Chapter 1. Sets and Mappings

§1. Sets

A set is considered to be a collection of objects (elements). If \( A \) is a set and \( x \) is an element of the set \( A \), we say \( x \) is a member of \( A \) or \( x \) belongs to \( A \), and we write \( x \in A \). If \( x \) does not belong to \( A \), we write \( x \notin A \). A set is thus determined by its elements.

Let \( A \) and \( B \) be sets. We say that \( A \) and \( B \) are equal, if they consist of the same elements; that is,

\[ x \in A \iff x \in B. \]

The set with no elements is called the empty set and is denoted by \( \emptyset \). For any object \( x \), there is a set \( A \) whose only member is \( x \). This set is denoted by \( \{x\} \) and called a singleton. For any two objects \( x, y \), there is a set \( B \) whose only members are \( x \) and \( y \). This set is denoted by \( \{x, y\} \) and called a doubleton. Note that \( \{y, x\} = \{x, y\} \).

Let \( A \) and \( B \) be sets. The set \( A \) is called a subset of \( B \) if every element of \( A \) is also an element of \( B \). If \( A \) is a subset of \( B \), we write \( A \subseteq B \). Further, if \( A \) is a subset of \( B \), we also say that \( B \) includes \( A \), and we write \( B \supseteq A \).

It follows immediately from the definition that \( A \) and \( B \) are equal if and only if \( A \subseteq B \) and \( B \subseteq A \). Thus, every set is a subset of itself. Moreover, the empty set is a subset of every set.

If \( A \subseteq B \) and \( A \neq B \), then \( A \) is a proper subset of \( B \) and written as \( A \subset B \).

Let \( A \) be a set. A (unary) condition \( P \) on the elements of \( A \) is definite if for each element \( x \) of \( A \), it is unambiguously determined whether \( P(x) \) is true or false. For each set \( A \) and each definite condition \( P \) on the elements of \( A \), there exists a set \( B \) whose elements are those elements \( x \) of \( A \) for which \( P(x) \) is true. We write

\[ B = \{x \in A : P(x)\}. \]

Let \( A \) and \( B \) be sets. The intersection of \( A \) and \( B \) is the set

\[ A \cap B := \{x \in A : x \in B\}. \]

The sets \( A \) and \( B \) are said to be disjoint if \( A \cap B = \emptyset \). The set difference of \( B \) from \( A \) is the set

\[ A \setminus B := \{x \in A : x \notin B\}. \]

The set \( A \setminus B \) is also called the complement of \( B \) relative to \( A \).

Let \( A \) and \( B \) be sets. There exists a set \( C \) such that

\[ x \in C \iff x \in A \text{ or } x \in B. \]

We call \( C \) the union of \( A \) and \( B \), and write \( C = A \cup B \).
Theorem 1.1. Let $A$, $B$, and $C$ be sets. Then

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

Proof. We shall prove $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ only. Suppose $x \in A \cap (B \cup C)$. Then $x \in A$ and $x \in B \cup C$. Since $x \in B \cup C$, either $x \in B$ or $x \in C$. Consequently, either $x \in A \cap B$ or $x \in A \cap C$, that is, $x \in (A \cap B) \cup (A \cap C)$.

Conversely, suppose $x \in (A \cap B) \cup (A \cap C)$. Then either $x \in A \cap B$ or $x \in A \cap C$. In both cases, $x \in A$ and $x \in B \cup C$. Hence, $x \in A \cap (B \cup C)$. 

Theorem 1.2. (DeMorgan’s Rules) Let $A$, $B$, and $X$ be sets. Then

1. $X \setminus (X \setminus A) = X \cap A$.
2. $X \setminus (A \cup B) = \left( X \setminus A \right) \cap \left( X \setminus B \right)$.
3. $X \setminus (A \cap B) = \left( X \setminus A \right) \cup \left( X \setminus B \right)$.

Proof. (1) If $x \in X \setminus (X \setminus A)$, then $x \in X$ and $x \notin X \setminus A$. It follows that $x \in A$. Hence, $x \in X \cap A$. Conversely, if $x \in X \cap A$, then $x \in X$ and $x \notin X \setminus A$. Hence, $x \in X \setminus (X \setminus A)$.

(2) Suppose $x \in X \setminus (A \cup B)$. Then $x \in X$ and $x \notin A \cup B$. It follows that $x \notin A$ and $x \notin B$. Hence, $x \in X \setminus A$ and $x \in X \setminus B$, that is, $x \in (X \setminus A) \cap (X \setminus B)$. Conversely, suppose $x \in (X \setminus A) \cap (X \setminus B)$. Then $x \in X \setminus A$ and $x \in X \setminus B$. It follows that $x \in X$, $x \notin A$ and $x \notin B$. Hence, $x \notin A \cup B$, and thereby $x \in X \setminus (A \cup B)$.

(3) Its proof is similar to the proof of (2).

Let $A$ and $B$ be sets. The set of all ordered pair $(a, b)$, where $a \in A$ and $b \in B$, is called the Cartesian product of $A$ and $B$, and is denoted by $A \times B$.

For each set $A$, there exists a set $B$ whose members are subsets of $A$. We call $B$ the power set of $A$ and write $B = \mathcal{P}(A)$. Note that $\mathcal{P}(\emptyset)$ is the singleton $\{\emptyset\}$.

§2. Mappings

By a mapping $f$ from a set $X$ to a set $Y$ we mean a specific rule that assigns to each element $x$ of $X$ a unique element $y$ of $Y$. The element $y$ is called the image of $x$ under $f$ and is denoted by $f(x)$. The set $X$ is called the domain of $f$. The range of $f$ is the set $\{y \in Y : y = f(x) \text{ for some } x \in X\}$.

Two mappings $f : X \to Y$ and $g : A \to B$ are said to be equal, written $f = g$, if $X = A$, $Y = B$, and $f(x) = g(x)$ for all $x \in X$. 

2
Let $f$ be a mapping from a set $X$ to a set $Y$. If $A$ is a subset of $X$, then the mapping $g$ from $A$ to $Y$ given by $g(x) = f(x)$ for $x \in A$ is called the restriction of $f$ to $A$ and is denoted by $f|_A$.

Let $f$ be a mapping from a set $X$ to a set $Y$. If $A \subseteq X$, then $f(A)$, the image of $A$ under $f$, is defined by

$$f(A) := \{f(x) : x \in A\}.$$ 

If $B \subseteq Y$, the inverse image of $B$ is the set

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

The graph of $f$ is the set

$$G := \{(x, y) \in X \times Y : y = f(x)\}.$$ 

Thus, a subset $G$ of $X \times Y$ is the graph of a mapping from $X$ to $Y$ if and only if for every $x \in X$, there is one and only one element $y \in Y$ such that $(x, y) \in G$.

If $X$ is nonempty and $Y$ is the empty set, then there is no mapping from $X$ to $Y$. On the other hand, if $X$ is the empty set, then the empty set, considered as a subset of $X \times Y$, is the graph of a mapping from $X$ to $Y$. Hence, if $X$ is the empty set, there is one and only one mapping from $X$ to $Y$. Note that this is true even if $Y$ is also empty.

Let $f$ be a mapping from a set $X$ to a set $Y$. If for every $y \in Y$ there exists some $x \in X$ such that $f(x) = y$, then $f$ is called surjective (or $f$ is onto). If $f(x) \neq f(x')$ whenever $x, x' \in X$ and $x \neq x'$, then $f$ is called injective (or $f$ is one-to-one). If $f$ is both surjective and injective, then $f$ is bijective.

Let $X$ be a nonempty set. The identity mapping on $X$ is the mapping $i_X : X \to X$ defined by $i_X(x) = x$ for $x \in X$. Clearly, $i_X$ is bijective.

Let $f : X \to Y$ and $g : Y \to Z$ be mappings. Let $h$ be the mapping from $X$ to $Z$ given by $h(x) = f(g(x))$ for $x \in X$. Then $h$ is called the composite of $f$ followed by $g$ and is denoted by $g \circ f$.

Let $f$ be a bijective mapping from $X$ to $Y$. The inverse of $f$ is the mapping $g$ from $Y$ to $X$ given by $g(y) = x$ for $y \in Y$, where $x$ is the unique element in $X$ such that $f(x) = y$. Evidently, $g \circ f = i_X$ and $f \circ g = i_Y$. We write $f^{-1}$ for the inverse of $f$.

Let $I$ and $X$ be nonempty sets. A mapping from $I$ to $\mathcal{P}(X)$ sends each element $i$ in $I$ to a subset $A_i$ of $X$. Then $I$ is called an index set and $(A_i)_{i \in I}$ is called a family of subsets of $X$ indexed by $I$. We define the union and intersection of the family of sets $(A_i)_{i \in I}$ as follows:

$$\cup_{i \in I} A_i := \{x \in X : x \in A_i \text{ for some } i \in I\}; \quad \cap_{i \in I} A_i := \{x \in X : x \in A_i \text{ for all } i \in I\}.$$
The sets $A_i$ ($i \in I$) are said to be **mutually disjoint** if $A_i \cap A_j = \emptyset$ whenever $i, j \in I$ and $i \neq j$.

**Theorem 2.1.** Let $(A_i)_{i \in I}$ be a family of sets, and let $B$ be a set. Then

1. $B \cup (\cap_{i \in I} A_i) = \cap_{i \in I} (B \cup A_i)$.
2. $B \cap (\cup_{i \in I} A_i) = \cup_{i \in I} (B \cap A_i)$.
3. $B \setminus (\cup_{i \in I} A_i) = \cap_{i \in I} (B \setminus A_i)$.
4. $B \setminus (\cap_{i \in I} A_i) = \cup_{i \in I} (B \setminus A_i)$.

**Proof.** We shall prove (2) only. If $x \in B \cap (\cup_{i \in I} A_i)$, then $x \in B$ and $x \in \cup_{i \in I} A_i$. The latter implies $x \in A_i$ for some $i \in I$. Thus, $x \in B \cap A_i$ for this $i$. Hence, $x \in \cup_{i \in I} (B \cap A_i)$.

Conversely, if $x \in \cup_{i \in I} (B \cap A_i)$, then $x \in B \cap A_i$ for some $i \in I$. It follows that $x \in B$ and $x \in \cup_{i \in I} A_i$. Consequently, $x \in B \cap (\cup_{i \in I} A_i)$. □

**Theorem 2.2.** Let $f$ be a mapping from a set $X$ to a set $Y$.

1. If $(A_i)_{i \in I}$ is a family of subsets of $X$, then

   \[ f(\cup_{i \in I} A_i) = \cup_{i \in I} f(A_i) \quad \text{and} \quad f(\cap_{i \in I} A_i) \subseteq \cap_{i \in I} f(A_i). \]

2. If $(B_i)_{i \in I}$ is a family of subsets of $Y$, then

   \[ f^{-1}(\cup_{i \in I} B_i) = \cup_{i \in I} f^{-1}(B_i) \quad \text{and} \quad f^{-1}(\cap_{i \in I} B_i) = \cap_{i \in I} f^{-1}(B_i). \]

**Proof.** (1) We shall prove $f(\cup_{i \in I} A_i) = \cup_{i \in I} f(A_i)$ only. If $y \in f(\cup_{i \in I} A_i)$, then $y = f(x)$ for some $x \in \cup_{i \in I} A_i$. Hence, $x \in A_i$ for some $i \in I$. Consequently, $y = f(x) \in f(A_i)$ for this $i$. It follows that $y \in \cup_{i \in I} f(A_i)$. Conversely, if $y \in \cup_{i \in I} f(A_i)$, then $y \in f(A_i)$ for some $i \in I$. Hence, $y = f(x)$ for some $x \in A_i$. Then $x \in \cup_{i \in I} A_i$ and $y = f(x) \in f(\cup_{i \in I} A_i)$.

(2) We shall prove $f^{-1}(\cap_{i \in I} B_i) = \cap_{i \in I} f^{-1}(B_i)$ only. If $x \in f^{-1}(\cap_{i \in I} B_i)$, then $f(x) \in \cap_{i \in I} B_i$. It follows that $f(x) \in B_i$ for all $i \in I$. Hence, $x \in f^{-1}(B_i)$ for all $i \in I$. Consequently, $x \in \cap_{i \in I} f^{-1}(B_i)$. Conversely, if $x \in \cap_{i \in I} f^{-1}(B_i)$, then $x \in f^{-1}(B_i)$ for all $i \in I$. Hence, $f(x) \in B_i$ for all $i \in I$. Consequently, $f(x) \in \cap_{i \in I} B_i$ and thereby $x \in f^{-1}(\cap_{i \in I} B_i)$. □

In the relation $f(A_1 \cap A_2) \subseteq f(A_1) \cap f(A_2)$, equality need not hold. For example, let $f$ be the mapping from a doubleton $\{a, b\}$ to a singleton $\{y\}$ given by $f(a) = y$ and $f(b) = y$. Then $f(\{a\} \cap \{b\}) = \emptyset$ but $f(\{a\}) \cap f(\{b\}) = \{y\}$.

**Theorem 2.3.** Let $(A_i)_{i \in I}$ and $(B_i)_{i \in I}$ be two families of sets indexed by the same set $I$. Suppose that $(f_i)_{i \in I}$ is a family of mappings such that $f_i$ maps $A_i$ to $B_i$ for each $i \in I$. If
\( f_i(x) = f_j(x) \) whenever \( i, j \in I \) and \( x \in A_i \cap A_j \), then there exists a unique mapping from \( A := \bigcup_{i \in I} A_i \) to \( B := \bigcup_{i \in I} B_i \) such that \( f|_{A_i} = f_i \) for every \( i \in I \).

**Proof.** For each \( i \in I \), let \( G_i := \{(x, y) \in A_i \times B_i : y = f_i(x)\} \) be the graph of \( f_i \).

Let \( G := \bigcup_{i \in I} G_i \subseteq A \times B \). Suppose \( x \in A \). Then \( x \in A_i \) for some \( i \). It follows that \( (x, f_i(x)) \in G_i \subseteq G \). Moreover, if \( (x, y) \) and \( (x, y') \) lie in \( G \), then \( (x, y) \in G_i \) for some \( i \in I \) and \( (x, y') \in G_j \) for some \( j \in I \). Consequently, \( x \in A_i \cap A_j \) and \( y = f_i(x) = f_j(x) = y' \).

Let \( f \) be the mapping from \( A \) to \( B \) given by \( f(x) = y \) for \( x \in A \), where \( y \) is the unique element in \( B \) such that \( (x, y) \in G \). Then \( f|_{A_i} = f_i \) for every \( i \in I \). The uniqueness of such a mapping is obvious. \( \square \)

### §3. The Natural Numbers

There is a set \( \mathbb{N}_0 \), called the set of **natural numbers** and a mapping \( \sigma : \mathbb{N}_0 \to \mathbb{N}_0 \), satisfying the following (Peano) axioms:

1. There is an element \( 0 \in \mathbb{N}_0 \);
2. For each \( n \in \mathbb{N}_0 \), \( \sigma(n) \neq 0 \);
3. The mapping \( \sigma : \mathbb{N}_0 \to \mathbb{N}_0 \) is injective;
4. For each subset \( X \) of \( \mathbb{N}_0 \), if \( 0 \in X \), and if \( n \in X \) implies \( \sigma(n) \in X \), then \( X = \mathbb{N}_0 \).

The last axiom is called the axiom of induction or the **first principle of induction**. It can be restated as follows:

**Theorem 3.1.** Let \( P \) be a definite condition on the elements of \( \mathbb{N}_0 \) such that \( P(0) \) is true and that \( P(n) \) is true implies that \( P(n') \) is also true. Then \( P(n) \) is true for all \( n \in \mathbb{N}_0 \).

We often denote \( \sigma(n) \) by \( n' \) and think of \( n' \) as the successor of \( n \). We write \( 0' = 1 \), \( 1' = 2 \), \( 2' = 3 \), \( 3' = 4 \), and so on.

The following recursion theorem is a fundamental result of axiomatic number theory.

**Theorem 3.2.** Suppose that \( (\mathbb{N}_0, 0, \sigma) \) satisfies the Peano axioms. Let \( E \) be a set, \( a \in E \), and \( h : E \to E \) a mapping. Then there is exactly one mapping from \( \mathbb{N}_0 \) to \( E \) such that \( f(0) = a \) and \( f(\sigma(n)) = h(f(n)) \) for every \( n \in \mathbb{N}_0 \).

Let \( m, n \in \mathbb{N}_0 \). The **addition** \( m + n \) is defined by the recursion:

\[
    m + 0 := m \quad \text{and} \quad m + n' := (m + n)'.
\]

We call \( m + n \) the **sum** of \( m \) and \( n \). The **multiplication** \( m \cdot n \) is also defined by the recursion:

\[
    m \cdot 0 := m \quad \text{and} \quad m \cdot n' := m \cdot n + m.
\]
We often write $m \cdot n$ simply as $mn$ and call it the **product** of $m$ and $n$. The addition is commutative and associative:

$$m + n = n + m \quad \forall m, n \in \mathbb{N}_0,$$

$$(m + n) + k = m + (n + k) \quad \forall m, n, k \in \mathbb{N}_0.$$ 

The multiplication is also commutative and associative:

$$mn = nm \quad \forall m, n \in \mathbb{N}_0,$$

$$(mn)k = m(nk) \quad \forall m, n, k \in \mathbb{N}_0.$$ 

Moreover, the multiplication is distributive with respect to the addition:

$$m(n + k) = mn + mk \quad \forall m, n, k \in \mathbb{N}_0.$$ 

Let $m, n, k \in \mathbb{N}_0$. If $m + n = m + k$, then $n = k$. In other words, the **cancellation law** holds for addition. Moreover, if $mn = mk$ and $m \neq 0$, then $n = k$. So the cancellation law also holds for multiplication. In particular, if $mn \neq 0$, then $m \neq 0$ and $n \neq 0$.

Let $m, n \in \mathbb{N}_0$. If there exists some $k \in \mathbb{N}_0$ such that $n = m + k$, then we write $m \leq n$ or $n \geq m$. If $m \leq n$ and $m \neq n$, we write $m < n$ or $n > m$. The following properties can be easily established:

1. $m \leq m$ for all $m \in \mathbb{N}_0$;
2. For $m, n, k \in \mathbb{N}_0$, if $m \leq n$ and $n \leq k$, then $m \leq k$;
3. For $m, n \in \mathbb{N}_0$, if $m \leq n$ and $n \leq m$, then $m = n$.
4. For $m, n \in \mathbb{N}_0$, either $m \leq n$ or $n \leq m$.

Let $X$ be a nonempty subset of $\mathbb{N}_0$. An element $m \in X$ is said to be a **least element** of $X$ if $m \leq x$ for all $x \in X$.

**Theorem 3.3. (well-ordering property of natural numbers)** Every nonempty subset $X$ of $\mathbb{N}_0$ has a least element.

**Proof.** Let $Y := \{n \in \mathbb{N}_0 : n \leq x \text{ for all } x \in X\}$. Then $0 \in Y$ and $Y \neq \mathbb{N}_0$. By the Peano axiom (4), there exists $m \in Y$ such that $m + 1 \notin Y$. Then $m$ lies in $X$, for otherwise $m + 1$ would lie in $Y$. Therefore, $m$ is a least element of $X$.

The following theorem, called the **second principle of induction**, follows from the above theorem.
Theorem 3.4. Let $P$ be a definite condition on the elements of $\mathbb{N}_0$. If $P(n)$ is true for each $n \in \mathbb{N}_0$ whenever $P(m)$ is true for all $m < n$, then $P(n)$ is true for all $n \in \mathbb{N}_0$.

Proof. Let $X$ be the set of the elements $x$ in $\mathbb{N}_0$ such that $P(x)$ is false. It suffices to show that $X$ is empty. If $X$ is not empty, then $X$ as a least element $n$, by Theorem 3.3. If $m \in \mathbb{N}_0$ and $m < n$, then $m \notin X$, and hence $P(m)$ is true. By the assumption, $P(n)$ is also true. This contradiction shows that $X$ is empty. In other words, $P(n)$ is true for all $n \in \mathbb{N}_0$. □

§4. Integers

A mapping $*: S \times S \to S$ is called a binary operation on the set $S$. Binary operations are usually represented by symbols like $\ast$, $\cdot$, $+$, $\circ$. Moreover, the image of $(x, y)$ under a binary operation $*$ is written $x*y$.

Addition and multiplication on the set $\mathbb{N}_0$ are familiar examples of binary operations. If $A$ and $B$ are subsets of a set $X$, then $A \cup B$ and $A \cap B$ are also subsets of $X$. Hence, union and intersection are binary operations on the set $\mathcal{P}(X)$.

A binary operation $*$ on a set $S$ is said to be associative if $(x*y)*z = x*(y*z)$ for all $x, y, z \in S$. A binary operation $*$ on a set $S$ is said to be commutative if $x*y = y*x$ for all $x, y \in S$.

Let $*$ be a binary operation on a set $S$. If $e$ is an element of $S$ such that $e*x = x*e = x$ for all $x \in S$, then $e$ is called an identity element for the binary operation $*$. An identity element, if it exists, is unique.

A semigroup is a nonempty set $S$ with one associative operation on $S$. A semigroup with an identity element is called a monoid. A semigroup or a monoid is commutative when its operation is commutative.

For example, $(\mathbb{N}_0, +)$ is a commutative monoid with 0 as its identity element, and $(\mathbb{N}_0, \cdot)$ is a commutative monoid with 1 as its identity element.

A set $G$ with a binary associative operation $*$ is called a group if the operation has an identity element $e$ and every element $x$ in $G$ has an inverse (there is an element $y \in G$ such that $x*y = y*x = e$). If, in addition, the operation $*$ is commutative, then $(G, *)$ is called an abelian group. For an abelian group, the symbol $+$ is often used to denote the operation, and the identity element is denoted by 0.

For the commutative monoid $(\mathbb{N}_0, +)$, if $n \in \mathbb{N}_0 \setminus \{0\}$, then $n$ has no inverse in $\mathbb{N}_0$. Thus, it is desirable to extend $(\mathbb{N}_0, +)$ to an abelian group.

We use $\mathbb{N}$ to denote the set $\mathbb{N}_0 \setminus \{0\}$. For each $n \in \mathbb{N}$, we introduce a new symbol denoted by $-n$. We identify $-0$ with 0. Let $\mathbb{Z}$ be the set consisting of all the elements of
and all the symbols $−n$ for $n \in \mathbb{N}$. The addition $+$ on $\mathbb{Z}$ is defined as follows. Suppose $x, y \in \mathbb{Z}$. If both $x$ and $y$ are elements of $\mathbb{N}_0$, we use the same addition as before. For all $x \in \mathbb{Z}$, we define $x + 0 = 0 + x = x$. Let $m, n \in \mathbb{N}$. If $m = n + k$ for some $k \in \mathbb{N}_0$, we define

$$m + (−n) := (−n) + m := k,$$
$$−m) + n := n + (−m) := −k,$$
$$(−m) + (−n) := −(m + n).$$

It is then routine to verify that $(\mathbb{Z}, +)$ is an abelian group.

Next we define multiplication on $\mathbb{Z}$. If $x, y \in \mathbb{N}_0$, we use the same multiplication as before. For all $x \in \mathbb{Z}$, we define $x0 = 0x = 0$. Let $m, n \in \mathbb{N}$. We define $$(−m)n := n(−m) := −(mn) \quad \text{and} \quad (−m)(−n) := mn.$$ A ring is an ordered triple $(R, +, ·)$ of a nonempty set $R$ and two binary operations on $R$, an addition $+$ and a multiplication $·$, such that

1. $(R, +)$ is an abelian group;
2. $(R, ·)$ is a semigroup;
3. the multiplication is distributive: $x(y + z) = xy + xz$ and $(y + z)x = yx + zx$ for all $x, y, z \in R$.

A ring with identity is a ring whose multiplicative semigroup $(R, ·)$ has an identity element. A ring is said to be commutative if its multiplication is commutative. It is easily verified that $(\mathbb{Z}, +, ·)$ is a commutative ring with 1 as its identity element.

A ring $R$ is called an integral domain if $xy = 0$ for $x, y \in R$ implies $x = 0$ or $y = 0$. Clearly, $(\mathbb{Z}, +, ·)$ is an integral domain.

Let $x, y \in \mathbb{Z}$. If there exists some $k \in \mathbb{N}_0$ such that $y = x + k$, then we write $x \leq y$ or $y \geq x$. If $x \leq y$ and $x \neq y$, we write $x < y$ or $y > x$. The following properties can be easily established:

1. $x \leq x$ for all $x \in \mathbb{Z}$;
2. For $x, y, z \in \mathbb{Z}$, if $x \leq y$ and $y \leq z$, then $x \leq z$;
3. For $x, y \in \mathbb{Z}$, if $x \leq y$ and $y \leq x$, then $x = y$.
4. For $x, y \in \mathbb{Z}$, either $x \leq y$ or $y \leq x$.

If $x, y \in \mathbb{Z}$ and $x \leq y$, then $x + z \leq y + z$ for all $z \in \mathbb{Z}$. Moreover, $xz \leq yz$ for all $z \geq 0$.

§5. Relations

A relation on a set $A$ is a subset $R$ of the Cartesian product $A \times A$. If $(x, y) \in R$, then $x$ is said to be in relation $R$ to $y$, written $xRy$. 

8
An **equivalence relation** on a set $A$ is a relation $\sim$ on $A$ having the following properties:

1. (Reflexivity) $x \sim x$ for every $x \in A$;
2. (Symmetry) If $x \sim y$, then $y \sim x$;
3. (Transitivity) If $x \sim y$ and $y \sim z$, then $x \sim z$.

Let $E$ be an equivalence relation on a set $X$. For $a \in X$, the set of all elements in $X$ that are in relation $E$ to $a$ is called the **equivalence class** of $a$ under $E$ and is denoted by $E(a)$. Two equivalence classes are either disjoint or equal. The union of all equivalence classes is $X$. In other words, equivalence classes form a **partition** of $X$. The set of all equivalence classes of $E$ in $X$ is called the **quotient set** of $X$ by $E$ and is written $X/E$.

For example, the relation congruence modulo $n$ on $\mathbb{Z}$ is defined as follows. Let $n$ be a fixed positive integer. For $x, y \in \mathbb{Z}$, $x$ is said to be congruent to $y$ (modulo $n$), written $x \equiv y \pmod{n}$, if $n$ divides $x - y$. It can be easily verified that this is an equivalence relation. The quotient set of $\mathbb{Z}$ by this equivalence relation is denoted by $\mathbb{Z}/n\mathbb{Z}$ or $\mathbb{Z}_n$.

For $x \in \mathbb{Z}$, the equivalence class containing $x$ is the set

$$[x]_n := \{x + kn : k \in \mathbb{Z}\}.$$ 

For $x, y \in \mathbb{Z}$, $[x]_n = [y]_n$ if and only if $n$ divides $x - y$. Thus, $\mathbb{Z}_n$ is the set

$$\{[0]_n, [1]_n, \ldots, [n-1]_n\}.$$ 

We may define addition and multiplication on $\mathbb{Z}_n$ as follows:

$$[x]_n + [y]_n := [x + y]_n \quad \text{and} \quad [x]_n [y]_n := [xy]_n, \quad x, y \in \mathbb{Z}.$$ 

It can be easily verified that $(\mathbb{Z}_n, +, \cdot)$ is a commutative ring with identity.

A relation $\leq$ on a set $X$ is called a **partial ordering** if it has the following properties:

1. (Reflexivity) $x \leq x$ for every $x \in X$;
2. (Antisymmetry) If $x \leq y$ and $y \leq x$, then $x = y$;
3. (Transitivity) If $x \leq y$ and $y \leq z$, then $x \leq z$.

For $x, y \in X$, $y \geq x$ means $x \leq y$. If $x \leq y$ and $x \neq y$, we write $x < y$ or $y > x$.

A partial ordering $\leq$ is called a **linear** or **total ordering** if it has the additional property

4. (Comparability) For $x, y \in A$, either $x \leq y$ or $y \leq x$.

For example, $\subseteq$ is a partial ordering on $\mathcal{P}(X)$, the power set of a set $X$. If $X$ has at least two elements, then $\subseteq$ is not a linear ordering. On the other hand, the usual ordering $\leq$ on the set $\mathbb{Z}$ is a linear ordering.
Suppose that \( \leq \) is a partial ordering on a nonempty set \( X \). Let \( A \) be a subset of \( X \). We say that \( A \) is **bounded above** if there is an element \( b \) of \( X \) such that \( x \leq b \) for all \( x \in A \); the element \( b \) is called an **upper bound** of \( A \). Similarly, we say that \( A \) is **bounded below** if there is an element \( a \) of \( X \) such that \( a \leq x \) for all \( x \in A \); the element \( a \) is called a **lower bound** of \( A \). If \( A \) is both bounded above and bounded below, then \( A \) is said to be **bounded**. If \( b \in A \) and there is no \( x \in A \) such that \( x > b \), then \( b \) is called a **maximum** or a **maximal element** of \( A \). If \( a \in A \) and there is no \( x \in A \) such that \( x < a \), then \( a \) is called the **least** element of \( A \). Maximal (minimal) elements may or may not exist, and they need not be unique. If \( b \in A \) and \( b \geq x \) for all \( x \in A \), then \( b \) is called the **greatest** element of \( A \). If \( b \) is the greatest element of \( A \), then \( b \) is the unique maximum of \( A \) and we write \( b = \max A \). If \( a \) is the least element of \( A \), then \( a \) is the unique minimum of \( A \) and we write \( a = \min A \).

If the set of all upper bounds for \( A \) has the least element, that element is called the **least upper bound** or the **supremum** of \( A \). It is denoted by \( \sup A \). Similarly, If the set of all lower bounds for \( A \) has the greatest element, that element is called the **greatest lower bound** or the **infimum** of \( A \). It is denoted by \( \inf A \). The set \( X \) is said to have the **least upper bound property** if every nonempty subset of \( X \) that is bounded above has a least upper bound. Analogously, the set \( X \) is said to have the **greatest lower bound property** if every nonempty subset of \( X \) that is bounded below has a greatest lower bound.

Let \( A \) be a nonempty subset of \( \mathbb{Z} \). If \( A \) is bounded above, then \( A \) has the greatest element. If \( A \) is bounded below, then \( A \) has the least element. Consequently, the set \( \mathbb{Z} \) with its natural ordering has the least upper bound property and the greatest lower bound property.

Let \( \leq \) be a linear ordering on a nonempty set \( X \). If every nonempty subset of \( X \) has a least element, then \( X \) is said to be **well ordered** by \( \leq \), or \( (X, \leq) \) has the **well-ordering** property.

For example, \( \mathbb{N}_0 \) is well ordered by its natural ordering. But \( \mathbb{Z} \) is not well ordered by its natural ordering.

§6. Countable Sets

Two sets \( A \) and \( B \) are called **equinumerous**, and we write \( A \sim B \), if there exists a bijective mapping from \( A \) onto \( B \).

For \( n \in \mathbb{N} \), let \( J_n \) be the set \( \{1, \ldots, n\} := \{m \in \mathbb{N} : 1 \leq m \leq n\} \). A set \( S \) is said to be **finite** if \( S = \emptyset \) or if there exists some \( n \in \mathbb{N} \) and a bijection \( f : S \to J_n \). If a set is not...
finite, it is said to be infinite.

A set $S$ is said to be denumerable if there exists a bijection $f : S \rightarrow \mathbb{N}$. If a set is finite or denumerable, it is called countable. If a set is not countable, it is uncountable.

**Theorem 6.1.** Let $S$ be a finite set. Then the following statements are true:

(a) Every subset of $S$ is finite.

(b) If $f$ is an injective mapping from $S$ to $S$, then $f(S) = S$.

**Proof.** The statements (a) and (b) are true if $S$ is the empty set. If $S$ is a nonempty finite set, then $S \sim J_n$ for some $n \in \mathbb{N}$. Thus, it suffices to prove (a) and (b) for $S = J_n$. This will be done by induction on $n$.

(a) For $n = 1$, $J_1$ has only two subsets: the empty set and $J_1$ itself. Hence, the statement (a) is true for $S = J_1$. Suppose it is true for $S = J_n$. Let $A$ be a nonempty subset of $J_{n+1}$, and let $B := A \setminus \{n + 1\}$. If $B = \emptyset$, then $A = \{n + 1\}$ and the mapping $g : n + 1 \mapsto 1$ is a bijection from $A$ to $J_1$. If $B$ is not empty, then there exists a bijection $f$ from $B$ to $J_k$ for some $k \in J_n$. If $n + 1 \notin A$, then $A = B$ and we are done. Otherwise, let $g$ be the mapping from $A$ to $J_{k+1}$ given by $g(m) = f(m)$ for $m \in B$ and $g(n + 1) = k + 1$. Thus, $g$ is a bijective mapping from $A$ to $J_{k+1}$. This completes the induction procedure.

(b) For $n = 1$, if $f$ is a mapping from $J_1$ to $J_1$, then we have $f(J_1) = J_1$. Thus the statement (b) is true for $S = J_1$. Suppose it is true for $S = J_n$. We wish to prove that the statement is true for $S = J_{n+1}$. Let $f$ be an injective mapping from $J_{n+1}$ to $J_{n+1}$. If $f(n + 1) = n + 1$, then $f|J_n$ is an injective mapping from $J_n$ to $J_n$. Consequently, $f(J_n) = J_n$ and $f(J_{n+1}) = J_{n+1}$. Otherwise, $f(n + 1) = k \neq n + 1$. Let $h := g \circ f$, where $g$ is the mapping from $J_{n+1}$ to $J_{n+1}$ given by $g(m) = m$ for $m \in J_{n+1} \setminus \{k, n + 1\}$, $g(k) = n + 1$, and $g(n + 1) = k$. Clearly, $g$ is bijective, $h$ is injective, and $h(n + 1) = n + 1$. By what has been proved, $h$ is surjective. Therefore, $f$ is also surjective. This completes the induction procedure.

If $S$ is a nonempty finite set, then $S \sim J_n$ for some $n \in \mathbb{N}$. By the above theorem, the number $n$ is uniquely determined by $S$. We call $n$ the cardinal number of $S$ and say that $S$ has $n$ elements. The cardinal number of the empty set is 0. We use $\#S$ to denote the cardinal number of $S$.

By Theorem 6.1, there is no injective mapping from a finite subset to its proper subset. The mapping $f$ from $\mathbb{N}$ to $\mathbb{N}$ given by $f(n) = n + 1$ for $n \in \mathbb{N}$ is injective. But $f(\mathbb{N})$ is a proper subset of $\mathbb{N}$. Therefore, $\mathbb{N}$ is an infinite set.

**Theorem 6.2.** A subset of $\mathbb{N}$ is bounded if and only if it is finite. A subset of $\mathbb{N}$ is unbounded if and only if it is denumerable. Consequently, every subset of $\mathbb{N}$ is countable.
Proof. Let $S$ be a subset of $\mathbb{N}$. If $S$ is bounded and $n$ is an upper bound, then $S \subseteq J_n$, and hence $S$ is finite, by Theorem 6.1. Conversely, if $S$ is finite, then either $S = \emptyset$ or $S \sim J_n$ for some $n \in \mathbb{N}$. In the former case, $S = \emptyset$ is bounded. In the latter case, we can use induction on $n$ to show that $S$ is bounded.

If $S$ is denumerable, then $S$ must be unbounded, by what has been proved. Suppose that $S$ is an unbounded subset of $\mathbb{N}$. We construct a mapping $f$ from $\mathbb{N}$ to $S$ as follows: Let $f(1)$ be the least element of $S$. Suppose $f(1), \ldots, f(n)$ have been constructed. Then $f(\{1, \ldots, n\})$ is a finite subset of $\mathbb{N}$ and hence it is bounded. But $S$ is unbounded. Consequently, the set $S \setminus f(\{1, \ldots, n\})$ is nonempty. Let $f(n+1)$ be the least element of this set.

In light of our construction, $f(m) < f(n)$ whenever $m < n$. Hence, $f$ is injective. Let us show that $f$ is also surjective. Suppose $x$ is an element of $S$. There exists some $m \in \mathbb{Z}$ such that $x \leq f(m)$. Let $n$ be the least element of the set $\{m \in \mathbb{N} : x \leq f(m)\}$. In particular, $x \leq f(n)$. If $n = 1$, then $f(1) = x$. If $n > 1$, then $f(m) < x$ for $m = 1, \ldots, n-1$. Since $f(n)$ is the least element of the set $S \setminus f(\{1, \ldots, n-1\})$, we have $f(n) \leq x$. It follows from $x \leq f(n)$ and $f(n) \leq x$ that $f(n) = x$. Therefore, $f$ is a bijective mapping. This shows that $S$ is denumerable.

A principle of recursive definition has been used in the above proof.

**Theorem 6.3.** Let $S$ be a nonempty set. The following three conditions are equivalent.

(a) $S$ is countable;

(b) There exists a surjection $f : \mathbb{N} \to S$;

(c) There exists an injection $g : S \to \mathbb{N}$.

Proof. Suppose that $S$ is countable. Then there exists a mapping $h$ from $\mathbb{N}$ or $J_n$ $(n \in \mathbb{N})$ onto $S$. In the latter case, we fix an element $x$ of $S$ and let $f$ be the mapping from $\mathbb{N}$ to $S$ given by $f(m) = h(m)$ for $m \in J_n$ and $f(m) = x$ for $m > n$. Then $f$ is a surjection from $\mathbb{N}$ to $S$. Thus (a) implies (b).

Suppose that $f$ is a surjection from $\mathbb{N}$ to $S$. Since $f$ is surjective, $f^{-1}(\{x\})$ is nonempty for every $x \in S$. For each $x \in S$, let $g(x)$ be the least element of $f^{-1}(\{x\})$. If $x, y \in S$ and $x \neq y$, then the sets $f^{-1}(\{x\})$ and $f^{-1}(\{y\})$ are disjoint, and hence $g(x) \neq g(y)$. The mapping $g$ is an injection from $S$ to $\mathbb{N}$. Thus (b) implies (c).

Suppose that $g$ is an injective mapping from $S$ to $\mathbb{N}$. Then $g$ is a bijective mapping from $S$ onto $g(S) \subseteq \mathbb{N}$. By Theorem 6.2, $g(S)$ is a countable set. Therefore, $S$ is also a countable set. Thus (c) implies (a).

The set $\mathbb{Z}$ is denumerable. Indeed, let $f$ be the mapping from $\mathbb{Z}$ to $\mathbb{N}$ given by

$$f(m) := \begin{cases} 2m + 1 & \text{for } m \geq 0, \\ -2m & \text{for } m < 0. \end{cases}$$
Then \( f \) is a bijection.

Let \( g \) be the mapping from \( \mathbb{N} \times \mathbb{N} \) to \( \mathbb{N} \) given by
\[
g(m, n) = m + \frac{(m + n - 2)(m + n - 1)}{2}, \quad (m, n) \in \mathbb{N} \times \mathbb{N}.
\]

Then \( g \) is a bijection. Consequently, \( \mathbb{N} \times \mathbb{N} \) is denumerable.

**Theorem 6.4.** A countable union of countable sets is countable.

**Proof.** Let \((A_i)_{i \in I}\) be a family of countable sets indexed by a countable set \( I \). The case \( I = \emptyset \) is trivial; so let us assume that \( I \neq \emptyset \). Without loss of any generality, we also assume that \( A_i \neq \emptyset \) for every \( i \in I \).

Since \( I \) is countable, there exists a surjective mapping \( g \) from \( \mathbb{N} \) to \( I \). Moreover, for each \( i \in I \), since \( A_i \) is countable, there exists a surjective mapping \( f_i \) from \( \mathbb{N} \) to \( A_i \). Let \( h \) be the mapping from \( \mathbb{N} \times \mathbb{N} \) to \( A := \bigcup_{i \in I} A_i \) given by
\[
h(m, n) = f_{g(m)}(n), \quad (m, n) \in \mathbb{N} \times \mathbb{N}.
\]

Then \( h \) is a surjective mapping. But there exists a bijection from \( \mathbb{N} \) to \( \mathbb{N} \times \mathbb{N} \). Hence, there exists a surjective mapping from \( \mathbb{N} \) to \( A \). By Theorem 6.3, \( A \) is countable. \( \square \)

The countable principle of choice has been used in the proof of the above theorem.

The power set \( \mathcal{P}(\mathbb{N}) \) is uncountable. This fact is a consequence of the following more general result.

**Theorem 6.5.** Let \( X \) be a set. Then any mapping \( f \) from \( X \) to \( \mathcal{P}(X) \) is not surjective.

**Proof.** Suppose that \( f \) is a mapping from \( X \) to \( \mathcal{P}(X) \). Let
\[
A := \{ x \in X : x \notin f(x) \}.
\]

If \( a \in A \), then \( a \notin f(a) \); hence \( f(a) \neq A \). If \( a \in X \setminus A \), then \( a \in f(a) \); hence \( f(a) \neq A \). Therefore, \( A \neq f(a) \) for every \( a \in X \). This shows that \( f \) is not surjective. \( \square \)

Since there is no surjective mapping from \( \mathbb{N} \) to \( \mathcal{P}(\mathbb{N}) \), \( \mathcal{P}(\mathbb{N}) \) is uncountable, by Theorem 6.3.