exists an injection $f : X \to Y$.

Remark. Since $|X|$ can be infinite, the symbol "$\leq$" is not the typical ordering in $\mathbb{N}$.

Similarly, one could define $|X| \leq |Y|$ by surjections, that is, if there exists a surjection $f : Y \to X$, then $|X| \leq |Y|$.

Theorem 4.3 (Cantor's theorem)**.** Let S be a set. Then any function $f : S \to \mathscr{P}(S)$ is not bijective.

Proof. Suppose there exists a bijection $f : S \to \mathscr{P}(S)$.

□

Problem 4.9. Use Cantor's theorem, prove that for all set S, there exists some element $x \notin S$. Moreover, if S is finite, let $x \notin S$, prove that $|S \cup \{x\}| = |X| + 1$.

Proposition. Let X be a finite set and let $Y \subset X$, then Y is finite and $|Y| \leq |X|$.

Proof. Let $f : Y \to X$ defined by $f(y) = y \in Y \subset X$, this map is trivially injective. □

Proposition. Let X be a finite set and let $f : X \to Y$ be a function, then $\mathrm{im}(f)$ is finite and $|f(X)| \leq |X|$.

Proof. Consider the function $g : x \to f(x)$, □

Problem 4.10. Let X and Y be finite sets. Prove that $|X \cup Y| \leq |X| + |Y|$.

Proposition. The finite Cartesian product of countable set is countable.

Proof. Let $A_1, A_2, \ldots, A_n$ be countable sets, then each A_i itself is in bijection with $\mathbb{N}$. Suppose we pick n distinct prime numbers $p_1, p_2, \ldots, p_n$ and define $f : \prod_{i=1}^{n} \mathbb{N} \to \mathbb{N}$ such that $f(a_1, a_2, \ldots, a_n) = p_1^{a+1} \cdot p_2^{a+2} \cdots p_n^{a+n}$. Notice f is injective as the prime factorization of a number is unique. Hence $\prod_{i=1}^{n} \mathbb{N}$ is countable then $\prod_{i=1}^{n} A_i$ is countable. □

Proposition. The sets $\mathbb{Z}$ and $\mathbb{Q}$ are countable.

Proof. Since $\mathbb{N}$ is countable, $\mathbb{N} \times \mathbb{N}$ is countable. The natural projection $\mathbb{N} \times \mathbb{N} \to \mathbb{Z}$ is surjective, so $\mathbb{Z}$ is countable. Since $\mathbb{Z}$ is countable, $\mathbb{Z} \setminus \{0\} \subset \mathbb{Z}$, so $\mathbb{Z}$ is countable. Now $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ is countable, since $\mathbb{Q}$ is a quotient set of $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$, $\mathbb{Q}$ is countable. □

Theorem 4.4 (Cantor-Schröder-Bernstein theorem)**.** Let $f : A \to B$ and $g : B \to A$ be injective functions. Then there exists a bijective function $h : A \to B$.

Remark. Cantor-Schröder-Bernstein theorem was originally proved as a consequence of the axiom of choice. However, it is provable in ZF,

Proof. □

By Cantor-Schröder-Bernstein theorem, we conclude that if $|A| \leq |B|$ and $|B| \leq |A|$, then $|A| = |B|$.

5 Vector Spaces and Linear Maps

6 Functions as Morphisms

Functions are the bridges to "connect" sets. According to the definition, a function takes a set to another while preserving the structure of sets. This property allows us to consider functions as morphisms between sets.

Definition 6.1. A *category* C consists of:

1. a class, denoted $\mathrm{Ob}(\mathsf{C})$, of *objects*;
2. for each pair of objects X and Y, there exists a class of *morphisms* $f : X \to Y$, where X is called the *domain* and Y is called the *codomain*;
3. a *composition operation*, which gives, for each pair of morphisms $f : X \to Y$ and $g : Y \to Z$, a morphism $g \circ f : X \to Z$,

such that

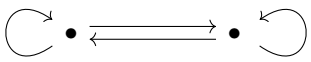
1. given any $f : X \to Y$, $g : Y \to Z$, and $h : Z \to W$, we have the identity $(h \circ g) \circ f = h \circ (g \circ f)$;
2. for each object X, there exists an *identity morphism* $\mathrm{id}_X : X \to X$ with the property that $f \circ \mathrm{id}_X = f$ and $\mathrm{id}_X \circ g = g$ for any $f : X \to Y$ and $g : Z \to X$.

The class of morphisms from X to Y is denoted by $\mathrm{Hom}_{\mathsf{C}}(X, Y)$. A category C is a *locally small category* if $\mathrm{Hom}_{\mathsf{C}}(A, B)$ is a set for all objects A and B.

Let sets be objects, let functions be morphisms, and let the composition of morphisms be the composition of functions. By our previous observations, it suffices to check this defines a category of sets, denoted Set. Moreover, Set is locally small.

Example. Let X be a set, then $\mathrm{Hom}_{\mathsf{Set}}(X, X)$ is a category.

Example. In the following diagram, let each vertex be an object and let each arrow be a morphism. This defines a category.



Problem 6.1. Morphisms are not guaranteed to be functions. Let $(S, \leq)$ be a partially ordered set. Let $\mathrm{Ob}(\mathsf{C}) = S$ and $x \to y$ be a morphism if $x \leq y$. Prove that C is a category.

Problem 6.2. Let C be a category. Let a system C^{op} consists of all objects in C and all morphisms $f : A \to B$ if $B \to A$ is a morphism in C. Prove that C^{op} is indeed a category. This is called the *dual category* of C.

Problem 6.3. Prove that the identity morphism is unique for each object A in a category C.

Definition 6.2. Let C be a category. A morphism $f : A \to B$ is a *monomorphism* if for all $g, h : C \to A$, $f \circ g = g \circ h$ implies $g = h$. A morphism $f : A \to B$ is called an *epimorphism* if for all $i, j : B \to D$, $i \circ f = j \circ f$ implies $i = j$.

Proposition. A function is injective if and only if it is a monomorphism in Set.

Proof. ($\Rightarrow$) Let $f : A \to B$ be injective. If $A = \varnothing$, then f is the unique empty function. If $A \neq \varnothing$, let $g : B \to A$ be the left inverse of f, ($\Leftarrow$) Let $f : A \to B$ be a monomorphism. If $A = \varnothing$, then $f : \varnothing \to B$ is vacuously injective. If $A \neq \varnothing$, □

Problem 6.4. Prove that a function is surjective if and only if it is an epimorphism in Set.

Definition 6.3. Let C be a category. A morphism $f : A \to B$ is an *isomorphism* if there exists $g \in \mathrm{Hom}_{\mathsf{C}}(B, A)$ such that $f \circ g = \mathrm{id}_B$ and $g \circ f = \mathrm{id}_A$. If there exists a morphism between two objects A and B, then we say they are *isomorphic*, denoted $A \approx B$.

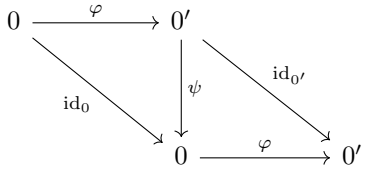
Problem 6.5. Prove that a morphism is an isomorphism if and only if it is a monomorphism and an epimorphism.

Problem 6.6. Let A and B be well-ordered sets that are isomorphic to each other. Prove that the isomorphism between them is unique.

Definition 6.4. An *initial object* 0 of a category C is an object in C such that for any object A, there is a unique morphism $0 \to A$. A *terminal object* 1 of a category C is an object of C such that for any object B, there is a unique morphism $A \to 1$. We say an object is a *zero object* if it is both an initial object and a terminal object.

Proposition. Initial objects of a category C are unique up to isomorphism.

Proof. Let 0 and $0'$ be initial objects in some category C. Let $\varphi : 0 \to 0'$ and $\psi : 0' \to 0$. Since the diagram is commutative, we have $\varphi \circ \psi \circ \varphi = \varphi \circ \mathrm{id}_0$, hence φ is an isomorphism. □



Problem 6.7. Prove that the initial object in a category C is a terminal object in C^{op}, conclude that terminal objects in C are unique up to isomorphism.

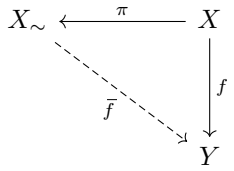
Problem 6.8. Prove that the initial object in the category of partially ordered set is the $\leq$-minimal element. Prove that the terminal object in this category is the $\leq$-maximal element.

Proposition. The initial object in Set is $\varnothing$ and the terminal object in Set is a singleton.

Proof. □

We conclude that there is no zero object in Set.

Universal property for quotient sets. Let X and Y be sets. Let $\sim$ be an equivalence relation on X and let $f : X \to Y$ be invariant under the equivalence relation. Then there exists a unique function $\overline{f} : X_\sim \to Y$ such that $f = \overline{f} \circ \pi$.



The proof of the universal property has two parts. We first verify that a quotient set has the universal property. Then we verify that a quotient set is characterized by this universal property, that is, any set satisfying the universal property must be a quotient set. This proof will be separated into multiple propositions.

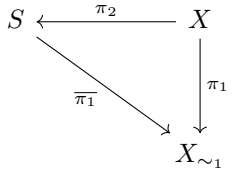
Proposition. There exists a map $\overline{f} : X_\sim \to Y$ such that $f = \overline{f} \circ \pi$.

Proof. Define the relation $\{([x], y)\} \in \overline{f}$ if and only if $f(x) = y$. Let $[x] \in X$ and $y, z \in Y$. Suppose $([x], y), ([x], z) \in \overline{f}$, then there exists $u, v \in X$ such that $[u] = [v] = [x]$, $f(u) = y$, and $f(v) = z$, so $u \sim v$. Since f is invariant under $\sim$, $y = z$. Hence $\overline{f}$ is well-defined. For all $x \in X$, $(\overline{f} \circ \pi)(x) = \overline{f}([x]) = f(y)$. □

Problem 6.9. Prove that such $\overline{f}$ is unique.

Problem 6.10. Let $f : X \to Y$ be a function that induces an equivalence relation $\sim_f$ on X. Since $X_{\sim_f}$ satisfies the universal property, prove that $X_{\sim_f} \approx \mathrm{im}(f)$.

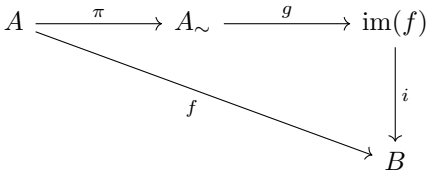
Proposition. Let S be any set satisfying the universal property for quotient sets. Let X be a set and let $\sim_1$ be an equivalence relation on X, so $\pi_1 : X \to X_{\sim_1}$ is invariant under the equivalence relation induced by $\pi_2 : X \to S$. Then $S \approx X_{\sim_1}$.



Proof. Since π_1 is surjective, $\overline{\pi_1}$ is surjective. Suppose $\overline{\pi_1}(s_1) = \overline{\pi_1}(s_2)$, where $s_1, s_2 \in S$. Since π_2 is surjective, there exist $x, y \in X$ with $s_1 = \pi_2(x)$ and $s_2 = \pi_2(y)$. Then $\overline{\pi_1}(s_1) = \overline{\pi_1}(\pi_2(x)) = \pi_1(x)$ and $\overline{\pi_1}(s_2) = \overline{\pi_1}(\pi_2(y)) = \pi_1(y)$, which implies $\pi_1(x) = \pi_1(y)$, so $[x] = [y]$. Since π_2 is invariant under $\sim_1$, it follows that $\pi_2(x) = \pi_2(y)$, so $s_1 = s_2$. Hence $S \approx X_{\sim_1}$. □

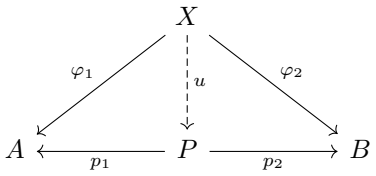
This completes our proof of the universal property for quotient sets.

Theorem 6.1 (canonical decomposition of functions)**.** Let $f : A \to B$ be a function, then there exists a surjective function π, a bijective function g, and an injective function i, such that $f = i \circ g \circ \pi$.



Proof. Let $i : \mathrm{im}(f) \to B$ be the identity map, then i is injective. The function $f : A \to B$ induces an equivalence relation $\sim$ on X, so g is bijective. The projection map $\pi : A \to A_\sim$ is surjective. For all $a \in A$, $(i \circ g \circ \pi)(a) = (i \circ g)([a]) = i(f(a)) = f(a)$. □

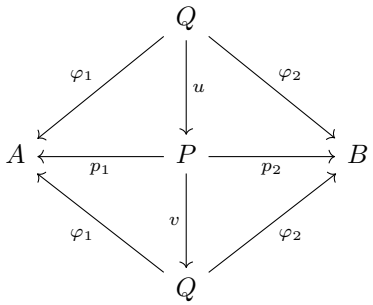
Definition 6.5. Let C be a category. Let A and B be objects in C. The *product* of A and B, denoted P, is an object in C with morphisms $p_1 : P \to A$ and $p_2 : P \to B$ such that for all object X in C with morphisms $\varphi_1 : X \to A$ and $\varphi_2 : X \to B$, there exists a unique $u : X \to P$ such that the following diagram commutes.



Problem 6.11. Prove that in Set, the Cartesian product satisfies the universal property for product.

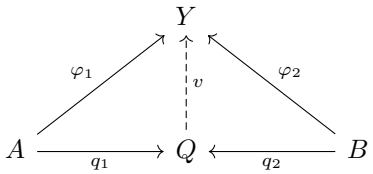
Proposition. Let A and B be objects in a category C. Then their product is unique up to isomorphism.

Proof. Let P with $p_1 : P \to A$ and $p_2 : P \to B$ be the product of A and B. Suppose Q with $\varphi_1 : Q \to A$ and $\varphi_2 : Q \to B$ is another product of A and B. By the universal property, $u : Q \to P$ and $v : P \to Q$ are unique such that $p_1 \circ u = \varphi_1$, $p_2 \circ u = \varphi_2$, $\varphi_1 \circ v = p_1$, and $\varphi_2 \circ v = p_2$. We have $\varphi_1 \circ u \circ v = \varphi_1$ and $\varphi_2 \circ u \circ v = \varphi_2$, then $u \circ v = \mathrm{id}_Q$. Similarly, $v \circ u = \mathrm{id}_P$. Hence $P \approx Q$. □



Problem 6.12. Let A, B, and C be sets. Use the universal property to prove that $A \times (B \times C) = (A \times B) \times C$.

Definition 6.6. Let C be a category. Let A and B be objects in C. The *coproduct* of A and B, denoted Q, is an object in C with morphisms $q_1 : A \to Q$ and $q_2 : B \to Q$ such that for all object Y in C with morphisms $\varphi_1 : A \to Y$ and $\varphi_2 : B \to Y$, there exists a unique $v : Q \to Y$ such that the diagram commutes.



Problem 6.13. Prove that in Set, the disjoint union satisfies the universal property for coproduct. Prove that the universal property characterizes disjoint union.

Problem 6.14. Describe the product and coproduct in $\mathsf{Set}^{\mathsf{op}}$.

7 Groups

We have shown a general constructions of mathematical entities use the convention of categories. Groups are the ne

Definition 7.1. A *monoid* is a triple $(M, \circ, e)$, where M is a set and $\circ : M \times M \to M$, called a *composition*, is a function such that:

1. for all $a, b, c \in M$, $a \circ (b \circ c) = (a \circ b) \circ c$; (associative)
2. there exists $e \in M$ such that for all $x \in M$, $e \circ x = x = x \circ e$. (identity)

Here e is called an *identity*.

Remark. We will use gf for $g \circ f$.

Example. The natural numbers $(\mathbb{N}, +, 0)$ is a monoid.

Example. Given a set S, the set of all functions $f : S \to S$ forms a monoid with the usual composition of function, denoted $M(S)$.

Problem 7.1. Prove that the identity element is unique in any monoid.

Definition 7.2. A triple $(G, \circ, e)$ is a *group* if it is a monoid and for all $g \in G$, there exists $h \in G$ such that $h \circ g = e = g \circ h$. Such f is called an *inverse* of g. If $\circ$ is commutative, then the group is *abelian*. The cardinality of a group is called its *order*, denoted $|G|$.

Example. The *trivial group* is the group of one element, that is, $\{e\}$.

Example. Consider the set D_n of all rotations and reflections that map a regular n-gon into itself, this is called the *dihedral group*. Let the reflection be S, which gives $S^2 = e$. Multiply the rotations by S on the left, then we have n distinct rotational symmetries. Hence there are n rotations and n reflections, that is, $|D_n| = 2n$.

Example. Let S be a set. All bijections $S \to S$ forms a group called the *symmetric group*, denoted $\mathrm{Sym}(S)$. If S is finite, then the symmetric group is denoted by $\mathfrak{S}_n$, where $|S| = n$. An element of the symmetric group is called a *permutation*. Let π be a permutation of $\mathfrak{S}_n$. A permuatation can be expressed by two-line notation, that is,

$$\pi = \begin{pmatrix} 1 & 2 & \cdots & n \\ \pi(1) & \pi(2) & \cdots & \pi(n) \end{pmatrix}.$$

The first line can be dropped to from an one-line notation. Given $1 \leq i \leq n$ and $i \in \mathbb{Z}$, we can write π in cycle notation, that is, $(i, \pi(i), \pi^2(i), \ldots, \pi^{p-1}(i))$, where p is the first integer such that $\pi^p(i) = i$. Such a cycle means that π sends i to $\pi(i)$, $\pi(i)$ to $\pi^2(i)$, and eventually, $\pi^{p-1}(i)$ back to i. The cycle type of a permutation is an expression of the form $(1^{m_1}, 2^{m_2}, \ldots, n^{m_n})$, where m_k is the number of cycles of length k in π. An *involution* is a permutation π such that $\pi^2 = e$.

Problem 7.2. Prove that every permutation in $\mathfrak{S}_n$ can be written as a finite composition of involutions.

Definition 7.3. A permutation $\pi \in \mathfrak{S}_n$ is called an *odd permutation* if it can be written as the composition of an odd number of involutions. A permutation $\pi \in \mathfrak{S}_n$ is called an *even permutation* if it can be written as the composition of an even number of involutions.

Proposition. Let G be a group. For all $g \in G$, the inverse of g is unique.

Proof. For all $g \in G$, let f and h be two inverses, then $f = f(gh) = (fg)h = h$, hence the inverse of g is unique. □

For all $g \in G$, where G is a group, the inverse of g is denoted by g^{-1}.

Problem 7.3. Let G be a group and let $g \in G$. A *left inverse* of g is an element $h \in G$ such that $hg = e$. Similarly, we can define a right inverse of g. Prove that a left inverse is equivalent to a right inverse, conclude that any left or right inverse of an element is its inverse.

Problem 7.4. Let G be a group and let $g, h \in G$. Prove that $(-e)g = -g$, $-(-g) = g$, and $-g(-h) = gh$.

Problem 7.5. Let G be a group and let $g_i \in G$, define the composition of finitely many elements by $\prod_{i=1}^{n} x_1 \cdots x_n = (x_1 \cdots x_{n-1})x_n$, prove that $\prod_{i=1}^{m} x_i \prod_{j=1}^{n} x_{m+j} = \prod_{j=1}^{m+n} x_j$. This is called the *generalized associativity*.

Problem 7.6. Prove that $|\mathfrak{S}_n| = n!$

Definition 7.4. Let G and H be monoids. A function $\varphi : G \to H$ is called a *monoid homomorphism* if $\varphi(ab) = \varphi(a)\varphi(b)$ and $\varphi(e_G) = e_H$ for all $a, b \in G$. Let G and H be groups, a function $\varphi : G \to H$ is called a *group homomorphism* if φ is a monoid homomorphism. The set $\{x \in G \mid \varphi(x) = e_H\}$ is called the *kernel* of φ, denoted $\ker(\varphi)$.

Problem 7.7. Let the objects be monoids and let the morphisms be monoid homomorphisms. Prove that this defines a category, denoted Mon. Let the objects be groups and let the morphisms be group homomorphisms. Prove that this defines a category, denoted Grp.

Problem 7.8. Prove that a group isomorphism is a bijective group homomorphism.

Definition 7.5. A *subgroup* of a group $(G, \circ, e)$ is a subset $H \subset G$ containing e such that $(H, \circ, e)$ is a group. If H is a subgroup of G, we write $H \leq G$.

Proposition. Let G be a group, then $H \leq G$ if and only if $e \in H \subset G$, for all $g, f \in H$, $gf^{-1} \in H$.

Proof. ($\Rightarrow$) Trivial. ($\Leftarrow$) Since $H \subset G$, the composition is associative. For all $g \in H$, take $f = g$, then $gg^{-1} = e$; take $f = e$, then $g^{-1} \in H$ for all g, hence $H \leq G$. □

Problem 7.9. If $A \leq B \leq G$, prove that $A \leq G$.

Problem 7.10. Let G be a group. Prove that all subgroups of G form a set. Prove that the set is partially ordered under $\subset$. This is called the *lattice of subgroups* of G.

Definition 7.6. Let $H \leq G$ be a subgroup. A *left coset* for H is a set of the form $xH = \{xh \mid h \in H\}$. A right coset is of the form $Hx = \{hx \mid h \in H\}$. The set of all left cosets for H is denoted by G/H and the set of all right cosets for H is denoted by $H \setminus G$. The *index* of H in G, denotd $[G : H]$, is defined to be $|G/H|$.

Problem 7.11. Prove that there exists a bijection $\psi : G/H \to H \setminus G$, conclude that $[G : H] = |H \setminus G|$.

Problem 7.12. Let $H \leq G$ be a subgroup. Prove that there exists a set of left cosets that partitions G.

Proposition. Let $H \leq G$ be a subgroup and let 1 be the trivial group, then $[G : 1] = [G : H][H : 1]$.

Proof. □

Let $H, K \leq G$ and let $H \subset K$, then $[G : 1] = [G : H][H : 1] =$

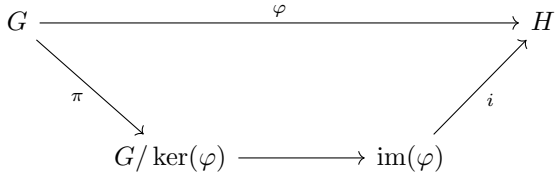
Theorem 7.1 (Cayley's theorem)**.** Any monoid is isomorphic to a submonoid of $M(S)$ for some set S. Any group is isomorphic to a subgroup of some symmetric group.

Proof. Let M be a monoid. For all $\alpha \in M$, let $\alpha_l(\alpha) = \alpha x$ for all $x \in M$, then α_l maps M to itself. Consider $S = \{\alpha_l \mid \alpha \in M\}$, which is a subset of $M(S)$. The identity map $\alpha_e \in S$. For all $\alpha, \beta \in M$, $\alpha_l(\beta_l(x)) = \alpha_l(\beta x) = \alpha\beta x = (\alpha\beta)x = (\alpha\beta)_l(x)$, so S is a submonoid of $M(S)$. Consider the map $\varphi(\alpha) = \alpha_l$. For all $\alpha, \beta \in M$, $\varphi(\alpha)\varphi(\beta) = \alpha_l\beta_l = (\alpha\beta)_l = \varphi(\alpha\beta)$. The map is trivially surjective. Let $\varphi(\alpha) = \varphi(\beta)$, then $\alpha_l = \beta_l$, that is, $\alpha x = \beta x$. Consider $x = 1$, then $\alpha = \beta$, hence φ is an isomorphism. Now consider a group G and construct the same set S. For all α_l, the inverse is $(\alpha^{-1})_l$. We have $\alpha_l(\alpha^{-1})_l(x) = \alpha_l(\alpha^{-1}x) = x$ and $(\alpha^{-1})_l\alpha_l(x) = (\alpha^{-1})_l(\alpha x) = x$, hence S is a subgroup of some symmetric group. Construct the same φ, then S is isomorphic to G. □

Definition 7.7. Let $H \leq G$ be a subgroup. We say H is a *normal subgroup* of G, denoted $H \trianglelefteq G$, if $xH = Hx$ for all $x \in G$. A non-trivial group G is said to be *simple* if the only normal subgroups of G are G and the trivial subgroup.

Problem 7.13. Prove that $H \trianglelefteq G$ if and only if $gHg^{-1} = H$ for all $g \in G$.

Theorem 7.2 (first isomorphism theorem)**.** Let $\varphi : G \to H$ be a group homomorphism. Then $\ker(\varphi) \trianglelefteq G$, $\mathrm{im}(\varphi) \leq H$, and $G/\ker(\varphi) \approx \mathrm{im}(\varphi)$.



Proof. Let $a, b \in \ker(\varphi)$, then $\varphi(ab^{-1}) = \varphi(a)\varphi(b) = \varphi b^{-1}$. We have $e_H = \varphi(bb^{-1}) = e_H\varphi(b^{-1})$, then $\varphi(b^{-1}) = e_H$, which implies $\ker(\varphi) \leq G$. For all $g \in G$, $\varphi(g\ker(\varphi)g^{-1}) = \varphi(g)\varphi(\ker(\varphi))\varphi(g^{-1}) = e_H$, then $g\ker(\varphi)g^{-1} \subset \ker(\varphi)$. For all $g \in \ker(\varphi)$, we have □

Problem 7.14. Let C_2 be a group of order 2. Prove that C_2 is unique up to isomorphism. Define the set $C_2 = \{1, -1\}$. Let 1 be the identity. For all $\pi \in \mathfrak{S}_n$, define $\varphi : \mathfrak{S}_n \to C_2$ by $\varphi(\pi) = -1$ if π is an odd permutation and $\varphi(\pi) = 1$ if it is an even permutation. Prove that sgn is a well-defined group homomorphism. The kernel of φ is called the *alternating group*, denoted $\mathfrak{A}_n$. Prove that $\mathfrak{A}_n$ is abelian if $n \leq 3$ and prove that $\mathfrak{A}_n$ is not abelian for $n > 3$, conclude that a normal subgroup is not necessarily abelian.

Now we show some categorical properties of Grp.

Problem 7.15. Prove that in Grp, the product of two groups G and H is the Cartesian product $G \times H$.

Proposition. In Grp, the trivial group 1 is the zero object.

Proof. □

Definition 7.8. Let C be a category. A category D is a *subcategory* of C if the Ob(D) and the $\mathrm{Hom}_{\mathsf{D}}(X,Y)$ are subcollections of Ob(C) and $\mathrm{Hom}_{\mathsf{C}}(X,Y)$, respectively, for all objects X and Y in C. The subcatgeory D is said to be *full* if for all objects X and Y in D, $\mathrm{Hom}_{\mathsf{D}}(X,Y)$ is exactly $\mathrm{Hom}_{\mathsf{C}}(X,Y)$.

Example. The category Grp is a subcategory of Set.

Problem 7.16. Let the objects be abelian groups, let the morphisms be group homomorphisms, and let the composition of morphisms be the composition of functions. Prove that this defines a category, denoted Ab. Prove that Ab is a full subcategory of Grp.

Definition 7.9. Let C be a category with zero object 0. The *zero morphism* $0_{A,B}$ between objects A and B is the unique morphism that factors through 0.

Let G and H be groups. Since 0 is the zero object, there exists unique group homomorphisms $\varphi : G \to 0$ and $\psi : 0 \to H$, $\psi \circ \varphi$ factors through the zero morphism $0_{G,H}$ and it is the desired zero morphism.

Definition 7.10. Let C be a category and let $f : X \to Y$ be a morphism in C. An object $\ker(f)$ is said to be the *kernel* of f if for every object Z and $h : Z \to X$ such that $f \circ g = 0$, where 0 is the zero morphism, there is a unique morphism $\varphi : Z \to \ker(f)$ such that $h = g \circ \varphi$.

Problem 7.17. Prove the set-theoretic definition of kernels coincide with the categorical definition of kernels.

Proposition. A group homomorphism is an epimorphism if and only if it is surjective.

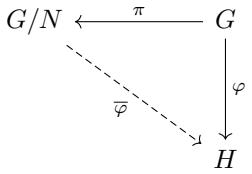
Proof. ($\Rightarrow$) Let $f : H \to K$ be an epimorphism. Let $X = K/f(H)$ be the set of right cosets of $f(H)$ in K. Let α not in X. Consider the set $Y = X \cup \{\alpha\}$

($\Leftarrow$) Let $f : G \to H$ be surjective, so f is an epimorphism in Set, since Grp is a subcategory of Set and f is a group homomorphism, f is a group epimorphism. □

Definition 7.11. Let G be a group and $N \trianglelefteq G$. The *quotient group* G/N is defined to be $\{aN \mid a \in G\}$.

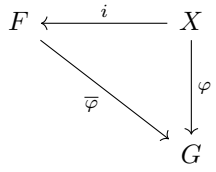
Problem 7.18. Prove that $(aN)(bN) = (ab)N$. Take this as a composition on G/N, prove that G/N is indeed a group. Prove that $\pi : G \to G/N$ defined by $\pi(g) = gN$ is a group epimorphism.

Universal property for quotient groups. Let G be a group and let $N \trianglelefteq G$. Let $\pi : G \to G/N$ be the quotient epimorphism. For all group homomorphism $\varphi : G \to H$ with $N \subset \ker(\varphi)$, there exists a unique group homomorphism $\overline{\varphi} : G/N \to H$ such that $\overline{\varphi} \circ \pi = \varphi$.



Problem 7.19. Prove that the quotient groups satisfy the universal property and the universal property characterizes quotient groups.

Universal property for free groups. Let X be a subset of a group F. Then F is a *free group* with *basis* X if for any function $\varphi : X \to G$, where G is a group, there exists a unique extension $\overline{\varphi} : F \to G$.



We will first show the existence of free groups. Consider

Problem 7.20. Let X_1 and X_2 be bases of free groups F_1 and F_2, respectively. Prove that if $X_1 \approx X_2$ as sets, then $F_1 \approx F_2$ as groups.

Definition 7.12. A *group presentation* is a pair (S, R) consisting of a set S and a subset $R \subset F(S)$. The group presented by (S, R) is defined to be $\langle S \mid R\rangle = F(S)/N$, where N is the smallest normal subgroup of $F(S)$ containing R.

Definition 7.13. Let G be a group with the presentation $\langle S \mid R\rangle$, then the elements are called the *generators* of G. A group is said to be *cyclic* if $|S| = 1$.

Example. The presentation of a free group $F(S)$ is $\langle S \mid \varnothing\rangle$.

Problem 7.21. Prove that $\langle a, b \mid abaa^{-1}b^{-2}, bab^{-1}a^{-2}\rangle$ is the trivial group.

Group presentation of a group is not unique.

Problem 7.22. Prove that a finite cyclic group of order n admits the presentation $\langle a \mid a^n\rangle$. We denote it by C_n.

Problem 7.23. A *Tarski monster group* is a finitely generated infinite group, where every proper non-trivial subgroup is cyclic of order a fixed prime p. Prove that every Tarski monster group is simple.

Problem 7.24. A *Baumslag-Solitar group* $BS(m, n)$ admits the presentation $\langle a, b \mid ba^m b^{-1} = a^n\rangle$. Prove that $BS(1, 1) \approx C_2$. Prove that $\varphi : BS(2, 3) \to BS(2, 3)$ defined by $\varphi(a) = a^2$ and $\varphi(b) = b$ is a group epimorphism.

Definition 7.14.

8 Rings

Definition 8.1. A triple $(G, \circ, e)$ is a *group* if it is a monoid and for all $g \in G$, there exists $h \in G$ such that $h \circ g = e = g \circ h$. Such f is called an *inverse* of g. If $\circ$ is commutative, then the group is *abelian*.

Definition 8.2. A *ring* is an abelian group $(R, +, 0)$ with the operation $\cdot$ satisfies the following properties.
1. For all $a, b, c \in R$, $a \cdot (b \cdot c) = (a \cdot b) \cdot c$. (associativity)
2. For all $a, b, c \in R$, $a \cdot (b + c) = a \cdot b + a \cdot c$ and $(a + b) \cdot c = a \cdot c + b \cdot c$. (distributivity)
3. There exists $1_R \in R$ such that $1R \cdot a = a = a \cdot 1_R$ for all $a \in R$. We call 1_R the *multiplicative identity* of R.

Definition 8.3 (unit)**.** An element u of a ring R is a unit if there exists $v \in R$ such that $u \cdot v = 1_R = v \cdot u$.

Remark. A ring is a data $(R, +, \cdot)$ consisting of a set R, with two binary operations $+, \cdot$. R is an abelian group under $+$. We use $R^\times$ to represent units of R.

Definition 8.4 (commutative ring)**.** A ring R is commutative if and only if for all $a, b \in R$, $a \cdot b = b \cdot a$.

Example (ring of functions)**.** Let X be a set, R be a ring, R^X is the set of all functions from X to R. R^X is a ring. For any $f, g \in R^X$, addition and multiplication are defined as

$$\begin{aligned} + : &\quad (f + g)(x) = f(x) + g(y) \\ \cdot : &\quad (f \cdot g)(x) = f(x) \cdot g(x) \end{aligned}$$

for all $x \in X$.

Definition 8.5 (polynomial ring)**.** Let R be a commutative ring. Define $R[x] = \{a_0 + a_1 x + ...a_n x^n | n \geq 0, a_i \in R\}$. Assuming $n > m$, we define addition and multiplication as follows:

$$\begin{aligned} + : &\quad \sum_{i=0}^{n} a_i x^i + \sum_{i=0}^{n} b_i x^i = \sum_{i=0}^{n} (a_i + b_i) x^i \\ \cdot : &\quad \left(\sum_{i=0}^{n} a_i x^i\right)\left(\sum_{j=0}^{m} b_j x^j\right) = \sum_{i=0}^{n+m} \left(\sum_{i+j=k} a_i b_j x^k\right) \end{aligned}$$

Problem 8.1. Prove that a polynomial ring is indeed a ring. What are 0 and 1 in $R[x]$?

Definition 8.6 (degree)**.** The degree of $0 \neq p(x) = a_0 + a_1 x + ... + a_n x^n \in R[x]$ is $\deg(p) = \max\{n | a_n \neq 0\}$. If $p(x) = 0$, $\deg(p) = -\infty$.

Definition 8.7 (field)**.** A commutative (non-zero) ring F in which any non-zero element is a unit (i.e. for all $a \in F$, there exists $b \in F$ such that $a \cdot b = 1 = b \cdot a$).

Lemma. For any ring R, any $a \in R$ satisfies $a \cdot 0 = 0 = 0 \cdot a$.

Proof. $0 = 0 + 0$ implies $a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0$. Since R is also a group, there exists an additive inverse $-a \cdot 0 \in R$. We can add $-a \cdot 0$ to both sides:

$$a \cdot 0 - a \cdot 0 = (a \cdot 0 + a \cdot 0) - a \cdot 0 = a \cdot 0 + (a \cdot 0 - a \cdot 0) \iff 0 = a \cdot 0$$

Doing a symmetric argument to $0 \cdot a = (0 + 0) \cdot a = 0 \cdot a + 0 \cdot a$, we conclude that $a \cdot 0 = 0 = 0 \cdot a$. □

Corollary. If R is a ring and $0 = 1$, then $R = \{0\}$, we call $\{0\}$ the *zero ring*.

Proof. For all $a \in R$, $a = a \cdot 1 = a \cdot 0 = 0$. □

Proposition. Let R be a commutative ring. If there exist nonzero $x, y \in R$ such that $x \cdot y = 0$, then $x, y \notin R^\times$.

Proof. Assume x is a unit. Then there exists some $x^{-1} \in R$ and $0 = x^{-1} \cdot 0 = x^{-1} \cdot (x \cdot y) = (x^{-1} \cdot x) \cdot y = 1 \cdot y = y$. This contradicts the assumption that $y \neq 0$. Therefore, x is not a unit. By a similar argument, we can show y is also not a unit. □

Lemma. Let R be a commutative ring. For any $f, g \in R[x]$, $\deg(f \cdot g) \leq \deg(f) + \deg(g)$.

Proof. If f or g is 0, then $f \cdot g = 0$. And $\deg(f \cdot g)$ is $-\infty$. $\deg(f) + \deg(g)$ is either $-\infty$ or $-\infty$ plus something. If $f, g \neq 0$, then let $f = \sum_{i=0}^{n} a_i x^i$ and $g = \sum_{j=0}^{m} b_j x^j$. Their product would be $\sum_{i=0}^{n+m} \left(\sum_{i+j=k} a_i b_j x^k \right)$. If $a_n, b_m \neq 0$, then $\deg(f \cdot g) = n + m$ which is equal to $\deg(f) + \deg(g)$. Otherwise $\deg(f \cdot g) < n + m = \deg(f) + \deg(g)$. □

Remark. $0(x) = 0_R, \forall x \in X$ is the zero function. Constant function $1(x) = 1_R, \forall x \in X$ is the unity.

Definition 8.8 (subring)**.** A subset S of a ring R is a subring if and only if
1. S is a subgroup of $(R, +, 0)$.
2. for all $a, b \in S$, $ab \in S$. This means S is closed under multiplication.

Definition 8.9 (ring homomorphism)**.** Let R, R' be two rings, a map $f : R \to R'$ is a ring homomorphism if and only if f preserves both $+$ and $\cdot$:

$$f(a + b) = f(a) + f(b)$$
$$f(ab) = f(a)f(b)$$

for all $a, b \in R$.

Definition 8.10 (unital homomorphism)**.** A ring homomorphism $f : R \to R'$ is unital if $f(1_R) = 1_{R'}$.

Definition 8.11 (ideal)**.** An ideal I in a ring is a subgroup of $(R, +, 0)$ so that ① $I \neq \emptyset$ and ② for all $a, b \in I, ab^{-1} = a - b \in I$. Or equivalently, for any $r \in R$ and $i \in I$, $ri \in I$.

9 Real Numbers

Definition 9.1 ($\mathbb{Q}$ norm)**.** $|a|$ is a norm on $\mathbb{Q}$: $|a| = \begin{cases} a & (a \geq 0) \\ -a & (a < 0) \end{cases}$ that satisfies

1. $|a| \geq 0$.
2. $|a| = 0$ implies $a = 0$.
3. $|ab| = |a||b|$.
4. Triangle inequality: $|a + b| \leq |a| + |b|$.

Definition 9.2 (convergence)**.** A sequence $\{a_n\}_{n=1}^{\infty}$ of rationals converges to $a \in \mathbb{Q}$ if for all $r \in \mathbb{Q}^+$, there exists $N \in \mathbb{N}$, such that for all $n \geq N$, $|a_n - a| < r$. A sequence $\{a_n\}$ converging to a is denoted as $\{a_n\} \to a$. a is called the limit (in rational) of $\{a_n\}$. In other words, the sequence $\{a_n\}$ can get arbitrarily close to a. A bad example would be $\{3.1, 3.14, 3.141, 3.1415...\} \to \pi \notin \mathbb{Q}$.

Definition 9.3 (Cauchy sequence)**.** A sequence $\{a_n\}_{n=1}^{\infty}$ of rationals is Cauchy if for all $r \in \mathbb{Q}^+$, there exists $N \in \mathbb{N}$ such that for all $n, m \geq N$ and $|a_n - a_m| < r$. As the sequence proceeds, the difference between two elements gets smaller and smaller.

Proposition. If $\{a_n\}$ converges in $\mathbb{Q}$, then $\{a_n\}$ is Cauchy.

Proof. Take some arbitrary $r \in \mathbb{Q}^+$. Since $\{a_n\} \to a$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$, $|a_n - a| < r/2$ and for all $m \geq N$, $|a_m - a| < r/2$. Then

$$\begin{aligned} |a_n - a_m| &= |a_n - a + a - a_m| = |(a_n - a) + (a - a_m)| \\ &\leq |a_n - a| + |a_m - a| && \text{(triangle inequality)} \\ &< \frac{r}{2} + \frac{r}{2} < r \end{aligned}$$

Therefore, $\{a_n\}$ is Cauchy. □

Remark. The converse of the above proposition is false.

Definition 9.4 (bounded)**.** A rational sequence $\{a_n\}$ is bounded if $S = \{a_n \mid n \in \mathbb{N}\}$ is bounded in $\mathbb{Q}$: If S is bounded, then there exist $a, b \in \mathbb{Q}$ such that for all $a_n \in S$, $a \leq a_n$ and $b \geq a_n$.

Lemma. If $\{a_n\}$ is a Cauchy sequence in $\mathbb{Q}$, then it is bounded in $\mathbb{Q}$.

Proof. Let $r = 347$. If the sequence $\{a_n\}$ is cauchy, then there exists $N \in \mathbb{N}$ such that for all $m, n \geq N$, $|a_n - a_m| < 347$. If $m = N$, then for all $n \geq N$, $|a_n - a_N| < 347$. By definition, showing $\{a_n\}$ is bounded is the same as showing $|a_n| \leq M \in \mathbb{Q}$ for all n:

$$\begin{aligned} |a_n| = |a_n - a_N + a_N| &\leq |a_n - a_N| + |a_N| \\ \Rightarrow |a_n| &\leq |a_n - a_N| + |a_N| \end{aligned}$$

Since $|a_n - a_N|$ is strictly smaller than 347, we can simply replace $|a_n - a_N|$ with 347. When $n \geq N$, $|a_n| < 347 + |a_N|$. Let $U = \max\{|a_1|, |a_2|, ..., |a_{N-1}|, 347 + |a_N|\}$. We have that $|a_n| \leq U$ for all $n \in \mathbb{N}$. From there:

① If $n \geq N$, then use $347 + |a_N|$ as the bound.

② If $n < N$, then there finitely many values, and they are bounded by $\max\{|a_1|, |a_2|, ..., |a_{N-1}|\}$.

□

Proposition. The set of all rational Cauchy sequences is a commutative ring with 1.

Proof. We prove this proposition by checking ring axioms one by one. All indices are natural numbers.

A1 Assume $\{a_n\}$, $\{b_n\}$ are Cauchy. Fix $r \in \mathbb{Q}^+$. Since $\{a_n\}$ is Cauchy, there exists N_1 such that for all $n, m \geq N_1$, $|a_n - a_m| < r/2$. Since $\{b_n\}$ is Cauchy, there exists N_2 such that for all $n, m \geq N_2$, $|b_n - b_m| < r/2$. Let $N = \max(N_1, N_2)$. For all $n, m \geq N$,

$$|(a_m + b_m) - (a_n + b_n)| = |a_m - a_n + b_m - b_n| \leq |a_m - a_n| + |b_m - b_n| < r/2 + r/2 = r$$

Therefore, $\{a_n\} + \{b_n\}$ is Cauchy.

A2 $\{a_n\} + (\{b_n\} + \{c_n\}) = \{a_n + (b_n + c_n)\} = \{(a_n + b_n) + c_n\} = (\{a_n\} + \{b_n\}) + \{c_n\}$.

A3 $\{a_n\} + \{b_n\} = \{a_n + b_n\} = \{b_n + a_n\} = \{b_n\} + \{a_n\}$.

A4 There exist a zero sequence $\{0\}$ such that $\{a_n\} + \{0\} = \{a_n + 0\} = \{a_n\}$ and $\{0\} + \{a_n\} = \{0 + a_n\} = \{a_n\}$.

A5 For any $\{a_n\}$, there exist $-\{a_n\} = \{-a_n\}$ such that $\{a_n\} + (-\{a_n\}) = \{a_n - a_n\} = \{0\}$ and $(-\{a_n\}) + \{a_n\} = \{-a_n + a_n\} = \{0\}$.

M1 Let $\{a_k\}$ and $\{b_k\}$ be Cauchy sequences in $\mathbb{Q}$. Then

$$\begin{aligned} |a_n b_n - a_m b_m| &= |a_n b_n - a_n b_m + a_n b_m - a_m b_m| \leq |a_n||b_n - b_m| + |b_m||a_n - a_m| \\ &\leq A|b_n - b_m| + B|a_n - a_m| \end{aligned}$$

where A and B are upper bounds for sequences $\{|a_k|\}$ and $\{|b_k|\}$. Since $\{|a_k|\}$ and $\{|b_k|\}$ are Cauchy sequences, for all $\epsilon(2A)^{-1}, \epsilon(2B)^{-1} > 0$, there exists $N \in \mathbb{N}$ such that for all $m, n \geq N$, $|a_m - a_n| < \epsilon(2B)^{-1}$ and $|b_m - b_n| < \epsilon(2A)^{-1}$. So $|a_m b_m - a_n b_n| < A\epsilon(2A)^{-1} + B\epsilon(2B)^{-1} = \epsilon$. Therefore, $\{a_n\}\{b_n\}$ is Cauchy.

M2 $\{a_n\} \cdot (\{b_n\} \cdot \{c_n\}) = \{a_n \cdot (b_n \cdot c_n)\} = \{(a_n \cdot b_n) \cdot c_n\} = (\{a_n\} \cdot \{b_n\}) \cdot \{c_n\}$.

M3 $\{a_n\} \cdot \{b_n\} = \{a_n \cdot b_n\} = \{b_n \cdot a_n\} = \{b_n\} \cdot \{a_n\}$.

M4 There exist $\{1\}$ such that $\{a_n\} \cdot \{1\} = \{a_n \cdot 1\} = \{a_n\}$ and $\{1\} \cdot \{a_n\} = \{1 \cdot a_n\} = \{a_n\}$.

D $\{a_n\} \cdot (\{b_n\} + \{c_n\}) = \{a_n \cdot (b_n + c_n)\} = \{a_n \cdot b_n + a_n \cdot c_n\} = \{a_n\} \cdot \{b_n\} + \{a_n\} \cdot \{c_n\}$.

□

Notation. The ring of rational Cauchy sequences is denoted as $\mathcal{C}$.

Remark. Is $\mathcal{C}$ an integral domain? That is: does $ab = 0$ imply $a = 0$ or $b = 0$? **No**, a counter example would be $a = (0, 0.1, 0, 0.01, 0, 0.001....)$ and $b = (0.1, 0, 0.01, 0, 0.001, 0....)$, but $ab = (0, 0, 0, 0, ...)$.

Proposition. Let $I := \{\text{Cauchy sequences that converge to } 0\}$. $I \subset \mathcal{C}$ is an ideal.

proof sketch. Let I be the set of sequences $\{a_k\}$ in $\mathcal{C}$ with the property that, given any rational $r > 0$, there exists an integer N such that if $n \geq N$, then $|a_n| < r$. For any $c \in \mathcal{C}$ and $i \in I$, the sequence ci can be eventually scaled down to 0 because c is bounded. For any $i' \in I$, obviously $i + i' \in I$. □

Proposition. If $\{a_k\} \in \mathcal{C} \setminus I$, then there exist a positive rational number r and an integer N so that $|a_k| \geq r$ for all $n \geq N$. In other words, if a Cauchy sequence $\{a_k\}$ does not converge to 0, then after certain point $n \geq N$ the sequence stays at least r distance away from 0.

Proof. If $\{a_k\} \nrightarrow 0$, then there exists $r \in \mathbb{Q}^+$ such that for all $N \in \mathbb{N}$ there exists $k \geq N$ such that $|a_k| \geq 2r$ (the negation of converging to 0: there are some r that a_k is $2r$ away from 0). Since $\{a_k\}$ is Cauchy, for all $n, m \geq \mathbb{N}$ there exists $N \in \mathbb{N}$ such that $|a_n - a_m| < r$.

$$\begin{aligned} &\Rightarrow |a_m - a_n| < r && (\text{ property of absolute value}) \\ &\Rightarrow |a_m| - |a_n| \leq |a_m - a_n| < r && (\text{reverse triangle inequality}) \\ &\Rightarrow -|a_n| < -|a_m| + r \Rightarrow |a_n| > |a_m| - r \end{aligned}$$

Now fix $m \geq N$ and let $|a_m| \geq 2r$: $|a_n| > |a_m| - r \geq 2r - r$. This implies $|a_n| > r$ and holds for all $n \geq N$. □

Corollary. if a Cauchy sequence does not converge to 0, all terms of the sequence eventually have the same sign.

Proof. If $\{a_k\}$ is a Cauchy sequence that doesn't converge to 0, then $\{a_k\} \notin I$. By the above proposition, there exists some $r > 0$ and N so that $|a_k| \geq r$ for all $n \geq N$. Since $\{a_k\}$ is Cauchy and $r > 0$, there exists M such that if $m, n \geq M$, then $|a_n - a_m| < r$. Assume without loss of generality that there is some $k > \max(M, N)$ such that a_k is positive. Fix a_k, pick any $n \geq k$ from the sequence:

① If $a_n > a_k$, then we are done (a_n is for sure positive).

② If $a_n < a_k$, then $a_k - a_n < r$ which implies a_n is positive for all n.

□

Definition 9.5 (equivalence classes of Cauchy sequences)**.** Let $\{a_k\}$ and $\{b_k\}$ be Cauchy sequences in $\mathbb{Q}$. We say that $\{a_k\}$ is equivalent to $\{b_k\}$, denoted by $\{a_k\} \sim \{b_k\}$, if $\{c_k\} = \{a_k - b_k\}$is in I.

Proposition. The "$\sim$" defined above gives an equivalence relation.

Proof. We need to check "$\sim$" is reflexive, symmetric, and transitive.

Reflexive: For all $\{a_k\}$, $\{a_k\} - \{a_k\} = \{0\} \in I$.

Symmetric: If $\{a_k - b_k\} \in I$, then for all $\epsilon > 0$, there exists some $N \in \mathbb{N}$ such that for all $n \geq N$, $|a_n - b_n - 0| < \epsilon$. This means that $|b_n - a_n - 0| < \epsilon$, so $\{b_k - a_k\} \in I$.

Transitive: If $\{a_k - b_k\} \in I$, then for all $\epsilon/2 > 0$, there exists $N_1 \in \mathbb{N}$ such that for all $n \geq N_1$, $|a_n - b_n - 0| < \epsilon/2$. If also $\{b_k - c_k\} \in I$, then for all $\epsilon > 0$, there exists $N_2 \in \mathbb{N}$ such that for all $n \geq N_2$, $|b_n - c_n - 0| < \epsilon/2$. Let $N = \max(N_1, N_2)$. Then for all $n > N$, $|a_n - b_n - 0| < \epsilon/2$ and $|b_n - c_n - 0| < \epsilon/2$. It follows that $|a_n - b_n - 0 + b_n - c_n - 0| = |a_n - c_n - 0| < \epsilon$. So $\{a_k - c_k\} \in I$.

□

Definition 9.6 (set of equivalence classes)**.** Define $\mathbf{R} := \mathcal{C}/I$ as the set of all equivalence classes in $\mathcal{C}$.

Proposition. Term-wise addition and multiplication are well-defined on $\mathbf{R}$: If $\{a_k\}$ is a Cauchy sequence, denote its equivalence class by $[a_k]$, then the sum and product of equivalence classes are defined as $[a_k] + [b_k] = [a_k + b_k]$ and $[a_k][b_k] = [a_k b_k]$.

Proposition. Assume $\mathbf{R} = \mathcal{C}/I$ is a commutative ring[1]. It turned out $\mathbf{R}$ is a field.

Proof. We only need to show that multiplicative inverses exist for nonzero elements of $\mathbf{R}$. Consider some arbitrary $[a_k] \neq I$. By Proposition ???, since $[a_k] \nrightarrow 0$, the absolute value of $[a_k]$ is eventually bounded below. In other words, there exists $N \in \mathbb{N}$ and $r > 0$, such that $|a_n| > r$ for all $n \geq N$. Therefore, we define $[b_k]$ as follows:

$$\{b_k\} = \begin{cases} 1 & (k < N) \\ \frac{1}{a_k} & (k \geq N) \end{cases}$$

We include the $k < N$ case since a_k may equal to 0 when $k < N$. Now show $\{b_k\}$ is cauchy. Take some arbitrary $\epsilon > 0$. Since $\{a_k\}$ is cauchy, there exists M such that for all $n, m \geq M, |a_n - a_m| < \epsilon r^2$. Moreover,

$$|b_n - b_m| = |\frac{1}{a_n} - \frac{1}{a_m}| = \frac{|a_n - a_m|}{|a_n a_m|} \leq \frac{1}{r^2}|a_n - a_m| \leq \frac{1}{r^2}\epsilon r^2 = \epsilon$$

This implies that $\{b_k\}$ is Cauchy. Furthermore, $\{a_k\}\{b_k\} \to 1$, which means that $\{b_k\}$ is the multiplicative inverse of $\{a_k\}$. Therefore, $\mathbf{R}$ is a field. □

[1]This is not hard to show. But for right now, we do not have the construction needed to prove this fact.

Definition 9.7 (order of **R**)**.** Let $a = [a_k]$,$b = [b_k]$ be distinct elements of **R**. We define $a < b$ if $a_k < b_k$ eventually and $b < a$ if $b_k < a_k$ eventually.

Definition 9.8 (order axioms)**.** Here are the order axioms

O1 Trichotemy: Either $a = b$, $a < b$ or $b < a$;

O2 Transitivity: If $a < b$ and $b < c$, then $a < c$;

O3 Addition Law: $a < b$ if and only if $a + c < b + c$;

O4 Multiplication Law: If $c > 0$, then $ac < bc$ if and only if $a < b$. If $c < 0$, then $ac < bc$ if and only if $b < a$.

Proposition. The order relation on **R** is well-defined and makes **R** an ordered field.

Proof. Suppose $\{a_k\} \sim \{a'_k\}$ and $\{b_k\} \sim \{b'_k\}$. We want to show $a_k < b_k$ implies $a'_k < b'_k$. Since $\{a_k\} \sim \{a'_k\}$ and $\{b_k\} \sim \{b'_k\}$, for all $r > 0$, there exists N_1 such that for all $n > N$, $|a_n - a'_n| < r$. This gives ① $-r < a_n - a'_n < r$. Similarly, for all $r > 0$, there exists N_2 such that for all $n > N$, $|b_n - b'_n| < r$. This gives ② $r < b_n - b'_n < r$. Notice that if we subtract inequality ① from ②, we get $a'_k - b'_k < a_k - b_k$. Since $a_k < b_k$, $a_k - b_k < 0$. Therefore, $a'_k < b'_k$. Now we check all the order axioms:

O1 Only one of $a_k < b_k$, $a_k = b_k$, and $a_k > b_k$ holds. So only one of $[a_k] < [b_k]$, $[a_k] = [b_k]$, and $[a_k] > [b_k]$ holds.

O2 If $[a_k] < [b_k]$ and $[b_k] < [c_k]$, then eventually $a_k < b_k$ and $b_k < c_k$, so $a_k < c_k$. Finally, $[a_k] < [c_k]$.

O3 If $[a_k] < [b_k]$, then eventually $a_k < b_k$, so $a_k + c_k < b_k + c_k$. Hence $[a_k] + [c_k] < [b_k] + [c_k]$.

O4 If $[a_k] < [b_k]$, then eventually $a_k < b_k$, so $a_k c_k < b_k c_k$ if c_k. Hence $[a_k][c_k] < [b_k][c_k]$.

□

Definition 9.9 (Archimedean)**.** For all $x \in F$ in an ordered field, there exists $N \in \mathbb{Z}$ such that $x < N$.

Proposition. **R** is an Archimedean ordered field.

Proof. Consider an arbitrary $[a_k] \in \mathbf{R}$. Since $\{a_k\}$ is Cauchy, it is bounded in $\mathbb{Q}$. Furthermore, $\mathbb{Q}$ is Archimedean. Therefore, there exists $U \in \mathbb{Q}$ and $N \in \mathbb{N}$, such that for all $n \in \mathbb{N}, a_n \leq U < N$. Therefore $[a_k] < [N, N, N, ...]$, which means that **R** is Archimedean. □

Definition 9.10 (least upper bound)**.** Let F be an ordered field, and let A be a nonempty subset of F that is bounded above. We say that $L \in F$ is a least upper bound for A if the following two conditions hold:

1. L is an upper bound for A;
2. If M is any upper bound for A, then $L \leq M$.

Definition 9.11 (least upper bound property)**.** Let $S \subseteq F$ be a nonempty set that is bounded above. Then S has a least upper bound bound in F.

Proposition. The least upper bound property holds in **R**.

proof sketch. Let $\emptyset \neq A \subseteq \mathbf{R} = \mathcal{C}/I$, that is bounded above by m. Then there exists $M \in \mathbb{Z}$ such that $m \leq M$. Since $A \neq \emptyset$, we can choose some $a \in A$. Hence there exists $n \in \mathbb{Z}$ such that $n < a$. Consider

$$S_p = \left\{k2^{-p} \middle| k \in \mathbb{Z}, n \leq k2^{-p} \leq M\right\}$$

Note that S_p is nonempty and finite. Let $a_p = \min\{x | x \in S_p$ and x is an upper bound of $A\}$. If $q > p$ implies $a_q \leq a_p$, then $a_p - 2^{-p} < a_q$. (Notice that a_p is not an upper bound of A, and a_q is an upper bound of A). Therefore, $|a_p - a_q| \leq 2^{-p}$, for all $p < q$. In other words, $\{a_p\}$ is Cauchy. So $[a_p] = \text{lub}(A) \in \mathcal{C}/I = \mathbf{R}$. □

Definition 9.12 (real number)**.** Real number is **R**. Now denoted as $\mathbb{R}$.

Proposition. Real numbers are not countable.

Proof. Use Cantor's diagonal argument on real numbers. □

Proposition. Real numbers are Archimedean. That is for any $a, b \in \mathbb{R}^+$, there exists some $n \in \mathbb{N}$ such that $an > b$.

Proof. If $a > b$, this is trivial. If $a < b$, assume for contradiction that we do not have an $n \in \mathbb{N}$ such that $an > b$. This is the same as saying for all $n \in \mathbb{N}$, $an < b$. Consider the set

$$S = \{an | n \in \mathbb{N}\}$$

Since all $an < b$, S is bounded above by b. By the least upper bound property, S has a least upper bound denoted as L. Now consider $L - a$. Because $a > 0$, $L - a < L$. Since L is the **least** upper bound, $L - a$ is not an upper bound of S. In other words, $L - a < an_0$, for some $n_0 \in \mathbb{N}$. Therefore

$$L < an_0 + a \qquad \text{and} \qquad L < a(n_0 + 1)$$

Since $n_0 \in \mathbb{N}$, $n_0 + 1 \in \mathbb{N}$. Therefore, $a(n_0 + 1) \in S$. But $a(n_0 + 1) > L$ is a contradiction. We conclude that there must exist some $n \in \mathbb{N}$ such that $an > b$. □

Remark. Ordered fields without the least upper bound property may not be Archimedean.

Proposition. The least upper bound is unique.

Proof. Suppose L_1 and L_2 are least upper bounds for $\emptyset \nsubseteq A \subseteq F$. By definition 6.1, $L_1 \leq L_2$ and $L_1 \geq L_2$. By order axiom **O1**, $L_1 = L_2$. □

Proposition. Every real is between two consecutive integers. Formally, if $a \in \mathbb{R}$, there exists $N \in \mathbb{Z}$ such that $N - 1 \leq a \leq N$.

Proof. Let $S = \{n \in \mathbb{Z} | n > a\}$. Then by the Archimedean property, $S \neq \emptyset$ and S is bounded below. By the well-ordering principle, S has a least element. Then $N - 1 \notin S$, so $N - 1 \leq a \leq N$. □

Theorem 9.1 (rationals are dense)**.** Let $a, b \in \mathbb{R}$. If $a < b$, there exists some $r \in \mathbb{Q}$ such that $a < r < b$.

Proof. Intuitively, if you multiply a and b by a large enough integer q, then the gap between them will be big enough to contain another integer, p. Let a, b be arbitrary real numbers such that $a < b$.

Step 1: We claim that there exists a $q \in \mathbb{N}$ such that $qb - qa > 1$. Notice that $qb - qa = q(b - a)$. Since $b - a > 0$ and they are both reals, by the Archimedean property, there exists $q \in \mathbb{N}$ such that $q(b - a) > 1$.
Step 2: We claim that there exists a $p \in \mathbb{Z}$, such that $qa < p < qb$. Since every real is between two consecutive integers, we know that there exists $p \in Z$ such that $p - 1 \leq qa < p$. Since $qb - qa > 1$, we know that $qb > qa + 1$. And because $p - 1 \leq qa$, we know that $p \leq qa + 1$. This means that $qa < p < qb$, which is what we needed to show.
Step 3: Since $qa < p < qb$, if we divide the inequality by q, we get $a < p/q < b$. Moreover, since p, q are both integers, $p/q \in \mathbb{Q}$. This is all we need to show. □

Corollary. The irrationals are dense over the reals. In other words, if $a < b \in \mathbb{R}$, there exists $r \in \mathbb{Q}^c$ such that $a < r < b$.

Proof. Let a, b be arbitrary real numbers such that $a < b$. Since $\sqrt{2} > 0$, we know that $a/\sqrt{2} < b/\sqrt{2}$. Because the rational numbers are dense, we know that there exists some $p/q \in \mathbb{Q}$ such that $a/\sqrt{2} < p/q < b/\sqrt{2}$. Therefore, $a < \sqrt{2}(p/q) < b$. Because p/q is rational and $\sqrt{2}$ is irrational, we know that their product $\sqrt{2}(p/q)$ is irrational. Let $r = \sqrt{2}(p/q)$. We have $a < r < b, r \in \mathbb{Q}^c$, which is what we needed to show. □

Proposition. Let $([a_n, b_n])_{n\in\mathbb{N}}$ be a nested sequence of closed bounded intervals in $\mathbb{R}$. That is, for any n, we have $[a_{n+1}, b_{n+1}] \subseteq [a_n, b_n]$, or equivalently, $a_n \leq a_{n+1} \leq b_{n+1} \leq b_n$ for all n. Then $\bigcap_{n\in\mathbb{N}}[a_n, b_n] \neq \emptyset$.

Proof. Let $A = \{a_n \mid n \in \mathbb{N}\}$. A is bounded above by some b_1, which implies for every $a_n \in A$, $a_n \leq b_1$. If a_m=lub(A), $a_n \leq a_m \leq b_1$. $a_m \in [a_n, b_1]$ and $a_m \in \bigcap_{n\in\mathbb{N}}[a_n, b_n]$. □

Remark. What happens if we replace $[a_n, b_n]$ with $(a_n, b_n]$ and assume $a_n < b_n$? Does the proposition still hold? No. An example would be $a_n = 0$ and $b_n = 1/n$ for all $n \in \mathbb{N}$. The intersection is empty (this is "clear" but can be proved using the Archimedean property).

10 General Topology

We have seen analytic properties of real numbers $\mathbb{R}$, now we offer the topological approach.

Definition 10.1. A *metric* on a set X is a function $d : X \times X \to [0, \infty)$ satisfying:

1. $d(x, y) = 0$ if and only if $x = y$ for all $x, y \in X$;
2. $d(x, y) = d(y, x)$ for all $x, y \in X$;
3. $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in X$.

If d is a metric on X, we say (X, d) is a *metric space*.

Definition 10.2.

Definition 10.3. A *topology* on a set X is a collection $\mathcal{T}$ of subsets of X such that:

1. $\varnothing$ and X are in $\mathcal{T}$;
2. the union of the elements of any subcollection of $\mathcal{T}$ is in $\mathcal{T}$;
3. the intersection of the elements of any finite subcollection of $\mathcal{T}$ is in $\mathcal{T}$.

If $\mathcal{T}$ is a topology on X, we say the pair $(X, \mathcal{T})$ is a *topological space*.

Example. Let X be a set. The set $\mathscr{P}(X)$ is a topology on X, called the *discrete topology*. The set $\{X, \varnothing\}$ is also a topology on X, called the *trivial topology*.

Proposition. Any metric space is a topological space.

Suggested Readings

We tried to include many results in different area. However, this is just a short note, if you are interested in a certain area, we here recommend some textbooks as supplementary sources.

We have mentioned several results of axiom of choice: well-ordering theorem, Tikhonov theorem. They are equivalence to axiom of choice. By equivalence of axiom, we mean that if axiom of choice is replaced with any of those results, we still obtain the ZFC system. Moreover, notice that this entire note is based on ZFC, but ZFC has its weaknesses. There are different systems of set theory, such as DB, ZF, ... The provability between systems is another essential idea in set theory. Our treatments on sets largely based on [3] and we encourage the readers to check further results. Another good source for learning functions and sets is [4], which is comprehensive and terse.

References

[1] Awodey, S. *Introduction to Category Theory*. 2nd ed. Oxford, Oxford University Press. 2008.
[2] Bradley, Tai-Danae, et al. *Topology*. 1st ed. MIT Press. 2020.
[3] Jech, T. *Set Theory*. 3rd ed. Berlin ; London, Springer. 2011.
[4] Kaplansky, I. *Set Theory and Metric Spaces*. 1st ed. Chelsea Publishing Co., New York. 1977.
[5] Lang, S. *Algebra*. 3rd ed. New York, Springer. 2002.
[6] Tao, T. *Analysis I*. 3rd ed. Singapore Springer Singapore, Imprint: Springer. 2016.

Alphabetical Index