An Introduction to Proofs

Hassium, Fung San Gaan

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 elementary math to proof-based formal mathematics. The beauty of mathematics is hidden behind the craft of abstraction and the search to common forms between entities. It is not merely about techniques of solving problems, but about a way of thinking.

The first section is a quick review of basic logic, the language we use in proofs. Sections 2-4 cover set-theoretic constructions that are essential in any kind of math. Those sections discuss some axioms in ZFC and their consequences. This part largely follows from *Set Theory* by Thomas Jech. Those properties will be frequently used later on. Sections 5-9 are selected topics covering linear algebra, abstract algebra, real analysis, and general topology. Those sections correspond to the first several weeks of core courses in undergraduate mathematics. Section 10 introduce the basic notions in category theory. Our goal is to reformulate some results we proved before in a more elegant and general way by using categorical definitions. You may skip this section if you think it is too formal. Section 11 is an appendix for further readings in different area of mathematics.

Many people have contributed to this note. Special thanks to my collaborator `Fung San Gaan`, who wrote the entire section 8 and part of section 7. 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. Consider the set of all even numbers, each even number can be uniquely expressed by $2n$, where $n\in\mathbb{N}$ and $\mathbb{N}$ is the set of natural numbers. This is the defining property of the set, so the set of even numbers is $\{2n \mid n\in\mathbb{N}\}$.

Example. Here are several examples of sets.

1. The set $S = \{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}\}$ has three elements.
2. The set $\mathbb{Q} = \{a/b \mid a, b \in \mathbb{Z}, b \neq 0\}$ is the set of rational numbers, where $\mathbb{Z}$ is the set of integers.
3. The set $\mathbb{Z}$ can be written as $\{-z \mid z \in \mathbb{Z}\}$. Why?

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 $\{x \mid x \in A, x \in B\}$. We say A and B are *disjoint* if $A \cap B = \varnothing$. The *complement* of A in B is the set $\{x \mid x \in B, x \notin A\}$, 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 the 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 $b \in Y$.

Remark. Let S be a set. We denote $S \times \cdots \times S$, the Cartesian product of n times of S, by S^n.

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 $x \in X, y \in Y$ 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 the 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*.

Let $(X, \leq)$ be a partially ordered set. Let $Y \subset X$ and $Y \neq \varnothing$. We say $a \in Y$ is the greatest element of Y if for all $y \in Y$, $a \geq y$. Similarly, we could define the smallest element of a nonempty subset of X.

Definition 2.13. Let $(X, \leq)$ be a partially ordered set and let $Y \subset X$ be nonempty. An element $a \in X$ is said to be an *upper bound* of Y if for all $y \in Y$, $y \leq a$. An element $a \in X$ is said to be a *lower bound* of Y if for all $y \in Y$, $a \leq y$. We call the least upper bound of Y the *supremum* and we call the greatest lower bound of Y the *infimum*.

Problem 2.20. Prove that if supremum and infimum exist, then they are unique.

Definition 2.14. 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 2.21. Let $(X, \leq)$ be a linearly ordered set and let $Y \subset X$, prove that Y is linearly ordered.

Problem 2.22. 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 2.15. 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 $s \in S$, $a \leq s$.

Proposition. The set $\mathbb{N}$ is well-ordered under $\subset$.

Proposition. Every well-ordered set has a $\leq$-minimal element.

Proof. Suppose for any a, there exists □

Definition 2.16. Let $(X, \leq)$ be a well-ordered set. Let $x \in X$. Then the set $\{a \mid a \leq x\}$ is called an *initial segment* of X given by x.

It is trivial that any initial segment of a well-ordered set is well-ordered.

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 in the next section. 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. □

The well-ordering principle is an equivalent form of finite induction, but we will not show the converse.

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 2.23. 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 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 $\mathrm{ran}(f) \subset \mathrm{cod}(f)$.

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 $h : Z \to S$, $h \circ (g \circ f) = (h \circ g) \circ f$.

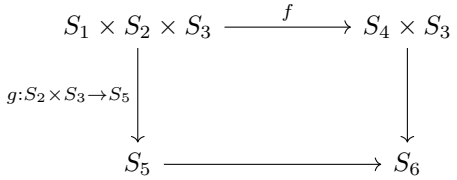
Proof. For all $x \in X$, $(h \circ g \circ f)(x) = h(g(f(x))) = (h \circ g)(f(x)) = ((h \circ g) \circ f)(x)$. □

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 $\psi(-x) = (f(-x) + f(x))/2 = \psi(x)$. Also, $\varphi(x) + \psi(x) = (f(x) - f(-x) + f(x) + f(-x))/2 = f(x)$. □

Problem 3.7. Given an example of a function that is neither odd nor even. Decompose it as a sum of an odd function and an even 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 examples if such functions exist.

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.

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 $\sim$. □

Recall that we have defined partial ordering by "$\leq$". From now on, we require that every partial-ordering to be a strict ordering.

Definition 3.12. Let $(P, <)$ and $(Q, <)$ be partially ordered set and let $f : P \to Q$ be a function. We say f is *order-preserving* if $x < y$ implies $f(x) < f(y)$. If f is also bijective, then we say P and Q have the same *order-type*, denoted $P \approx Q$.

Problem 3.19. Prove that if $f : P \to Q$ is order-preserving and bijective, then f^{-1} is also order-preserving.

Problem 3.20. Let $f : X \to Y$ be order-preserving. Prove that if X is well-ordered, then Y is well ordered.

Proposition. Let X be well-ordered and let $f : X \to X$ be order-preserving. Then $f(x) \geq x$ for all $x \in X$.

Proof. Suppose $\{x \in X \mid f(x) < x\} \neq \varnothing$, then it is well-ordered. Let z be the least element of this set and let $f(z) = w$, then $f(w) < w$, a contradiction. □

Problem 3.21. Prove that any well-ordered set cannot have the same order-type as any of its initial segment.

Definition 3.13. 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. Let C is a class, we write $x \in C$ if x satisfies the formula φ correspond to C.

Example. The collection of all sets is a class.

Problem 3.22. Prove that every set is a class.

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

You may notice that the axiom schema of replacement seems to "imply" the axiom schema of separation. However, the axiom schema of separation guarantees the existence of functions, so we need both axioms.

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 the axiom of choice as: for every nonempty set X, there exists a choice function. Here are some equivalent statements of the axiom of choice:

1. Well ordering theorem: Every set is well-orderable.
2. Zorn's lemma: If a partially ordered set has the property that every chain has an upper bound, then the poset contains at least one maximal element.

Remark. Throughout the notes, we shall see many equivalent forms of the axiom of choice. It is routine and technical to prove the equivalences so we will not provide those proofs.

4 Integers and Cardinality

Recall that we have constructed $\mathbb{N}$ by axioms. By the convention of quotient sets, we could further construct the set of 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 $\{[(a, b)] \in \mathbb{Z} \mid a \subset b\}$ and $\{[(b, a)] \in \mathbb{Z} \mid b \subset a\}$, conclude that the inverse of an integer is unique.

Let $n \in \mathbb{Z}$, the inverse of n is denoted by $-n$. The analogy of $\sim_+$ is the "subtraction". We will use a similar trick to define "division" later.

Problem 4.4. Prove that there exists a bijection $f : \{[(a, b)] \in \mathbb{Z} \mid a \subset b\} \to \mathbb{N}$.

For two sets, if there exists a bijection between them, we say they are isomorphic as sets, which means they have exactly the same structure. Now $\{[(a, b)] \in \mathbb{Z} \mid a \subset b\} \subset Z$ has the same structure as $\mathbb{N}$. We only care about the structures here, so please consider $\mathbb{N} \subset \mathbb{Z}$. The idea of structures and isomorphisms will be explained later.

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

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 $a + d < b + c$.

Definition 4.5. Let $p \in \mathbb{Z}$. We say p is a *prime* if $p \neq a \cdot b$ for all $a, b \in \mathbb{Z}$ and $a, b \neq p$.

Problem 4.6. Let $n \in \mathbb{N}$ and $n \neq 0$. Prove that if there exists $a, b \in \mathbb{N} \setminus \{0\}$ such that $a \cdot b = n$, then $a, b \leq n$.

Theorem 4.2 (fundamental theorem of arithmetic)**.** For all $n \in \mathbb{N}$ and $n \geq 2$, n can be factored as a product of prime numbers and the factorization is unique.

Proof. It is trivial that $2 = 2$, $3 = 3$, $4 = 2 \cdot 2$, ... Suppose every number $n \leq k$ for a given k admits a prime factorization. If $k + 1$ is prime, then $k + 1 = k + 1$. If $k + 1$ is not prime, then there exist $n_1, n_2 \in \mathbb{Z}$ such that $n_1 n_2 = k + 1$. Since n_1 and n_2 admits prime factorization, $k + 1$ admits prime factorization. We have the trivial case $2 = 2$. Assume the factorization is unique for all m with $2 \leq m < n$. Suppose $\prod p_i$ and $\prod q_j$ are prime factorizations of n. Rewrite those factorizations such that $p_i \leq p_{i+1}$ and $q_j \leq q_{j+1}$. Since p_1 divides $\prod q_j$ and p_1 is prime, it must divide one of the q_j, so $p_1 \leq q_1 \leq q_j$. Apply the same argument, $q_1 \leq p_1$, so $q_1 = p_1$. Now $p_2 \cdots p_r = q_2 \cdots q_s < n$. Hence, by induction, $p_i = q_i$, the factorization is unique. □

Theorem 4.3 (Euclid's theorem)**.** There are infinitely many primes.

Proof. Suppose there are finitely many primes, denoted $p_1, \ldots, p_n$, where $p_i \leq p_{i+1}$. Define $x = 1 + \prod_1^n p_i$, then N is not prime. Let p_x be a prime in the factorization, then p_i lies in the factorization of $N - \prod p_i = 1$, a contradiction. □

Problem 4.7. 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 4.8. Find the subset of $S \subset \mathbb{Q}$ such that there exists a bijection from S to $\mathbb{Z}$. Prove your choice of S is the desired set.

Similarly, we take this as $\mathbb{Z} \subset \mathbb{Q}$.

Problem 4.9. Prove that addition and multiplication on $\mathbb{Q}$ are well-defined.

Problem 4.10. 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 $a/b = a \cdot b^{-1}$.

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 4.11. 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 4.4 (Cantor's theorem)**.** Let S be a set. Then any function $f : S \to \mathscr{P}(S)$ is not surjective.

Proof. Suppose there exists a bijection $f : S \to \mathscr{P}(S)$. Define $X = \{s \in S \mid s \notin f(x)\}$, this is a subset of S, so $X \in \mathscr{P}(S)$ and $X = f(x)$ for some $x \in S$. If $x \in X$, then $x \notin f(x) = X$, so $x \notin X$, a contradiction. If $x \notin X$, then $x \in f(x) = X$, so $x \in S$, a contradiction. □

Problem 4.12. 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. □

Problem 4.13. Let X be a finite set and let $f : X \to Y$ be a function. Prove that $\mathrm{im}(f)$ is finite and $|f(X)| \leq |X|$.

Problem 4.14. 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 $f : \mathbb{N}^n \to \mathbb{N}$ 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 $\mathbb{N}^n$ is countable then $\prod_1^n A_i$ 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 4.5 (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, the system without the axiom of choice. In the next proof, assume we have all axioms except for the axiom of choice. This proof is technical and you may skip it.

Proof. Let $f : A \to B$ and $g : B \to A$ be injections. Define $A_0 = A \setminus g(B)$. For any $n \geq 0$, let $B_n = f(A_n)$ and let $A_{n+1} = g(B_n)$. Since g is injective, $g(B_n)$ are pairwise disjoint. It is clear that by induction, A_n are pairwise disjoint. Similarly, B_n are pairwise disjoint. Let $U = \bigcup A_n$ and $V = A \setminus U$. We have $B \setminus B_0 = B \setminus f(A_0)$, $V \subset A \setminus A_1 = g(B \setminus f(A_0))$. Here there exists a bijection between $B \setminus f(A_0)$ and $g(B \setminus f(A_0))$. The restriction $g|_V : V \to A$ is a bijection between $B \setminus B_0$ and V. Now let $h : A \to B$ be a function defined by $h(x) = f(x)$ if $x \in U$ and $h(x) = g^{-1}(x)$ if $x \in V$. Suppose $h(x_1) = h(x_2)$. If $x_1, x_2 \in U$, then $h(x_i) = f(x_i) \in B_n$ for some n. The restriction of f on every A_n is injective, so $x_1 = x_2$. Similarly, if $x_1, x_2 \in V$, $x_1 = x_2$. If $x_1 \in U$ and $x_2 \in V$, then $h(x_1) \subset B_n$ for some n, where $B_n \cap (B \setminus B_0) = \varnothing$. Hence h is injective. Pick some $y \in B$. If $y \in B_n$ for some n, then there is a unique $x \in A_n$ such that $f(x) = y$, so $h(x) = f(x) = y$. If $y \notin B_n$, then $y \in B \setminus B_0$. Consider $x = g(y) \in V$, then $h(x) = h(g(y)) = g^{-1}(g(y)) = y$. Hence h is surjective, conclude that h is a bijection. □

Problem 4.15. Prove the Cantor-Schröder-Bernstein theorem using the axiom of choice.

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 Groups

Definition 6.1. 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 6.1. Prove that the identity element is unique in any monoid.

Definition 6.2. 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 $hg = e = gh$. Such f is called an *inverse* of g. If $\circ$ is commutative, then the group is *abelian*. The *order* of a group G. denoted $|G|$, is the cardinality of the set 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 6.2. Prove that every permutation in $\mathfrak{S}_n$ can be written as a finite composition of involutions. Is this expression unique?

Definition 6.3. 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 6.3. 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 6.4. Let G be a group and let $g, h \in G$. Prove that $(-e)g = -g$, $-(-g) = g$, and $-g(-h) = gh$.

Problem 6.5. 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 6.6. Prove that $|\mathfrak{S}_n| = n!$

Definition 6.4. 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 $\ker(\varphi)$. If two groups G and H are isomorphic to each other, then we write $G \approx H$.

Problem 6.7. Prove that a group isomorphism is a bijective group homomorphism.

Definition 6.5. 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 $H \leq G$.

Definition 6.6. Let G be a group. Every homomorphism $f : G \to G$ is called an *automorphism*. The set of all automorphisms onto a group G is denoted by $\mathrm{Aut}(G)$.

Problem 6.8. Let G be a group and let $\circ$ be the composition of functions. Prove that $(\mathrm{Aut}(G), \circ, \mathrm{id}_G)$ is a group. This is called the *automorphism group*.

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 6.9. If $A \leq B \leq G$, prove that $A \leq G$.

Problem 6.10. 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 6.7. 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 6.11. Prove that there exists a bijection $\psi : G/H \to H \setminus G$, conclude that $[G : H] = |H \setminus G|$.

Problem 6.12. 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. □

Problem 6.13. Based on the previous proposition, prove a more general form:

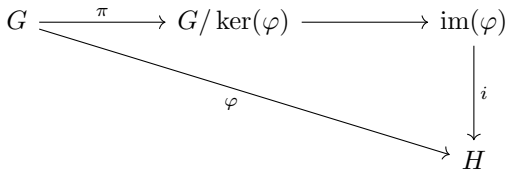
Theorem 6.1 (Cayley's theorem)**.** Any monoid is isomorphic to a submonoid of $M(S)$ for some set S. Any group is isomorphic to a subgroup of some symmetric group.

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 6.8. 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 normal subgroups of G are G and the trivial subgroup.

Problem 6.14. Prove that $H \trianglelefteq G$ if and only if $gHg^{-1} = H$ for all $g \in G$.

Theorem 6.2 (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 6.15. 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.

Definition 6.9. Let G be a group and $N \trianglelefteq G$. The *quotient group* G/N is defined to be $\{aN \mid a \in G\}$.

Problem 6.16. Take $(aN)(bN) = (ab)N$ as the composition on G/N, prove that G/N is indeed a group.

Definition 6.10. Let S be a set and let G be a group.

Definition 6.11. Let G be a group and let $S \subset G$. The set *generated* by S is defined to be $\{\prod s_i^{\pm 1} \mid s_i \in S\}$. The set S is called the *generating set* and the elements of S are called the *generators*. If $|S|$ is finite, then we say the group generated by S is *finitely-generated*. A finitely generated group is said to be *cyclic* if $|S| = 1$.

You may notice that everytime we define a group generated by a set, we need a larger group containing this set. Among a fixed number of generators, the $\subset$-maximal group is known as a free group. Every group admits a group presentation, check section 11 for more information about free groups and group presentations.

Definition 6.12. The

Let G be a group and let $S \subset G$. If $\langle S \rangle = G$, then we say G is *generated by* S.

Example. The presentation of the trivial group is $\langle a, b \mid abaa^{-1}b^{-2}, bab^{-1}a^{-2} \rangle$.

Proposition. The group $\langle X \mid \varnothing \rangle$ is the smallest group that contains X.

Problem 6.17. The *Heisenberg group* H admits the presentation $\langle a, b, c \mid cac^{-1}a^{-1}, cbc^{-1}b^{-1}, bab^{-1}a^{-1}c^{-1}\rangle$. Prove $a^m b^n c^p a = a^m(b^n a)c^p = a^{m+1}b^n c^{p+n}$, $a^m b^n c^p b = a^m b^{n+1} c^p$, and $a^m b^n c^p c = a^m b^n c^{p+1}$. Can we rewrite any element in H by $a^m b^n c^p$? Prove or disprove it.

Problem 6.18. A *Tarski monster group* is a finitely generated infinite group where every proper non-trivial subgroup is cyclic of order p, a fixed prime. Prove that every Tarski monster group is simple.

Problem 6.19. A *Baumslag-Solitar group* $BS(m, n)$ admits the presentation $\langle a, b \mid ba^m b^{-1} = a^n\rangle$. Recall that C_2 is the cyclic group of order 2. Prove that $BS(1, 1) \approx C_2 = \langle a \mid a^2\rangle$. This shows that the group presentation of a group is not necessarily unique. Furthermore, prove $\varphi : BS(2, 3) \to BS(2, 3)$ defined by $\varphi(a) = a^2$ and $\varphi(b) = b$ is a surjective group homomorphism.

7 Rings

Definition 7.1. 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.

Remark. We will use ab to denote $a \cdot b$ for convenience.

Definition 7.2. An element u of a ring R is a *unit* if there exists $v \in R$ such that $uv = 1_R = vu$. The set of all units in R is denoted by $R^\times$.

Definition 7.3. A ring R is a *commutative ring* if for all $a, b \in R$, $ab = ba$.

Example. Let S be a set and let R be a ring. Denote the set of functions from S to R by R^S. For any $f, g \in R^S$ and $x \in S$, the addition and multiplication are defined as $+ : (f+g)(x) = f(x)+g(x)$ and $\cdot : (fg)(x) = f(x) \cdot g(x)$. It it trivial that $(R^S, +, \cdot)$ is a ring.

Definition 7.4. Let R be a commutative ring. Define $R[x] = \{a_0 + a_1x + \cdots + a_nx^n \mid n \geq 0, a_i \in R\}$. Assuming $n > m$, we define addition and multiplication to be $+ : \sum_{i=0}^n a_ix^i + \sum_{i=0}^n b_ix^i = \sum_{i=0}^n (a_i + b_i)x^i$ and $\cdot : (\sum_{i=0}^n a_ix^i)(\sum_{j=0}^m b_jx^j) = \sum_{i=0}^{n+m}(\sum_{i+j=k} a_ib_jx^k)$. We say $R[x]$ is a *polynomial ring*.

Problem 7.1. Prove that a polynomial ring is indeed a ring. What are 0 and 1 in $R[x]$?

Definition 7.5. The *degree* of $0 \neq p(x) = a_0 + a_1x + \cdots + a_nx^n \in R[x]$ is $\deg(p) = \max\{n \mid a_n \neq 0\}$. If $p(x) = 0$, $\deg(p) = -\infty$.

Definition 7.6. Let F be a non-zero commutative ring. We say F is a *field* if for all $u \in F$ with $u \neq 0$, u is a unit.

Problem 7.2. Let R be a ring. Prove that for all $a \in R$, $a \cdot 0 = 0 = 0 \cdot a$.

Problem 7.3. If R is a ring and $0 = 1$, prove that $R = \{0\}$. We call $\{0\}$ the *zero ring*.

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 $0 = x^{-1} \cdot 0 = x^{-1} \cdot (x \cdot y) = (x^{-1} \cdot x) \cdot y = 1 \cdot y = y$, a contradiction. Therefore, x is not a unit. Similarly, y is not a unit. □

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

Proof. If f or g is 0, then $fg = 0$. We have $\deg(fg)$ is $-\infty$. So $\deg(f)+\deg(g)$ is either $-\infty$ or $-\infty$ plus something. If $f, g \neq 0$, then let $f = \sum_{i=0}^n a_ix^i$ and $g = \sum_{j=0}^m b_jx^j$. The product is $\sum_{i=0}^{n+m}(\sum_{i+j=k} a_ib_jx^k)$. If $a_n, b_m \neq 0$, then $\deg(fg) = n + m$ which is equal to $\deg(f) + \deg(g)$. Otherwise $\deg(fg) < n + m = \deg(f) + \deg(g)$. □

Definition 7.7. A subset S of a ring $(R, +, \cdot)$ is a *subring* if the following properties hold.

1. The group $(S, +, 0)$ is a subgroup of $(R, +, 0)$.
2. For all $a, b \in S$, $a \cdot b \in S$.

Definition 7.8. Let R and R' be rings. A map $f : R \to R'$ is a *ring homomorphism* if for all $a, b \in R$, $f(a + b) = f(a) + f(b)$ and $f(ab) = f(a) \cdot f(b)$.

Definition 7.9. A ring homomorphism $f : R \to R'$ is *unital* if $f(1_R) = 1_{R'}$.

Definition 7.10. An *ideal* I in a ring R is a subgroup of $(R, +, 0)$ such that $I \neq \varnothing$ and for all $a, b \in I$, $ab^{-1} = a - b \in I$.

Definition 7.11. Let $f : R \to S$ be a ring homomorphism. The *kernel* of f is the set $\ker(f) = \{r \in R \mid f(r) = 0\}$.

Proposition. Let R and S be commutative rings with identity. Suppose $\varphi : R \to S$ is a unital ring homomorphism. For each $c \in S$, there exists a unique ring homomorphism $\varphi_c : R[x] \to S$, such that $\varphi_c(x) = c$ and $\varphi_c(r) = \varphi(r)$ if r is a constant polynomial. The formula for φ_c is $\varphi_c(\sum a_k x^k) = \sum a_k c^k$.

Proof. If φ is unital, then $\varphi_c(1_R) = \varphi(1_R) = 1_S$ is unital. Assume a substitution morphism exists. (Existence) Define $\varphi_c : R[x] \to R'$ by $\varphi_c(a_0+a_1x+\cdots+a_nx^n) = \varphi(a_0)+\varphi(a_1)c+\cdots+\varphi(a_n)c^n$. Then $\varphi_c((\sum a_ix^i)(\sum b_jx^j)) = \varphi_c(\sum_k(\sum_{i+j=k}a_ib_j)x^k) = \sum_k(\sum_{i+j=k}\varphi(a_i)\varphi(b_j)c^k) = \sum_k(\sum_{i+j=k}\varphi(a_i)\varphi(b_j)c^ic^j) = (\sum_i\varphi(a_i)c^i)(\sum_j\varphi(b_j)c^j) = \varphi_c(\sum a_ix_i)\cdot\varphi_c(\sum b_jx_j)$. This verified φ_c preserves $\cdot$. It is easy to check that φ_c preserves $+$. (Uniqueness) Suppose $\psi : R[x] \to R'$ is another homomorphism with $\psi(r) = \varphi_c(r) = \varphi(r)$ for $r \in R$ and $\psi(1 \cdot x) = \psi(1)\psi(x) = c$. Then for all $p(x) = a_0 + a_1x + \cdots + a_nx^n \in R[x]$, $\psi(p(x)) = \psi(a_0 + \cdots + a_nx^n) = \psi(a_0) + \cdots + \psi(a_n)(\psi(x))^n = \varphi(a_0) + \cdots + \varphi(a_n)c^n = \varphi_c(a_0 + a_1x + \cdots + a_nx^n) = \varphi_c(p(x))$. Hence φ_c with is unique. □

This process is called the "substitution principle" because the formula for φ_c is literally "substitute x for c".

Problem 7.4. Let $\varphi \circ \psi$ denote the composite of two ring homomorphisms φ and ψ. Prove that $\varphi \circ \psi$ is also a ring homomorphism.

Example. Now it is time to introduce special cases of the substitution principle.

Evaluation function: Suppose $R = S$ and $\varphi : R \to S$ is the identity map. Then we may write $\mathbf{ev}_c$ for φ_c, and call it the evaluation function. It is given by

$$\mathrm{ev}_c(a_0 + a_1x + \ + a_nx^n) = a_0 + a_1c + \ + a_nc^n.$$

We have another notation for this: if $f = \sum a_kx^k$, we write $f(c) := \mathrm{ev}_c(f)$. The point is that this function ev_c is an unital ring homomorphism: $\mathrm{ev}_c(1) = 1$, $\mathrm{ev}_c(f + g) = \mathrm{ev}_c(f) + \mathrm{ev}_c(g)$, and $\mathrm{ev}_c(fg) = \mathrm{ev}_c(f)\mathrm{ev}_c(g)$. These are just obscure ways of writing the identities:

$$s(0) = 1, \quad (f + g)(c) = f(c) + g(c), \quad (fg)(c) = f(c)g(c)$$

where s is the constant polynomial corresponding to $1 \in R$.

Change of coefficients: Let $\varphi : R \to S$ be a ring homomorphism, the inclusion $i : S \hookrightarrow S[x]$ $i(a_0) = a_0$ is also a ring homomorphism. Hence $i \circ \varphi : R \to S[y]$ is a ring homomorphism. Let $c = y = 1_R \cdot y \in S[y]$. Then $\varphi_c : R[x] \to S[y]$ is given by $\varphi_c\left(\sum a_ix^i\right) = \sum \varphi(a_i)y^i$. φ_c is unital because φ is.

Problem 7.5. Prove that the kernel of a ring homomorphism $f : R \to S$ is an ideal.

Definition 7.12. Let R be a commutative ring. An element $\alpha \in R$ is a root of $p(x) \in R[x]$ if $p(\alpha) = 0_R$ (i.e. $p \in \ker(\mathrm{ev}_\alpha)$).

Lemma. I is an ideal of R, then $1_R \in I$ if and only if $I = R$.

Proof. ($\Rightarrow$) If $1_R \in I$, then for all $u \in R$, $1_Ru = u \in I$. It follows that $R \subseteq I$. Since $I \subseteq R$, $I = R$. ($\Leftarrow$) If $I = R$, then $1_R \in I = R$. □

Corollary. $I \subseteq R$ be an ideal. I contains a unit if and only if $I = R$.

Proof. ($\Rightarrow$) Assume I contains a unit v. There exists some $u \in R$ such that $uv = 1_R \in I$ if and only if $I = R$. ($\Leftarrow$) If $I = R$, then $1, -1 \in R$. □

Corollary. Let F be a field. Only ideals in F are $\{0\}$ and F.

Proof. Assume $I \subseteq F$ is an ideal. $\{0\}$ is an ideal because for all $u \in F$, $u \cdot 0 = 0 \cdot u = 0 \in I$. If $I \neq \{0\}$, then it must contain a unit. By the Corollary above, $I = F$. □

Theorem 7.1. Let R be a ring. $I \subseteq R$ is an ideal. The quotient group $(R/I, +, \cdot, 0+I, 1+I)$ is a ring and $\pi : R \to R/I$ given by $\pi(a) = a + I$ is a surjective ring homomorphism.

Proof. We need to show multiplication is well-defined, R/I has a ring structure, and π is surjective.
Well-defined multiplication: For any $a + I = a' + I \in R/I$ and $b + I = b' + I \in R/I$, there exist $i_1, i_2 \in I$ such that $a = a' + i_1, b = b' + i_2$. Then

$$ab = (a' + i_1)(b' + i_2) = a'b' + a'i_1 + b'i_2 + i_1i_2$$

By the definition of an ideal, $a'i_1, b'i_2 \in I$. Because an ideal is a subgroup of R under addition, $a'i_1 + b'i_2 + i_1i_2 \in I$. So $ab = a'b' + i_3$ for some $i_3 \in I$. Conclude that $ab + I = a'b' + I$.
Ring structure: Since I is a subgroup of R, $(R/I, +, 0+I)$ must be a group. Because addition in R is commutative, it is an abelian group. For any $a + I, b + I, c + I \in R/I$, the group is associative:

$$((a+I)(b+I))(c+I) = ((ab)+I)(c+I) = ((ab)c)+I = (a(bc))+I = (a+I)((bc)+I) = (a+I)((b+I)(c+I)).$$

It is also distributive:

$$(a+I)((b+I)+(c+I)) = (a+I)((b+c)+I) = (a(b+c))+I = ((ab)+I)+((ac)+I) = (a+I)(b+I)+(a+I)(c+I)$$

Conclude that multiplication is associative and distributes over addition. Moreover, $(a + I)(1 + I) = (a1) + I = a + I = (1a) + I = (1 + I)(a + I)$ implies that $1 + I$ is the unity in R/I. Therefore, $(R/I, +, \cdot, 0 + I, 1 + I)$ is a ring.
Surjective homomorphism: For all $a, b \in R$, $\pi(ab) = (ab) + I = (a + I)(b + I) = \pi(a)\pi(b)$. And for all $a, b \in R$, $\pi(a+b) = (a+b)+I = (a+I)+(b+I) = \pi(a)+\pi(b)$. Therefore, π is a ring homomorphism. For all $a+I \in R/I$, there exists some $a \in R$ such that $\pi(a) = a + I$. Conclude that π is a surjective homomorphism. □

Theorem 7.2 (first ring isomorphism theorem)**.** Let $\varphi : R \to S$ be a unital ring homomorphism. Let $I = ker(\psi)$. Then $\widetilde{\psi} : R/I \to S$ given by $\widetilde{\psi}(a + I) = \psi(a)$ is a well-defined injective ring homomorphism. In particular, $\widetilde{\psi} : R/I \to \psi(R)$ is a ring isomorphism.

Proof. From first isomorphism theorem of groups, we know that $\widetilde{\psi}$ is a group isomorphism. We need to check if it preserves multiplication. For all $a+I, b+I \in R/I$, $\widetilde{\psi}((a+I)(b+I)) = \widetilde{\psi}((ab)+I) = \psi(ab) = \widetilde{\psi}(a+I)\widetilde{\psi}(b+I)$. □

Lemma. R is a commutative ring. For all $a \in R$, $aR = \{ar | r \in R\}$ is an ideal in R.

Proof. We first show it is a subgroup. $0 = a \cdot 0 \in aR$. Then $aR \neq \emptyset$. For all $x, y \in aR$, there exist $r, r' \in R$ such that $x = ar$ and $y = ar'$ satisfying $x - y = a(r - r') \in aR$. Therefore, aR is a subgroup of R. Now we show aR is an ideal. For all $ar \in aR$, $arr' = a(rr') \in aR$ for some $r' \in R$. And $r'ar = ar'r = a(r'r) \in aR$. □

Definition 7.13 (generated ideal)**.** Let R be a ring, $S \subseteq R$ be a subset, the ideal generated by S is $< S >= \bigcap_{I:\text{ideal}; S \subseteq I} I$. This is the smallest ideal containing S.

Definition 7.14 (zero divisor)**.** Let R be a ring, $b \in R$ with $b \neq 0$ is a zero divisor if and only if there exists $a \in R$ with $a \neq 0$ such that $ab = 0$ or $ba = 0$.

Definition 7.15 (principal ideal)**.** An ideal $I \subseteq R$ is a principal if $I = \langle a \rangle$, where $\langle a \rangle >= \{ar | r \in R\} = Ra$. This is when $S = \{a\}$. If $S = \{a, b\}$, then $\langle S \rangle = \{ra + sb | r, s \in R\} = Ra + Rb$.

Definition 7.16 (integral domain)**.** A ring is an integral domain if and only if it ① is commutative and it ② has no zero divisors.

Lemma. Let R be a ring, $\{a \in R | \text{a is a unit}\} \cap \{a \in R | \text{a is a zero divisor}\} = \emptyset$.

Proof. Assume there exists u in both sets. This implies the existence of $v, b \in R$ such that $uv = vu = 1$, and $bu = 0$ or $ub = 0$ while $b \neq 0$. Assume $bu = 0$. Without loss of generality,

$$0 = 0 \cdot v = (bu)v = b(uv) = b \cdot 1 = b$$

Since $b \neq 0$, u does not exist. □

Remark. Things to keep in mind:

1. For any integral domain R and any $0 \neq a \in R$, $ab = ac$ implies $b = c$.
2. Any subring of an integral domain is an integral domain.
3. Any field is an integral domain since any nonzero element is a field is a unit.

Lemma. Any finite integral domain is a field.

Proof. Let D be a finite integral domain. Choose any $0 \neq a \in D$. Consider $f_a : D \to D$ that is given by $f_a(b) = ab$. Then $ab = ac$ implies $b = c$. This proves injectivity. Since D is finite, f_a has to be surjective. This implies there exists some $v \in D$ such that $1 = f_a(v) = av = va$. It follows that a is a unit. Since a is arbitrary, every nonzero element is a unit. Conclude that D is a field. □

Lemma. $\mathbb{Z}/n\mathbb{Z}$ is an integral domain if and only of n is prime.

Proof. If $\mathbb{Z}/n\mathbb{Z}$ is an integral domain, then for all $k, l \in \mathbb{Z}$, $[k][l] = [0]$ if and only of $[k] = 0$ or $[l] = 0$. This is saying $p|kl$ if and only if $p|k$ or $p|l$. By definition, p is a prime. □

Proposition. Let D be an integral domain. Then for all $f, g \in D[x]$, $\deg(fg) = \deg(f) + deg(g)$.

Proof. For all $f, g \in D[x]$, if f or g is 0, then $fg = 0$. $\deg(fg) = \deg(f) + \deg(g) = -\infty$. Now suppose $f \neq 0$ and $g \neq 0$. Let $f = \sum_{i=0}^{m} a_i x^i$ and $g = \sum_{j=0}^{n} b_j x^j$ where $m, n > 0$. Since D in a domain, f, g do not have zero divisors:

$$fg = a_m b_n x^{m+n} + (\text{terms with lower power of } x) \neq 0$$

Therefore, $\deg(fg) = m + n = \deg(f) + \deg(g)$. □

Problem 7.6. Let $S = \{2, x\} \subseteq \mathbb{Z}[x]$. Show that $\langle S \rangle$ is not principal by

1. describing $\langle S \rangle$ in set theoretical terms, and
2. showing $2, x \in \langle S \rangle S$ according to the description in (1.).

Suppose $\langle S \rangle$ is principal. Let $\langle S \rangle = \langle p(s) \rangle$ for some $p(x) \in \mathbb{Z}[x]$. Then $2 = p(x)h(x)$ and $x = p(x)m(x)$ for some $h(x), m(x) \in \mathbb{Z}[x]$. Since $\mathbb{Z}$ is an integral domain, $\deg(2) = \deg(p(x)h(x)) = \deg(p(x)) + \deg(r(x)) = 0$. Therefore, $\deg(p(x)) = \deg(h(x)) = 0$. And since 2 is prime, $p(x) = \pm 2$ or ± 1. Finally, show that

3. $p(x) = \pm 1$ will lead to a contradiction, and
4. $p(x) = \pm 2$ will lead to a contradiction as well.

8 Real Numbers

Consider the function $|\cdot| : \mathbb{Q} \to \mathbb{Q}$ defined by $|a| = a$ if $a \geq 0$ and $|a| = -a$ if $a < 0$.

Problem 8.1. Use the definition of $|\cdot|$, prove that for all $a, b \in \mathbb{Q}$, the following conditions hold.

1. $|a| \geq 0$.
2. If $|a| = 0$, then $a = 0$.
3. $|ab| = |a||b|$.
4. $|a + b| \leq |a| + |b|$.

Definition 8.1. A *sequence* of rationals is a function $a : \{n \in \mathbb{Z} \mid n \geq 1\} \to \mathbb{Q}$. We denote a sequence by $\{a_n\}_{n=1}^{\infty}$.

Remark. We will use the notation $\{a_n\}$ for a sequence $\{a_n\}_{n=1}^{\infty}$.

Remark. For the rest of our construction, we use $\mathbb{Q}^+$ to denote the set $\{a \in \mathbb{Q} \mid a > 0\}$.

Definition 8.2. A sequence $\{a_n\}$ *converges* to $a \in \mathbb{Q}$ if for all $r \in \mathbb{Q}^+$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$, $|a_n - a| < r$. If a sequence $\{a_n\}$ converges to a, then we write $\{a_n\} \to a$. Here a is called the *limit* of $\{a_n\}$.

Proposition. If a sequence converges to a limit, then this limit is unique.

Proof. Suppose $\{a_n\} \to a$ and $\{a_n\} \to b$. Fix $r = |a - b|/3$, then for sufficiently large n, we have $|a_n - a| < r$ and $|a_n - b| < r$. Then $|a - b| \leq |a - a_n| + |a_n - b| < 2|a - b|/3$, a contradiction. □

Example. Consider the sequence $a_n = 1 + 1/n$. We claim it converges to 1. Choose some r, and we want $|a_n - 1| < r$ always hold. Notice that $|a_n - 1| = |1/n| = 1/n$. Let the integer part of $1/r$ be c, so $c \leq 1/r < c + 1$, pick $N = c + 1$.

Example. Not every sequence of rationals has a limit. Consider $\{3.1, 3.14, 3.141, 3.1415, \dots\} \to \pi \notin \mathbb{Q}$.

Problem 8.2. Find the limits of the following sequences. If some limit does not exist, prove it.

1. $a_n = 1/n$.
2. $a_n = \sum_{i=1}^{n} 2^{-k}$.
3. $a_n = (-1)^n$.

Definition 8.3. A sequence $\{a_n\}$ 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 $|a_n - a_m| < r$.

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 $|a_n - a_m| = |(a_n - a) + (a - a_m)| \leq |a_n - a| + |a_m - a| < r$. Therefore, $\{a_n\}$ is Cauchy. □

Problem 8.3. Not every Cauchy sequence in $\mathbb{Q}$ is bounded. Consider our previous example $\{3.1, 3.14, 3.141, \dots\}$, prove it is Cauchy but does not converge in $\mathbb{Q}$.

Definition 8.4. Let $S \subset \mathbb{Q}$. We say S is *bounded* if there exists $L > 0$ such that $|x| \leq M$ for all $x \in S$.

Definition 8.5. A rational sequence $\{a_n\}$ is *bounded* if $S = \{a_n \mid n \in \mathbb{N}\}$ is bounded in $\mathbb{Q}$.

Proposition. 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 $n \geq N$, $|a_n - a_N| < 347$. Since $|a_n - a_N| < 347$, replace $|a_n - a_N|$ with 347. When $n \geq N$, $|a_n| < 347 + |a_N|$. Let $U = \max\{|a_1|, \ldots, |a_{N-1}|, 347 + |a_N|\}$. We have that $|a_n| \leq U$ for all $n \in \mathbb{N}$. If $n \geq N$, then use $347 + |a_N|$ as the bound. If $n < N$, then there finitely many values, and they are bounded by $\max\{|a_1|, \ldots, |a_{N-1}|\}$. □

Let $\{a_n\}$ and $\{b_n\}$ be sequences. Define the addition as $\{a_n\} + \{b_n\} = \{a_n + b_n\}$. Define the multiplication as $\{a_n\} \cdot \{b_n\} = \{a_n b_n\}$.

Proposition. The set of all rational Cauchy sequence is closed under addition and multiplication.

Proof. Let $\{a_n\}$ and $\{b_n\}$ be Cauchy. Fix r. Since $\{a_n\}$ is Cauchy, there exists N_1 such that for all $n, m \geq N_1$, $|a_n - a_m| < r/2$. Similarly, we have N_2 for $\{b_n\}$. Let $N = \max(N_1, N_2)$. For all $n, m \geq N$, $|(a_m + b_m) - (a_n + b_n)|$ $\leq |a_m - a_n| + |b_m - b_n| = r$. Let $\{a_k\}$ and $\{b_k\}$ be Cauchy sequences in $\mathbb{Q}$. Then $|a_n b_n - a_m b_m| \leq |a_n||b_n - b_m| +$ $|b_m||a_n - a_m| \leq A|b_n - b_m| + B|a_n - a_m|$, where A and B are upper bounds for sequences $\{|a_k|\}$ and $\{|b_k|\}$. For all $\varepsilon(2A)^{-1}, \varepsilon(2B)^{-1} > 0$, there exists $N \in \mathbb{N}$ such that for all $m, n \geq N$, $|a_m - a_n| < \varepsilon(2B)^{-1}$ and $|b_m - b_n| <$ $\varepsilon(2A)^{-1}$. Now $|a_m b_m - a_n b_n| < A\varepsilon(2A)^{-1} + B\varepsilon(2B)^{-1} = \varepsilon$. □

Problem 8.4. Prove that the addition and multiplication satisfies commutativity, associativity, and distributivity.

Define the sequence 1 as $a_n = 1$ and define the sequence 0 as $a_n = 0$.

Problem 8.5. Prove 1 and 0 are the identities of multiplication and addition, respectively. Moreover, prove that additive inverse of a Cauchy sequence is Cauchy.

We conclude that the set of all rational Cauchy sequences is a commutative ring with 1, denoted $\mathcal{C}$.

Problem 8.6. Prove that $\mathcal{C}$ is not an integral domain.

Problem 8.7. Let I be the set of all Cauchy sequences in $\mathcal{C}$ that converge to 0. Prove I is an ideal of $\mathcal{C}$.

Proposition. If $\{a_k\}$ is a sequence in $\mathcal{C} \setminus I$, then there exist a positive rational number r and $N \in \mathbb{N}$ such that $|a_k| \geq r$ for all $n \geq N$.

Proof. If $\{a_k\}$ does not converge to 0, then there exists $r \in \mathbb{Q}^+$ such that for all $N \in \mathbb{N}$ there exists $k \geq N$ such that $|a_k| \geq 2r$. Since $\{a_k\}$ is Cauchy, for all $n, m \geq \mathbb{N}$, there exists $N \in \mathbb{N}$ such that $|a_n - a_m| < r$. Now we have $|a_m - a_n| < r$, then $|a_m - a_n| \leq |a_m - a_n| < r$, which implies $|a_n| > |a_m| - r$. Fix $m \geq N$ and let $|a_m| \geq 2r$, then $|a_n| > |a_m| - r \geq 2r - r = r$. □

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.

Proposition. If a Cauchy sequence does not converge to 0, then all terms of it eventually have the same sign.

Proof. If $\{a_k\}$ is a Cauchy sequence that does not converge to 0, then $\{a_k\} \notin I$. 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. If $a_n > a_k$, then we are done. If $a_n < a_k$, then $a_k - a_n < r$ which implies a_n is positive for all n. □

Let $\{a_k\}$ and $\{b_k\}$ be Cauchy sequences in $\mathbb{Q}$. Define an relation $\sim$ on $\mathcal{C}$ as follows: $\{a_k\} \sim \{b_k\}$ if and only if $\{a_k - b_k\} \in I$.

Problem 8.8. Prove that $\sim$ is indeed an equivalence relation.

Now we denote $\mathcal{C}/I$ by $\mathcal{R}$. If $\{a_k\}$ and $\{b_k\}$ are Cauchy sequences, define the addition and multiplication to be $[a_k] + [b_k] = [a_k + b_k]$ and $[a_k][b_k] = [a_k b_k]$.

Problem 8.9. Prove that the addition and multiplication are well-defined on $\mathcal{C}/I$.

Proposition. The quotient set $\mathcal{R}$ is a field.

Proof. Since $\mathcal{R} = \mathcal{C}/I$ and $\mathcal{C}$ is a commutative ring, $\mathcal{R}$ is a commutative ring. Now we are left to show that multiplicative inverses exist for nonzero elements of $\mathcal{R}$. Consider some arbitrary $[a_k] \neq I$. Since $[a_k] \to 0$, the absolute value of $[a_k]$ is eventually bounded below. There exists $N \in \mathbb{N}$ and $r > 0$ such that $|a_n| > r$ for all $n \geq N$. We define $b_k = 1$ if $k < n$ and $b_k = 1/a_k$ for $k \geq N$ for a sequence $\{b_k\}$. Fix $\epsilon > 0$. Since $\{a_k\}$ is Cauchy, there exists some M such that for all $m, n \geq M$, $|a_n - a_m| < \epsilon r^2$. We have $|b_n - b_m| = |a_n - a_m|/|a_n a_m| \leq |a_n - a_m|/r^2 \leq \epsilon$, so $\{b_k\}$ is Cauchy. Since $\{a_k\}\{b_k\} \to 1$, it is the desired multiplicative inverse. □

Let $a = [a_k], b = [b_k] \in \mathcal{R}$. We define $a \leq b$ if and only if $a_k \leq b_k$ eventually.

Definition 8.6. Let F be a ring and let $\leq$ be a partial-ordering on F. We say F is an *ordered ring* if for all $a, b, c \in F$, the following properties hold.

1. If $a \leq b$, then $a + c \leq b + c$.
2. If $0 \leq a, b$, then $ab \geq 0$.

If F is a field, then we say F is an *ordered field*.

We define the set of all real numbers, denoted $\mathbb{R}$, to be the ordered field $\mathcal{R}$. For all $x \in \mathbb{Q}$, x is considered to be the sequence $\{x_n\}$, where $x_n = x$ for all n. It is trivial that $\mathbb{Q} \subset \mathbb{R}$.

Problem 8.10. Prove that this ordering is well-defined on $\mathbb{R}$. Prove that $\mathbb{R}$ is an ordered field under the ordering.

Definition 8.7. Let R be an ordered field and let $\varphi : R \to \mathbb{Z}$ be an ordered-ring homomorphism. We say R is *Archimedean* if for all $x \in R$, there exists $N \in \mathbb{Z}$ such that $x < N$.

Problem 8.11. Prove that $\mathbb{Q}$ is Archimedean.

Proposition. The ordered field $\mathbb{R}$ is Archimedean.

Proof. Let $i : \mathbb{R} \to \mathbb{Z}$ be the identity function, it is trivial that i is an ordered-ring homomorphism. Consider an arbitrary $[a_k] \in \mathbb{R}$. Since $\{a_k\}$ is Cauchy, it is bounded in $\mathbb{Q}$. There exists $U \in \mathbb{Q}$ and $N \in \mathbb{N}$, such that for all $n \in \mathbb{N}$, $a_n \leq U < N$. Hence $[a_k] < [N, N, N, \dots]$ and $\mathbb{R}$ is Archimedean. □

Proposition. Let S be a bounded nonempty subset of $\mathbb{R}$. Then S has a least upper bound in $\mathbb{R}$.

Proof. Let $A \subset \mathcal{C}/I$ and $A \neq \varnothing$ that is bounded above by m. Then there exists $M \in \mathbb{Z}$ such that $m \leq M$. For all $a \in A$, then there exists $n \in \mathbb{Z}$ such that $n < a$. Consider the set □

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 $a_p = \min\{x | x \in S_p \text{ and } x \text{ is an upper bound of } A\}$. 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 $[a_p] = \text{lub}(A) \in \mathcal{C}/I = \mathbf{R}$. □

Ordered fields without the least upper bound property may not be Archimedean. Examples are beyond the scope of this note so we will not show any example here.

Proposition. If $a \in \mathbb{R}$, there exists $N \in \mathbb{Z}$ such that $N - 1 \leq a \leq N$.

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

Proposition. For all $x, y \in \mathbb{R}$ with $x < y$, there exists $q \in \mathbb{Q}$ such that $x < q < y$.

Proof. Let $a, b \in \mathbb{R}$ with $a < b$. 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 $q(b - a) > 1$. Now 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 $p \in \mathbb{Z}$ 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. Since $qa < p < qb$, $a < p/q < b$. Since $p, q \in \mathbb{Z}$, $p/q \in \mathbb{Q}$. □

We say $\mathbb{R}$ is dense over $\mathbb{Q}$. Now we define the set of *irrationals* to be $\mathbb{R} \setminus \mathbb{Q}$.

Definition 8.8. A number $a \in \mathbb{R}$ is *algebraic* if it satisfies a polynomial equation $c_n x^n + \cdots + c_1 x + c = 0$.

Problem 8.12. Prove that rational numbers are algebraic.

Proposition. If the coefficients of the monic polynomial equation $x^n + c_{n-1}x^{n-1} + \cdots + c_1 x + c = 0$ are integers, then any rational solution q of the polynomial is an integer that divides c.

Problem 8.13. Prove that $\sqrt{p}$ is irrational, where p is a prime.

Problem 8.14. Let $a \in \mathbb{R} \setminus \mathbb{Q}$. Let $c \in \mathbb{Q}$ and $c \neq 0$. Prove that $a + c, ac \in \mathbb{R} \setminus \mathbb{Q}$.

Proposition. Irrationals are dense over $\mathbb{Q}$.

Proof. Let a, b be arbitrary real numbers such that $a < b$. Since $\sqrt{2} > 0$, $a/\sqrt{2} < b/\sqrt{2}$. 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 $a < r < b, r \in \mathbb{R} \setminus \mathbb{Q}$. □

Definition 8.9.

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 $[a_{n+1}, b_{n+1}] \subset [a_n, b_n]$. Then $\bigcap_{n \in \mathbb{N}} [a_n, b_n] \neq \varnothing$.

Definition 8.10. 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 8.11. Let (X, d) be a metric space and let $x_0 \in X$. An *open ball* centered at x_0 with radius r, where $r \geq 0$, is the set $B(x_0, r) = \{x \mid x \in X, d(x, x_0) < r\}$. An *open set* is a subset of X that can be written as a union of open balls.

Problem 8.15. Prove that $(\mathbb{R}, |\cdot|)$ is a metric space.

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 $a_m = \inf(A)$, $a_n \leq a_m \leq b_1$. $a_m \in [a_n, b_1]$ and $a_m \in \bigcap_{n\in\mathbb{N}}[a_n, b_n]$. □

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).

9 Topological Spaces

Definition 9.1. A *topology* on a set X is a set $\mathcal{T}$ 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*.

It is trivial that $\mathcal{T}$ must be a subcollection of $\mathscr{P}(X)$, so $\mathcal{T}$ is a set.

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*.

Definition 9.2. Let $\mathcal{T}$ be a topology on X. A subset $U \subset X$ is said to be *open* if $U \in \mathcal{T}$. A subset V is said to be *closed* if $X \setminus V$ is open.

Definition 9.3. Let $\mathcal{T}$ and $\mathcal{T}'$ be topologies on X. If $\mathcal{T} \subset \mathcal{T}'$, $\mathcal{T}$ is said to be *coarser* than $\mathcal{T}'$; if $\mathcal{T}' \subset \mathcal{T}$, $\mathcal{T}$ is said to be *finer* than $\mathcal{T}'$.

Proposition. Every metric space is a topological space.

Proof. Denote the family of all open sets in X by $\mathcal{T}$. Fix $x_0 \in X$, then the open ball $B(x_0, 0) = \varnothing$. There exists an open set consisting of all open balls in X, which is X. Let U_1 and U_2 be open sets, we can express them as unions of open balls, and the union of open balls is again an open set, similarly, the finite intersection is an open set, so (X, d) induces a topology. □

The metric on the real line $\mathbb{R}$ induces a topology called the standard topology on $\mathbb{R}$.

Problem 9.1. Let $(X, \mathcal{T}_1)$ and $(X, \mathcal{T}_2)$ be two topological spaces. Are $(X, \mathcal{T}_1 \cap \mathcal{T}_2)$ and $(X, \mathcal{T}_1 \cup \mathcal{T}_2)$ topological spaces? Prove or state counterexamples.

Remark. Given a topological space $(X, \mathcal{T})$, we will denote it by X for convenience.

Problem 9.2. Let K be a field, let $\mathbb{A}^n$ be the set $\{(a_1, \ldots, a_n) \mid a_i \in K\}$, and let $A = K[x_1, \ldots, x_n]$ be the polynomial ring. The *zero locus* of $T \subset A$ is the set $Z(T) = \{P \in \mathbb{A}^n \mid f \in T, f(P) = 0\}$. A subset Y of $\mathbb{A}^n$ is an *algebraic set* if there exists a subset $T \subset A$ such that $Y = Z(T)$. Prove the following statements.

Let the complements of the algebraic sets be the open sets. This defines a topology, called the *Zariski topology* on $\mathbb{A}^n$.

Problem 9.3. A *topological group*

Definition 9.4. A function $f : X \to Y$ between two topological spaces is *continuous* if $f^{-1}(U)$ is open in X whenever U is open in Y. A continuous function with a left and right inverse is called a *homeomorphism*.

Problem 9.4. Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces. Let $f : X \to Y$ be a function, prove that f is continuous if and only if for any closed set $C \in Y$, $f^{-1}(C)$ is closed in X.

Problem 9.5. Not every continuous bijection between topological spaces is a homeomorphism. Consider the identity function $\mathrm{id} : (\mathbb{R}, \mathcal{T}_1) \to (\mathbb{R}, \mathcal{T}_2)$, where $\mathcal{T}_1$ is the discrete topology and $\mathcal{T}_2$ is the standard topology. Prove that id is a continuous bijection but not a homeomorphism.

Definition 9.5. Let $(X, \mathcal{T}_X)$ be a topological space and $Y \subset X$. The *subspace topology* on Y is $\mathcal{T}_Y = \{U \cap Y \mid U \in \mathcal{T}_X\}$.

Problem 9.6. Prove that a subspace topology is indeed a topology.

Problem 9.7. Let Y be a subspace of X. Prove that if U is open in Y and Y is open in X, then U is open in X.

Proposition. Let $(X, \mathcal{T}_X)$ be a topological space and $Y \subset X$. The subspace topology on Y is the coarsest topology on Y for which the canonical inclusion $i : Y \hookrightarrow X$ is continuous.

Proof. Given a topological space $(X, \mathcal{T})$ and an arbitrary set Y, consider the topology on Y that makes $f : Y \to X$ continuous. There exists such topology and the discrete topology is an example. Therefore, the intersection of topologies such that $f : Y \to X$ is continuous is trivially the coarsest such topology, denoted $\mathcal{T}_f$. Now let $Y \subset X$ and $f = i$, the inclusion map, then $\mathcal{T}_i = \{i^{-1}(U) \mid U \in \mathcal{T}_X\}$, where $\mathcal{T}_X$ is the topology on X and $i^{-1}(U) = U \cap Y$, hence $\mathcal{T}_i$ is the subspace topology. □

Definition 9.6. Let X be a topological space and let $x, y \in X$. A *path* in X from x to y is a continuous map $[0, 1] \to X$ such that $u(0) = x$ and $u(1) = y$.

Let X be a topological space and let $x, y \in X$.

Theorem 9.1 (fixed point theorem)**.** Any continuous map $f : [0, 1] \to [0, 1]$ has a fixed point, that is, there exists $x \in [0, 1]$ such that $f(x) = x$.

Proof. Suppose $f(x) \neq x$ for all $x \in [0, 1]$. Define $g : [0, 1] \to \{1, -1\}$ by $g(x) = (f(x) - x)/|f(x) - x|$. If $x = 0$, then $f(x) \neq 0$, so $f(0) = 1$. Similarly, $f(1) = -1$, so g is surjective. □

Definition 9.7. Let $(X, \mathcal{T}_X)$ be a topological space, let S be a set, and let $\pi : X \to S$ be surjective. The *quotient topology* on S is the finest topology for which π is continuous, and π is called a *quotient map*.

Example. A *torus* is

We denote a torus by T^2.

10 Functions as Morphisms

Definition 10.1. 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

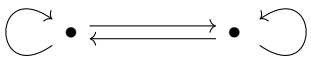
1. given any $f : X \to Y$, $g : Y \to Z$, and $h : Z \to W$, we have the identity $(h \circ g) \circ f = h \circ (g \circ f)$;
2. for each object X, there exists an *identity morphism* $\mathrm{id}_X : X \to X$ with the property that $f \circ \mathrm{id}_X = f$ and $\mathrm{id}_X \circ g = g$ for any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

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 10.1. 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. Let the objects be topological spaces and let the morphisms be continuous maps. Prove that this defines a category, denoted Top.

Problem 10.2. 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 10.3. 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 10.4. Prove that the identity morphism is unique for each object A in a category C.

Definition 10.2. Let C be a category. A category D is a *subcategory* of C if the $\mathrm{Ob}(\mathsf{D})$ and the $\mathrm{Hom}_{\mathsf{D}}(X, Y)$ are subcollections of $\mathrm{Ob}(\mathsf{C})$ 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 10.5. 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 10.3. 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 10.6. Prove that a function is surjective if and only if it is an epimorphism in Set.

Definition 10.4. 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 $A \approx B$.

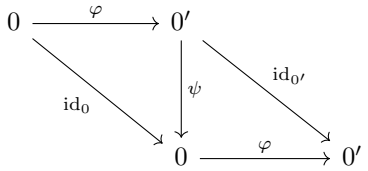
Problem 10.7. Prove that a morphism is an isomorphism if and only if it is a monomorphism and an epimorphism.

Problem 10.8. Let A and B be well-ordered sets that are isomorphic to each other. Prove that the isomorphism between them is unique.

Definition 10.5. 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 10.9. 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.

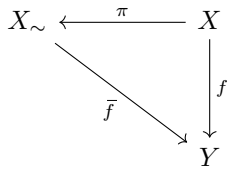
Problem 10.10. Prove that the initial object in the category of partially ordered set is the $\leq$-minimal element. Prove that the terminal object in this category is the $\leq$-maximal element.

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.

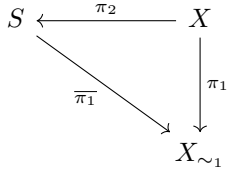
Proposition. There exists a map $\overline{f} : X_\sim \to Y$ such that $f = \overline{f} \circ \pi$.

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

Problem 10.12. 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 $X_{\sim_f} \approx \mathrm{im}(f)$.

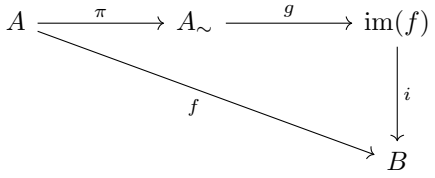
Proposition. Let S be any set satisfying the universal property for quotient sets. Let X be a set and let $\sim_1$ be an equivalence relation on X, so $\pi_1 : X \to X_{\sim_1}$ is invariant under the equivalence relation induced by $\pi_2 : X \to S$. Then $S \approx X_{\sim_1}$.



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 $S \approx X_{\sim_1}$. □

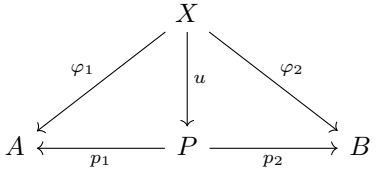
This completes our proof of the universal property for quotient sets.

Theorem 10.1 (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 $a \in A$, $(i \circ g \circ \pi)(a) = (i \circ g)([a]) = i(f(a)) = f(a)$. □

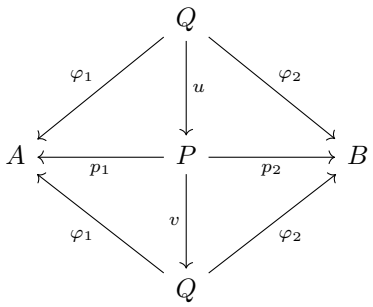
Definition 10.6. Let C be a category. Let A and B be objects in C. The *product* of A and B, denoted P, is an object in C with morphisms $p_1 : P \to A$ and $p_2 : P \to B$ such that for all object X in C with morphisms $\varphi_1 : X \to A$ and $\varphi_2 : X \to B$, there exists a unique $u : X \to P$ such that the following diagram commutes.



Problem 10.13. Prove that in Set, the Cartesian product satisfies the universal property for product.

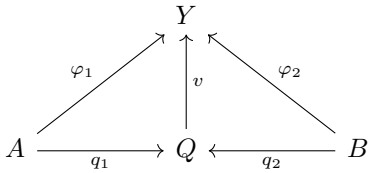
Proposition. Let A and B be objects in a category C. Then their product is unique up to isomorphism.

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 10.14. 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 10.7. 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 10.15. Prove that in Set, the disjoint union satisfies the universal property for coproduct. Prove that the universal property characterizes disjoint union.

Problem 10.16. Describe the product and coproduct in $\mathsf{Set}^{\mathsf{op}}$.

Recall that Set does not have a zero object. We will take Grp as an example to show more categorical definitions and reformulate our notions of Grp.

Problem 10.17. 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 10.8. 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 10.9. 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 $h = g \circ \varphi$.

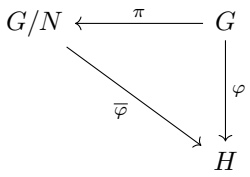
Problem 10.18. Prove the set-theoretic definition of kernels coincide with the categorical definition of kernels.

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. □

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 $\overline{\varphi} \circ \pi = \varphi$.

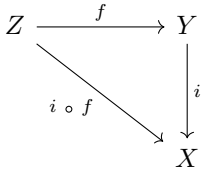


Problem 10.19. Prove that the quotient groups satisfy the universal property and the universal property characterizes quotient groups.

Proposition. Every group can be written as a quotient group of a free group.

Proof. □

Universal property for subspace topology. Let $(X, \mathcal{T}_X)$ be a topological space, let Y be a subset of X, and let $i : Y \hookrightarrow X$ be the natural inclusion. For every topological space $(Z, \mathcal{T}_Z)$ and every function $f : Z \to Y$, f is continuous if and only if $i \circ f : Z \to X$ is continuous.

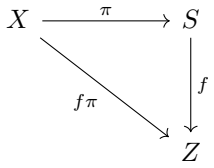


Problem 10.20. Prove that the composition of two continuous functions is continuous.

Proposition. Any topology satisfies this property is the subspace topology.

Proof. Let $i \circ f$ be a continuous function. Since i is continuous, for all $U \in \mathcal{T}_X$, there exists $V \in \mathcal{T}_Y$ and $W \in \mathcal{T}_Z$ such that $(i \circ f)^{-1}(U) = W$ and $i^{-1}(U) = V$, so $f^{-1}(V) = W$, so f is continuous. Let $\mathcal{T}$ be any topology on Y with the property and let $\mathcal{T}_Y$ be the subspace topology. Consider the left diagram, since id_Y is continuous, $i \circ \mathrm{id}_Y = i$ is continuous, then $\mathcal{T}_Y \subset \mathcal{T}$. Now consider the right diagram, since i is continuous, id_Y is continuous, then $\mathcal{T} \subset \mathcal{T}_Y$. Hence $\mathcal{T} = \mathcal{T}_Y$. □

Universal property for quotient topology. Let $(X, \mathcal{T}_X)$ be a topological space, let S be a set, and let $\pi : X \twoheadrightarrow S$ be surjective. For every topological space $(Z, \mathcal{T}_Z)$ and every function $f : S \to Z$, f is continuous if and only if $f\pi : X \to Z$ is continuous.



Problem 10.21. Prove that the quotient topology satisfies the universal property. Prove that any topology satisfies this property is the quotient topology.

11 Deep Water

> The furies are at home in the mirror; it is their address.
> Even the clearest water, if deep enough can drown.
>
> – R. S. Thomas

11.1 The Well-Ordering Theorem

We will explicitly give a proof for the well-ordering theorem using the axiom of choice.

Proposition. If W_1 and W_2 are well-ordered, then one of the following conditions hold.

1. W_1 and W_2 have the same order-type.
2. W_1 has the same order-type as an initial segment of W_2.
3. W_2 has the same order-type as an initial segment of W_1.

Proof. Denote the initial segment of W_i by x by $W_i(x)$. Let $f : W_1 \to W_2$ be a function defined by $f(x) = y$ if $W_1(x)$ has the same order-type as $W_2(y)$. It is trivial that this map is well-defined and order-preserving. Let $y_1 = y_2$, then there exist $x_1, x_2 \in W_1$ such that $W_1(x_1) \approx W_2(y_1) = W_2(y_2) \approx W_1(x_2)$, so $W_1(x_1) \approx W_1(x_2)$, which implies $x_1 = x_2$. Hence f is injective. Let $h : W_1(x) \to W_2(y)$ be an order-preserving bijection. If $x' < x$, since f is order-preserving, $W_1(x') \approx W_2(h(x'))$. If $\mathrm{dom}(f) = W_1$ and $\mathrm{ran}(f) = W_2$, then it is exactly case (i). If $\mathrm{ran}(f) \neq W_2$. Let y_0 be the least element in $W_2 \setminus \mathrm{ran}(f)$, then $W_2(y_0) = \mathrm{ran}(f)$. Suppose $\mathrm{dom}(f) \neq W_1$, then pick the least element of $W_1 \setminus \mathrm{dom}(f)$, denoted x_0. It is trivial that $f(x_0) = y_0$, a contradiction. This is case (ii). Similarly, if $\mathrm{dom}(f) \neq W_1$, define the same x_0. By contradiction, we have case (iii). □

Definition 11.1. 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 11.1. 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 11.2. The intersection of ordinals is an ordinal.

Definition 11.2. An ordinal λ is said to be a *limit ordinal* if $\lambda \neq \varnothing$ and for all $\alpha < \lambda$, $\alpha + 1 < \lambda$.

Problem 11.3. 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*.

Problem 11.4. Prove that two different ordinals do not have the same order-type.

Proposition. If a set is well-ordered, then it has the same order-type as an ordinal.

Proof. Let W be well-ordered. For all $x \in W$, consider the initial segment I_x given by x. If $I_x \approx \alpha_x$, where α_x is some ordinal, then define $F(x) = \alpha_x$. Suppose we have a set $M \subset W$ consisting of elements $x \in W$ such that I_x does not exist. Let m be the least element in M. Consider the set $\{\alpha_y \mid y < m\}$, this is a set of ordinals. Define $\alpha = \sup\{\alpha_y \mid y < m\}$. It is trivial that $I_m \approx \alpha$, a contradiction. Hence $F : W \to \text{On}$ is well-defined, by the axiom of replacement, $F(W) = \{\alpha_x\}$ is a set. For all $x, y \in W$ and $x < y$, $I_x \subsetneq I_y$, so $\alpha_x \in \alpha_y$, which is $F(x) < F(y)$. This also shows F is injective. Let $\gamma = \min(\text{On} \setminus F(W))$, then $F(W) \subset \gamma$. If $\beta < \gamma$ is an ordinal, then $\beta \in F(W)$, so $F(W) = \{\beta \mid \beta < \gamma\} = \gamma$. Hence $F : W \to \gamma$ is an order-preserving bijection. □

We can conclude that every well-ordered set has the same order-type as a unique ordinal.

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

Proof. Let λ be an ordinal and $\lambda \neq \varnothing$. Assume λ is not a successor ordinal, then $\lambda \neq \alpha + 1$ for any ordinal α. Now suppose λ is not a limit ordinal, then there exists β such that $\beta < \lambda \not< \beta + 1$. Since $\lambda \neq \alpha + 1$ for all α, $\lambda < \beta + 1$, then either $\lambda \in \beta$ or $\lambda = \beta$, a contradiction. □

Problem 11.5. 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 $\alpha \in C$.

Use the previous proposition, prove that C is the class of all ordinals. This is called the *transfinite induction*.

Theorem 11.1 (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. □

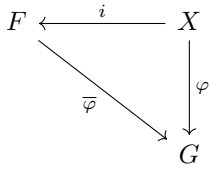
By the well-ordering theorem, $\mathbb{N}$ is well-ordered, this is called the *well-ordering principle*.

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 [?]

11.2 Free Groups and Group Presentations

Recall that to generate a group G, we need a larger group $S \supset G$ containing the generating set. Intuitively, adding relators

Definition 11.3. 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 $\overline{\varphi} : F \to G$ such that $i \circ \overline{\varphi} = \varphi$.



We will first show the existence of free groups. Consider an arbitrary set X. The elements of X are called *generators*. Let Y be a disjoint copy of X with a bijection $f : X \to Y$. For all $x \in X$, if $f(x) = y$, then we say y is the *inverse* of x, denoted x^{-1}. We can rewrite $Y = X^{-1}$ and denote $X^{\pm 1} = X \cup X^{-1}$.

Definition 11.4. Let X be a set. The elements of $X^{\pm 1}$ are called *letters*. A *word* w is a finite sequence $(l_1, \ldots, l_n)$ of letters. The *length* of $w = (l_1, \ldots, l_n)$, denoted $|w|$, is the value n. If $n = 0$, we call it the *empty word*, denoted 1. A word is *reduced* if it contains no part of the form aa^{-1}.

Let the set $F(X)$ be the set of all reduced words. Let $x = (x_1, \ldots, x_m), y = (y_1, \ldots, y_n) \in F(X)$. Pick the largest k such that $0 \leq k \leq \min(m, n)$ and $x_{m-i} = y_{i+1}^{-1}$ for all $i \leq k$. Define the composition $\cdot : F(X) \times F(X) \to F(X)$ to be the following function:

$$x \cdot y = \begin{cases} (x_1, \ldots, x_{m-k}, y_{k+1}, \ldots, y_n) & \text{if } k < \min(m, n) \\ (x_1, \ldots, x_{m-k}) & \text{if } k = n < m \\ (y_{k+1}, \ldots, y_n) & \text{if } k = m < n \\ 1 & \text{if } k = m = n \end{cases}$$

Problem 11.6. Prove that $\cdot$ is a well-defined composition and prove that 1 is the identity. Prove that the inverse of a reduced word is also reduced.

Proposition. The triple $(F(S), \cdot, 1)$ is a group.

Proof. It is left to show that $\cdot$ is associative. □

Proposition. The group $F(S)$ is satisfies the definition of a free group.

Proof. □

Problem 11.7. Prove that $F(S)$ is the unique group, up to group isomorphism, that satisfies the definition.

Definition 11.5. A *group presentation* is a pair (S, R), where S is a set and $R \subset F(S)$. 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.

11.3 More on Freeness

We have shown the existence of a free group. The idea of a free object can be generalized in every category. This leads to the question: is there a free set in Set?

References

Alphabetical Index