Problem Solving in Mathematics Solution Manual

Example Proofs and Solutions for Theorems and Problems

15-076 v0.1

이재호 2017-06-26

Contents

1 Sets

1.1 Sets

- A set is a collection of distinct objects with a precise description that provides a way of deciding whether given objects are in it.
- $x = y$: x and y are equal.
- $x \neq y$: x and y are not equal.

Definition 1.1. The objects in a set are its elements or members. When x is an element of A, we write $x \in A$ and say x belongs to A. When x is not in A, we write $x \notin A$.

We use the symbols $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ to denote the following sets of numbers:

- $\mathbb{N} = \{0, 1, 2, 3, \cdots\}$ is the set of natural numbers.
- $\mathbb{Z} = \{\cdots, -2, -1, 0, 1, 2, \cdots\}$ is the set of integers.
- $\mathbb{Q} = \left\{\frac{a}{b} \middle| a, b \in \mathbb{Z} \text{ and } b \neq 0\right\}$ is the set of rational numbers.
- $\mathbb{R}$ is the set of real numbers.
- $\mathbb{C} = \left\{a + ib \middle| a, b \in \mathbb{R} \text{ and } i = \sqrt{-1}\right\}$

Definition 1.2. Sets A and B are equal, written $A = B$, if they have the same elements. The empty set, written $\varnothing$, is the unique set with no elements.

1.2 Subsets

Definition 1.3. Let A and B be sets. We say that A is a subset of B, written $A \subseteq B$, if all elements of A are also elements of B. B is a proper subset of a set A, written $A \subset B$, if B is a subset of A that is not A itself.

Proposition 1.4. Let A, B, and C be sets. Then the following properties hold:

1. $A \subseteq A$

2. $A \subseteq B \land B \subseteq A \Rightarrow A = B$

3. $A \subseteq B \land B \subseteq C \Rightarrow A \subseteq C$

Proof. 1. $A \subseteq A$

$$\begin{aligned} A = A &\Rightarrow (x \in A \Leftrightarrow x \in A) && \text{by Definition 1.2} \\ &\Rightarrow (x \in A \Rightarrow x \in A) \\ &\Rightarrow (A \subseteq A) && \text{by Definition 1.3} \end{aligned}$$

2. $A \subseteq B \land B \subseteq A \Rightarrow A = B$

$$\begin{aligned} (A \subseteq B \land B \subseteq A) &\Rightarrow [(x \in A \Rightarrow x \in B) \land (x \in B \Rightarrow x \in A)] && \text{by Definition 1.3} \\ &\Rightarrow (x \in A \Leftrightarrow x \in B) \\ &\Rightarrow A = B && \text{by Definition 1.2} \end{aligned}$$

3. $A \subseteq B \land B \subseteq C \Rightarrow A \subseteq C$

$$\begin{aligned} (A \subseteq B \land B \subseteq C) &\Rightarrow [(x \in A \Rightarrow x \in B) \land (x \in B \Rightarrow x \in C)] && \text{by Definition 1.3} \\ &\Rightarrow (x \in A \Rightarrow x \in C) \\ &\Rightarrow A \subseteq C && \text{by Definition 1.3} \end{aligned}$$

□

Definition 1.5. The power set of a set A, denoted by $\mathcal{P}(A)$, is a set of all subsets of A.

Example 1.6. Let $x \neq y$ and $A = \{x, y\}$. Then,

$$\mathcal{P}(A) = \{\varnothing, \{x\}, \{y\}, \{x, y\}\}$$

1.3 Set Operators

Definition 1.7. Let A and B be sets. Their union, written $A \cup B$, consists of all elements in A or B. Their intersection, written $A \cap B$, consists of all elements in both A and B. Their difference, written $A - B$, consists of all elements of A that are not in B. Two sets are disjoint if their intersection is an empty set $\varnothing$. If a set A is contained in some universe $\mathcal{U}$ under discussion, then the complement A^c of A is the set of elements of $\mathcal{U}$ not in A.

Proposition 1.8. Let A and B be sets. Then,

$$A \subseteq B \Leftrightarrow A \cup B = B$$

Proof. 1. $A \subseteq B \Rightarrow A \cup B = B$

$$\begin{aligned} A \subseteq B &\Rightarrow (x \in A \Rightarrow x \in B) && \text{by Definition 1.3} \\ &\Rightarrow (x \in A \vee x \in B \Rightarrow x \in B) \\ &\Rightarrow (x \in A \cup B \Rightarrow x \in B) && \text{by Definition 1.7} \\ &\Rightarrow [(x \in A \cup B \Rightarrow x \in B) \wedge B \subseteq A \cup B] && \text{by Definition 1.7} \\ &\Rightarrow (A \cup B \subseteq B \wedge B \subseteq A \cup B) && \text{by Definition 1.3} \\ &\Rightarrow A \cup B = B && \text{by Proposition 1.4} \end{aligned}$$

2. $A \subseteq B \Leftarrow A \cup B = B$

$$\begin{aligned} A \cup B = B &\Rightarrow A \cup B \subseteq B && \text{by Proposition 1.4} \\ &\Rightarrow (A \cup B \subseteq B \wedge A \subseteq A \cup B) && \text{by Definition 1.7} \\ &\Rightarrow A \subseteq B && \text{by Proposition 1.4} \end{aligned}$$

□

Proposition 1.9. Let A, B, and C be sets. Then

1. $A \cup A = A$
2. $A \cup B = B \cup A$
3. $(A \cup B) \cup C = A \cup (B \cup C)$

Proof. 1. $A \cup A = A$

$$\begin{aligned} x \in A \cup A &\Leftrightarrow x \in A \vee x \in A && \text{by Definition 1.7} \\ &\Leftrightarrow x \in A \end{aligned}$$

Therefore, $A \cup A = A$.
2. $A \cup B = B \cup A$

$$\begin{aligned} x \in A \cup B &\Leftrightarrow (x \in A \vee x \in B) && \text{by Definition 1.7} \\ &\Leftrightarrow (x \in B \vee x \in A) \\ &\Leftrightarrow x \in B \cup A && \text{by Definition 1.7} \end{aligned}$$

Therefore, $A \cup B = B \cup A$.

3. $(A \cup B) \cup C = A \cup (B \cup C)$

$$\begin{aligned} x \in (A \cup B) \cup C &\Leftrightarrow x \in A \cup B \vee x \in C && \text{by Definition 1.7} \\ &\Leftrightarrow (x \in A \vee x \in B) \vee x \in C && \text{by Definition 1.7} \\ &\Leftrightarrow x \in A \vee (x \in B \vee x \in C) \\ &\Leftrightarrow x \in A \vee x \in B \cup C \\ &\Leftrightarrow x \in A \cup (B \cup C) \end{aligned}$$

□

Example 1.10. Let A and B be sets. Show that

$$A \subseteq B \Leftrightarrow A \cap B = A.$$

Proof. In the same way as in Proposition 1.8. □

Example 1.11. Let A, B, and C be sets. Then prove the following properties.

1. $A \cap A = A$
2. $A \cap B = B \cap A$
3. $(A \cap B) \cap C = A \cap (B \cap C)$

Proof. In the same way as in Proposition 1.9. □

Proposition 1.12. Let A, B, C be sets. Then

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

Proof. 1. $A \cup (A \cap B) = A$

$$\begin{aligned} x \in A \cap B &\Rightarrow (x \in A \wedge x \in B) && \text{by Definition 1.7} \\ &\Rightarrow x \in A \end{aligned}$$

Thus, $A \cap B \subseteq A$. Therefore, by Proposition 1.8, $A \cup (A \cap B) = A$.

2. $A \cap (A \cup B) = A$

$$\begin{aligned} x \in A &\Rightarrow (x \in A \vee x \in B) \\ &\Rightarrow x \in A \cup B && \text{by Definition 1.7} \end{aligned}$$

Thus, $A \subseteq A \cup B$. Therefore, by Example 1.10, $A \cap (A \cup B) = A$.

3. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

$$\begin{aligned} x \in A \cup (B \cap C) &\Leftrightarrow (x \in A \vee x \in B \cap C) && \text{by Definition 1.7} \\ &\Leftrightarrow [x \in A \vee (x \in B \wedge x \in C)] \\ &\Leftrightarrow [(x \in A \vee x \in B) \wedge (x \in A \vee x \in C)] \\ &\Leftrightarrow (x \in A \cup B \wedge x \in A \cup C) && \text{by Definition 1.7} \\ &\Leftrightarrow x \in (A \cup B) \cap (A \cup C) && \text{by Definition 1.7} \end{aligned}$$

Therefore, $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

4. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

$$\begin{aligned} x \in A \cap (B \cup C) &\Leftrightarrow (x \in A \wedge x \in B \cup C) && \text{by Definition 1.7} \\ &\Leftrightarrow [x \in A \wedge (x \in B \vee x \in C)] && \text{by Definition 1.7} \\ &\Leftrightarrow [(x \in A \wedge x \in B) \vee (x \in A \vee x \in C)] \\ &\Leftrightarrow (x \in A \cup B) \vee (x \in A \cup C) && \text{by Definition 1.7} \\ &\Leftrightarrow x \in (A \cap B) \cup (A \cap C) && \text{by Definition 1.7} \end{aligned}$$

Therefore, $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$. □

Example 1.13. Let A and B be sets. Show that

$$A \subseteq B \Leftrightarrow A - B = \varnothing.$$

Proof. We first show that $A \subseteq B \Rightarrow A - B = \varnothing$.

$$\begin{aligned} A \subseteq B &\Rightarrow (x \in A \Rightarrow x \in B) && \text{by Definition 1.3} \\ &\Rightarrow [(x \in A \wedge x \notin B) \Rightarrow (x \in B \wedge x \notin B)] \\ &\Rightarrow [(x \in A \wedge x \notin B) \Rightarrow x \in \varnothing] \\ &\Rightarrow (x \in A - B \Rightarrow x \in \varnothing) && \text{by Definition 1.7} \\ &\Rightarrow A - B \subseteq \varnothing && \text{by Definition 1.3} \\ &\Rightarrow A - B \subseteq \varnothing \wedge \varnothing \subseteq A - B \\ &\Rightarrow A - B = \varnothing && \text{by Proposition 1.4} \end{aligned}$$

We now show that $A - B = \varnothing \Rightarrow A \subseteq B$.

$$
\begin{aligned}
A - B = \varnothing &\Rightarrow A - B \subseteq \varnothing && \text{by Proposition 1.4} \\
&\Rightarrow (x \in A - B \Rightarrow x \in \varnothing) && \\
&\Rightarrow [(x \in A \wedge x \notin B) \Rightarrow x \in \varnothing] && \text{by Definition 1.7} \\
&\Rightarrow [(x \in A \wedge x \notin B) \vee x \in B \Rightarrow x \in \varnothing \vee x \in B] && \\
&\Rightarrow [(x \in A \vee x \in B) \wedge (x \notin B \vee x \in B) \Rightarrow x \in \varnothing \cup B \wedge \varnothing \subseteq B] && \text{by Definition 1.7} \\
&\Rightarrow (x \in A \cup B \wedge x \in \mathcal{U} \Rightarrow x \in B) && \text{by Def. 1.7 and Prop. 1.8} \\
&\Rightarrow (x \in A \cup B \Rightarrow x \in B) && \\
&\Rightarrow A \cup B \subseteq B && \text{by Definition 1.3} \\
&\Rightarrow A \cup B \subseteq B \wedge B \subseteq A \cup B && \text{by Definition 1.7} \\
&\Rightarrow A \cup B = B && \text{by Proposition 1.8}
\end{aligned}
$$

□

Example 1.14. Let A, B,and C be sets. Prove the following properties:

1. $A - (B \cup C) = (A - B) \cap (A - C)$
2. $A - (B \cap C) = (A - B) \cup (A - C)$

Proof. Let's first prove that $A - B = A \cap B^{\mathsf{c}}$.

$$
\begin{aligned}
x \in A - B &\Leftrightarrow x \in A \wedge x \notin B && \text{by Definition 1.7} \\
&\Leftrightarrow x \in A \wedge x \in B^{\mathsf{c}} && \text{by Definition 1.7} \\
&\Leftrightarrow x \in A \cap B^{\mathsf{c}} &&
\end{aligned}
$$

Thus, $A - B = A \cap B^{\mathsf{c}}$. Now we prove the following:

$$
\begin{aligned}
(A \cup B)^{\mathsf{c}} &= A^{\mathsf{c}} \cap B^{\mathsf{c}} \\
(A \cap B)^{\mathsf{c}} &= A^{\mathsf{c}} \cup B^{\mathsf{c}}.
\end{aligned}
$$

First show that $(A \cup B)^{\mathsf{c}} = A^{\mathsf{c}} \cap B^{\mathsf{c}}$.

$$
\begin{aligned}
x \in (A \cup B)^{\mathsf{c}} &\Leftrightarrow \neg(x \in A \cup B) && \text{by Definition 1.7} \\
&\Leftrightarrow \neg(x \in A \vee x \in B) && \text{by Definition 1.7} \\
&\Leftrightarrow (\neg x \in A) \wedge (\neg x \in B) && \\
&\Leftrightarrow x \in A^{\mathsf{c}} \wedge x \in B^{\mathsf{c}} && \text{by Definition 1.7} \\
&\Leftrightarrow x \in A^{\mathsf{c}} \cap B^{\mathsf{c}} && \text{by Definition 1.7}
\end{aligned}
$$

Showing $(A \cap B)^{\mathsf{c}} = A^{\mathsf{c}} \cup B^{\mathsf{c}}$ is analogous.

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

$$\begin{aligned} A - (B \cup C) &= A \cap (B \cup C)^{\mathsf{c}} \\ &= A \cap (B^{\mathsf{c}} \cap C^{\mathsf{c}}) && \text{by Definition 1.7} \\ &= (A \cap B^{\mathsf{c}}) \cap (A \cap C^{\mathsf{c}}) && \text{by Example 1.11} \\ &= (A - B) \cap (A - C) \end{aligned}$$

2. $A - (B \cap C) = (A - B) \cup (A - C)$
It is analogous to the above. □

Definition 1.15. A list with entries in A consists of elements of A in a specified order, with repetition allowed. A k-tuple is a list with k entries. We write A^k for the set of k-tuples with entries in A. An ordered pair is a list with two entries. The Cartesian product of sets S and T, written $S \times T$, is the set $\{(x, y) | x \in S, y \in T\}$. Two ordered pairs (a, b) and (c, d) are equal, written $(a, b) = (c, d)$, if $a = c$ and $b = d$.

2 Language of Mathematics

2.1 Mathematical Statement

- A proposition or mathematical statement is a sentence which is either true or false, but not both.
  - "Every even integer greater than 2 may be written as the sum of two prime numbers."
- A predicate is a statement which can not be determined true or false due to free variables.
  - "n is a prime number."
- A statement is either a proposition or a predicate. A single capital letter P, Q, etc. will be used to indicate a statement, or sometiems $P(m, n)$ to indicate a predicate with free variables listed.
  - "π is a special number" is not a statement.

2.2 Logical Connectives

Or

"P or Q" is called the disjunction of the two statements P and Q. It is denoted as $P \vee Q$. $P \vee Q$ corresponds to the following truth table:

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

And

"P and Q" is called the conjunction of the two statements P and Q. It is denoted as $P \wedge Q$. $P \wedge Q$ corresponds to the following truth table:

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

Not

"Not P" is called the negation of the statement P. It is denoted as $\neg P$. $\neg P$ corresponds to the following truth table:

P	$\neg P$
T	F
F	T

2.3 Quantifiers and Logical Statements

Quantifiers

Definition 2.1. In the statement "For all x in S, $P(x)$ is true," the variable x is universally quantified. It can be written:

$$(\forall x \in S)P(x),$$

where $\forall$ is a universal quantifier. In "There exists an x in S such that $P(x)$ is true," x is existentially quantified. It can be written:

$$(\exists x \in S)P(x)$$

where $\exists$ is an existential quantifier. The set of allowed values for a variable is its universe.

Order of Quantifiers

Consider the sentence "Ther is a real number y such that $x = y^3$ for every real number. It seems to say that some number y is the cube root of all numbers, which is obviously false. To say that every number has a cube root, we write "For every real number x, there is a real number y such that $x = y^3$.

Compare the two:

$$(\forall x \in A)(\exists y \in B)P(x,y) \qquad (\exists y \in B)(\forall x \in A)P(x,y)$$

Ther first is true if for each x we can take y that works. For the second statement to be true, there must be a single y that will always work, no matter which x is chosen.

Negation of Quantified Statements

$$\neg[(\forall x)P(x)] \Leftrightarrow (\exists x)[\neg P(x)]$$
$$\neg[(\exists x)P(x)] \Leftrightarrow (\forall x)[\neg P(x)]$$

Example 2.2. "Every classroom has a chair that is not broken" can be negated as following. Let R be a set of all classrooms, and let $C(r)$ be a set of all chairs in a classroom $r \in R$. Suppose a statement $P(c)$ be true if a chair c is broken and flse if it is not broken. Then, the given statement can be written:

$$(\forall r \in R)(\exists c \in C(r))(\neg P(c)).$$

The negation of this is then:

$$\begin{aligned}&\neg[(\forall r \in R)(\exists c \in C(r))(\neg P(c))]\\ \Leftrightarrow&(\exists r \in R)(\forall c \in C(r))P(c),\end{aligned}$$

which is "There is a classroom in which all chairs in it are broken."

2.4 Compound Statements

Definition 2.3. Let P and Q be statements. The logical connective conditional, written by $P \Rightarrow Q$, means that P implies Q. Biconditional, written by $P \Leftrightarrow Q$, means that P iff Q. In the conditional statement $P \Rightarrow Q$, we call P the hypothesis and Q the conclusion. The statement $Q \Rightarrow P$ is the converse of $P \Rightarrow Q$.

The following is a truth table for $P \Rightarrow Q$.

P	Q	$P \Rightarrow Q$	$(\neg P) \vee Q$
T	T	T	T
T	F	F	F
F	T	T	T
F	F	T	T

Note that $(P \Rightarrow Q) \Leftrightarrow [(\neg P) \vee Q]$. The following is a truth table for $P \Leftrightarrow Q$.

P	Q	$P \Leftrightarrow Q$
T	T	T
T	F	F
F	T	F
F	F	T

Logical Connectives and Membership in Sets

Let $P(x)$ and $Q(x)$ be statements about an element x from a universe $\mathcal{U}$. We often write a conditional statement $(\forall x \in \mathcal{U})(P(x) \Rightarrow Q(x))$ as $P(x) \Rightarrow Q(x)$, or simply $P \Rightarrow Q$ with an implicit universal quantifier.

The hypothesis $P(x)$ can be interpreted as a universal quantifier in another way. With $A = \{x \in \mathcal{U} \mid P(x) \text{ is true}\}$, the statement $P(x) \Rightarrow Q(x)$ can be written as $(\forall x \in A)Q(x)$. Another interpretation of $P(x) \Rightarrow Q(x)$ uses set inclusion. With $B = \{x \in \mathcal{U} \mid Q(x)\text{is true}\}$, the conditional statement has the same meaning as the statement $A \subseteq B$. The converse statement $Q(x) \Rightarrow P(x)$ is $B \subseteq A$. Thus the biconditional $P \Leftrightarrow Q$ is equivalent to $A = B$.

2.5 Proofs

A proof of a mathematical statement is a logical argument which shows the truth of the statement. The logical argument consists of several steps provided by implications. In this chapter, a variety of methods of proof will be described.

Direct Proofs

Most of theorems are of the form $P \Rightarrow Q$. Since the statement is necessarily true if P is false, we only need to consider the case when P is true. Thus to prove $P \Rightarrow Q$, it is sufficient ot assume that P is true and deduce Q is true by logical arguments. This is the direct proof.

Example 2.4. Show that for positive real numbers a and b, $a < b \Rightarrow a^2 < b^2$.

Proof.

$$\begin{aligned} a < b &\Rightarrow a \times a < b \times a \wedge a \times b < b \times b \\ &\Rightarrow a^2 < ba \wedge ab < b^2 \\ &\Rightarrow a^2 < ab \wedge ab < b^2 \\ &\Rightarrow a^2 < b^2 \end{aligned}$$

□

Example 2.5 (Proof by Cases)**.** Prove that $a^2 > 0$ for non-zero real number a.

Proof. For the case $a > 0$, $a \times a > 0 \times a = 0$. Hence, $a^2 > 0$ for positive a.

For the case $a < 0$, $a \times a > 0 \times a = 0$, since a is negative. Therfore, $a^2 > 0$ for any non-zero real number. □

Example 2.6 (Constructing Proofs Backwards)**.** Prove that $a < b \Rightarrow 4ab < (a+b)^2$ for real numbers a and b.

Proof.

$$\begin{aligned} 4ab < (a+b)^2 &\Leftarrow 4ab < a^2 + 2ab + b^2 \\ &\Leftarrow 0 < a^2 - 2ab + b^2 \\ &\Leftarrow 0 < (a-b)^2 \\ &\Leftarrow a - b \neq 0 \\ &\Leftarrow a - b < 0 \vee a - b > 0 \\ &\Leftarrow a < b \vee a > b \\ &\Leftarrow a < b \end{aligned}$$

□

Contrapositive

The direct method can be inconvenient and does not always work. In this section we will consider a logically equivalent but very commonand useful method of proof. The contrapositive of $P \Rightarrow Q$ is $\neg Q \Rightarrow \neg P$. The equivalence between a conditional and its contrapositive allows us to prove $P \Rightarrow Q$ by proving $\neg Q \Rightarrow \neg P$. This is the contrapositive method.

Example 2.7. Let $f(x) = mx + b$. Show that if $x \neq y$, then $f(x) \neq f(y)$.

Proof. The contrapositive of the given statement is:

$$f(x) = f(y) \Rightarrow x = y,$$

where $f(x) = mx + b$.

When $f(x) = f(y)$, we obtain $mx + b = my + b$. If $b \neq 0$, then $x = y$. For the case when $b = 0$, the statement is false. Therefore, the statement requires a condition $b \neq 0$ to be true. □

A universally quantified statement like $(\forall x \in \mathcal{U})[P(x) \Rightarrow Q(x)]$ can be disproved by finding an element x in $\mathcal{U}$ such that $P(x)$ is true but $Q(x)$ is false. Such an element is a counterexample.

Example 2.8. Prove that if a is less than or equal to every real number greater than b, then $a \leq b$.

Proof. The given statement can be written:

$$(\forall r > b : a \leq r) \Rightarrow a \leq b,$$

and its contrapositive is:

$$a > b \Rightarrow (\exists r > b : a > r).$$

To show this, it is sufficient to find such r.

Let $r = \frac{a+b}{2}$. Then,

$$\begin{aligned} a > b &\Rightarrow a + a > a + b \\ &\Rightarrow 2a > a + b \\ &\Rightarrow \frac{2a}{2} > \frac{a+b}{2} \\ &\Rightarrow a > \frac{a+b}{2} \\ &\Rightarrow a > r. \end{aligned}$$

Since the contrapositive of the given statement is true, the original statement is also true. □

Indirect Proof

Negating both sides $(P \Rightarrow Q) \Leftrightarrow \neg[P \wedge (\neg Q)]$. Hence we can prove $P \Rightarrow Q$ by proving P and $\neg Q$ cannot both be true. We do this by obtaining a contradiction after assuming both P and $\neg Q$. This is the method of contradiction or indirect proof.

Example 2.9. Show that among the numbers $y_1, \ldots, y_n$, some number is as large as the average.

Proof. Suppose, for the sake of contradiction, that there is no number as large as the average, given n numbers $y_1, \ldots, y_n$. That is,

$$\begin{aligned} &\neg \left[(\exists y \in \{y_1, \ldots, y_n\})\ y \geq \frac{y_1 + \cdots + y_n}{n} \right] \\ \Leftrightarrow &(\forall y \in \{y_1, \ldots, y_n\})\ y < \frac{y_1 + \cdots + y_n}{n}. \end{aligned}$$

The sum of all numbers of $\{y_1, \ldots, y_n\}$ be S. Then,

$$\begin{aligned} S &= \sum_{i=1}^{n} y_i \\ &< \sum_{i=1}^{n} \frac{y_1 + \cdots + y_n}{n} \\ &= n \cdot \frac{y_1 + \cdots + y_n}{n} \\ &= y_1 + \cdots + y_n \\ &= S \ \lightning \end{aligned}$$

Hence the assumption that no number is as large as the average must be false. Thus some number is as large as the average. □

Example 2.10. Show that there is no largest real number.

Proof. For the sake of contradiction, suppose there is the largest real number α. Since the set of real numbers $\mathbb{R}$ is closed under addition, $\alpha + 1 \in \mathbb{R}$. However, $\alpha + 1 > \alpha$, which is greater than α. $\lightning$
Hence the assumption that there exists the largest real number is not true. Therefore, there is no largest real number. □

2.6 Exercises for Chapter 1 to 2

Example 2.11. Prove that, if $a \le b$, then $[a, b] \subseteq (c, d)$ iff $c < a$ and $b < d$.

Proof. First, prove that $c < a \wedge b < d \Rightarrow [a, b] \subseteq (c, d)$ when $a \le b$. Choose x such that $a \le x \le b$, i.e., $x \in [a, b]$.

$$\begin{aligned} a \le x \le b \wedge c < a \wedge b < d &\Rightarrow c < a \le x \le b < d \\ &\Rightarrow c < x < d \\ &\Rightarrow x \in (c, d) \end{aligned}$$

Now we prove that $[a, b] \subseteq (c, d) \Rightarrow c < a \wedge b < d$. To do so, we suppose $c \ge a \vee b \ge d$ for the sake of contradiction.

We consider the case $c \ge a$. Since $c \ge a \wedge c < d$, the following relationship holds:

$$a \le c < d.$$

Choose any x such that $a \le x \le \min\{b, c\}$, which implies $x \in [a, b]$. By Definition 1.3, $x \in (c, d)$, which implies $c < x < d$. However, x is chosen so that $x \le c$. The case when $b \ge d$ can be shown to be contradictive in a similar manner. $\lightning$ Thus, our assumption that $c \ge a \vee b \ge d$ is false.

Therefore, $[a, b] \subseteq (c, d) \Leftrightarrow c < a \wedge b < d$. □

Example 2.12. Prove that, if $A \cap B \subseteq C$ and $x \in B$, then $x \notin A - C$.

Proof. What we want to prove is the following statement:

$$A \cap B \subseteq C \wedge x \in B \Rightarrow x \notin A - C.$$

Since $A \cap B \subseteq C$, $x \in A \cap B \Rightarrow x \in C$.

$$\begin{aligned}
x \in B &\Rightarrow x \in B \wedge x \in \mathcal{U} \\
&\Rightarrow x \in B \wedge x \in A \cup A^{\mathsf{c}} \\
&\Rightarrow x \in B \wedge (x \in A \vee x \in A^{\mathsf{c}}) \\
&\Rightarrow (x \in B \wedge x \in A) \vee (x \in B \wedge x \in A^{\mathsf{c}}) \\
&\Rightarrow x \in B \cap A \vee x \in B \cap A^{\mathsf{c}} \\
&\Rightarrow x \in C \vee x \in A^{\mathsf{c}} \\
&\Rightarrow x \in C \cup A^{\mathsf{c}} \\
&\Rightarrow x \notin (C \cup A^{\mathsf{c}})^{\mathsf{c}} \\
&\Rightarrow x \notin C^{\mathsf{c}} \cap A \\
&\Rightarrow x \notin A - C
\end{aligned}$$

□

3 Inequalitites

3.1 The Real Number System

Axiom 1 (Field Axioms)**.** A set $\mathcal{F}$ with operation $+$, $\cdot$, and distinguished elements 0 and 1 with $0 \neq 1$ is a field if the follwoing properties hold for all $a, b, c \in \mathcal{F}$.

A0:	$a + b \in \mathcal{F}$	**M0**:	$a \cdot b \in \mathcal{F}$	Closure
A1:	$(a + b) + c = a + (b + c)$	**M1**:	$(a \cdot b) \cdot c = a \cdot (b \cdot c)$	Associativity
A2:	$a + b = b + a$	**M2**:	$a \cdot b = b \cdot a$	Commutativity
A3:	$a + 0 = a$	**M3**:	$a \cdot 1 = a$	Identity
A4:	$(\forall a \in \mathcal{F})(\exists x \in \mathcal{F})\ a + x = 0$	**M4**:	$(\forall a \in \mathcal{F} \backslash \{0\})(\exists y \in \mathcal{F})\ a \cdot y = 1$	Inverse
		DL:	$a \cdot (b + c) = a \cdot b + a \cdot c$	Distributive Law

Example 3.1. Show that the set of all rational numbers is a field but the set of all natural number and integers are not fields.

Proof. Since $\mathbb{Q}$ satisfies all the axioms given in Axiom 1, it is a field. However, $\mathbb{N}$ does not satisfy **A4**, e.g. $(\nexists x \in \mathbb{N})\ x + 1 = 0$, since such x is uniquely determined as -1–even if we consider $0 \in \mathbb{N}$. For $\mathbb{Z}$, it does not satisfies **M4**, e.g. $(\nexists x \in \mathbb{Z})\ x \cdot 2 = 1$, since such x is uniquely $\frac{1}{2}$. □

Theorem 3.2. *Additive identity and multiplicative identity are unique.*

Proof. For the sake of contradiction, suppose there are two distinct additive identities, x

and y. Then,

$$\begin{aligned} x + x = x & \qquad \text{by Axiom 1 } \mathbf{A3} \\ = x + y & \qquad \text{by Axiom 1 } \mathbf{A3} \\ = y + x & \qquad \text{by Axiom 1 } \mathbf{A2} \\ = y & \qquad \text{by Axiom 1 } \mathbf{A3} \end{aligned}$$

We deduced that $x = y$, which contradicts the assumption that x and y are distinct. $\lightning$ Therefore, additive identity is unique.

Showing the uniqueness of the multiplicative identity is exactly analogous. □

Theorem 3.3. *Additive inverse and multiplicative inverse are unique.*

Proof. For the sake of contradiction, suppose there are two distinct additive inverses, x and y, for an element a. Then,

$$\begin{aligned} x &= x + 0 & \qquad \text{by Axiom 1 } \mathbf{A3} \\ &= x + (a + y) & \qquad \text{by Axiom 1 } \mathbf{A4} \\ &= (x + a) + y & \qquad \text{by Axiom 1 } \mathbf{A1} \\ &= (a + x) + y & \qquad \text{by Axiom 1 } \mathbf{A2} \\ &= 0 + y & \qquad \text{by Axiom 1 } \mathbf{A4} \\ &= y & \qquad \text{by Axiom 1 } \mathbf{A3} \end{aligned}$$

Thus, $x = y$, which contradicts the assumption that x and y are distinct. $\lightning$

Showing the uniqueness of the multiplicative inverse in exactly analogous. □

The additive inverse of a is called the negative of a and denoted by $-a$. The multiplicative inverse of a is called reciprocal of a and denoted by a^{-1}.

Let the set of natural numbers $\mathbb{N} = \{1, 2, 3, \ldots, n, \ldots\}$. Then since $\mathbb{N}$ is a subset of $\mathbb{R}$, the operation addtion is defined. Now let $-\mathbb{N} = \{-n \mid n \in \mathbb{N}\}$. Then the set of integers $\mathbb{Z}$ can be defined as

$$\mathbb{Z} = \mathbb{N} \cup \{0\} \cup -\mathbb{N}.$$

Similarly, the set of rational numbers $\mathbb{Q}$ is defined as

$$\mathbb{Q} = \{ab^{-1} \mid a, b \in \mathbb{Z}\}.$$

Axiom 2 (Order Axioms). A positive set in a field $\mathcal{F}$ is a set $P \subseteq \mathcal{F}$ such that for $a, b \in \mathcal{F}$,

P1:	$a, b \in P \Rightarrow a + b \in P$	Closure under Addition
P2:	$a, b \in P \Rightarrow a \cdot b \in P$	Closure under Multiplication
P3:	$a \in \mathcal{F}$ implies exactly one of $a = 0$, $a \in P$, or $-a \in P$.	Trichotomy

The ordered field is a field with a positive set P. In an ordered field, we define $a < b$ by $b - a \in P$. The relations $\leq$, $>$, and $\geq$ have analogous definitions in terms of P.

Proposition 3.4. Let a, b, and c be element of an ordered field. Then,

O1:	$a \leq a$	Reflextivity
O2:	$a \leq b \wedge b \leq a \Rightarrow a = b$	Antisymmetry
O3:	$a \leq b \wedge b \leq c \Rightarrow a \leq c$	Transitivity
O4:	$a \leq b \vee b \leq a$ is true.	Total Ordering

Proof. **O1**

Since $a = a$, $a \leq a$.

O2

For the sake of contradiction, suppose $a \neq b$. Let $c = a - b$. Then,

$$\begin{aligned} c + (b - a) &= (a - b) + (b - a) \\ &= a + (-b + b) - a && \text{by Axiom 1 } \textbf{A1} \\ &= a + 0 - a && \text{by Axiom 1 } \textbf{A4} \\ &= a - a && \text{by Axiom 1 } \textbf{A3} \\ &= 0 && \text{by Axiom 1 } \textbf{A4} \end{aligned}$$

Thus, by Axiom 1 **A4**, $-c = b - a$.

$$\begin{aligned} a \neq b \wedge a \leq b \wedge a \leq b &\Rightarrow a < b \wedge b < a \\ &\Rightarrow b - a \in P \wedge a - b \in P \\ &\Rightarrow -c \in P \wedge c \in P \;\lightning && \text{by Axiom 2 } \textbf{P3} \end{aligned}$$

Therefore, our assumption that $a \neq b$ is wrong, so $a = b$.

O3

$$\begin{aligned} a \leq b \wedge b \leq c &\Rightarrow (a = b \vee b - a \in P) \wedge (b = c \vee c - b \in P) \\ &\Rightarrow (c - b) + (b - a) \in P \vee (c - b) + (b - a) = 0 && \text{by Ax. 1 } \textbf{A3} \text{ \& Ax. 2 } \textbf{P1} \\ &\Rightarrow c - a \in P \vee c - a = 0 && \text{by Axiom 1} \\ &\Rightarrow a \leq c \end{aligned}$$

O4

For the sake of contradiction, suppose none of $a \leq b$ and $b \leq a$ holds. Then, neither of $a - b = 0, a - b \in P, b - a \in P$ holds, which contradicts Axiom 2 **P3**. $\lightning$ □

Definition 3.5. If $S \subseteq \mathcal{F}$, then $\beta \in \mathcal{F}$ is an upper bound for S if $x \leq \beta$ for all $x \in S$. An upper bound α for S is the least upper bound or supremum of S if S has no upper bound less than α.

Similarly, $\beta \in \mathcal{F}$ is a lower bound for S if $x \geq \beta$ for all $x \in S$, and a lower bound α for S is the greates lower bound or infimum of S if S has no lower bound greater than α.

We use $\sup S$ and $\inf S$ for S to denote the supremum and infimum of S, if they exist.

Axiom 3 (Completeness Axiom)**.** An ordered field $\mathcal{F}$ is complete if every nonempty subset of $\mathcal{F}$ that has an upper bound in $\mathcal{F}$ has a least upper bound in $\mathcal{F}$.

The set of real numbers $\mathbb{R}$ can be defined by the complete ordered field, because any two complete ordered field is isomorphic to each other.

Example 3.6. Prove the existence of $\sqrt{2}$.

Proof. Let $S = \{x \mid x^2 < 2 \wedge x > 0\} \subset \mathbb{R}$. Since $1 \in S$ and $S \subset [0, 2]$, it is both bounded and nonempty. Thus, by Axiom 3, $\exists \sup S \in \mathbb{R}$. Let $\alpha = \sup S$.

To show $\alpha = \sqrt{2}$, we show $\neg(\alpha^2 > 2 \vee \alpha^2 < 2)$. For the sake of contradiction, suppose $\alpha^2 > 2$.

$$\begin{aligned} \alpha^2 > 2 &\Rightarrow \frac{1}{\alpha^2} < \frac{1}{2} \\ &\Rightarrow \left(\frac{2}{\alpha}\right)^2 < 2 \end{aligned}$$

Thus, we get $\alpha^2 > 2 > \left(\frac{2}{\alpha}\right)^2$. From AM–GM inequality,

$$\beta = \frac{1}{2}\left(\alpha + \frac{2}{\alpha}\right) > \sqrt{\alpha \cdot \frac{2}{\alpha}} = \sqrt{2}$$

Hence, $\beta^2 > 2$. Since β is an arithmetic mean of α and $\frac{2}{\alpha}$, and $\alpha > \frac{2}{\alpha}$, we see that $\alpha > \beta > \frac{2}{\alpha}$. Therfore, $\alpha^2 > \beta^2 > 2$. Then, β is smaller than $\alpha = \sup S$ yet it is still an upper bound of S. $\lightning$ Therefore, the assumption that $\alpha^2 > 2$ is wrong.

We now check $\alpha^2 < 2$ is also not true. Again, suppose $\alpha^2 < 2$ for the sake of contradiction.

$$\begin{aligned} \alpha^2 < 2 &\Rightarrow \frac{1}{\alpha^2} > \frac{1}{2} \\ &\Rightarrow \left(\frac{2}{\alpha}\right)^2 > 2 \end{aligned}$$

From similar argument, $\beta = \frac{1}{2}\left(\alpha + \frac{2}{\alpha}\right) > \sqrt{2}$. Thus, we have

$$\begin{aligned} \beta > \sqrt{2} \Rightarrow \beta^2 &> 2 \\ \Rightarrow \frac{1}{\beta^2} &< \frac{1}{2} \\ \Rightarrow \left(\frac{2}{\beta}\right)^2 &< 2 \end{aligned}$$

Hence, $\frac{2}{\beta} \in S$. Also,

$$\begin{aligned} \beta < \frac{2}{\alpha} \Rightarrow \frac{1}{\beta} &> \frac{\alpha}{2} \\ \Rightarrow \frac{2}{\beta} &> \alpha \\ \Rightarrow \left(\frac{2}{\beta}\right)^2 &> \alpha^2 \end{aligned}$$

Then, $\frac{2}{\beta}$ is larger than $\alpha = \sup S$ which is the upper bound. $\lightning$

Therefore, $\alpha^2 = 2$, which implies $\sqrt{2}$ exists. □

3.2 Properties of Real Numers

In this section, $a, b, c, \ldots, x, y, z$ are all real numbers.

Theorem 3.7 (Cancellation Law for Addition). *$a + b = a + c \Rightarrow b = c$.*

Proof.

$$\begin{aligned} a + b = a + c &\Rightarrow b + a = c + a && \text{by Axiom 1 } \mathbf{A2} \\ &\Rightarrow b + a + (-a) = c + a + (-a) && \text{by Axiom 1 } \mathbf{A4} \\ &\Rightarrow b + 0 = c + 0 \\ &\Rightarrow b = c && \text{by Axiom 1 } \mathbf{A3} \end{aligned}$$

□

Theorem 3.8 (Property of Subtraction). *Let a and b be given. Then there is only one x such that $a + x = b$.*

Proof. Let $c = a - b$. Then, from Axiom 1 **A4**, $(\exists y \in \mathbb{R})\ c + y = 0$.

$$\begin{aligned} c + y = 0 &\Leftrightarrow (a - b) + y = 0 \\ &\Leftrightarrow y + a - b = 0 && \text{by Axiom 1 } \mathbf{A2} \\ &\Leftrightarrow y + a - b + b = 0 + b && \text{by Axiom 1 } \mathbf{A4} \\ &\Leftrightarrow y + a = b && \text{by Axiom 1 } \mathbf{A3} \\ &\Leftrightarrow a + y = b && \text{by Axiom 1 } \mathbf{A2} \end{aligned}$$

Thus, we see that such x is uniquely determined as the additive inverse of $b - a$, which is $a - b$. □

Proposition 3.9.

1. $b - a = b + (-a)$
2. $-(-a) = a$
3. $a \cdot (b - c) = a \cdot b - a \cdot c$
4. $0 \cdot a = a \cdot 0 = 0$

Proof. 4.
Choose any $x \in \mathbb{R}$. Then,

$$\begin{aligned} 0 &= a \cdot x - a \cdot x \\ &= a \cdot (x - x) \\ &= a \cdot 0 \end{aligned}$$

□

Theorem 3.10 (Cancellation Law for Multiplication)**.** *If $a \cdot b = a \cdot c$ and $a \neq 0$, then $b = c$.*

Theorem 3.11 (Possibility of Division)**.** *Given a and b with $a \neq 0$, there is only one x such that $ax = b$.*

In Theorem 3.11, x is denoted by b/a or $\frac{b}{a}$ and called the quotient of b and a. In particular, $1/a$ is the multiplicative inverse a^{-1} called the reciprocal of a.

Proposition 3.12.

1. $x \cdot 0 = 0$
2. $(-x)y = -(xy)$
3. $-x = (-1)x$
4. $(-x)(-y) = xy$

Proposition 3.13.

1. If $a \neq 0$, then $b/a = b \cdot a^{-1}$.
2. If $a \neq 0$, then $(a^{-1})^{-1} = a$.

3. If $ab = 0$, then $a = 0$ or $b = 0$.

Proposition 3.14. 1. If $b \neq 0$ and $d \neq 0$, then $\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd}$.

2. If $b \neq 0$ and $d \neq 0$, then $\left(\frac{a}{b}\right)\left(\frac{c}{d}\right) = \frac{ac}{bd}$.

3. If $b \neq 0$, $c \neq 0$, and $d \neq 0$, then $\frac{a/b}{c/d} = \frac{ad}{bc}$.

Proposition 3.15. Properties of ordered field:
F1: $x \leq y \Rightarrow x + z \leq y + z$
F2: $x \leq y \wedge 0 \leq z \Rightarrow xz \leq yz$
F3: $x \leq y \wedge u \leq v \Rightarrow x + u \leq y + v$
F4: $0 \leq x \leq y \wedge 0 \leq u \leq v \Rightarrow xu \leq yv$

Proposition 3.16. 1. $x \leq y \Rightarrow -y \leq -x$

2. $x \leq y \wedge z \leq 0 \Rightarrow yz \leq xz$

3. $x \leq y \wedge u \leq v \Rightarrow x + u \leq y + v$

4. $0 \leq x \leq y \wedge 0 \leq u \leq v \Rightarrow xu \leq yv$

Theorem 3.17 (The Archimedean Propoerty)**.** *Given positive real numbers a and b, there exists a natural number n such that $na > b$. That is, no real number is an upper bound for the set $\mathbb{N}$.*

Proof. Suppose there is an upper bound for $\mathbb{N}$ for the sake of contradiction. From Axiom 3, there is a supremum $\alpha = \sup \mathbb{N}$, since $\mathbb{N} \subset \mathbb{R}$ and $\mathbb{N} \neq \varnothing$. Since $1 \in \mathbb{N}$, $\alpha \geq 1 > 0$. Then, $(\exists m \in \mathbb{N})\ \alpha \geq m > \alpha - 1$, since α the lowest upper bound. Thus, $m + 1 > \alpha$. However, $m + 1 \in \mathbb{N}$ is greather then $\alpha = \sup \mathbb{N}$. ↯

Therefore, $\mathbb{N}$ is not bounded above. Specifically, there always exists $n \in \mathbb{N}$ larger than $\frac{b}{a} \in \mathbb{R}$, i.e., $(\exists n \in \mathbb{N})\ na > b$. □

3.3 Elementary Inequalities

Proposition 3.18.

$$0 < a < b \Rightarrow a^2 < ab < b^2 \wedge 0 < \sqrt{a} < \sqrt{b}$$

Proof. We obtain $a^2 < ab$ and $ab < b^2$, since both a and b are positive. Thus, $a^2 < ab < b^2$.

Now, for the sake of contradiction, suppose $\sqrt{b} < \sqrt{a}$. Then, from $0 < a < b \Rightarrow a^2 < ab < b^2$, we see that $b < \sqrt{ab} < a$. However, $a < b$. ↯

Therefore, $0 < \sqrt{a} < \sqrt{b}$. □

Definition 3.19. The absolute value of a real number x, denoted by $|x|$, is defined by

$$|x| = \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x \leq 0 \end{cases}$$

Proposition 3.20 (Triangle Inequality)**.** If x and y are real numbers, then

$$|x+y| \leq |x| + |y|.$$

Proof.

$$\begin{aligned} xy \leq |x||y| &\Rightarrow 2xy \leq 2|x||y| \\ &\Rightarrow x^2 + y^2 + 2xy \leq x^2 + y^2 + 2|x||y| \\ &\Rightarrow (x+y)^2 \leq |x|^2 + |y|^2 + 2|x||y| \\ &\Rightarrow (x+y)^2 \leq (|x| + |y|)^2 \\ &\Rightarrow \sqrt{(x+y)^2} \leq \sqrt{(|x|+|y|)^2} \\ &\Rightarrow |x+y| \leq ||x| + |y|| \\ &\Rightarrow |x+y| \leq |x| + |y| \end{aligned}$$

□

The arithmetic mean (or average) of x and y is $\frac{x+y}{2}$. The geometric mean of nonnegative numbers x and y is $\sqrt{xy}$. The term AGM Inequality stands for Arithmetic Mean–Geometric Mean Inequality given by the following proposition.

Proposition 3.21. If x and y are real numbers, then

$$2xy \leq x^2 + y^2 \quad \text{and} \quad xy \leq \left(\frac{x+y}{2}\right)^2.$$

If x and y are also nonnegative, then

$$\sqrt{xy} \leq \frac{x+y}{2}$$

Equality holds in each inequality iff $x = y$.

Proof.

$$\begin{aligned} (x-y)^2 \geq 0 &\Leftrightarrow x^2 + y^2 \geq 2xy \\ &\Leftrightarrow x^2 + y^2 + 2xy \geq 4xy \\ &\Leftrightarrow (x+y)^2 \geq 4xy \\ &\Leftrightarrow \frac{1}{4}(x+y)^2 \geq xy \\ &\Leftrightarrow \left(\frac{x+y}{2}\right)^2 \geq xy \end{aligned}$$

Equality holds iff $(x-y)^2 = 0$, i.e., $x = y$. □

Corollary 3.22. If $x > 0$ and $y > 0$, then

$$\frac{2xy}{x+y} \leq \sqrt{xy} \leq \frac{x+y}{2}.$$

Equality holds in each inequality iff $x = y$.

Proof. From Proposition 3.21, $\sqrt{xy} \leq \frac{x+y}{2}$ where equality holds iff $x = y$. Substitute $\frac{1}{x}$ and $\frac{1}{y}$ to Proposition 3.21:

$$\begin{aligned}\sqrt{\frac{1}{x}\frac{1}{y}} \leq \frac{\frac{1}{x}+\frac{1}{y}}{2} &\Leftrightarrow \frac{1}{\sqrt{xy}} \leq \frac{x+y}{2xy} \\ &\Leftrightarrow \sqrt{xy} \geq \frac{2xy}{x+y}\end{aligned}$$

Equality holds iff $\frac{1}{x} = \frac{1}{y}$, i.e., $x = y$. □

The expression $\frac{2xy}{x+y}$ is the harmonic mean of x and y. General arithmetic, geometric, and harmonic mean of the numbers $a_1, a_2, \ldots, a_n$ are given by

$$\begin{aligned}\text{AM} &= \frac{a_1 + a_2 + \cdots + a_n}{n}, \\ \text{GM} &= \sqrt[n]{a_1 a_2 \ldots a_n}, \\ \text{HM} &= \frac{n}{\frac{1}{a_1} + \frac{1}{a_2} + \cdots + \frac{1}{a_n}}.\end{aligned}$$

3.4 Rearrangement Inequality

Definition 3.23. Let $a_1 \leq a_2 \leq \cdots \leq a_n$ and $b_1 \leq b_2 \leq \cdots \leq b_n$ be any real numbers.

1. $S_n = a_1b_1 + a_2b_2 + \cdots + a_nb_n$ is called the sorted sum of the numbers.
2. $R_n = a_1b_n + a_2b_{n-1} + \cdots + a_nb_1$ is called the reversed sum of the numbers.
3. Let $c_1, c_2, \ldots, c_n$ be any permutation of the numbers $b_1, b_2, \ldots, b_n$.
   $P_n = a_1c_1 + a_2c_2 + \cdots + a_nc_n$ is called the permuted sum of the numbers.

Theorem 3.24 (Rearrangement Inequality)**.** *Let S_n, R_n, and P_n as above. Then*

$$S_n \geq P_n \geq R_n.$$

If a_i are strictly increasing, then equality holds iff b_i are all equal. And unlike most inequalities, we do not require the numbers involved to be positive.

Proof. We first show that $S_n \geq P_n$.

Use induction on n. When $n = 1$, $S_n = P_n = a_1b_1$, so $S_n \geq P_n$ holds.

Suppose as an induction hypothesis that $S_n \geq P_n$ for some $n = k \in \mathbb{N}$. Consider the case $n = k+1$. Let

$$\begin{aligned} S_n &= a_1 b_1 + \cdots + a_k b_k + a_{k+1} b_{k+1} \\ P_n &= a_1 c_1 + \cdots + a_k c_k + a_{k+1} c_{k+1} \end{aligned}$$

When $b_{k+1} = c_{k+1}$, it is trivial from the induction hypothesis that $S_n \geq P_n$. Let $b_{k+1} = c_i$, $c_{k+1} = b_j$. Since $(a_{k+1} - a_i)(b_{k+1} - b_j) \geq 0$, we see that $a_{k+1} b_{k+1} + a_i b_j \geq a_{k+1} b_j + a_i b_{k+1}$.

$$\begin{aligned} P_n &= a_1 c_1 + \cdots + a_i c_i + \cdots + a_{k+1} c_{k+1} \\ &= a_1 c_1 + \cdots + a_i b_{k+1} + \cdots + a_{k+1} b_j \\ &\leq a_1 c_1 + \cdots + a_i b_j + \cdots + a_{k+1} b_{k+1} \\ &= P_k + a_{k+1} b_{k+1} \\ &\leq S_k + a_{k+1} b_{k+1} \\ &= S_{k+1} \\ &= S_n \end{aligned}$$

For the equality to hold, $(a_{k+1} - a_i)(b_{k+1} - b_j) \geq 0$ should hold for any i, j. If a_i's are strictly increasing, $b_{k+1} - b_j = 0$ for any j. Thus, equality holds iff all b_i's are equal.

By induction principle, $S_n \geq P_n$ for all $n \in \mathbb{N}$, where its equality holds iff all b_i's are equal when a_i's are strictly increasing.

Now to show $P_n \geq R_n$, define new sequences $b_i' = -b_i$ and $c_i' = -c_i$. Sorting b_i' and c_i' leads to a reverse of b_i and c_i, respectively. Let

$$\begin{aligned} S_n' &= a_1 b_n' + \cdots + a_n b_1' \\ &= a_1(-b_n) + \cdots + a_n(-b_1) \\ &= -(a_1 b_n + \cdots + a_n b_1) \\ &= -R_n \\ P_n' &= a_1 c_1' + \cdots + a_n c_n' \\ &= a_1(-c_1) + \cdots + a_n(-c_n) \\ &= -(a_1 c_1 + \cdots + a_n c_n) \\ &= -P_n \end{aligned}$$

Since $S_n' \geq P_n'$, it follows that $-R_n \geq -P_n$. Therefore, $P_n \geq R_n$, and its equality holds iff b_i's are same when a_i's are strictly increasing. □

Corollary 3.25. Let $a_1, a_2, \ldots, a_n$ be real numbers and $c_1, c_2, \ldots, c_n$ be its permutation. Then

$$a_1^2 + a_2^2 + \cdots + a_n^2 \geq a_1 c_1 + a_2 c_2 + \cdots + a_n c_n$$

Proof. Let α_i be a sorted sequence of a_i in an ascending order. Then $\sum a_i a_i = \sum \alpha_i \alpha_i$ is a sorted sum whereas $\sum a_i c_i = \sum \alpha_i c'_i$ is a permuted sum, where $a_i c_i = \alpha_j c'_j$ for some $j \in [n]$. From Theorem 3.24, $\sum a_i^2 \geq \sum a_i c_i$. □

Corollary 3.26. Let $a_1, a_2, \ldots, a_n$ be positive real numbers and $c_1, c_2, \ldots, c_n$ be its permutation. Then

$$\frac{c_1}{a_1} + \frac{c_2}{a_2} + \cdots + \frac{c_n}{a_n} \geq n.$$

Proof. Let α_i be a sorted sequence of a_i in an ascending order. Note that $\frac{1}{\alpha_i}$ is in a descending order. Then $\sum \frac{c_i}{a_i} = \sum \frac{1}{a_i} \cdot c_i = \sum \frac{1}{\alpha_i} \cdot c'_i$ is a permuted sum whereas $n = \sum \frac{1}{\alpha_i} \cdot \alpha_i$ is a reveresed sum, where $\frac{1}{a_i} \cdot c_i = \frac{1}{\alpha_j} \cdot c'_j$ for some $j \in [n]$. From Theorem 3.24, $\sum \frac{c_i}{a_i} \geq n$. □

Theorem 3.27 (Arithmetic Mean–Geometric Mean Inequality)**.** *Let $x_1, x_2, \ldots, x_n$ be positive real numbers. Then*

$$\frac{x_1 + x_2 + \cdots + x_n}{n} \geq \sqrt[n]{x_1 x_2 \ldots x_n}.$$

Equality holds iff $x_1 = x_2 = \cdots = x_n$.

Proof. Let $G = \sqrt[n]{x_1 x_2 \ldots x_n}$. Define a sequence $\alpha_i = \frac{x_1 \ldots x_i}{G^i}$. From Corollary 3.26, the following holds:

$$\sum_{i=1}^{n} \frac{\alpha_i}{\alpha_{i-1}} \geq n,$$

where $\alpha_0 = \alpha_n = \frac{x_1 x_2 \ldots x_n}{G^n} = 1$.

$$\begin{aligned} \sum_{i=1}^{n} \frac{\alpha_i}{\alpha_{i-1}} &= \sum_{i=1}^{n} \frac{\frac{x_1 \ldots x_i}{G^i}}{\frac{x_1 \ldots x_{i-1}}{G^{i-1}}} \\ &= \sum_{i=1}^{n} \frac{x_i}{G} \end{aligned}$$

Therefore, $\sum \frac{x_i}{G} \geq n$, which implies $\frac{x_1 + x_2 + \cdots + x_n}{n} \geq G = \sqrt[n]{x_1 x_2 \ldots x_n}$ □

Proposition 3.28 (Geometric Mean–Harmonic Mean Inequality)**.** Let $x_1, x_2, \ldots, x_n$ be positive real numbers. Then

$$\sqrt[n]{x_1 x_2 \ldots x_n} \geq \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n}}.$$

Proof. Substituting $\frac{1}{x}_i$ to x_i in Theorem 3.27 leads to the desired inequality. □

Proposition 3.29 (Root Mean Square–Arithmetic Mean Inequality)**.** Let $x_1, x_2, \ldots, x_n$ be positive real numbers. Then

$$\sqrt{\frac{x_1^2 + x_2^2 + \ldots x_n^2}{n}} \geq \frac{x_1 + x_2 + \cdots + x_n}{n}.$$

Proof. Let

$$\begin{aligned} S &= \sum_{i=1}^{n} x_i^2 \\ P_k &= \sum_{i=1}^{n} x_i x_{i+k}, \end{aligned}$$

where $x_{i+k} = x_{i+k-n}$ if $i + k > n$, and $k \in [n-1]$. Now expand $(x_1 + x_2 + \cdots + x_n)^2$.

$$\begin{aligned} (x_1 + x_2 + \cdots + x_n)^2 &= x_1^2 + x_2^2 + \cdots + x_n^2 \\ &\quad + x_1x_2 + x_2x_3 + \cdots + x_nx_1 \\ &\quad + x_1x_3 + x_2x_4 + \cdots + x_nx_2 \\ &\quad + \cdots \\ &\quad + x_1x_{1+k} + x_2x_{2+k} + \cdots + x_nx_{n+k} \\ &\quad + \cdots \\ &\quad + x_1x_n + x_2x_1 + \cdots + x_nx_{n-1} \\ &= S + \sum_{i=1}^{n-1} P_i \\ &\le S + \sum_{i=1}^{n-1} S \\ &= nS \\ &= n(x_1^2 + x_2^2 + \cdots + x_n^2) \end{aligned}$$

Thus, $\left(\sum x_i\right)^2 \le n \sum x_i^2$.

$$\begin{aligned} \left(\sum x_i\right)^2 \le n\left(\sum x_i^2\right) &\Leftrightarrow \frac{\left(\sum x_i\right)^2}{n^2} \le \frac{\sum x_i^2}{n} \\ &\Leftrightarrow \frac{\sum x_i}{n} \le \sqrt{\frac{\sum x_i^2}{n}} \end{aligned}$$

Therefore, $\sqrt{\frac{x_1^2+x_2^2+\cdots+x_n^2}{n}} \ge \frac{x_1+x_2+\cdots+x_n}{n}$. □

Theorem 3.30 (Chebyshev's Inequality). *Let $x_1 \le x_2 \le x_n$ and $y_1 \le y_2 \le \cdots \le y_n$ be any real numbers. Then*

$$\begin{aligned} x_1y_1 + x_2y_2 + \cdots + x_ny_n &\ge \frac{(x_1 + x_2 + \cdots + x_n)(y_1 + y_2 + \cdots + y_n)}{n} \\ &\ge x_1y_n + x_2y_{n-1} + \cdots + x_ny_1. \end{aligned}$$

Proof. The first part of the inequality is a sorted sum, whereas the last part of the inequality is a reversed sum. The one in the middle is an average of a sorted sum, a reversed sum, and permuted sums. □

3.5 Cauchy-Schwarz Inequality

Theorem 3.31 (Cauchy-Schwarz Inequality). *Let $a_1, a_2, \dots, a_n$ and $b_1, b_2, \dots, b_n$ be two sequences of real numbers, then*

$$\left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{i=1}^{n} b_i^2\right) \geq \left(\sum_{i=1}^{n} a_i b_i\right)^2.$$

with equality iff the sequences $a_1, a_2, \dots, a_n$ and $b_1, b_2, \dots, b_n$ are proportional, i.e., there is a constant λ such that $a_i = \lambda b_i$ for each $i = 1, \dots, n$.

Proof. Let $c_{ij} = a_i b_j$. Then,

$$\begin{aligned}\left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{i=1}^{n} b_i^2\right) &= a_1^2b_1^2 + a_1^2b_2^2 + \cdots + a_1^2b_n^2 \\ &\quad + a_2^2b_1^2 + a_2^2b_2^2 + \cdots + a_2^2b_n^2 \\ &\quad + \cdots \\ &\quad + a_n^2b_1^2 + a_n^2b_2^2 + \cdots + a_n^2b_n^2 \\ &= c_{11}^2 + c_{12}^2 + \cdots + c_{1n}^2 \\ &\quad + c_{21}^2 + c_{22}^2 + \cdots + c_{2n}^2 \\ &\quad + \cdots \\ &\quad + c_{n1}^2 + c_{n2}^2 + \cdots + c_{nn}^2\end{aligned}$$

Thus the left side of the inequality is a sorted sum. Note that $c_{ij}c_{kl} = (a_ib_j)(a_kb_l) = (a_ib_l)(a_kb_j) = c_{il}c_{kj}$.

$$\begin{aligned}\left(\sum_{i=1}^{n} a_i b_i\right)^2 &= (a_1b_1)^2 + (a_1b_1)(a_2b_2) + \cdots + (a_1b_1)(a_nb_n) \\ &\quad + (a_2b_2)(a_1b_1) + (a_2b_2)^2 + \cdots + (a_2b_2)(a_nb_n) \\ &\quad + \cdots \\ &\quad + (a_nb_n)(a_1b_1) + (a_nb_n)(a_2b_2) + \cdots + (a_nb_n)^2 \\ &= c_{11}^2 + c_{11}c_{22} + \cdots + c_{11}c_{nn} \\ &\quad + c_{22}c_{11} + c_{22}^2 + \cdots + c_{22}c_{nn} \\ &\quad + \cdots \\ &\quad + c_{nn}c_{11} + c_{nn}c_{22} + \cdots + c_{nn}^2 \\ &= c_{11}c_{11} + c_{12}c_{22} + \cdots + c_{1n}c_{n1} \\ &\quad + c_{21}c_{12} + c_{22}c_{22} + \cdots + c_{2n}c_{n2} \\ &\quad + \cdots \\ &\quad + c_{n1}c_{1n} + c_{n2}c_{2n} + \cdots + c_{nn}c_{nn}\end{aligned}$$

Thus the right side of the inequality is a permuted sum. From Theorem 3.24, the left side–the sorted sum–is greater or equal to the right side–the permuted sum. □

3.6 Exercises for Chapter 3

Exercise 3.1. Find the minimum of

$$\frac{\sin^3 x}{\cos x} + \frac{\cos^3 x}{\sin x},\ 0 < x < \frac{\pi}{2}.$$

Solution. From Theorem 3.27,

$$\frac{\sin^3 x}{\cos x} + \frac{\cos^3 x}{\sin x} \geq 2\sqrt{\frac{\sin^3 x}{\cos x}\frac{\cos^3 x}{\sin x}} = 2\sin x\cos x$$

Since equality holds iff $\frac{\sin^3 x}{\cos x} = \frac{\cos^3 x}{\sin x}$, $\sin x = \cos x$ as $x \in \left(0, \frac{\pi}{2}\right)$. Thus, $x = \frac{\pi}{4}$ is the value of x at the minimum of the given expression. Therefore, the minimum is $2\sin\frac{\pi}{4}\sin\frac{\pi}{4} = 2 \cdot \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} = 1$. ■

Exercise 3.2. Prove that

1. $a^2 + b^2 + c^2 \geq ab + bc + ca$
2. $a^n + b^n + c^n \geq ab^{n-1} + bc^{n-1} + ca^{n-1}$

Proof. 1. Use Theorem 3.24 for a, b, c and a, b, c.
2. Use Theorem 3.24 for a, b, c and $a^{n-1}, b^{n-1}, c^{n-1}$. □

Exercise 3.3. Prove that

$$\frac{1}{a^2}+\frac{1}{b^2}+\frac{1}{c^2}\geq\frac{a+b+c}{abc}.$$

Proof. Note that $\frac{a+b+c}{abc}=\frac{1}{bc}+\frac{1}{ca}+\frac{1}{ab}$. Use Theorem 3.24 for $\frac{1}{a},\frac{1}{b},\frac{1}{c}$ and $\frac{1}{a},\frac{1}{b},\frac{1}{c}$. □

Exercise 3.4. Prove that

$$\frac{a^2}{b^2}+\frac{b^2}{c^2}+\frac{c^2}{a^2}\geq\frac{b}{a}+\frac{c}{b}+\frac{a}{c}.$$

Proof. Note that $\frac{b}{a}+\frac{c}{b}+\frac{a}{c}=\frac{b}{c}\frac{c}{a}+\frac{c}{a}\frac{a}{b}+\frac{a}{b}\frac{b}{c}$. Use Theorem 3.24 for $\frac{b}{a},\frac{c}{b},\frac{a}{c}$ and $\frac{b}{a},\frac{c}{b},\frac{a}{c}$. □

Exercise 3.5. Prove that

$$\frac{a^2}{b}+\frac{b^2}{c}+\frac{c^2}{a}\geq a+b+c.$$

Proof. Note that $\frac{a^2}{b}+\frac{b^2}{c}+\frac{c^2}{a}=a\cdot\frac{a}{b}+b\cdot\frac{b}{c}+c\cdot\frac{c}{a}$. Use Theorem 3.24 for a,b,c and $\frac{a}{b},\frac{b}{c},\frac{c}{a}$. □

Exercise 3.6. Prove that

$$\frac{a^n}{b+c}+\frac{b^n}{c+a}+\frac{c^n}{a+b}\geq\frac{a^{n-1}+b^{n-1}+c^{n-1}}{2}.$$

Proof 1. Without loss of generality, suppose $a\geq b\geq c$. From Theorem 3.24 and Theorem 3.30,

$$\begin{aligned}
a^n\cdot\frac{1}{b+c}+b^n\cdot\frac{1}{c+a}+c^n\cdot\frac{1}{a+b}&\geq a^n\cdot\frac{1}{c+a}+b^n\cdot\frac{1}{a+b}+c^n\cdot\frac{1}{b+c}\\
&=a^{n-1}\cdot\frac{a}{c+a}+b^{n-1}\cdot\frac{b}{a+b}+c^{n-1}\cdot\frac{c}{b+c}\\
&\geq\frac{1}{3}\left(a^{n-1}+b^{n-1}+c^{n-1}\right)\left(\frac{a}{c+a}+\frac{b}{a+b}+\frac{c}{b+c}\right)\\
a^n\cdot\frac{1}{b+c}+b^n\cdot\frac{1}{c+a}+c^n\cdot\frac{1}{a+b}&\geq a^n\cdot\frac{1}{a+b}+b^n\cdot\frac{1}{b+c}+c^n\cdot\frac{1}{c+a}\\
&=a^{n-1}\cdot\frac{a}{a+b}+b^{n-1}\cdot\frac{b}{b+c}+c^{n-1}\cdot\frac{c}{c+a}\\
&\geq\frac{1}{3}\left(a^{n-1}+b^{n-1}+c^{n-1}\right)\left(\frac{a}{a+b}+\frac{b}{b+c}+\frac{c}{c+a}\right).
\end{aligned}$$

Thus, we obtain

$$\begin{aligned}
a^n\cdot\frac{1}{b+c}+b^n\cdot\frac{1}{c+a}+c^n\cdot\frac{1}{a+b}&\geq\frac{1}{2}\cdot\frac{1}{3}\left(a^{n-1}+b^{n-1}+c^{n-1}\right)\left[\left(\frac{a}{c+a}+\frac{b}{a+b}+\frac{c}{b+c}\right)\right.\\
&\qquad\left.+\left(\frac{a}{a+b}+\frac{b}{b+c}+\frac{c}{c+a}\right)\right]\\
&=\frac{1}{2}\cdot\frac{1}{3}\left(a^{n-1}+b^{n-1}+c^{n-1}\right)\cdot 3\\
&=\frac{1}{2}\left(a^{n-1}+b^{n-1}+c^{n-1}\right).
\end{aligned}$$

□

Proof 2. Without loss of generality, suppose $a \geq b \geq c$. From Theorem 3.30,

$$\frac{a^n}{b+c} + \frac{b^n}{c+a} + \frac{c^n}{a+b} \geq \frac{1}{3}\left(a^{n-1} + b^{n-1} + c^{n-1}\right)\left(\frac{a}{b+c} + \frac{b}{c+a} + \frac{c}{a+b}\right).$$

Note that

$$\begin{aligned} \frac{a}{b+c} + \frac{b}{c+a} + \frac{c}{a+b} &\geq \frac{1}{3}(a+b+c)\left(\frac{1}{b+c} + \frac{1}{c+a} + \frac{1}{a+b}\right) && \text{by Theorem 3.30} \\ &\geq \frac{1}{3}\left(\frac{a}{b+c} + \frac{b}{c+a} + \frac{c}{a+b} + 3\right) \end{aligned}$$

Hence

$$\frac{a}{b+c} + \frac{b}{c+a} + \frac{c}{a+b} \geq \frac{3}{2}.$$

Therefore,

$$\frac{a^n}{b+c} + \frac{b^n}{c+a} + \frac{c^n}{a+b} \geq \frac{1}{2}\left(a^{n-1} + b^{n-1} + c^{n-1}\right)$$

□

Exercise 3.7. Prove that if a, b, c are nonnegative numbers, then

$$a^a b^b c^c \geq (abc)^{\frac{a+b+c}{3}}.$$

Proof. Take log on each side of the inequality:

$$a \log a + b \log b + c \log c \geq \frac{a+b+c}{3}(\log a + \log b + \log c).$$

This a direct result of Theorem 3.30. □

4 Induction

4.1 The Principle of Induction

Definition 4.1. The set $\mathbb{N}$ of natural numbers is the intersection of all sets $S \subset \mathbb{R}$ that have following properties:

1. $1 \in S$

2. $x \in S \Rightarrow x + 1 \in S$.

Theorem 4.2 (Principle of Induction). *For each natural number n, let $P(n)$ be a mathematical statement. Then for each $n \in \mathbb{N}$ the statement $P(n)$ is true, if the following properties hold:*

1. *$P(1)$ is true.*

2. *For $k \in \mathbb{N}$, if $P(k)$ is true, then $P(k+1)$ is true.*

Proving property 1 of Theorem 4.2 is the basis step, and proving property 2 is the induction step.

Proposition 4.3. For $n \in \mathbb{N}$, the formula $1 + 2 + \cdots + n = \frac{n(n+1)}{2}$ holds.

Proposition 4.4. If $n \in \mathbb{N}$ and $q \geq 2$, then $n < q^n$.

Proof. We use induction on n. For the case $n = 1, 1 < 2 \leq q$, so the statement holds.

Suppose now as an induction hypothesis that $k < q^k$ for some $k \in \mathbb{N}$ where $q \geq 2$.

$$\begin{aligned} k + 1 < 2k &\Rightarrow k + 1 < qk \\ &\Rightarrow k + 1 < q \cdot q^k \\ &\Rightarrow k + 1 < q^{k+1} \end{aligned}$$

Hence, by the induction principle (Theorem 4.2), the given statement is true for all $n \in \mathbb{N}$. □

Proposition 4.5. If $x_1, x_2, \ldots, x_n$ are numbers in the interval $[0, 1]$, then

$$\prod_{i=1}^{n}(1 - x_i) \geq 1 - \sum_{i=1}^{n} x_i.$$

Proof. We use induction on n. For the case $n = 1$, the given inequality holds since $1 - x_1 \geq 1 - x_1$.

Suppose now as an induction hypothesis that $\prod_{i=1}^{k}(1 - x_i) \geq 1 - \sum_{i=1}^{k} x_i$ for some $k \in \mathbb{N}$ where $k \geq 2$. Let $P = \prod_{i=1}^{k}(1 - x_i)$ and $S = 1 - \sum_{i=1}^{k} x_i$. Since P is a product of numbers in $[0, 1]$, $P \in [0, 1]$. Choose any x_{k+1} from $[0, 1]$.

$$\begin{aligned} P \leq 1 &\Rightarrow x_{k+1}P \leq x_{k+1} \\ &\Rightarrow -x_{k+1}P \geq -x_{k+1} \end{aligned}$$

From the induction hypothesis $P \geq S$,

$$\begin{aligned} -x_{k+1}P \geq -x_{k+1} \wedge P \geq S &\Rightarrow P - x_{k+1}P \geq S - x_{k+1} \\ &\Rightarrow P(1 - x_{k+1}) \geq S - x_{k+1} \\ &\Rightarrow \prod_{i=1}^{k+1} x_i \geq 1 - \sum_{i=1}^{k+1} x_i. \end{aligned}$$

Hence, by the induction principle, the given inequality is true for all $n \in \mathbb{N}$. □

Corollary 4.6. If $0 \leq a \leq 1$ and $n \in \mathbb{N}$, then $(1 - a)^n \geq 1 - na$.

Proof. This is the case when $x_1 = x_2 = \cdots = x_n = a$ in Theorem 4.5. □

Proposition 4.7. If $n \in \mathbb{N}$ and $n \geq 4$, then $n^2 \leq 2^n$.

Proof. We use induction on n. For the case $n = 4$, the given inequality holds since $4^2 = 16 \leq 2^4 = 16$.

Suppose now as an induction hypothesis that $k^2 \leq 2^k$ for some $k \in \mathbb{N}$ where $k \geq 5$.

$$\begin{aligned} k \geq 5 &\Rightarrow k-1 \geq 4 \\ &\Rightarrow (k-1)^2 \geq 16 \\ &\Rightarrow k^2 - 2k + 1 \geq 16 \\ &\Rightarrow k^2 - 2k - 1 \geq 14 \\ &\Rightarrow k^2 - 2k - 1 \geq 0 \\ &\Rightarrow k^2 \geq 2k + 1 \\ &\Rightarrow 2k^2 \geq k^2 + 2k + 1 \\ &\Rightarrow 2k^2 \geq (k+1)^2 \end{aligned}$$

Since $k^2 \leq 2^k$ from the induction hypothesis, we have the following:

$$\begin{aligned} k^2 \leq 2^k \wedge 2k^2 \geq (k+1)^2 &\Rightarrow 2 \cdot 2^k \geq (k+1)^2 \\ &\Rightarrow 2^{k+1} \geq (k+1)^2. \end{aligned}$$

Hence, by the induction principle, the given inequality is true for all $n \in \mathbb{N} \backslash \{1, 2, 3\}$. □

4.2 Applications

Lemma 4.8. If f is a polynomial of degree d, then a is a zero of f iff $f(x) = (x - a)h(x)$ for some polynomial h of degree $d - 1$.

Proof. We use strong induction on d. For the case $d = 1$, $f(x) = c_1 x + c_0 = \left(x - \frac{c_0}{c_1}\right) \cdot c_1$, and as $x = \frac{c_0}{c_1}$ is its zero, the given statement holds.

Suppose now as an induction hypothesis that a polynomial f with a degree i, a is a zero of f iff $f(x) = (x - a)h(x)$ for some polynomial h of degree $i - 1$ for all positive integer i less or equal to $k \geq 2$. Now consider a polynomial g with a degree $k + 1$ which has a as its zero. We can represent $g(x) = \sum_{i=0}^{k+1} b_i x^i$. Then the following holds:

$$\begin{aligned} g(x) = \sum_{i=0}^{k+1} b_i x^i \Leftrightarrow g(x) - g(a) &= \sum_{i=0}^{k+1} b_i x^i - \sum_{i=0}^{k+1} b_i a^i \\ \Leftrightarrow g(x) - 0 &= \sum_{i=0}^{k+1} b_i (x^i - a^i). \end{aligned}$$

Since $x^i - a^i$ is a polynomial of degree less or equal to k, the induction hypothesis holds. $x = a$ is a zero, it can be factorized as $x^i - a^i = (x - a)h_i(x)$ where h_i is a polynomial of degree $i-1$. Thus we have

$$\begin{aligned} g(x) = \sum_{i=0}^{k+1} b_i(x-a)h_i(x) &\Leftrightarrow g(x) = (x-a)\sum_{i=0}^{k+1} h_i(x) \\ &\Leftrightarrow g(x) = (x-a)r(x) \end{aligned}$$

where r is a polynomial of degree $(k+1)-1 = k$. Hence the given statement holds for all $d \in \mathbb{N}$ by the induction principle. □

Theorem 4.9. *Every polynomial of degree d has at most d zeros.*

Proof. For the sake of contradiction, suppose there are more than d zeros for a polynomial of degree d. Let the zeros be $a_1, a_2, \ldots, a_d, a_{d+1}, \ldots$. From Lemma 4.8, $f(x) = (x-a_1)(x-a_2)\ldots(x-a_d)(x-a_{d+1})h(x)$. However, the right hand side results in a polynomial of degree greater than or equal to $d+1$. $\lightning$

Therefore, every polynomial of degree d has at most d zeros. □

Problem 4.10 (The Handshake Problem)**.** Consider n married couples at a party. Suppose that no person shakes hands with his or her spouse, and total $2n-1$ people other than the host shake hands with different numbers of people. With how many people does the hostess shake hands?

Solution. We want to show that the spouse of the one with the maximum handshakes has the minimum handshake. $2n-1$ people should all have distinct numbers of handshakes, and the only possible numbers of handshakes are $0, 1, \ldots, 2n-2$. The one with two handshakes–the maximum handshakes–must shake hand with everyone except his/her spouse. Now, everyone except his/her spouse has at least one handshake. Therefore, the one with no handshake must be the spouse of the one with the maximum handshakes.

We now show that the number of the handshakes the hostess has is $n-1$. We use strong induction on n. For the case $n = 2$, it is trivial that the hostess has one handshake, which follows from the argument above.

Suppose now as an induction hypothesis that the hostness has $i-1$ handshakes when there are i couples for all positive integer $i \leq k$ where $k \geq 2$. Now consider the case of $k+1$ couples. The one with the maximum handshake, A, must have handshakes with everyone except his/her spouse A^*, and his/her spouse has no handshake as shown above. Remove A and A^*. There are k couples, all with one handshake removed attributed from A. Removing A^* does not affect other people's number of handshakes. This is exactly the case of k couples, where the hostess has $k-1$ handshakes. She previously had one more handshake from A. Thus, the hostess has $k = (k+1)-1$ handshakes in the case of $k+1$ couples. Therefore, the hostess has $n-1$ handshakes when there are n couples for all $n \in \mathbb{N}$ by the induction principle. ■

Problem 4.11 (The L-Tiling Problem)**.** There are a large number of L-shaped tiles as illustrated on page 31 of the textbook. Is it possible to form the large similar region on the right with non-overlapping copies of this tile?

Solution. It is possible. Classify the large region by the length of its edge in modulo 3. It is simple to come up with a band that surrounds the edge of the region with a thickness of two. Recursively construct the region. ■

4.3 Strong Induction

Sometimes a proof of $P(k)$ in the induction step needs the hypothesis that $P(i)$ is true for all $i < k$ (which is already used above). By assuming more in the induction hypothesis, we make the condition statement in the induction step weaker. Nevertheless, this weaker implication suffices to complete the proof, so we call the method strong induction.

Theorem 4.12 (Strong Induction Principle)**.** *Let $\{P(n)|n \in \mathbb{N}\}$ be a sequence of mathematical statements. If properties 1 and 2 below hold, the for every $n \in \mathbb{N}$, $P(n)$ is true.*

1. *$P(1)$ is true.*
2. *For $k \geq 2$, if $P(i)$ is true for all $i < k$, then $P(k)$ is true.*

Proof. Let the statement $Q(n)$ be "$P(i)$ hold for all $i \leq n$." From property 1, $Q(1)$ is true. Then, if the property 2 holds, it can be said that if $Q(k-1)$ holds, then $Q(k)$ holds. From the principle of induction, $Q(n)$ is true for all $n \in \mathbb{N}$. Therefore, $P(n)$ is true for all $n \in \mathbb{N}$. □

Problem 4.13. Suppose that n coins are arranged in a row. We remove head-up coins, one by one. Each time we remove a coin, we must flip the coins still present in the (at most) two positions surrounding it. For which arrangements of heads and tails can we remove all the coin? For example, $THTHT$ fails, but $THHHT$ succeeds. Using $\circ$ to denote gaps due to removed coins, we remove $THHHT$ via

$$THHHT \to H \circ THT \to \circ \circ THT \to \circ \circ H \circ H \to \circ \circ \circ \circ H \to \circ \circ \circ \circ \circ.$$

Solution. From observation, we can find out an arrangement with odd numbers of head-up coins can always be removed, whereas the one the even numbers of head-up coins cannot be removed. We show that the observation above is true. We use strong induction on n, where n is the length of an arrangement. Case $n = 1$ is trivially true.

Suppose as an induction hypothesis that all arrangements of length k or smaller is removable iff there are odd numbers of H's for some $k \in \mathbb{N}$. Consider an arrangement of length $k+1$.

Case 1: There are no H's in the middle of the arrangement.

The arrangement must be either $HT \ldots TT$, $TT \ldots TH$, or $HT \ldots TH$. Only consider $HT \ldots TT$ and $HT \ldots TH$ without loss of generality.

Case 1-1: $HT \dots TT$

Only choice is to remove the H at the start of the arrangement. Removing it results in $\circ HT \dots TT$ with a k-long arrangement not regarding $\circ$. From the induction hypothesis, it is removable.

Case 1-2: $HT \dots TH$

Removing either the first or the last H results in the identical type–$\circ HT \dots TH$ or HT-$\dots TH\circ$. From the induction hypothesis, it is unremovable.

Case 2: There are H's in the middle of the arrangement.

Case 2-1: There is an odd number of H's

Find the first H that occurs in an arrangement. There are two possible cases:

Case 2-1-1: $T \dots THT \dots$

When we remove the first H, the arrangement becomes $T \dots TH \circ H \dots$, where the first chunk contains 1 (odd) H and the latter chunk contains odd numbers of H's. From the induction hypothesis, the two chunks are removable.

Case 2-1-1: $T \dots THH \dots$

When we remove the first H, the arrangement becomes $T \dots TH \circ T \dots$. Again, both chunks contains odd numbers of H's. From the induction hypothesis, the two chunks are removable.

Case 2-2: There is an even numbers of H's

There are three possible scenarios to remove H's:

Case 2-2-1: Remove H in between H's

The two chunks each contains even and odd numbers of H's, which makes it impossible to remove.

Case 2-2-2: Remove H after T, but not between T's

It is impossible to remove.

Case 2-2-2: Remove H between T's

It is impossible to remove.

Therefore, by the induction principle, an arrangement is removable iff it contains an odd number of H's. ■

Proposition 4.14 (Well-Ordering Property)**.** Every nonempty subset of $\mathbb{N}$ has a least element.

Proof. The given statement is equivalent to "Every subset of $\mathbb{N}$ containing any element $n \in \mathbb{N}$ has a least element." We show this using strong induction on n. The case $n = 1$ holds as 1 is the least element of $\mathbb{N}$.

Suppose now as an induction hypothesis that a subset with an element i has a least element for all $i \leq k$ such that $k \geq 2$. Consider a subset with an element $k + 1$. If the subset has a smaller element than $k + 1$, it must contain at least one of $1, 2, \dots, k$. From the induction hypothesis, the subset has a least element. If not, $k+1$ is the least element. Hence, from the induction principle, the given statement is true. □

Suppose that $S \subset \mathbb{N}$. By the well-ordering property, S^c has a least element. Thus when $P(n)$ fails for some $n \in \mathbb{N}$, there is a least n where it fails. This yields yet another approach to induction, called the method of decent. We can prove $P(n)$ for all $n \in \mathbb{N}$ by proving that there is no least n where $P(n)$ fails. To do this, we suppose that $P(n)$ fails for some n and show that $P(k)$ must fail for some k less than n. The existence of k implies that $n > 1$, and thus we have proved the contrapositive of property 2 from Theorem 4.12.

Theorem 4.15. *$\sqrt{2}$ is irrational.*

Proof. The proof using an argument, "Assume $\sqrt{2} = \frac{p}{q}$ where $\frac{p}{q}$ is a reduced form," and showing that it contradicts the fact that p and q are relatively prime is well known. We approach this in a slightly different manner, using the method of decent introduced above.

For the sake of contradiction, suppose $\sqrt{2} = \frac{m}{n}$ where $m, n \in \mathbb{N}$. The fraction needs not necessarily be reduced. We see that $2 = \frac{m^2}{n^2}$, so $2n^2 = m^2$. Then,

$$\begin{aligned}\frac{m}{n} &= \frac{m(m-n)}{n(m-n)} \\ &= \frac{m^2 - mn}{mn - n^2} \\ &= \frac{2n^2 - mn}{mn - n^2} \\ &= \frac{2n - m}{m - n}\end{aligned}$$

Since $1 < \sqrt{2} = \frac{m}{n} < 2$, $n < m < 2n$. Thus $0 < m - n < n$. Now consider a set $S = \left\{ n \middle| \sqrt{2} = \frac{m}{n} \right\} \subset \mathbb{N}$. We just showed that if $n \in S$, $m - n < n$ is also an element of S. This contradicts the well-ordering principle. $\lightning$

Therefore, our assumption that $\sqrt{2}$ can be represented as $\frac{m}{n}$ is wrong, leading to the fact that $\sqrt{2} \notin \mathbb{Q}$. From Example 3.6, $\sqrt{2} \in \mathbb{R}$, so $\sqrt{2}$ is irrational. □

Proposition 4.16. Any $n \in \mathbb{N}$ can be expressed in exactly one way as the product of an odd number and a power of 2.

Proof. The given statement can be rewritten:

$$(\forall n \in \mathbb{N})(\exists! k, m \in \mathbb{N})\ n = 2^k(2m+1).$$

If n is odd, it immediately follows that $k = 0$ and $m = \frac{n-1}{2}$. Consider the case n is even. Define a set $S = \left\{ a \middle| a = \frac{n}{2^k} \wedge a, k \in \mathbb{N} \right\}$. Then $S \subset N$ from the definition. From the well-ordering principle, the exists a least element α of S. Let l be a natural number such that $\frac{n}{2^l} = \alpha$. Since α is the least element of S, l is the maximum natural number such that

n is divisible by 2^l. The remainder of $n \div 2^l$ is thus an odd number, i.e. $2m+1$. Therefore, any even number can be written in the form of $2^k(2m+1)$.

We now show the uniqueness of such expression. Suppose $n = 2^k(2m+1) = 2^{k'}(2m'+1)$. When $k = k'$, it immediately follows that $m = m'$. When $k \neq k'$, let $k > k'$ without loss of generality. Then $2^{k-k'}(2m+1) = 2m'+1$. Since $k - k' > 0$, $2^{k-k'}$ is even, whereas $2m'+1$ is odd. ↯ Therefore, such expression is unique. □

Problem 4.17 (Sums of Consecutive Positive Integers)**.** Prove that a natural number n is a sum of consecutive smaller natural numbers iff n is not a power of 2.

Proof. We first show that if n is a sum of consecutive smaller natural numbers, n is not a power of 2. Let $n = a + (a+1) + \cdots + (a+k) = \frac{(2a+k)(k+1)}{2}$. If k is even, $k+1$ is not. If k is odd, $2a+k$ is not. Therefore, we see that $(2a+k)(k+1)$ is not a power of 2.

We now show that if n is not a power of 2, it is a sum of consecutive smaller natural numbers. From Proposition 4.16, $n = 2^\alpha(2\beta+1)$, where $\beta \neq 0$. There are two cases:

Case 1: $2^\alpha \geq \beta \geq 1$

$2^\alpha - \beta \geq 0$, so

$$\begin{aligned} n &= 2^\alpha(2\beta+1) \\ &= \frac{\left[2\left(2^\alpha-\beta\right)+2\beta\right](2\beta+1)}{2} \\ &= (2^\alpha-\beta) + (2^\alpha-\beta+1) + \cdots + (2^\alpha+\beta)\,. \end{aligned}$$

Case 2: $\beta \geq 2^\alpha \geq 1$

$\beta - 2^\alpha \geq 0$, so $\beta - 2^\alpha + 1 \geq 1$. Then,

$$\begin{aligned} n &= 2^\alpha(2\beta+1) \\ &= \frac{\left[2\left(\beta-2^\alpha+1\right)+\left(2^{\alpha+1}-1\right)\right]\left[\left(2^{\alpha+1}-1\right)+1\right]}{2} \\ &= (\beta-2^\alpha+1) + (\beta-2^\alpha+2) + \cdots + (\beta+2^\alpha)\,. \end{aligned}$$

Therefore, the given statement holds. □

4.4 Exercises for Chapter 4

Exercise 4.1. Prove that $n^3 + 20 > n^2 + 15n$ for all $n \in \mathbb{N}$.

Proof. We first show that $4k^2 + 18k + 1 \geq (k+1)^2 + 15(k+1)$ for $k \geq 3$. Simplifying the inequality yields $\left(k+\frac{1}{6}\right)^2 - \left(5+\frac{1}{36}\right) \geq 0$. The inequality is true for all $k \geq 3$.

We use induction on n. Substituting $n = 1, 2$ shows that the inequality given in the problem holds. Suppose now as an induction hypothesis that $k^3 + 20 > k^2 + 15k$ for some

$k \geq 3$. Then the following holds.

$$\begin{aligned}(k+1)^3 + 20 &= k^3 + 3k^2 + 3k + 1 + 20 \\ &= (k^3 + 20) + (3k^2 + 3k + 1) \\ &> (k^2 + 15k) + (3k^2 + 3k + 1) && \text{From the induction hyp.} \\ &= 4k^2 + 18k + 1 \\ &\geq (k+1)^2 + 15(k+1) && \text{As shown above}\end{aligned}$$

Hence, by the induction principle, the given inequality holds for any $n \in \mathbb{N}$. □

Exercise 4.2. For $n \in \mathbb{N}$, when does $3^n > n^4$ hold?

Solution. It can be checked that $n = 1$ satisfies the given inequality, but $n = 2, 3, 4, 5, 6, 7$ do not by substituting the values. We will show that $n \geq 8$ satisfies the given inequality using induction on n. $3^8 = 65651 > 8^4 = 4096$, so $n = 8$ satisfies the given inequality.

Suppose now as an induction hypothesis that $3^k > k^4$ for some $k \geq 9$. We need to show that the following two inequalities hold where $k \geq 9$::

$$\begin{aligned}k^4 &> 4k^3 - 6k^2 + 4k - 1 \\ k^4 &> 12k^2 + 2\end{aligned}$$

The first inequality immediately follows from the fact that $(k-1)^4 > 0$. The second inequality can be modified: $(k^2 - 6)^2 > 38$. It is obviously true. Add the two inequalities:

$$2k^4 > 4k^3 + 6k^2 + 4k + 1.$$

Then,

$$\begin{aligned}3^{k+1} &= 3 \cdot 3^k \\ &> 3k^4 && \text{From the induction hyp.} \\ &= k^4 + 2k^4 \\ &> k^4 + 4k^3 + 6k^2 + 4k + 1 && \text{As shown above} \\ &= (k+1)^4\end{aligned}$$

Hence, by the induction principle, the given inequality holds for any $n \geq 9$.

Therefore, the given inequality holds for $n = 1$ and $n \geq 9$. ■

Exercise 4.3. If $n \in \mathbb{N}$ and $x, y \geq 0$, then $\left(\frac{x+y}{2}\right)^n \leq \frac{x^n + y^n}{2}$.

Proof. We use induction on n. When $n = 1$, $\left(\frac{x+y}{2}\right)^1 \leq \frac{x^1 + y^1}{2}$, so the given inequality holds.

Suppose now as induction hypothesis that $\left(\frac{x+y}{2}\right)^k \leq \frac{x^k+y^k}{2}$ for some $k \geq 2$. Then,

$$\begin{aligned}
\frac{x+y}{2}\left(\frac{x+y}{2}\right)^k &\leq \frac{x+y}{2}\cdot\frac{x^k+y^k}{2} && \text{From the induction hyp.}\\
&= \frac{1}{4}\left[(x^{k+1}+y^{k+1})+(x^ky+xy^k)\right]\\
&\leq \frac{1}{4}\left[(x^{k+1}+y^{k+1})+(x^{k+1}+y^{k+1})\right]. && \text{by Theorem 3.24}\\
&= \frac{x^{k+1}+y^{k+1}}{2}.
\end{aligned}$$

Hence, from the induction principle, the given inequality holds for all $n \in \mathbb{N}$. □

Exercise 4.4. Two players move alternately in a game that starts with two equal-sized piles of coins. One move consists of removing some positive number of coins from one pile. The winner is the player who removes the last coin. Who will be the winner?

Solution. The second player will win. The following is the winning strategy.
Case 1: The first player removes all coins from one pile

Remove all coins from the other pile.
Case 2: The first player removes some coins from one pile

Remove all coins except for one from the pile the first player removed some coins. Now there are two piles–one with one coin and one without a coin removed. Let the first pile be A, and the latter be B. The first player will now try to remove some coins from B, since taking a coin from A would lead to the second player taking all coins from B, making the second player the winner. Now, the second player should remove all coins from B except one, similarly as the last move. It is evident that the second player will win after this move. ■

Exercise 4.5. Let $\{a_n\}$ be a sequence satisfying $a_1 = 2, a_2 = 8$ and $a_n = 4(a_{n-1} - a_{n-2})$ for $n \geq 3$. Find a formula for a_n.

Solution. Define a new sequence $b_n = a_n - 2a_{n-1}$ for $n \geq 2$. From the recurrence relation of a_n, it can be implied that $b_n = 2b_{n-1}$ for $n \geq 3$. Thus, $b_n = 2^{n-2}b_2 = 2^n$. Now we have

$$a_n - 2a_{n-1} = 2^n.$$

Then,

$$\begin{aligned}
a_n - 2a_{n-1} &= 2^n,\\
2(a_{n-1} - 2a_{n-2}) &= 2^n,\\
&\vdots\\
2^{n-2}(a_2 - 2a_1) &= 2^n.
\end{aligned}$$

Add all the equations:

$$a_n - 2^{n-2}a_1 = 2^n(n-1).$$

Therefore, $a_n = 2^n n$. ■

Exercise 4.6. Show that the Fibonacci sequence $\{f_n\}$ have the following formula:

$$f_n = \frac{\alpha^n - \beta^n}{\sqrt{5}}$$

where $\alpha = \frac{1+\sqrt{5}}{2}$ and $\beta = \frac{1-\sqrt{5}}{2}$.

Proof.

$$\begin{aligned}
f_1 &= \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2} - \frac{1-\sqrt{5}}{2}\right) = 1 \\
f_2 &= \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^2 - \left(\frac{1-\sqrt{5}}{2}\right)^2\right) \\
&= \frac{1}{\sqrt{5}}\left(\frac{6+2\sqrt{5}}{4} - \frac{6-2\sqrt{5}}{4}\right) = 1
\end{aligned}$$

Thus, we see that the initial conditions match that of the Fibonacci sequence.

We will show that $f_{n+1} = f_n + f_{n-1}$ for $n \geq 2$.

$$\begin{aligned}
f_{n+1} &= \frac{\alpha^{n+1} - \beta^{n+1}}{\sqrt{5}} \\
&= \frac{\frac{1+\sqrt{5}}{2}\alpha^n - \frac{1-\sqrt{5}}{2}\beta^n}{\sqrt{5}} \\
&= \frac{\frac{1}{2}(\alpha^n - \beta^n) + \frac{\sqrt{5}}{2}(\alpha^n + \beta^n)}{\sqrt{5}} \\
&= \frac{(\alpha^n - \beta^n) - \frac{1}{2}(\alpha^n - \beta^n) + \frac{\sqrt{5}}{2}(\alpha^n + \beta^n)}{\sqrt{5}} \\
&= f_n + \frac{\frac{\sqrt{5}-1}{2}\alpha^n + \frac{\sqrt{5}+1}{2}\beta^n}{\sqrt{5}} \\
&= f_n + \frac{\frac{\sqrt{5}-1}{2}\cdot\frac{1+\sqrt{5}}{2}\alpha^{n-1} + \frac{\sqrt{5}+1}{2}\cdot\frac{1-\sqrt{5}}{2}\beta^{n-1}}{\sqrt{5}} \\
&= f_n + \frac{\alpha^{n-1} - \beta^{n-1}}{\sqrt{5}} \\
&= f_n + f_{n-1}
\end{aligned}$$

Therefore, the Fibonacci sequence follows the given formula. □

5 Bijections and Cardinality

5.1 Representation of Natural Numbers

Definition 5.1. Let q be a natural number greater than 1. A q-ary or base q representation of n is a list $a_m, \ldots, a_0$ of integers, each in $\{0, 1, \ldots, q-1\}$, such that $a_m > 0$ and $n = \sum_{i=0}^{m} a_i q^i$. For clarity, we may use a subscript q to indicate that the base is q. We call representation in base 2, 3, and 10 as binary, ternary, and decimal, respectively.

Theorem 5.3 allows us to use the term **the** base q representation, instead of **a** base q representation.

Example 5.2. The ternary representation for the first ten natural numbers in order are 1, 2, 10, 11, 12, 20, 21, 22, 100, 101. The corresponding representation in base 4 are 1, 2, 3, 10, 11, 12, 13, 20, 21, 22.

Theorem 5.3. *Let q be a natural number greater than 1. Every natural number has a unique base q representation with no leading zeros.*

Proof. We use induction on n. For the case $n = 1$, $1 = 1 \cdot q^0$ is a base q representation of 1.

Suppose now as an induction hypothesis that n has a base q representation $n = \sum_{i=0}^{m} a_i q^i$. We consider two cases:

Case 1: $a_0 = \cdots = a_m = q - 1$

$$\begin{aligned} n &= \sum_{i=0}^{m} a_i q^i \\ &= \sum_{i=0}^{m} (q-1) q^i \\ &= (q-1) \sum_{i=0}^{m} q^i \\ &= (q-1) \frac{q^{m+1} - 1}{q - 1} \\ &= q^{m+1} - 1 \end{aligned}$$

Thus, $n + 1 = 1 \cdot q^{m+1}$.

Case 2: Otherwise

Let t be the smallest index such that $a_t < q - 1$, i.e.,

$$a_0 = \cdots = a_{i-1} = q - 1, a_t < q - 1, \ldots$$

Define b_j as following:

$$b_j = \begin{cases} a_j & \text{if } j > t \\ a_t + 1 & \text{if } j = t \\ 0 & \text{if } j < t \end{cases}$$

Then,

$$\begin{aligned} \sum_{i=0}^{m} b_i q^i &= 0 + \cdots + 0 + (a_t + 1)q^t + a_{t+1}q^{t+1} + \cdots + a_m q^m \\ &= q^t + a_t q^t + a_{t+1}q^{t+1} + \cdots + a_m q^m \\ &= q^t - 1 + 1 + a_t q^t + a_{t+1}q^{t+1} + \cdots + a_m q^m \\ &= (q-1)\sum_{i=0}^{t-1} q^i + 1 + a_t q^t + a_{t+1}q^{t+1} + \cdots + a_m q^m \\ &= \sum_{i=0}^{t-1} (q-1)q^i + 1 + a_t q^t + a_{t+1}q^{t+1} + \cdots + a_m q^m \\ &= \sum_{i=0}^{t-1} a_i q^i + 1 + a_t q^t + a_{t+1}q^{t+1} + \cdots + a_m q^m \\ &= \sum_{i=0}^{m} a_i q^i + 1 \\ &= n + 1 \end{aligned}$$

Hence, from the induction principle, there is a q-ary representation for all $n \in \mathbb{N}$.

Now we show the uniqueness of the base q represenation. For the sake of contradiction, suppose there are two distinct q-ary representation of n. Let the two be $\sum_{i=0}^{s} a_i q^i$ and $\sum_{i=0}^{t} b_i q^i$. If $s \neq t$, let $s > t$ without loss of generality. Since $s, t \in \mathbb{N}$, $s \geq t + 1$ can be implied. Then,

$$\begin{aligned} n &= \sum_{i=0}^{s} a_i q^i > q^s \\ n &= \sum_{i=0}^{t} b_i q^i < q^{t+1} \leq q^s \end{aligned}$$

We deduced $n > q^s \wedge n \leq q^s$. $\lightning$

Therefore, $s = t$. Let $S = \{n | n \text{ has more than one distinct base } q \text{ representations.}\}$. Since n has two distinct base q representation, $n - q^s$ also has two distinct representations. Thus, for any element e in S, where $q^m \leq e < q^{m+1}$, $e - q^m$ is also an element of S. As $S \subset \mathbb{N}$, this contradicts the well-ordering principle. □

From the following, we define $[n] = \{m | m \in \mathbb{N} \wedge m \leq n\} = \{1, 2, \ldots, n\}$.

Problem 5.4 (The Weights Problem). A balance scale has left and right pans; we can place objects in each pan and test whether the total weight is the same on each side. Suppose that five objects of known integer weight can be selected. How caan we choose the weights to guarantee being able to check all integer weights form 1 through 121? Given an object believed to have weight $n \in [121]$, how should we place the known weights to chekc it? Is it possible to choose five values to check more weights?

Solution. Intuitively, we can think of putting weights on the same side as the object as a negative weight. Thus, each weight can attribute 1, 0, -1, which corresponds to putting the weight on the other side of the pan, not putting the weight at all, and putting the weight on the same side, respectively. Since there are three possible attributes, we can come up with a ternary representation of the weight. We therefore choose $3^0 = 1, 3^1 = 3, 3^2 = 9, 3^3 = 27, 3^4 = 81$ as weights.

Now show that our choice can actually weigh any object with weight $w \in [121]$, i.e., $(\forall w \in [121])(\exists (a_0, a_1, a_2, a_3, a_4) \in \{-1, 0, 1\}^5)\ w = \sum_{i=0}^{4} a_i 3^i$. From Theorem 5.3, there is a unique ternary representation of $u \in \{121, 122, \dots, 242\}$,

$$u = \sum_{i=0}^{4} b_i 3^i$$

where $b_i \in \{0, 1, 2\}$. Substracting $3^0 + 3^1 + \cdots + 3^4 = 121$ to each side results in

$$u - 121 = \sum_{i=0}^{4} (b_i - 1) 3^i.$$

Note that $b_i - 1 \in \{-1, 0, 1\}$ and $u - 121 \in [121]$. Hence we can choose $a_i = b_i - 1$. Therefore, there always is a representation of w as $\sum_{i=0}^{4} a_i 3^i$. ■

5.2 Bijections

Definition 5.5. A function $f : A \to B$ is a bijection if for every $b \in B$ there is exactly one $x \in A$ such that $f(x) = b$. Such f is called a one-to-one correspondence between A and B.

Example 5.6. There is a bijection between $\mathbb{N}$ and $\mathbb{Z}$.

Proof. We give an explicit one-to-one correspondance from $\mathbb{N}$ to $\mathbb{Z}$. Correspond all even numbers to non-negative integers and odd numbers to negative integers:

$$f(n) = \begin{cases} \frac{n}{2} - 1 & \text{if } n \text{ is even.} \\ -\frac{n+1}{2} & \text{if } n \text{ is odd.} \end{cases}$$

We show that f is a bijection. Note that $\frac{n}{2} - 1 \geq 0$ for all $n \in \mathbb{N}$, and $-\frac{n+1}{2} < 0$ for all $n \in \mathbb{N}$. Choose any integer $m \in \mathbb{Z}$. If $m \leq 0$, the only possible n that corresponds to it must have a relation $-\frac{n+1}{2} = m$, and such n exists uniquely as $-2m - 1$. If $m > 0$, the only possible n that correponds to it must have a relation $\frac{n}{2} - 1 = m$. We can determine n uniquely as $2m + 2$. Therefore, f is a bijection. □

Example 5.7. Given constants $a, b, c, d \in \mathbb{R}$, let $f : \mathbb{R}^2 \to \mathbb{R}^2$ is defined by $f(x, y) = (ax + by, cx + dy)$. When is f a bijection?

Solution. For each $(\alpha, \beta) \in \mathbb{R}^2$, there needs to be $(x, y) \in \mathbb{R}^2$ such that $\alpha = ax + by$ and $\beta = cx + dy$. Using a matrix, we can rewrite the system as

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

For x and y to exist uniquely, the inverse of

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

should exist. Therefore, the determinant being nonzero–$ad - bc \neq 0$–is both the necessary and sufficient condition for f to be bijective. ■

Definition 5.8. If f is a bijection from A to B, then the inverse of f is a function $g : B \to A$ such that for each $b \in B$, $g(b)$ is a unique element $x \in A$ such that $f(x) = b$. We write f^{-1} for the function g.

Example 5.9. If f is a bijection and g is the inverse of f, then g is also a bijection and f is the inverse of g. Thus $(f^{-1})^{-1} = f$.

Proof. Let f be a function from X to Y. Choose any $x_1 \in X$. Then there is a unique $y_1 \in Y$ such that $f(x_1) = y_1$, from the definition of a function (which was never formally stated). For such y_1, since g is an inverse of f, i.e.,

$$(\forall y \in Y)(\exists! x \in X)\ g(y) = x,$$

there is a unique $g(y_1) = x_2 \in X$ such that $f(x_2) = y_1 \in Y$. Since f is a bijection, i.e.,

$$(\forall y \in Y)(\exists! x \in X)\ f(x) = y,$$

and as $f(x_1) = f(x_2) = y_1$, it follows that $x_1 = x_2$.

Therefore, g is also a bijection as there is a unique $y \in Y$ for any $x \in X$. □

Showing that $f : A \to B$ is a bijection means showing that for each $b \in B$, the equation $f(x) = b$ has a unique solution in A. Solving the equation to write a formula hat determines x in terms of b obtains a formula for f^{-1}. We must check that the formula is valid on all of B.

Example 5.10. Define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = 5x - 2|x|$. Show that f is a bijection.

Proof. Note that when $x \geq 0$, $f(x) = 3x \geq 0$, and when $x < 0$, $f(x) = 7x < 0$. Thus, $f(x)$ follows the sign of x and vice versa. Choose any $y \geq 0$. If there is any, possible x such that

$f(x) = y$ is also greater or equal to 0. Hence we solve $3x = y$. x is uniquely determined as $\frac{y}{3}$.

Now choose any $y < 0$. Only possible x such that $f(x) = y$ is negative. Thus solving $7x = y$, $x = \frac{y}{7}$.

x is uniquely determined for any $y \in \mathbb{R}$. Therefore, f is a bijection. □

A bijection transforms elements of one set into elements of another, allowing us to work in either context. For example, we can encode a subset S of $[n]$ by recording the presence or absence of an element i as 1 or 0 in position i of an n-tuple $m(S)$. An n-tuple with entries in $\{0,1\}$ is a binary n-tuple; we call $m(s)$ the binary encoding of S. From a binary n-tuple b, we will uniquely retrieve S such that $m(S) = b$. Thus binary encoding is a bijection from the power set of $[n]$ to the set of binary n-tuples.

Example 5.11. Consider the set [3]. For $S = \varnothing$, $m(S) = (0,0,0)$, and for $S = \{1,3\}$, $m(S) = (1,0,1)$. Similarly, $S = \{1,2,3\}$ encodes to $m(S) = (1,1,1)$.

Proposition 5.12. Binary encoding establishes a bijection from the power set of $[n]$ to the set of binary n-tuples.

Proof. We show that

$$m : 2^{[n]} \to \mathbb{Z}_2^n$$

is a bijection.

Choose any $(b_1, \dots, b_n) \in \mathbb{Z}_2^n$. We can construct $S \subseteq [n]$ such that $i \in S$ iff $b_i = 1$. Thus, for any $(b_1, \dots, b_n) \in \mathbb{Z}_2^n$, we can find $S \in 2^{[n]}$.

Now show that there is a unique subset of $[n]$ for a binary n-tuple. Suppose there are two distinct subsets S and T of $[n]$ for a binary encoding b for the sake of contradiction. From a simple argument, we can show that

$$\neg(p \Leftrightarrow q) \Leftrightarrow (p \wedge \neg q) \vee (q \wedge \neg p)$$

Then, since $S \neq T$,

$$(\exists i \in [n])\ (i \in S \wedge i \notin T) \vee (i \notin S \wedge i \in T).$$

Without loss of generality, we only need to check the case $(\exists i \in [n])\ i \in S \wedge i \notin T$. Thus, $[m(S)]_i = 1$ whereas $[m(T)]_i = 0$. This contradicts the prior assumption that $b = m(S) = m(T)$. $\lightning$

Therefore, m is a bijection. □

5.3 Injections and Surjections

Definition 5.13. A function $f : A \to B$ is injection if for each $b \in B$, there is at most one $x \in A$ such that $f(x) = b$. A function $f : A \to B$ is surjective if for each $b \in B$, there is at least one $x \in A$ such that $f(x) = b$. The corresponding nouns are injection and surjection.

Proposition 5.14. Let f be a real-valued function defined on a subset of $\mathbb{R}$. If f is strictly monotone, then f is injective.

Proof.

$$x_1 > x_2 \Rightarrow f(x_1) > f(x_2),$$

since f is strictly monotone. Then,

$$\begin{aligned} x_1 \neq x_2 &\Rightarrow x_1 > x_2 \wedge x_1 < x_2 \\ &\Rightarrow f(x_1) > f(x_2) \wedge f(x_1) < f(x_2) \\ &\Rightarrow f(x_1) \neq f(x_2). \end{aligned}$$

Therefore, f is injective. □

Example 5.15. What are the solutions to the equation below?

$$x^4 + x^3y + x^2y^2 + xy^3 + y^4 = 0$$

Solution. Multiply each side of the given equation with $x - y$ yields:

$$x^5 - y^5 = 0$$

Since $f(x) = x^5$ is strictly monotone, from Proposition 5.14, f is an injection. Thus, if $x^5 = y^5$, $x = y$. Hence, the only possible solution for $(x-y)(x^4 + x^3y + x^2y^2 + xy^3 + y^4) = 0$ is $x = y$. Now substitute $x = y$ to the given equation:

$$5x^4 = 0.$$

Therefore, $x = y = 0$. ■

Problem 5.16. Show that if $n \in \mathbb{N}$ is an odd number, then $f(x) = x^n$ is strictly increasing.

Proof. Choose any odd number $n \in \mathbb{N}$. Then $(\exists m \in \mathbb{N} \cup \{0\})\ n = 2m + 1$. There are five cases:

Case 1: $0 < x_1 < x_2$

If $0 < x_1 < x_2$, $0 < x_1^n < x_2^n$ follows since the two are in a positive set $(0, \infty) \subset \mathbb{R}$.

Case 2: $x_1 < 0 < x_2$

x_2^n is still positive since it is an element of a positive set. Since $-x_1 > 0$, $(-x_1)^{2m} = x_1^{2m} > 0$. Thus, $x_1^n = x_1^{2m} \cdot x_1 < 0$. Therefore, $x_1^n < x_2^n$.

Case 3: $x_1 < x_2 < 0$

$0 < -x_2 < -x_1$, so from **Case 1**, $0 < (-x_2)^n < (-x_1)^n$. Since $(-x_1)^n = (-x_1)^{2m+1} = (-x_1)^{2m} \cdot (-x_1) = -x_1^n$, and $(-x_2)^n$ is similarly $-x_2^n$, $0 < -x_2^n < -x_1^n$. Therefore, $x_1^n < x_2^n < 0$.

Other two cases are when either x_1 or x_2 is 0, when $x_1 < x_2$. The two cases can be shown similarly.

Therefore, $f(x) = x^n$ is strictly increasing. □

5.4 Composition of Functions

Definition 5.17. If $f : A \to B$ and $g : B \to C$, then the composition of g with f is a function $h : A \to C$ defined by $h(x) = g\left(f(x)\right)$ for $x \in A$. When h is a composition of g with f, we write $h = g \circ f$.

Proposition 5.18. 1. The composition of two injections is an injection.

2. The composition of two surjections is a surjection.

3. If f and g are bijections, then $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

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

1.

$$\begin{aligned}(g \circ f)(x_1) = (g \circ f)(x_2) &\Rightarrow g(f(x_1)) = g(f(x_2)) \\ &\Rightarrow f(x_1) = f(x_2) \\ &\Rightarrow x_1 = x_2\end{aligned}$$

Thus, $g \circ f$ is also an injection.

2. For each $c \in C$, there is $b \in B$ such that $g(b) = c$. Also, for each $b \in B$, there is $a \in A$ such that $f(a) = b$. Thus, For each $c \in C$, there is $a \in A$ such that $(g \circ f)(a) = g(f(a)) = g(b) = c$. Therefore, $g \circ f$ is also a surjection.

3.

$$\begin{aligned}\left(f^{-1} \circ g^{-1}\right) \circ (g \circ f)(x) &= \left(f^{-1} \circ g^{-1}\right)((g \circ f)(x)) \\ &= \left(f^{-1} \circ g^{-1}\right)(g(f(x))) \\ &= f^{-1}\left(g^{-1}(g(f(x)))\right) \\ &= f^{-1}(f(x)) \\ &= x\end{aligned}$$

Therefore, $f^{-1} \circ g^{-1} = (g \circ f)^{-1}$. □

Proposition 5.19 (Associativity of Composition)**.** If $f : A \to B$, $g : B \to C$, and $h : C \to D$, then $h \circ (g \circ f) = (h \circ g) \circ f$.

Proof.

$$\begin{aligned}(h \circ (g \circ f))(x) &= h((g \circ f)(x)) \\ &= h(g(f(x))) \\ &= (h \circ g)(f(x)) \\ &= ((h \circ g) \circ f(x))\end{aligned}$$

Therefore, $h \circ (g \circ f) = (h \circ g) \circ f$. □

Example 5.20. Let S be the set of polynomials in one variable. Given the polynomial f defined by $f(x) = \sum_{i=0}^{k} a_i x^i$, let Df denote the polynomial whose value x is $\sum_{i=1}^{k} a_i i x^{i-1}$. The operator D is a function $D : S \to S$. It is surjective; the polynomial with coefficients $\{a_k\}$ is the image of the polynomial $\sum_{i=0}^{k} \frac{a_i x^{i+1}}{i+1}$. The operator D is not injective; the polynomials f and g defined by $f(x) = x + 1$ and $g(x) = x + 2$ have the same image.

We define another operator $J : S \to S$. For $f(x) = \sum_{i=0}^{k} a_i x^i$, let Jf denote the polynomial whose value at x is $\sum_{i=0}^{k} \frac{a_i x^{i+1}}{i+1}$. If $Jf = Jg$, then term-by-term comparison of coefficients show that $f = g$; hence J is injective. On the other hand, J is not surjective, because there is no polynomial f such that Jf is a nonzero polynomial of degree 0.

We can compose operators. We have $D(J(f)) = f$ for all $f \in S$, but $J(D(f))$ does not equal f when $f(0) \neq 0$.

5.5 Cardinality

Definition 5.21. A set A is finite if there is a bijection from A to $[k]$ for some $k \in \mathbb{N} \cup \{0\}$. A set is infinite if there is no such bijection.

Proposition 5.22. If there is a bijection $f : [m] \to [n]$, then $m = n$.

Proof. We use induction on n. Consider the case $n = 1$. Since f is a bijection, there should exist $x \in [m]$ such that $f(x) = 1$. For any $e \in [m]$, it should correspond with some element in $[1]$, which can only be 1. Thus $f(e) = 1$. Since f is injective, $f(e) = f(1) = 1$ implies $e = 1$ for any $e \in [m]$. Therefore, $m = 1$.

Suppose now as an induction hypothesis that if there is a bijection $f : [m] \to [k]$, then $m = k$ for some $k \geq 2$. Let $g : [m] \to [k+1]$ be a bijection.

Case 1: $g(m) = k + 1$

Then $g|_{[m-1]} : [m-1] \to [k]$ is a bijection. Then from the induction hypothesis, $m - 1 = k$. Hence $m = k + 1$.

Case 2: $g(m) \in [k]$

Let $g(m) = a \in [k]$. Now define $h : [k+1] \to [k+1]$ as the following:

$$h(x) = \begin{cases} k+1 & \text{if } x = a \\ a & \text{if } x = k+1 \\ x & \text{otherwise.} \end{cases}$$

Then we see that $h \circ g$ is still a bijection, as h and g are both bijective. Note that $(h \circ g)(m) = h(g(m)) = h(a) = k+1$. Since this was the case in **Case 1**, $m = k+1$.

Therefore, from the induction principle, existence of a bijection from $[m]$ to $[n]$ implies $m = n$. □

Corollary 5.23. If A is finite, then for exactly one n, there is a bijection from A to $[n]$.

Proof. Since A is finite, there is a bijection $f : A \to [k]$ for some $k \in \mathbb{N} \cup \{0\}$. From Proposition 5.22, $n = k$. □

Definition 5.24. The size of a finite set A, written $|A|$, is the unique n such that there is a bijection from A to $[n]$. A set of size n is called n-element set or n-set.

Corollary 5.25. If A and B are disjoint finite sets, then

$$|A \cup B| = |A| + |B|.$$

Proof. From Corollary 5.23, there are bijections $f : A \to [n], g : B \to [m]$ for some $n, m \in \mathbb{N}$. Now construct $h : A \cup B \to [m+n]$ as following:

$$h(x) = \begin{cases} f(x) & \text{if } x \in A \\ g(x) + n & \text{if } x \in B \end{cases}.$$

We need not consider the case $x \in A \wedge x \in B$ as A and B are disjoint. Note that $1 \leq f(x) \leq n$ and $n+1 \leq g(x) + n \leq m+n$. We shall show that h is a bijection.

Choose any $k \in [m+n]$. If $k \leq m$, the only possible $x \in A \cup B$ to make $h(x) \leq n$ is the case when $x \in A$. Since f is a bijection, there exists a unique $x \in A$ such that $f(x) = h(x) = k$.

Similarly, if $m < k \leq m+n$, the only possible case is when $x \in B$. There exists $l \in [n]$ such that $m + l = n$. Also, there exists a unique $x \in B$ such that $g(x) = l$ since g is a bijection. Thus, such x satisfies the condition $h(x) = l + m = k$.

Therefore, there is a unique $x \in A \cup B$ such that $h(x) = k$ implying h is a bijection. □

Corollary 5.26. Every nonempty finite set of real numbers has both a maximum element and a minimum element.

Proof. We use induction on the size of a set S, $|S| = n$. For the case $n = 1$, the only element of S should both be the maximum and the minimum.

Now as an induction hypothesis, suppose that a set of k elements has both a maximum element and a minimum element, for some $k \geq 2$. Cosider a set S where $|S| = k+1$. Choose any element $e \in S$. Then from the induction hypothesis, there is both a maximum element M and a minimum element m of $S \backslash \{e\}$. We see that $\max(\{e, M\})$ and $\min(\{e, m\})$ are the maximum and the minimum element of S, respectively. Thus, there is both a maximum and a minimum element of any nonempty finite set of real numbers. □

Definition 5.27. An infinite set A is countably infinite (or countable) if there is a bijection from A to $\mathbb{N}$; otherwise A is uncountably infinite (or uncountable). Sets A and B have the same cardinality if there is a bijection from A to B.

Example 5.28. $\mathbb{Z}$ and $\mathbb{Q}$ are countable, whereas $\mathbb{R}$ is not.

Proof. From Example 5.6, $\mathbb{Z}$ is countable. Now consider the following sequence:

$$\frac{0}{1}, \frac{1}{1}, -\frac{1}{1}, \frac{1}{2}, \frac{2}{1}, -\frac{1}{2}, -\frac{2}{1}, \frac{1}{3}, \frac{2}{3}, \frac{3}{1}, \frac{3}{2}, -\frac{1}{3}, -\frac{2}{3}, -\frac{3}{1}, -\frac{3}{2}, \frac{1}{4}, \cdots$$

It is evident that all rational numbers will appear at the above sequence. Since we can see a sequence as a bijective function from $\mathbb{N}$ to the elements of the sequence, the above sequence shows the bijection from $\mathbb{N}$ to $\mathbb{Q}$. Thus $\mathbb{Q}$ is countable.

We now show the uncountability of $\mathbb{R}$ using the famous diagonal argument from Cantor. Every real number in $(0, 1]$ can be represented in the form $0.a_1a_2a_3\ldots$ uniquely. Note that real number like 0.1 should be written in form $0.999\cdots$ to ensure the uniqueness of such representation. For the sake of contradiction, suppose there is a bijection f from $\mathbb{N}$ to $\mathbb{R}$:

$$\begin{aligned} f(1) &= 0.a_{11}a_{12}a_{13}\ldots \\ f(2) &= 0.a_{21}a_{22}a_{23}\ldots \\ &\vdots \\ f(k) &= 0.a_{k1}a_{k2}a_{k3}\ldots \\ &\vdots \end{aligned}$$

Now consider the number $b \in (0, 1]$ defined as the following:

$$b = 0.b_1b_2b_3\ldots, \text{ where } b_i = \begin{cases} 1 & \text{if } a_{ii} \neq 1 \\ 0 & \text{if } a_{ii} = 1 \end{cases}$$

Such b is never equal with $f(k)$ for all $k \in \mathbb{N}$, as $b_i \neq a_{ii}$ for all $i \in \mathbb{N}$. However, f is a bijection. $\lightning$

Therefore, no bijection from $\mathbb{N}$ to $\mathbb{R}$ exists, so $\mathbb{R}$ is uncountably infinite. □

Theorem 5.29. *The sets $\mathbb{N} \times \mathbb{N}$ and $\mathbb{N}$ has the same cardinality. That is, $\mathbb{N} \times \mathbb{N}$ is countable.*

Proof. We can construct the following sequence:

$$(1,1),(2,1),(1,2),(3,1),(2,2),(1,3),(4,1),(3,2),\ldots$$

It is evident that all elements in $\mathbb{N} \times \mathbb{N}$ are included in the above sequence. Since it is a sequence, there is a bijection from $\mathbb{N}$ to $\mathbb{N} \times \mathbb{N}$. □

Problem 5.30. Show that $(0,1)$ and $\mathbb{R}$ have the same cardinality.

Proof. We can construct a bijection from $(0,1)$ to each point of $x^2 + (y-1)^2 = 1$ where $y < 1$. Simply consider the angle created by the point (x,y) on the half-circumference, $(0,1)$, and $(1,1)$, which in $(0,\pi)$. Thus, consider the bijection $f(x) = \pi x$, which is a bijection between the angle and $(0,1)$. Then the angle makes a bijection to the half-circumference. Now consider the intersection of the half-line connecting (x,y) from $(0,1)$ and the x-axis. We see that it creates a bijection from each point of the half-circumference and the point on the x-axis. Thus, there is a bijection from the half-circumference to $\mathbb{R}$. Therefore, there is a bijection from $(0,1)$ to $\mathbb{R}$, implying that the cardinalities of the two sets are equal. □

Lemma 5.31. For two sets A and B, if $B \subseteq A$ and $f : A \to B$ is injective, there exists a bijection between A and B.

Proof. Let $Y = A - B$ and $X = Y \cup f(Y) \cup f^2(Y) \cup \ldots$. Then $Y \cap B = \varnothing$ and $f^k(Y) \subseteq B$. Thus,

$$(\forall k \geq 1)\ Y \cap f^k(Y) = \varnothing.$$

Since f is an injection, f^m is also an injection for any $m \geq 1$. Then we have

$$(\forall k, m \geq 1)\ f^m(Y) \cap f^m\left(f^k(Y)\right) = f^m(Y) \cap f^{m+k}(Y) = \varnothing.$$

Thus, all $f^k(Y)$'s are disjoint for all $k \geq 0$.

From $X = Y \cup f(Y) \cup f^2(Y) \cup \ldots$, we have $f(X) = f(Y) \cup f^2(Y) \cup f^3(Y) \cup \cdots \subseteq B$. Since f is injective, $f|_X : X \to f(X)$ is a bijection. Also, we have

$$A - X = (B \cup Y) - (Y \cup f(X)) = B - f(X).$$

Define $h : A \to B$ by:

$$h(z) = \begin{cases} f(z) & \text{if } z \in X \\ z & \text{if } z \in A - X \end{cases}$$

By definition, h is injective. Let $b \in B$ be given. If $b \in f(X)$, then $f(z) = b$ for some $z \in X$. Then $h(z) = f(z) = b$. If $b \in B - f(X)$, then $b \in A - X$, implying $h(b) = b$. Hence h is a surjection. Therefore, h is a bijection from A to B □

Theorem 5.32 (Schröder-Bernstein Theorem). *If $f : A \to B$ and $g : B \to A$ are injections, then there exists a bijection $h : A \to B$, and hence A and B have the same cardinality.*

Proof. $g : B \to A$, so $g(B) \subseteq A$. Now consider $g \circ f : A \to g(B)$. Since both f and g are injections, $g \circ f$ is also an injection. From Lemma 5.31, there is a bijection between A and $g(B)$. Let such bijection be h'. Now define $h = g^{-1} \circ h' : A \to B$. Note that $g^{-1} : g(B) \to B$ is a bijection as g is an injection. Thus, h, a composition of two bijections is also a bijection. □

5.6 Exercises for Chapter 5

Exercise 5.1. Let $f : A \to B$ be a function. Prove that there is a function $g : B \to A$ such that $f \circ g = I_B$ iff f is a surjection.

Proof. We need to show that for functions $f : A \to B$,

$$(\exists g : B \to A)(\forall x \in B)\ (f \circ g)(x) = x \Leftrightarrow (\forall y \in B)(\exists x \in A)\ f(x) = y.$$

We first show the $\Rightarrow$ direction of the statement. Choose any $y \in B$. Since $(\forall x \in B)\ (f \circ g)(x) = x$, $(f \circ g)(y) = f(g(y)) = y$. Thus we have $x = g(y) \in A$ such that $f(x) = y$. Therefore,

$$(\exists g : B \to A)(\forall x \in B)\ (f \circ g)(x) = x \Rightarrow (\forall y \in B)(\exists x \in A)\ f(x) = y.$$

Now we show the $\Leftarrow$ direction of the statement. Since there is always a corresponding $x \in A$ to any $y \in B$ since f is surjective, we can define a nonempty set $S(y) \subset A$ such that $S(y) = \{e \in A | f(e) = y\}$. We construct a function $g : B \to A$:

$$g(x) = a,$$

where a is an arbitrary element from $S(x) \neq \varnothing$. Then,

$$\begin{aligned}(f \circ g)(x) &= f(g(x)) \\ &= f(a) && \text{where } a \in S(x) = \{e \in A | f(e) = x\} \\ &= x.\end{aligned}$$

Thus, $f \circ g = I_B$. Therefore,

$$(\exists g : B \to A)(\forall x \in B)\ (f \circ g)(x) = x \Leftarrow (\forall y \in B)(\exists x \in A)\ f(x) = y.$$

We showed both direction of the given statement. □

Exercise 5.2. Suppose that $f : X \to Y$ is a function. Define $\overrightarrow{f} : \mathcal{P}(X) \to \mathcal{P}(Y)$ by

$$\overrightarrow{f}(A) = \{f(x) | x \in A\},$$

for $A \in \mathcal{P}(X)$. Similarly, define $\overleftarrow{f} : \mathcal{P}(Y) \to \mathcal{P}(X)$ by

$$\overleftarrow{f}(B) = \{x \in X | f(x) \in B\},$$

for $B \in \mathcal{P}(Y)$.

Prove that the followings are equivalent:

1. f is injective.
2. $\overrightarrow{f}$ is injective.
3. $\overleftarrow{f}$ is surjective.

Proof. We first show that if f is injective, then $\overrightarrow{f}$ is injective. Choose any two distinct X_1 and X_2 from $\mathcal{P}(X)$. If one of X_1 and X_2 is an empty set, let it be X_1, without loss of generality. Then $\overrightarrow{f}(X_1) = \overrightarrow{f}(\varnothing) = \varnothing$, and $\overrightarrow{f}(X_2) \neq \varnothing$, so $f(X_1) \neq f(X_2)$. Now consider the case when both X_1 and X_2 are not empty. Note that $X_1 \neq X_2$ implies the following:

$$X_1 - X_2 \neq \varnothing \vee X_2 - X_1 \neq \varnothing.$$

Without loss of generality, assume $X_1 - X_2 \neq \varnothing$. Then, there is $x_1 \in X_1 - X_2$. There is no other element in X other than x_1 such that $f(x) = f(x_1)$, since f is injective. Thus $f(x_1) \in \overrightarrow{f}(X_1) - \overrightarrow{f}(X_2)$. Therefore, $\overrightarrow{f}(X_1) \neq \overrightarrow{f}(X_2)$, implying $\overrightarrow{f}$ is injective.

Now show that if $\overrightarrow{f}$ is injective, then f is injective as well. Choose any two distinct x_1 and x_2 from X. Consider $\{x_1\}$ and $\{x_2\}$ from $\mathcal{P}(X)$. Since $\overrightarrow{f}$ is injective,

$$\begin{aligned}\overrightarrow{f}(\{x_1\}) \neq \overrightarrow{f}(\{x_2\}) &\Rightarrow \{f(x_1)\} \neq \{f(x_2)\} \\ &\Rightarrow f(x_1) \neq f(x_2)\end{aligned}$$

Therefore, f is injective.

We shall show that if f is injective, then $\overleftarrow{f}$ is surjective. Choose any Y_1 from $\mathcal{P}(Y)$. If $Y_1 = \varnothing$, then $\overleftarrow{f}(Y_1) = \overleftarrow{f}(\varnothing) = \varnothing \in \mathcal{P}(X)$. Now consider the case when Y_1 is nonempty. For any $y \in Y_1$, there are two cases: there is a unique $x \in X$ such that $f(x) = y$, or there is no $x \in X$ at all such that $f(x) = y$. For the either case, $\{x | f(x) \in Y\} \in \mathcal{P}(X)$. Thus, $\overleftarrow{f}$ is surjective.

Finally, show that $\overleftarrow{f}$ being a surjection implies that f is injective. For the sake of contradiction, suppose f is not injective. Then for some $y_1 \in Y$, there are two distinct x_1 and x_2 in X such that $f(x_1) = f(x_2) = y_1$. Since $\overleftarrow{f}$ is surjective, there is $Y_1, Y_2, Y_3 \in \mathcal{P}(Y)$ such that $\overleftarrow{f}(Y_1) = \{x_1\}$ and $\overleftarrow{f}(Y_2) = \{x_2\}$. Only $y \in Y$ such that $y = f(x_1) = f(x_2)$ is $y = y_1$ from the definition of a function. Thus, $y_1 \in Y_1$ and $y_1 \in Y_2$. However, if $y_1 \in Y_1$, then $x_2 \in \overleftarrow{f}(Y_1)$ from definition of $\overleftarrow{f}$. $\lightning$ Therefore, f is injective if $\overleftarrow{f}$ is a surjection.

We now have that all three expressions are equivalent. □

Exercise 5.3. Let B be a proper subset of a set A, and let f be a bijection from A to B. Prove that A is an infinite set.

Proof. For the sake of contradiction, suppose that A is finite, i.e., there is a bijection $h : A \to [n]$ for some $n \in \mathbb{N}$. Note that the existance of a proper subset of A ensures that A is nonempty.

B is a proper subset of A, so $\exists a \in A - B \neq \varnothing$. Then we see that there is no bijection from B to $[n]$. However, we do have a bijection f from A to B and a bijection h from A to $[n]$. $\lightning$

Hence the assumption that A is finite is wrong. Therefore, A is an infinite set. □

Exercise 5.4. Let f be a function from A to itself. Prove that f is injective iff f is surjective. Prove that this equivalence fails when A is infinite.

Proof. Consider the case when A is finite.

We first show that if f is injective, then f is surjective. Define a function $g : A \to f(A)$, where $g(x) = f(x)$. Then g is a bijection, since f is injective and $f(A)$ is now the codomain of g. Thus, $|A| = |f(A)|$. We also have $f(A) \subseteq A$, so $A = f(A)$. Therefore $f = g$, implying f is also a bijection, thus surjective as well.

Now show that if f is surjective, then f is injective. For the sake of contradiction, suppose

$$(\exists y_1 \in A)(\exists x_1, x_2 \in A)\ f(x_1) = f(x_2) = y_1.$$

Then for a set $S = \{x | f(x) = y_1\}$, $|S| \geq 2$. Define a function $h : A \backslash S \to A \backslash \{y_1\}$, where $h(x) = f(x)$. Choose any $y^* \in A \backslash \{y_1\}$. Since f is surjective, there is $x^* \in A$ such that $f(x^*) = y^*$. $A \backslash S = \{x | f(x) \neq y_1\}$, so $x^* \in A \backslash S$. Thus,

$$(\forall y \in A \backslash \{y_1\})(\exists x \in A \backslash S)\ f(x) = h(x) = y$$

Hence h is surjective. However, $|A \backslash \{y_1\}| = |A| - 1$ whereas $|A \backslash S| \leq |A| - 2$. With a cardinality of the domain smaller than the codomain, constructing a surjection is impossible. $\lightning$

Hence our previous assumption that f is not an injection is wrong. Therefore, if f is surjective, then f is injective.

We have shown that when A is a finite set, f is injective iff f is surjective. However, for the case when A is an infinite set, the statement does not hold. For instance, consider the function $f : \mathbb{N} \to \mathbb{N}$, defined by $f(n) = 2n$. f is clearly an injection, yet it is not a surjection. □

Exercise 5.5. Let $A_1, A_2, \ldots$ be a sequence of sets, each of which is countable. Prove that the union of all the sets in the sequence is a countable set.

Proof. Before we prove the statement, we show that a union of two countable sets is countable as well.

Assume two sets A and B are countable. Then there exist bijections $f_A : A \to \mathbb{N}$ and $f_B : B \to \mathbb{N}$. Now define a function $f : A \cup B \to \mathbb{N}$:

$$f(x) = \begin{cases} 2f_A x & \text{if } x \in A \\ 2f_B(x) + 1 & \text{if } x \in B \backslash A \end{cases}$$

f is injective, as $x \in A$ corresponds to even numbers and f_A is injective, whereas $x \in B\backslash A$ corresponds to odd numbers and f_B is injective.

On the other hand, $f_A : A \to \mathbb{N}$ is a bijection. Hence, $f_A^{-1} : \mathbb{N} \to A$ is an injection. Let $g : \mathbb{N} \to A \cup B$ be defined as $g(x) = f_A^{-1}(x)$. Then g is an injection.

We have shown that there are injections both from $A \cup B$ to $\mathbb{N}$ and $\mathbb{N}$ to $A \cup B$. From Theorem 5.32, $A \cup B$ and $\mathbb{N}$ have the same cardinality, i.e., $A \cup B$ is countable.

Now we prove the given statement, and we use induction on the length of the sequence n. For the case $n = 1$, when A_1 is countable, union of all sets in the sequence, which is A_1 itself, is countable. Suppose as an induction hypothesis that for some $k \geq 2$, if A_i is countable for all $i \in [k]$, then $\bigcup_{i=1}^{n} A_i$ is countable. Consider $\bigcup_{i=1}^{n+1} A_i$ when all A_i's are countable for $i \in [k+1]$. From the induction hypothesis, $\bigcup_{i=1}^{n} A_i$ is countable. We have shown above that a union of two countable sets is countable. $\bigcup_{i=1}^{n+1} A_i = \bigcup_{i=1}^{n} \cup A_{n+1}$ is thus countable. Therefore, from the induction principle, a union of all the sets in a sequence of coutable sets is countable. □

Exercise 5.6. Construct an explicit bijection from the open interval $(0, 1)$ to the closed interval $[0, 1]$.

Proof. Construct a sequence $\langle a_i \rangle$ of rational numbers in $(0, 1)$:

$$a_1 = \frac{1}{2}, a_2 = \frac{1}{3}, a_3 = \frac{2}{3}, a_4 = \frac{1}{4}, a_5 = \frac{3}{4}, a_6 = \frac{1}{5}, a_7 = \frac{2}{5}, a_8 = \frac{3}{5}, a_9 = \frac{4}{5}, a_{10} = \frac{1}{6}, \ldots$$

Any rational number in $(0, 1)$ is included in the above sequence only once. Now define a function $f : [0, 1] \to (0, 1)$:

$$f(x) = \begin{cases} a_1 & \text{if } x = 0 \\ a_2 & \text{if } x = 1 \\ a_{i+2} & \text{if } x = a_i \\ x & \text{if } x \notin \mathbb{Q} \end{cases}$$

For any $y \in (0, 1)$, y is either in $(0, 1) \cap \mathbb{Q}$ or $(0, 1) - \mathbb{Q}$. Choose any $y^* \in (0, 1) \cap \mathbb{Q}$. Then $(\exists! k \in \mathbb{N})\ a_k = y^*$. When $k = 1$, the only corresponding x such that $f(x) = y^*$ is

$x = 0$. When $k = 2$, the only corresponding x such that $f(x) = y^*$ is $x = 1$. When $k \geq 3$, corresponding x such that $f(x) = y^*$ is uniquely determined as $x = a_{k-2}$. Now when $y^* \in (0,1) - \mathbb{Q}$, only x such that $f(x) = y^*$ is $x = y^*$. Therefore, proposed f is bijective. □

6 Combinatorial Reasoning

6.1 Arrangements and Selection

Definition 6.1. A partition of a set A is a collection of pairwise disjoint subsets of A whose union is A. The rule of sum states that if A is finite and $B_1, \ldots, B_m$ is a partition of A, then $|A| = \sum_{i=1}^{m} |B_i|$.

Problem 6.2. Prove the rule of sum.

Proof. The rule of sum follows directly from Corollary 5.25. □

Definition 6.3. Let T be a set whose elements can be described using a procedure involving steps $S_1, \ldots, S_k$ such that step S_i can be performed in r_i ways, regardless of how steps $S_1, \ldots, S_{i-1}$ are performed. The rule of product states that $|T| = \prod_{i=1}^{k} r_i$.

Problem 6.4. Prove the rule of product.

Proof. We use induction n, the number of procedures needed to describe elements in T. For the case $n = 1$, $|T| = r_1$, so the given statement holds. Assume now as an induction hypothesis that $|T| = \prod_{i=1}^{k} r_i$ for some $k \geq 2$. Now when $k+1$ steps are needed to perform elements of T, k steps should be performed before the $(k+1)$-th one. Performing $k+1$ steps before the $(k+1)$-th one takes $R = \prod_{i=1}^{k} r_i$ ways from the induction hypothesis. Since each way of performing the $(k+1)$-th step is distinct, the rule of sum applies:

$$|T| = \sum_{i=1}^{R} r_{k+1} = Rr_{k+1} = \prod_{i=1}^{k+1} r_i.$$

Therefore, the rule of product holds for any number of steps from the induction principle. □

Definition 6.5. A permutation of a finite set S is a bijection from S to itself. The word form of a permutation of $[n]$ is the list obtained by writing the image of i in position i.

Example 6.6. The word form of a permutation simply records the function; for example, $f : [3] \to [3]$ defined by $f(1) = 2$, $f(2) = 3$, and $f(3) = 1$ is the permutation with word form 231. We often use the term permutation for botht the function and the word form.

In counting problems, we use the word arrangements to refer to lists formed from a specified set. We generalize permutations by considering arrangements without repetition.

Theorem 6.7. *An n-element set has $n!$ permutations (arrangements without repetition). In general, the number of arrangements of k distinct elements from a set of size n is $n(n-1)\ldots(n-k+1)$.*

Proof. This is a direct result of the rule of product. □

Definition 6.8. A selection of k elements from $[n]$ is a k-element subset of $[n]$. The number of such selection is n choose k, written as $\binom{n}{k}$. If $k < 0$ or $k > n$, $\binom{n}{k}$ is 0; in these cases, there are no selection of k elements from $[n]$.

Theorem 6.9. *For integer n, k with $0 \le k \le n$,*

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Proof. From Theorem 6.7, there are $n(n-1)\dots(n-k+1)$ permutations of length k from n elements. This is equal to selecting k elements from n elements and making a permutation out of the selection, i.e., $\binom{n}{k} \cdot k!$. Thus, we have the following relation:

$$n(n-1)\dots(n-k+1) = \frac{n!}{(n-k)!} = \binom{n}{k} \cdot k!.$$

Therefore, $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. □

Problem 6.10 (Comparison of Poker Hands)**.** A poker hand consists of five cards from an ordinary deck of cards. Find the number of cases for the following:

1. One pair
2. Two pair
3. Three of a kind
4. Straight
5. Flush
6. Full house
7. Four of a kind
8. Straight flush

Solution. **1. One pair**

Choose a rank of the pair: 13. Choose two suits of the pair: $\binom{4}{2}$. Choose ranks of the rest of the cards: $\binom{12}{3}$. Choose suits of the rest of the cards: 4^3. The overall number of cases is $\binom{13}{1}\binom{4}{2}\binom{12}{3}4^3 = 1,098,240$.

2. Two pair

Choose ranks of the two pairs: $\binom{13}{2}$. Choose two suits of the pairs: $\binom{4}{2}^2$. Choose a rank of the left card: 11. Choose a suit of the left card: 4. The overall number of cases is $\binom{13}{2}\binom{4}{2}^2 \cdot 11 \cdot 4 = 123,552$

3. Three of a kind

Choose a rank of the three: 13. Choose three suits of the three: $\binom{4}{3}$. Choose ranks of the other two cards: $\binom{12}{2}$. Choose suits of the other two cards: 4^2. The overall number of cases is $13 \cdot \binom{4}{3}\binom{12}{2} \cdot 4^2 = 54,912$.

4. Straight

Choose the consecutive ranks (ace to 5, 10 to ace): 10. Choose a suit for each card, but all are not the same: $4^5 - 4$. The overall number of cases is: $10(4^5 - 4) = 10,200$.

5. Flush

Choose the suit for the flush: 4. Choose a rank for each card, but not consecutive: $\binom{13}{5} - 10$. The overall number of cases is $4\left(\binom{13}{5} - 10\right) = 5,108$.

6. Full house

Choose ranks for the three and the pair: $13 \cdot 12$. Choose a suit for the three: $\binom{4}{3}$. Choose a suit for the pair: $\binom{4}{2}$. The overall number of cases is $13 \cdot 12 \cdot \binom{4}{3}\binom{4}{2} = 3,744$.

7. Four of a kind

Choose a rank for the four: 13. Choose the left card: 48. The overall number of cases is: $13 \cdot 48 = 624$.

8. Straight flush

Choose the consecutive ranks (ace to 5, 10 to ace): 10. Choose a suit for the flush: 4. The overall number of cases is $10 \cdot 4 = 40$. ■

The number $\binom{n}{k}$ is called the **binomial coefficient** due to its appearance as a coefficient in the n-th power of a sum of two terms.

Theorem 6.11 (Binomial Theorem)**.**

$$(x+y)^n = \sum_{i=0}^{n} \binom{n}{i} x^i y^{n-i}$$

Proof. We can choose either x or y from n terms of $x+y$. The number of cases to choose x, i times and y, $(n-1)$ times is $\binom{n}{i}\binom{n-i}{n-i} = \binom{n}{i}$. Therefore,

$$(x+y)^n = \sum_{i=0}^{n} \binom{n}{i} x^i y^{n-i}.$$

□

6.2 Binomial Coefficients

Lemma 6.12.

$$\binom{n}{k} = \binom{n}{n-k}$$

Proof. Consider a set S where $|S| = n$. Choosing a subset T with k elements is equal to choosing a subset $S - T$ with $n-k$ elements. Therefore, $\binom{n}{k} = \binom{n}{n-k}$. □

Definition 6.13. A lattice path in the plane is a path joining integer points via steps of unit length rightward or upward. Alternatively, it is a list of ordered pairs of integers, with each step increasing one coordinate by 1. The length of a path is the total number of steps.

Example 6.14 (Lattice Paths and Binary Lists)**.** We start lattice paths at the origin. Since each step increases a coordinate by 1, the length of the walk is the sum of the coordinates of the ending point. We can encode a path by recording in position i as 1, when the i-th step is rightward and 0, when the i-th step is upward. In a path of length n, if there are k steps to the right, we reach the point $(k, n-k)$, and the encoding has k 1's.

Furthermore, the actual path is determined by which steps are taken to the right. Thus the path is determined by the binary n-tuple. This establishes a one-to-one correspondence between the lattice paths reaching $(k, n-k)$ and the binary n-tuples with k 1's. Hence the number of a lattice path $(k, n-k)$ is $\binom{n}{k}$.

Proposition 6.15. For nonnegative integers a and b, the number of lattice paths from the origin to the point (a, b) is $\binom{a+b}{a}$.

Proof. Lattice paths from the origin to the point (a, b) is composed of a 1's and b 0's. Since there is a one-to-one correspondence between each path and binary $(a+b)$-tuples with a 1's, we can simply count the number of the binary tuples. We should choose a places to locate 1's out of $a+b$ places, which is $\binom{a+b}{a}$. □

The triangular array of number in which row n consists of all the binomial coefficients with n on top (starting with row 0) is called Pascal's triagle.

Lemma 6.16 (Pascal's Formula)**.** If $n \geq 1$, then

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}.$$

Proof. Consider a nonempty set S with n elements. Suppose we choose k elements from S to construct a subset T. Choose any element $a \in S$. Then either $a \in T$ or $a \notin T$. The two cases are disjoint. For the case $a \in T$, we need to choose other $k-1$ elements from $S \backslash \{a\}$, which has $\binom{n-1}{k-1}$ cases. For the case $a \notin T$, we need to choose k elements from $S \backslash \{a\}$, which has $\binom{n-1}{k}$ cases. From the rule of sum, the total number of cases is $\binom{n-1}{k-1} + \binom{n-1}{k}$. A simpler way of counting T's is just to choose k elements from n, which is $\binom{n}{k}$. Therefore, $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$. □

Theorem 6.17. *With repetition allowed, there are $\binom{n+k-1}{k-1}$ ways to select n objects from k types. This is also equal to the number of nonnegative integer solutions to $x_1 + \cdots + x_k = n$.*

Proof. We first show that choosing n objects from k types is equivalent to the number of nonnegative integer solutions to $x_1 + \cdots + x_k = n$. Let x_i be the number of objects chosen from the i-th type. Since total n objects are chosen, $\sum_{i=0}^{k} x_i = n$. Thus, we aim to show that the number of nonnegative integer solutions is $\binom{n+k-1}{k-1}$.

The problem of finding the number of nonnegative integer solutions of $\sum_{i=0}^{k} x_i = n$ corresponds to arranging n □'s and $k-1$ /'s:

$$\underbrace{\square\square\ldots\square}_{n}\underbrace{//\ldots/}_{k-1},$$

where the number of □'s in the i-th spacing created by /'s corresponds to x_i. The number of cases is to place $k-1$ /'s in $n+k-1$ positions: $\binom{n+k-1}{k-1}$. □

Problem 6.18 (Nonnegative Integer Solutions)**.** Suppose that each resident of New York City has 100 coins in a jar. The coins come in five types. We consider two jars of coins to be equivalent if they have same number of coins of each type. Is it possible that no two people have equivalent jars of coins?

Solution. Let the number of each type of coin is x_i. Then, the number of possible combinations of the coins is equal to the number of nonnegative integer solutions of $\sum_{i=1}^{5} x_i = 100$. The number of the solutions is $\binom{100+5-1}{5-1} = \binom{104}{4} = 4,598,126$. The population of NYC is 8.41 million as of 2013. From the pigeon hole principle, there are at least two people with the equivalent jars–though the principle is yet to be introduced. ■

Corollary 6.19. The expansion of $(\sum_{i=1}^{m} x_i)^d$ has $\binom{d+m-1}{m-1}$ terms.

Proof. Possible form of terms from the expansion of $(\sum_{i=1}^{m} x_i)^d$ is $x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_m^{\alpha_m}$, where $\sum_{i=1}^{m} \alpha_i = d$, and each α_i is a nonnegative integer. Thus, it has $\binom{d+m-1}{m-1}$ terms. □

Lemma 6.20 (The Chairperson Identity)**.**

$$k\binom{n}{k} = n\binom{n-1}{k-1}$$

Proof. Let S be a set with n elements. We shall choose a subset T of S and again choose an arbitrary element from T, where $|T| = k$. Then, choosing T has $\binom{n}{k}$ number of cases, and choose one element from it has k number of cases. Thus, the overall number of cases is $k\binom{n}{k}$.

We can construct T and the element from it in a slightly different manner: first choose any element from S and then choose $k-1$ elements from S to construct T. Choosing any element e from S has n number of cases, and choosing $k-1$ elements from $S\backslash\{e\}$ has $\binom{n-1}{k-1}$ number of cases. Thus, the number of cases of the whole procedure is $n\binom{n-1}{k-1}$. Therefore, we have the desired relation. □

Theorem 6.21 (The Summation Identity)**.**

$$\sum_{i=0}^{n} \binom{i}{k} = \binom{n+1}{k+1}$$

Proof. We use induction on n. Consider the base case $n = 0$. Then the left hand side becomes

$$\sum_{i=0}^{n} \binom{i}{k} = \binom{0}{k} = \begin{cases} 1 & \text{if } k = 0 \\ 0 & \text{otherwise} \end{cases},$$

whereas the right hand side becomes

$$\binom{1}{k+1} = \begin{cases} 1 & \text{if } k = 0 \\ 0 & \text{otherwise} \end{cases}.$$

Thus the given relationship holds for the base case.

Let us suppose the given relationship holds for some $n = l \geq 1$. Then,

$$\begin{aligned} \sum_{i=0}^{l+1} \binom{i}{k} &= \sum_{i=0}^{l} \binom{i}{k} + \binom{l+1}{k} \\ &= \binom{l+1}{k+1} + \binom{l+1}{k} && \text{by the induction hyp.} \\ &= \binom{l+2}{k+1} && \text{by Lemma 6.16} \end{aligned}$$

Therefore, from the induction principle, the given relationship holds for all $n \in \mathbb{N} \cup \{0\}$. □

Example 6.22 (Summation of Integer Power)**.** Formulas for sums of powers are easy to verify by induction but difficult to gues. Theorem 6.21 provides a method that automatically generates both the answer and the proof. Notice that $i = \binom{i}{1}$. Therefore, Theorem 6.21 proves the summation formula for the first n natural numbers by

$$\sum_{i=0}^{n} i = \sum_{i=0}^{n} \binom{i}{1} = \binom{n+1}{2} = \frac{n(n+1)}{2}.$$

For the squares, we rewrite i^2 using binomial coefficients. Since $i^2 = 2\binom{i}{2} + \binom{i}{1}$, we have

$$\begin{aligned} \sum_{i=0}^{n} i^2 &= 2\sum_{i=0}^{n} \binom{i}{2} + \sum_{i=0}^{n} \binom{i}{1} \\ &= 2\binom{n+1}{3} + \binom{n+1}{2} \\ &= \frac{n(n+1)(2n+1)}{6}. \end{aligned}$$

This approach yields $\sum_{k=0}^{n} f(k)$ for any polynomial f.

Theorem 6.23. *For $k \in \mathbb{N}$, the value of $\sum_{i=1}^{n} i^k$ is a polynomial in n with the leading term $\frac{1}{k+1}n^{k+1}$ and the next term $\frac{1}{2}n^k$.*

Proof. i^k can be written as:

$$i^k = a_k\binom{i}{k} + a_{k-1}\binom{i}{k-1} + \cdots + a_1\binom{i}{1}.$$

Thus we have

$$\begin{aligned}\sum_{i=0}^{n} i^k &= a_k\sum_{i=1}^{n}\binom{n+1}{k+1} + a_{k-1}\sum_{i=1}^{n}\binom{n+1}{k} + \cdots + a_1\sum_{i=1}^{n}\binom{n+1}{1}\\ &= a_k\binom{n+1}{k+1} + a_{k-1}\binom{n+1}{k} + \cdots + a_1\binom{n+1}{2}\\ &= a_kO(n^{k+1}) + a_{k-1}O(n^k) + O(n^{k-1}).\end{aligned}$$

To find the coefficient of the n^{k+1} and n^k, we only need to consider a_k and a_{k-1}. Before getting the exact form of a_k and a_{k+1}, we find the coefficients of n^{k+1} and n^k about a_k and a_{k+1}.

$$\begin{aligned}a_k\binom{n+1}{k+1} &= a_k\frac{(n+1)!}{(k+1)!(n-k)!}\\ &= a_k\frac{(n+1)n\ldots(n-(k-1))}{(k+1)!}\\ &= \frac{a_k}{(k+1)!}\left(n^{k+1} + (1-1-2-\cdots-(k-1))n^k + O(n^{k-1})\right)\\ &= \frac{a_k}{(k+1)!}\left(n^{k+1} + \left(1-\frac{k(k-1)}{2}\right)n^k + O(n^{k-1})\right)\\ &= \frac{a_k}{(k+1)!}n^{k+1} + \frac{a_k}{(k+1)!}\frac{-k^2+k+2}{2}n^k + O(n^{k-1})\\ a_{k-1}\binom{n+1}{k} &= a_{k-1}\frac{(n+1)!}{k!(n+1-k)!}\\ &= a_{k-1}\frac{(n+1)n\ldots(n-(k-2))}{k!}\\ &= \frac{a_{k-1}}{k!}n^k + O(n^{k-1})\end{aligned}$$

Thus, we have

$$\sum_{i=0}^{n} i^k = \frac{a_k}{(k+1)!}n^{k+1} + \left(\frac{a_k}{(k+1)!}\frac{-k^2+k+2}{2} + \frac{a_{k-1}}{k!}\right)n^k + O(n^{k-1}).$$

To get a_k and a_{k-1}, we expand $a_k\binom{i}{k} + a_{k-1}\binom{i}{k-1} + O(n^{k-1})$:

$$\begin{aligned}
a_k\binom{i}{k} + a_{k-1}\binom{i}{k-1} &= a_k\frac{i(i-1)\dots(i-(k-1))}{k!} + a_{k-1}\frac{i(i-1)\dots(i-(k-2))}{(k-1)!} \\
&= a_k\frac{i^k - \frac{k(k-1)}{2}i^{k-1} + O(i^{k-2})}{k!} + a_{k-1}\frac{i^{k-1} + O(i^{k-2})}{(k-1)!} \\
&= \frac{a_k}{k!}i^k + \left(\frac{a_{k-1}}{(k-1)!} - \frac{k(k-1)}{2}\frac{a_k}{k!}\right)i^{k-1} + O(i^{k-2})
\end{aligned}$$

Therefore, we have $a_k = k!$ and $a_{k-1} = (k-1)! \cdot \frac{k(k-1)}{2}$. Substitute these back to $a_k\binom{n+1}{k+1}$ and $a_{k-1}\binom{n+1}{k}$:

$$\begin{aligned}
\sum_{i=0}^{n} i^k &= \frac{a_k}{(k+1)!}n^{k+1} + \left(\frac{a_k}{(k+1)!}\frac{-k^2+k+2}{2} + \frac{a_{k-1}}{k!}\right)n^k + O(n^{k-1}) \\
&= \frac{k!}{(k+1)!}n^{k+1} + \left(\frac{k!}{(k+1)!}\frac{-k^2+k+2}{2} + \frac{(k-1)! \cdot \frac{k(k-1)}{2}}{k!}\right)n^k + O(n^{k-1}) \\
&= \frac{1}{k+1}n^{k+1} + \left(\frac{-k^2+k+2}{2(k+1)} + \frac{k(k-1)}{2k}\right)n^k + O(n^{k-1}) \\
&= \frac{1}{k+1}n^{k+1} + \left(\frac{-k^2+k+2}{2(k+1)} + \frac{k^2-1}{2(k+1)}\right)n^k + O(n^{k-1}) \\
&= \frac{1}{k+1}n^{k+1} + \frac{1}{2}n^k + O(n^{k-1}).
\end{aligned}$$

We have the desired result.

□

7 Introduction to Groups

7.1 Definitions and Examples

A binary operation should be defined before giving a definition of a group.

Definition 7.1. A binary operation on a set S is a mapping of $S \times S$ into S.

Definition 7.2. A group is a set G together with a binary operation defined on G as

$$(a, b) \mapsto a * b,$$

satisfying the following conditions:

1. $(\forall a, b, c \in G)\ (a * b) * c = a * (b * c)$ **associativity**

2. a.