# Chapter 1
# Background Results in Set Theory

This chapter summarizes the basic mathematical information required for studying this book, including definitions and results regarding operations on sets, functions, and matrices. The first section discusses operations on sets such as unions, intersections and differences. Moreover, the properties and fundamental results of these operations are presented. Section 1.2 describes the principle of mathematical induction. Section 1.3 is devoted to binary relations on sets and their properties while Sect. 1.4 classifies binary relations into different types. Equivalence and ordered relations are also discussed. The concept of a function, which is a special and important type of binary relations, is defined and examined in Sect. 1.5. Section 1.6 describes matrices and their operations. The last section contains results regarding the symmetries of the regular $n$-polygon.

## 1.1 Operations on Sets

Set theory is considered the foundation of most branches of mathematics. Abstract algebra, for example, focused on sets that are closed under one or more operations. This section discusses the basic operations on sets such as unions, intersections, and symmetric differences. The essential results related to these operations are presented. The reader may refer to (Printer, 2014) and (Halmos, 2013) for proofs of these results.

**Definition 1.1.1** A set is any collection of distinct objects, which are called elements of the set.

Sets are denoted by uppercase letters such as $A, B, C, \ldots$, while lowercase letters such as $x, y, a, b, c, \ldots$ are usually used to denote the elements. The notation $a \in A$ (read $a$ belongs to $A$, or $a$ in $A$) is used to express that $a$ is an element of $A$. If $a$ is not an element of $A$, the notation $a \notin A$ is used. The symbol $\emptyset$ denotes the empty set $\{\}$, which has no elements. The universal set that contains all the elements under consideration is denoted by $U$. A set can be described by two methods: One

 
B. Al Subaiei and M. Al Nuwairan, *A Gentle Introduction to Group Theory*,
https://doi.org/10.1007/978-981-99-0147-0_1

method is to list (if possible) all the elements of the set between two curly braces {}, this method is known as the roster method. To describe an infinite set in roster notation, some dots are placed at the end of the list, or at both ends, to indicate that the list continues forever. The other method, the descriptive method, involves stating a common characteristic of all elements of the set. The type of elements of a given set determines the method that is more appropriate. A set is called finite if it has a finite number of elements; otherwise, the set is called infinite. Figure 1.1 lists the most important infinite sets.

***Example 1.1.2***

1. The rainbow color set is a finite set that can be expressed by listing all its elements as {red, orange, yellow, green, blue, indigo, violet}.
2. The set of positive integers, denoted by $\mathbb{N}$, is an infinite set described using the roster method as $\{1, 2, 3, 4, 5, \ldots, \ldots\}$.
3. The finite set $\{2, 4, 6, 8\}$ can be expressed using the descriptive method as

$$\{x \in \mathbb{N} : x \text{ is even } \wedge 1 \leq x \leq 8\}$$

4. The set $\{7, 8, 9, 10\}$ can be expressed sing the descriptive method as

$$\{x \in \mathbb{N} : 7 \leq x \leq 10\}$$

5. The infinite set of all integers $\{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ is denoted by $\mathbb{Z}$, which can be described as

$$\{x : x \in \mathbb{N} \vee x = 0 \vee x = -y, y \in \mathbb{N}\}$$

6. The set of rational numbers, denoted by $\mathbb{Q}$, is expressed as

$$\{m/n : m, n \in \mathbb{Z} \wedge n \neq 0\}.$$

7. The set of all real numbers, denoted by $\mathbb{R}$, consisting of rational and irrational numbers, cannot be easily described using the abovementioned methods.
8. The set of complex numbers, denoted by $\mathbb{C}$, is formed using real numbers and described as

$$\left\{a + ib : a, b \in \mathbb{R}, i^2 = -1\right\}$$

9. Several sets are difficult or impossible to list using the roster method. These sets can be expressed using only the descriptive method. For example,

   a. The set of second-year students at King Faisal University (KFU) can be expressed as

$$\{x \in S\text{KFU} : x \text{ is a student in the second year}\}$$

**Fig. 1.1** Sets of numbers

$\mathbb{N}$ = The positive integers.

$\mathbb{Z}$= The integers.

$\mathbb{Q}$= The rational numbers.

$\mathbb{Q}^c$= The irrational numbers.

$\mathbb{R}$= The reals.

$\mathbb{C}$= The complex numbers.

where SKFU denotes all the students in King Faisal university.

b. The set of integers that are greater than 100 can be described as

$$\{x \in \mathbb{Z} : x > 100\}$$

c. The set of real numbers that lie between 0 and 1 can be described as

$$\{x \in \mathbb{R} : 0 < x < 1\}$$

In Fig 1.1, the main sets of numbers are listed.

**Definition 1.1.3** Let $A$ be any set. The cardinality of $A$, denoted by $|A|$, is defined as the number of its elements if $A$ is finite. If $A$ is an infinite set, the cardinality of $A$ is said to be infinite.

**Definition 1.1.4** Let $A$ and $B$ be two sets. The set $A$ is a subset of $B$, denoted by $A \subseteq B$, if each element in $A$ belongs to $B$, i.e., $A \subseteq B \Leftrightarrow (x \in A \Leftrightarrow x \in B)$. The sets $A$ and $B$ are said to be equal, denoted by $A = B$, if $A \subseteq B$ and $B \subseteq A$. The set of all subsets of a given set $A$ is called the power set of $A$ and is denoted by $\mathcal{P}(A)$.

***Example 1.1.5***

1. The empty set $\emptyset$ is a subset of any set, and any set is a subset of itself.
2. The sets $\{1, 2\}$, $\{1, 3\}$ and $\{3\}$ are subsets of $\{1, 2, 3\}$.
3. The set of positive integers $\mathbb{N}$, is a subset of $\mathbb{Z}$. As $\mathbb{Z} = \{m/1 : m \in \mathbb{Z}\}$, the set $\mathbb{Z}$ can be considered as a subset of the rational numbers $\mathbb{Q}$. Generally,

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}.$$

4. The set of all negative integers $-\mathbb{N}$ is a subset of $\mathbb{Z}$.
5. The set $\{-2, 0, 2\}$ is a subset of $\mathbb{Z}$ but not a subset of $\mathbb{N}$.
6. For any set $A$, the power set of $A$ is never empty as $\emptyset \subseteq A$ and $A \subseteq A$ for any $A$.

7. $\mathcal{P}(\{1\}) = \{\emptyset, \{1\}\}, \mathcal{P}(\emptyset) = \{\emptyset\}$.
8. $\mathcal{P}(\{1, 2, 3\}) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$.
9. $\mathcal{P}(\mathbb{Z})$ is an infinite set.
10. The sets $A = \{x \in \mathbb{Z} : |x| = 1\}$ and $B = \left\{x \in \mathbb{R} : x^2 - 1 = 0\right\}$ are equal. The integer solutions of $|x| = 1$ and the real solution of $x^2 - 1 = 0$ are the same, both sets are equal to $\{1, -1\}$.
11. The sets $A = \left\{x \in \mathbb{Q} : x^3 = x\right\}$ and $B = \left\{x \in \mathbb{R} : x^2 = x\right\}$ are not equal. The set $A$ is the rational solutions of the equation $x^3 = x$, specifically, $A = \{-1, 0, 1\}$, while $B$ is the real solutions of the equation $x^2 = x$, which are 0 and 1.
12. The sets $A = \left\{x \in \mathbb{R} : x^2 + 1 = 0\right\}$ and $B = \left\{x \in \mathbb{C} : x^2 + 1 = 0\right\}$ are not equal. The set $A$ is the empty set $\emptyset$, while $B$ equals to $\{i, -i\}$.

In the above examples, we saw that

- The power set of $\emptyset$ has only one element, i.e., $|\mathcal{P}(\emptyset)| = 1 = 2^0$.
- The power set of $\{1, 2, 3\}$ has eight elements, i.e., $|\mathcal{P}(\{1, 2, 3\})| = 8 = 2^3$.
- If $A$ has six elements, $|\mathcal{P}(A)| = 64 = 2^6$.

In general, the following proposition holds, the proof of which is presented in Exercise 1.5.

**Proposition 1.1.6** *Let $A$ be a finite set. The power set of $A$ is a finite set whose cardinality equals* $2^{|A|}$.

**Definition 1.1.7** Let be any set. The compliment of $A$, denoted by $A^c$, is defined as the set of all elements in the universal set $U$ that are not in $A$, i.e., $A^c = \{x \in U : x \notin A\}$.

**Definition 1.1.8** Let $A$ and $B$ be any two sets.

1. The union of $A$ and $B$, denoted by $A \cup B$, is the set of elements that are either in $A$, in $B$, or both, i.e., $A \cup B = \{x : x \in A \vee x \in B\}$.
2. The intersection of $A$ and $B$, denoted by $A \cap B$, is the set of elements that are in both $A$ and $B$, i.e., $A \cap B = \{x : x \in A \wedge x \in B\}$.
   – If $A \cap B = \emptyset$, we say that $A$ and $B$ are disjoint sets. The sets $A_1, A_2, \ldots, A_n$ are called mutually disjoint if $A_i \cap A_j = \emptyset$ for each $i \neq j$.
3. The difference of $A$ and $B$, denoted by $A \backslash B$ or $A - B$, is the set of elements that are in $A$ but not in $B$, i.e., $A \backslash B = \{x : x \in A \wedge x \notin B\}$.
4. The symmetric difference (disjunctive union) of $A$ and $B$, denoted by $A \Delta B$, is the set of elements in either $A$ or $B$ but not in their intersection, i.e., $A \Delta B = (A \backslash B) \cup (B \backslash A)$.

***Example 1.1.9***

1. Let $A = \{1, 2, 3, 4, 5, 6\}$ and $B = \{5, 6, 7, 8\}$.

$$A \cup B = \{1, 2, 3, 4, 5, 6, 7, 8\}, \; A \cap B = \{5, 6\}$$
$$A \backslash B = \{1, 2, 3, 4\}, \; B \backslash A = \{7, 8\},$$
$$\mathrm{A\Delta B} = \{1, 2, 3, 4, 7, 8\}.$$

Clearly, the sets $A$ and $B$ are not disjoint. Note that $A\backslash B$ and $B\backslash A$ are different sets.

2. The set of positive integers $\mathbb{N}$, and the set of negative integers $-\mathbb{N}$ are examples of disjoint sets.

**Proposition 1.1.10** *Let* $A$, $B$, and $C$ *be any sets.*

1. $A \subseteq A \cup B$, $B \subseteq A \cup B$, $A \cap B \subseteq A$, and $A \cap B \subseteq B$.
2. $A \cup (B\backslash A) = A \cup B$, and $A \cap (B\backslash A) = \emptyset$.
3. If $A \subseteq B$, then $A \cup B = B$ and $A \cap B = A$.
4. If $A \subseteq B$ and $C \subseteq D$, then $A \cup C \subseteq B \cup D$, and $A \cap C \subseteq B \cap D$.
5. $A\backslash B = A \cap B^c$.
6. $A \subseteq B$ if and only if $B^c \subseteq A^c$.

**Proposition 1.1.11** *Let* $A$, $B$, *and* $C$ *be sets, and* $U$ *be the universal set. The following identities hold.*

1. Complementation Law:$(A^c)^c = A$.
2. Idempotent Laws: $A \cup A = A$, and $A \cap A = A$.
3. Identity Laws: $A \cup \emptyset = A$, and $A \cap U = A$.
4. Domination Laws: $A \cup U = U$, and $A \cap \emptyset = \emptyset$.
5. $A\backslash\emptyset = A$, and $A\backslash A = \emptyset$.
6. $A \triangle A = \emptyset$, $A \cap A^c = \emptyset$, and $A \cup A^c = U$.
7. Commutative Laws: $A \cup B = B \cup A$, $A \cap B = B \cap A$ and $A\Delta B = B\Delta A$
8. Associative Laws:

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

9. Distributive Laws:

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

10. De Morgan's Laws:

$$(A \cup B)^c = A^c \cap B^c, \; (A \cap B)^c = A^c \cup B^c$$

11. Difference Laws:

$$A\backslash(B \cup C) = (A\backslash B) \cap (A\backslash C)$$
$$A\backslash(B \cap C) = (A\backslash B) \cup (A\backslash C)$$

**Definition 1.1.12** Let $I$ be an index set. For each $i \in I$, let $A_i$ be a set indexed by $i$. The union and intersection of $A_i$ for all values of $i$ are defined as.

$$\bigcup_{i=I} A_i = \{x : \exists\, i \in I \ni \mathrm{x} \in \mathrm{A_i}\}, \bigcap_{i \in I} A_i = \{x : x \in A_i \;\; \forall\, i \in I\}$$

If $I = \{k, k+1, \ldots, m\}$ is a finite set,

$$\bigcup_{i \in I} A_i = A_k \cup A_{k+1} \cup \cdots \cup A_m = \bigcup_{i=k}^{m} A_i$$

and

$$\bigcap_{i \in I} A_i = A_k \cap A_{k+1} \cap \cdots \cap A_m = \bigcap_{i=k}^{m} A_i$$

If $\mathcal{F}$ is a collection of sets,

$$\bigcup_{A \in \mathcal{F}} A = \{x : \exists\, A \in \mathcal{F} \ni x \in A\}$$

and

$$\bigcap_{A \in \mathcal{F}} A = \{x : x \in A \quad \forall A \in \mathcal{F}\}$$

**Proposition 1.1.13** *Let $I$ be an index set. For each $i \in I$, let $A_i$ be the set indexed by $i$.*

1. For each $j \in I$, $\bigcap_{i \in I} A_i \subseteq A_j$ and $A_j \subseteq \bigcup_{i \in I} A_i$.
2. If $B$ is a set such that for all $i \in I$, $A_i \subseteq B$, then $\bigcup_{i \in I} A_i \subseteq B$.
3. If $B$ is a set such that for all $i \in I$, $B \subseteq A_i$, then $B \subseteq \bigcap_{i \in I} A_i$.

If $I = \mathcal{F}$ is a collection of sets, the above statements can be restated as follows:

1. For all $B \in \mathcal{F}$, $\bigcap_{A \in \mathcal{F}} A \subseteq B$ and $B \subseteq \bigcup_{A \in \mathcal{F}} A$.

2. If $A \subseteq B$ for all $A \in \mathcal{F}$, then $\bigcup_{A\in\mathcal{F}} A \subseteq B$.
3. If $B \subseteq A$ for all $A \in \mathcal{F}$, then $B \subseteq \bigcap_{A\in\mathcal{F}} A$.

***Example 1.1.14***

1. Let $I = \{1, 2, 3\}$. If $A_1 = \{1, 2, 3, 5, 7\}$, $A_2 = \{3, 7, 8, 9\}$, and $A_3 = \{3, 4, 6, 7\}$, then

$$\bigcup_{i=1}^{3} A_i = \{1, 2, 3, 4, 5, 6, 7, 8, 9\} \text{ and } \bigcap_{i=1}^{3} A_i = \{3, 7\}.$$

2. For each $i \in \mathbb{N}$, let $A_i = \{m \in \mathbb{N} : m \geq i\}$. Since $A_i \subseteq \mathbb{N}$ for all $i \in I$,

$$\bigcup_{i=1} A_i \subseteq \mathbb{N} = A_1 \subseteq \bigcup_{i\in I} A_i.$$

Hence, $\bigcup_{i=1}^{\infty} A_i = \mathbb{N}$

Moreover, $\bigcap_{i=1}^{\infty} A_i = \emptyset$. For if not, then there exists $x \in \bigcap_{i=1}^{\infty} A_i$, i.e.,

$$x \in A_i = \{m \in \mathbb{N} : m \geq i\} \text{ for each } i \in \mathbb{N}.$$

That is, $x \geq i \ \forall i \in \mathbb{N}$. If $i = x+1$, then $x \geq x+1$, which is a contradiction.

3. Let $I = \{x \in \mathbb{R} : x > 0\}$ and for all $x \in I$, let $A_x = (-x, x)$. We show that $\bigcup_{x\in I} A_x = \mathbb{R}$ as follow: since $A_x \subseteq \mathbb{R}$ for each $x \in I$, then $\bigcup_{x\in I} A_x \subseteq \mathbb{R}$. For the other inclusion, let $y \in \mathbb{R}$ and pick $x = |y| + 1$, then $|y| < x$, i.e., $-x < y < x$. Hence, $y \in A_x$. Since $y \in \mathbb{R}$ is an arbitrary element, then $\mathbb{R} \subseteq \bigcup_{x\in I} A_x$.

   For the intersection, $\{0\} \subseteq (-x, x) = A_x$ for each $x \in I$, and thus, $\{0\} \subseteq \bigcap_{x\in I} A_x$. However, if y $\neq 0$, then $|y| > 0$. Let $x = \frac{|y|}{2} < |y|$, then $y \notin (-x, x)$. Hence, $y \notin \bigcap_{x\in I} A_x$ and $\bigcap_{x\in I} A_x \subseteq \{0\}$. Therefore, $\bigcap_{x\in I} A_x = \{0\}$.

**Proposition 1.1.15** *Let $I$ be an index set and $A_i$ be a set for all $i \in I$. The following identities hold:*

1. $(\bigcup_{i\in I} A_i)^c = \bigcap_{i\in I} A_i^c$ *and* $(\bigcap_{i\in I} A_i)^c = \bigcup_{i\in I} A_i^c$.
2. $B \cup (\bigcap_{i\in I} A_i) = \bigcap_{i\in I} (B \cup A_i)$ *and* $B \cap (\bigcup_{i\in I} A_i) = \bigcup_{i\in I} (B \cap A_i)$.

The equations in item (1) in Proposition 1.1.15 are known by Generalized De Morgan's Laws.

**Definition 1.1.16** (Partition of a set) Let $A$ be a nonempty set and $\mathcal{C} \subseteq \mathcal{P}(A)$. The set $\mathcal{C}$ is called a partition of $A$ if

1. $\emptyset \notin \mathcal{C}$ and $\bigcup_{E \in C} E = A$
2. For all $E, F \in \mathcal{C}$, either $E = F$ (equal) or $E \cap F = \emptyset$ (disjoint).

A partition of a set can be considered a split of the set into smaller separate and nonempty parts.

***Example 1.1.17*** Let $A = \{1, 2, 3, 4\}$. Each of the sets

$$\mathcal{C}_1 = \{\{1\}, \{2, 4\}, \{3\}\}, \mathcal{C}_2 = \{\{1, 2\}, \{3, 4\}\}, \text{ and } C_3 = \{\{1\}, \{2, 3, 4\}\}.$$

is an example of a partition of $A$, as they all satisfy the above two conditions. However, none of the sets

$$\mathcal{D}_1 = \{\{1, 2\}, \{2, 3, 4\}\}, \mathcal{D}_2 = \{\emptyset, \{1\}, \{2, 3, 4\}\}, \text{ and } \mathcal{D}_3 = \{\{1, 3\}, \{4\}\}$$

forms a partition of $A$ (Check!).

***Example 1.1.18***

1. Consider the set of integers $\mathbb{Z}$. Let $E_1$ and $E_2$ be the sets of positive and negative integers, respectively. The set $\mathcal{C} = \{E_1, E_2\}$ is not a partition of $\mathbb{Z}$ because $E_1 \cup E_2 \neq \mathbb{Z}$. The set $\mathcal{D} = \{E_1, E_2, \{0\}\}$ forms a partition of $\mathbb{Z}$.
2. Consider the set of positive integers $\mathbb{N}$. Let $E_1$, $E_2$ and $E_3$ be the sets of even positive integers, odd positive integers, and primes, respectively. The set $\mathcal{C} = \{E_1, E_2\}$ forms a partition of $\mathbb{N}$, and $\mathcal{D} = \{E_1, E_2, E_3\}$ does not, because $E_2 \cap E_3 \neq \emptyset$.
3. Consider the set of positive real numbers $\mathbb{R}^+$. For each $n \in \mathbb{N}$, let $E_n = (n - 1, n)$. The set $\mathcal{C} = \{E_n : n \in \mathbb{N}\}$ is not a partition of $\mathbb{R}^+$ since $\bigcup_{n \in \mathbb{N}} E_n = \mathbb{R}^+ \backslash \mathbb{Z} \neq \mathbb{R}^+$. If $E_n$ is replaced by $[n - 1, n]$, then $\mathcal{C}$ will not form a partition of $\mathbb{R}^+$ because $E_n \cap E_{n+1} \neq \emptyset$. If $E_n$ is replaced by $(n - 1, n]$, then $\mathcal{C}$ forms a partition of $\mathbb{R}^+$.

Another operation that is defined on sets is the Cartesian product, in which two sets form a new one using the notion of ordered pairs. An ordered pair of $a$ and $b$ is defined as the set $\{\{a\}, \{a, b\}\}$ and expressed as $(a, b)$, where $a$ and $b$ are called the first and second components of the pair. Two ordered pairs $(a, b)$ and $(c, d)$ are equal if and only if $a = c$ and $b = d$. In general, the $n$-tuple of $a_1, a_2, \ldots, a_n$ is the ordered list $(a_1, a_2, \ldots, a_n)$. The $j$th element in the $n$-tuple is called the $j$th component.

**Definition 1.1.19** (Cartesian product of sets) Let $A$ and $B$ be two sets. The Cartesian product of $A$ and $B$, denoted by $A \times B$, is defined as the set of all ordered pairs whose first and second components are elements of $A$ and $B$, respectively. That is,

$$A \times B = \{(a, b) : a \in A, b \in B\}.$$

In general, if $A_1, A_2, \ldots, A_n$ are sets, then their Cartesian product, denoted by $A_1 \times A_2 \times \ldots \times A_n$, is the set of $n$-tuples of which the $i$ th component belongs to $A_i$. That is

$$A_1 \times A_2 \times \ldots \times A_n = \{(a_1, a_2, \ldots, a_n) : a_i \in A_i, 1 \leq i \leq n\}.$$

**Proposition 1.1.20** *Let A and B be any sets, the following statements hold.*

1. $A \times \emptyset = \emptyset$.
2. $A \times B = \emptyset$ if and only if $A = \emptyset$ or $B = \emptyset$.
3. If $A$ and $B$ are nonempty sets, then $A \times B = B \times A \Leftrightarrow A = B$.
4. $|A \times B| = |B \times A| = |A||B|$.
5. If $A \times B \neq \emptyset$, then $A \times B \subseteq C \times D$ if and only if $A \subseteq C$ and $B \subseteq D$.
6. If $A \times B \neq \emptyset$, then $A \times B = C \times D$ if and only if $A = C$ and $B = D$.
7. $(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$.

***Example 1.1.21***

1. If $A = \{1, 2\}$ and $B = \{x, y, z\}$, then

$$A \times B = \{(1, x), (1, y), (1, z), (2, x), (2, y), (2, z)\}.$$

2. If $A = \{1, 2\}$ and $B = \{0\}$, then

$$A \times B = \{(1, 0), (2, 0)\}.$$

3. If $A = \mathbb{R}$, and $B = \{0\}$, then

$$A \times B = \{(x, y) : x \in \mathbb{R} \wedge y = 0\} = \{(x, 0) : x \in \mathbb{R}\}$$

is the set of points that represent the $x$-axis on the plane.
4. Consider the real numbers $\mathbb{R}$. Let $A = (0, 1)$ and $B = [0, 1)$, then

$$A \times B = \{(x, y) : 0 < x < 1 \wedge 0 \leq y < 1\}$$

is the set of points on the plane represented by a square bounded by the lines $x = 0, x = 1, y = 0$, and y $= 1$. The square's side on the $x$-axis ($y = 0$) is included (Fig. 1.2).
5. The set $\mathbb{Z} \times \mathbb{R} = \{(x, y) : x \in \mathbb{Z} \wedge y \in \mathbb{R}\}$ represents all vertical lines on the plane at which the $x$-coordinates are integers (Fig. 1.3)
6. The set $\mathbb{Z} \times \mathbb{Z} = \{(x, y) : x \in \mathbb{Z}, y \in \mathbb{Z}\}$ consists of points in the plane with both of coordinates integers. This set is represented on the plane as (Fig. 1.4):
7. The set $\mathbb{R} \times \mathbb{R} = \{(x, y) : x, y \in \mathbb{R}\} = \mathbb{R}^2$ represents the entire plane. Note that $\mathbb{R} \times \mathbb{R}^* = \{(x, y) : x, y \in \mathbb{R} \wedge y \neq 0\}$ represents the plane $\mathbb{R}^2$ except the line $y = 0$.

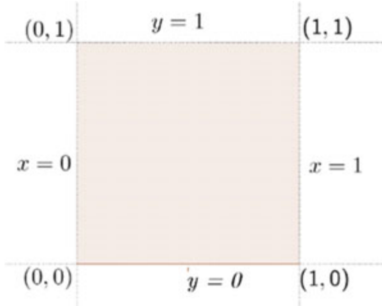



**Fig. 1.2** Graph of (0,1) × [0,1) in the plane

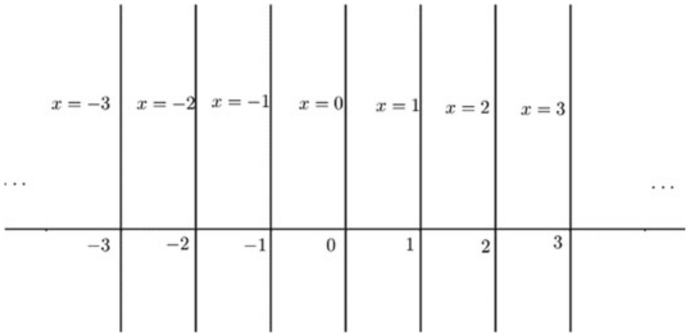



**Fig. 1.3** Graph of $\mathbb{Z} \times \mathbb{R}$

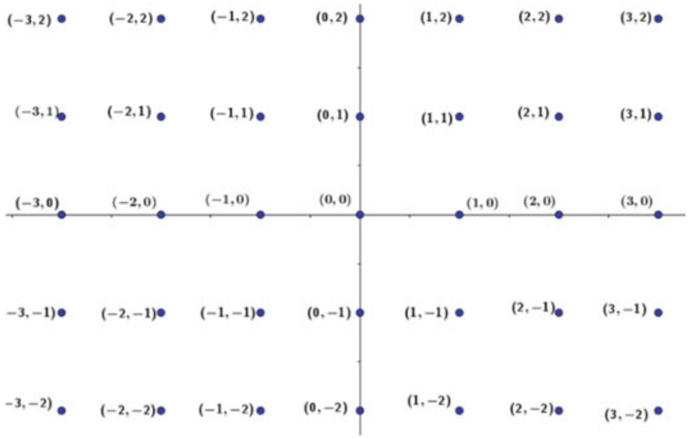



**Fig. 1.4** Graph of $\mathbb{Z} \times \mathbb{Z}$

## 1.2 Principle of Mathematical Induction

The principle of mathematical induction, or simply, mathematical induction, is a mathematical technique used to prove a statement defined for the set of integers $\mathbb{Z}$, or any of its subsets. Among the many forms of mathematical induction, we present two forms and examples of each form. For the proofs of the theorem and proposition provided in this section, see (Hammack, 2013).

**Theorem 1.2.1** (Principle of mathematical induction) *Let $n \in \mathbb{Z}$ and $P(n)$ be a mathematical statement that depends on $n$. If*

1. there exists $m \in \mathbb{Z}$ such that $P(m)$ is a true statement, and
2. for all $n \geq m$,

$$P(n) \text{ is true } \Rightarrow P(n+1) \text{ is true}$$

then $P(n)$ is a true statement for all $n \geq m$.

The statement in item (1), in above theorem, is called Base step, the statement in item (2) is called Inductive step. The process of the mathematical induction is intuitive, as the base step assumes that the statement is true for $m$; subsequently, the inductive step ensures that the statement is true for the next integer. Figure 1.5 provides an intuitive justification for the principle of mathematical induction.

$$P(m) \text{ true } \Rightarrow P(m+1) \text{ true } \Rightarrow P(m+2) \text{ true } \Rightarrow P(m+3) \text{ true } \cdots\cdots$$

***Remark 1.2.2*** The principle of mathematical induction can also be used to prove mathematical statements on a finite subset of $\mathbb{Z}$. If $A = \{m, m+1, m+2, \ldots, n\}$ is a subset of $\mathbb{Z}$ and $P(k)$ is a mathematical statement, then we can show that $P(k)$ is true for all $k \in A$ by demonstrating that $P(m)$ is true, and $P(k)$ is true $\Rightarrow P(k+1)$ is true for all $m \leq k < n$.

***Example 1.2.3*** Let $n \in \mathbb{N}$. Using the principle of mathematical induction, one can show that

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2} \quad \text{for all } n \geq 1$$

as follows:


**Fig. 1.5** Principle of mathematical induction

Base step: since $1 = 1(2)/2$, then the equality holds if $n = 1$.

Inductive step: Assume that the statement is true for $n$, i.e., $1 + 2 + \cdots + n = n(n + 1)/2$. We verify

$$1 + 2 + \cdots + n + (n + 1) = (n + 1)(n + 2)/2$$

as follows:

$$\begin{aligned} \text{L.H.S} &= 1 + 2 + \cdots + n + (n + 1) = n(n + 1)/2 + (n + 1) \\ &= (n(n + 1) + 2(n + 1))/2 = (n + 1)(n + 2)/2 = \text{ R.H.S.} \end{aligned}$$

As the two conditions are satisfied, according to the principle of mathematical induction, the statement is true for all $n \geq 1$.

In mathematical induction, the validity of the statement at $n + 1$ is derived from the validity of the statement at the previous value $n$. However, in several situations, the process might need information about several values to complete the proof. For example, if

$$a_1 = 1, a_2 = 4 \text{ and } a_n = 4a_{n-1} - 4a_{n-2} \text{ for all } n \geq 3,$$

then two values must be considered to show that $a_n = n2^{n-1}$ for all $n \geq 1$. Such a problem cannot be solved using the principle of mathematical induction in the abovementioned form. Other forms of the principle can be used to address such situations, one of which is called the principle of strong induction or simply strong induction.

**Proposition 1.2.4** (The principle of strong induction) *Let $n \in \mathbb{Z}$ and $P(n)$ be a mathematical statement that depends on $n$. If*

1. There exists $m \in \mathbb{Z}$ such that $P(m)$ is a true statement, and "Base step"
2. If $P(k)$ is true for all $k$ such that $m \leq k < n$ implies that $P(n)$ is true. "Inductive step"

then $P(n)$ is a true statement for all $n \geq m$.

***Example 1.2.5*** Let $a_1 = 1, a_2 = 4$, and $a_n = 4a_{n-1} - 4a_{n-2}$ for all $n \geq 3$. Using the principle of strong induction, we can show that

$$a_n = n2^{n-1} \text{ for all } n \geq 1$$

as follows:

Base step: The statement is true for $n = 1$ and $n = 2$.

Inductive step: Assume that $n > 2$ and the statement is true for all $1 \leq k < n$, as $1 \leq n - 2, n - 1 < n$, then

$$a_{n-1} = (n-1)2^{n-2}, \text{ and } a_{n-2} = (n-2)2^{n-3}.$$

Hence,

$$\begin{aligned} a_n &= 4(n-1)2^{n-2} - 4(n-2)2^{n-3} \\ &= 4(2(n-1) - (n-2))2^{n-3} = n2^{n-1}. \end{aligned}$$

i.e., the statement is true for $n$. According to the strong induction, the statement holds for all $n \geq 1$.

## 1.3 Binary Relations on Sets

As in any other field of study, objects in mathematics are related in various ways. For example, the relation of a point lying on a line, or the inclusion relation for sets. In mathematics, relations are usually represented by a set of ordered pairs in which the first and second components are related. For example, let $P$ be a set of points and $L$ be a set of lines on the plane. The set $\mathcal{J} = \{(p, l) \in P \times L : p \text{ lies on } l\}$ represents the relation that $p$ lies on $l$. The inclusion relation among a family of sets is represented by $\mathcal{S} = \{(A, B) \in \mathfrak{F} \times \mathfrak{F} : A \subseteq B\}$, where $\mathfrak{F}$ is a collection of sets. In general, the following definition can be stated.

**Definition 1.3.1** Let $A$ and $B$ be two sets. A binary relation (or simply, a relation) $\mathcal{R}$ from $A$ to $B$ is a subset of $A \times B$, i.e.,

$$\mathcal{R} \subseteq \{(a, b) : a \in A, b \in B\}$$

If $A = B$, we say $\mathcal{R}$ is a relation on $A$.

For an ordered pair, $(a, b) \in A \times B$, either $(a, b) \in \mathcal{R}$ or $a\mathcal{R}$b is used to denote that $a$ is related to $b$ through the relation $\mathcal{R}$. Both notations $(a, b) \notin \mathcal{R}$ and $a\not{\mathcal{R}}b$ are used to express that $a$ and $b$ are not related through the relation $\mathcal{R}$. The notion of a binary relation is generalized to more than two sets as follows:

Let $A_1, A_2, \ldots, A_n$ be any sets. A relation $\mathcal{R}$ of these sets is a subset of $A_1 \times A_2 \times \ldots \times A_n$, i.e., $\mathcal{R} \subseteq \{(a_1, a_2, \ldots, a_n) : a_1 \in A_1, a_2 \in A_2, \ldots, a_n \in A_n\}$. In this book, only the binary relation is considered unless otherwise stated.

For any two sets $A$ and $B$, two relations always exist from $A$ to $B$. Namely, $\mathcal{R} = \emptyset$ and $\mathcal{R} = A \times B$. More examples are presented below.

***Example 1.3.2***

1. The set $\mathcal{R} = \{(1, 7), (1, 8), (1, 9), (3, 7)\}$ is a relation from $\{1, 3\}$ to $\{7, 8, 9\}$.
2. The set $\mathcal{R} = \{(a, 7), (a, 8), (a, 9), (b, 7)\}$ is a relation from $\{a, b\}$ to $\{7, 8, 9\}$.
3. The set $\mathcal{R} = \{(1, 1), (2, 2), (1, 3)\}$ is a relation on $\{1, 2, 3\}$.
4. The set $\mathcal{S} = \{(A, B) \in \mathfrak{F} \times \mathfrak{F} : A \subseteq B\}$ is a relation on $\mathfrak{F}$, where $\mathfrak{F}$ is any family of sets.

5. The set $\mathcal{S} = \{(m, n) \in \mathbb{Z} \times \mathbb{Z} : n \text{ divides } m\}$ is a relation on $\mathbb{Z}$.
6. The set $\mathcal{R} = \left\{(x, y) \in \mathbb{N}^2 : y = x + 1\right\}$ is a relation on $\mathbb{N}$.
7. The set $\mathcal{R} = \left\{(x, y) \in \mathbb{R}^2 : y = x + 1\right\}$ is a relation on $\mathbb{R}$.
8. Let $\mathcal{R}_1 = \left\{(x, y) \in \mathbb{R} \times \mathbb{R} : x^2 + y^2 = 1\right\}$ and $\mathcal{R}_2 = \left\{(x, y) \in \mathbb{R} \times \mathbb{R} : x^2 + y^2 \leq 1\right\}$ be the unit circle and the unit disk in the plane, respectively. Both $\mathcal{R}_1$ and $\mathcal{R}_2$ are relations on $\mathbb{R}$. The sets $\mathcal{R}_1$ and $\mathcal{R}_2$ are not relations on $\mathbb{Z}$ since $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ belongs to $\mathcal{R}_1$ and $\mathcal{R}_2$ but it does not belong to $\mathbb{Z} \times \mathbb{Z}$.

**Definition 1.3.3** Let $A$ and $B$ be any two sets and $\mathcal{R}$ be a relation from $A$ to $B$.

1. The domain of $\mathcal{R}$, denoted by $D(\mathcal{R})$, is the set of elements of $A$ that appear as the first components in the elements of $\mathcal{R}$. i.e., $D(\mathcal{R}) = \{a \in A : \exists\, b \in B \wedge (a, b) \in \mathcal{R}\}$.
2. The range of $\mathcal{R}$, denoted by $Rang(\mathcal{R})$, is the set of elements of $B$ that appear as the second components in the elements of $\mathcal{R}$. i.e., $Rang(\mathcal{R}) = \{b \in B : \exists\, a \in A \wedge (a, b) \in \mathcal{R}\}$.
3. The set $B$ is called the codomain of $\mathcal{R}$.
4. For each $(a, b) \in \mathcal{R}$, the element $b$ is called the image of $a$ under $\mathcal{R}$.

For any sets $A$ and $B$, we have
$D(\emptyset) = Rang(\emptyset) = \emptyset$, $D(A \times B) = A$, and $Rang(A \times B) = B$.

***Example 1.3.4*** The domains and ranges of the relations in Example 1.3.2 are

1. $D(\mathcal{R}) = \{1, 3\}$, and $Rang(\mathcal{R}) = \{7, 8, 9\}$.
2. $D(\mathcal{R}) = \{a, b\}$, and $Rang(\mathcal{R}) = \{7, 8\}$.
3. $D(\mathcal{R}) = \{1, 2\}$, and $Rang(\mathcal{R}) = \{1, 2, 3\}$.
4. $D(\mathcal{R}) = \mathfrak{F}$, and $Rang(\mathcal{R}) = \mathfrak{F}$ (This is true because any set is a subset of itself).
5. $D(\mathcal{R}) = \mathbb{Z}$, and $Rang(\mathcal{R}) = \mathbb{Z}\backslash\{0\}$.
6. $D(\mathcal{R}) = \mathbb{N}$, and $Rang(\mathcal{R}) = \mathbb{N}\backslash\{1\}$.
7. $D(\mathcal{R}) = \mathbb{R}$, and $Rang(\mathcal{R}) = \mathbb{R}$.
8. $D(\mathcal{R}_1) = D(\mathcal{R}_2) = [-1, 1]$, and $Rang(\mathcal{R}_1) = Rang(\mathcal{R}_2) = [-1, 1]$.

In the following, we restrict our study to the relations in which $A = B$. We study the properties of these relations and discuss two types of relations that appear frequently in algebra. Recall that a relation from $A$ to $A$ is called a relation on $A$.

**Definition 1.3.5** (*Properties of a relation on a set*) Let $A$ be any set and $\mathcal{R}$ be a relation on $A$. The relation $\mathcal{R}$ is

1. Reflexive: if $(a, a) \in \mathcal{R}$ for all $a \in A$.
2. Symmetric: if for all $a, b \in A$, $(a, b) \in \mathcal{R} \Rightarrow (b, a) \in \mathcal{R}$.
3. Antisymmetric: if for all $a, b \in A$, $((a, b) \in \mathcal{R} \wedge (b, a) \in \mathcal{R}) \Rightarrow a = b$.
4. Transitive: if for all $a, b, c \in A$, $((a, b) \in \mathcal{R} \wedge (b, c) \in \mathcal{R}) \Rightarrow (a, c) \in \mathcal{R}$.

**Definition 1.3.6** Let $A$ be any set and $\mathcal{R}$ be a relation on $A$. The relation $\mathcal{R}$ is called a connex relation if either $(a, b) \in \mathcal{R}$ or $(b, a) \in \mathcal{R}$ for each $a, b \in A$.

It is clear that any connex relation is reflexive.

***Example 1.3.7***

1. Let $A = \emptyset$. As $A \times A$ is the empty set $\emptyset$, the only relation that can be defined on $A$ is the empty relation $\emptyset$. Clearly, $\emptyset$ is reflexive (if not, then there exists $a \in A = \emptyset$ such that $(a, a) \notin \mathcal{R}$, which is impossible). As $\emptyset$ does not contain any elements, the conditional statement for the symmetry is true, which implies that $\emptyset$ is symmetric. The same justification implies that $\emptyset$ is antisymmetric and transitive. Clearly, since there are no elements in $A$, the relation $\emptyset$ is a connex relation.
2. Let $A \neq \emptyset$. The set $\emptyset$ represents a relation on $A$ that is symmetric, antisymmetric, and transitive. It is not reflexive (pick $a \in A$, then $(a, a) \notin \emptyset$), therefore, it is not a connex relation.
3. Let $A = \{1, 2, 3\}$. The relation $\mathcal{R} = \{(1, 1), (2, 2), (1, 3)\}$ is not reflexive since $(3, 3) \notin \mathcal{R}$. It is not symmetric as $(1, 3) \in \mathcal{R}$ but $(3, 1) \notin \mathcal{R}$. The relation $\mathcal{R}$ is transitive since the only ordered pairs in the form $(a, b), (b, c)$ in $\mathcal{R}$ are $(1, 1), (1, 3)$ and $(a, c) = (1, 3) \in \mathcal{R}$. Moreover, $\mathcal{R}$ is antisymmetric, because no elements in $\mathcal{R}$ in the form $(a, b), (b, a)$ where $a \neq b$. As $\mathcal{R}$ is not reflexive, it is not a connex relation.
4. For any set $A$, define the relation $\Delta_A = \{(a, a) \in A \times A : a \in A\}$. It straightforward to check that $\Delta_A$ is reflexive, symmetric, antisymmetric, and transitive. Moreover, $\Delta_A$ is not a connex relation for any $A$ such that $|A| \geq 2$ (if $a \neq b$ are two elements in $A$, then neither $(a, b)$ nor $(b, a)$ belongs to $\Delta_A$). The relation $\Delta_A$ expresses the equality relation and is called the identity (or the diagonal) relation. The symbol $\Delta_A$ denotes the identity relation on $A$. If $A = \{1, 2, 3\}$, then $\Delta_A = \{(1, 1), (2, 2), (3, 3)\}$ is an example for the identity relation on a finite set. The line $y = x$ is a visualization for $\Delta_A = \{(x, x) \in \mathbb{R} \times \mathbb{R} : x \in \mathbb{R}\}$, an example of an infinite identity relation.
5. Let $A \neq \emptyset$ be any set. The relation $\mathcal{R} = A \times A = \{(a, b) \in A \times A : a, b \in A\}$ is reflexive, symmetric, and transitive. It is not antisymmetric for any $A$ having more than one element. Let $a \neq b$ be elements in $A$, according to the definition of $\mathcal{R}$, both $(a, b)$ and $(b, a)$ belong to $\mathcal{R}$, but $a \neq b$, which implies that $\mathcal{R}$ is not antisymmetric. Clearly, as $(a, b) \in \mathcal{R}$ for all $a, b \in A$, then $\mathcal{R}$ is a connex relation.
6. Let $A = \mathbb{N}$ and $\mathcal{R} = \left\{(x, y) \in \mathbb{N}^2 : y = x + 1\right\}$. The relation $\mathcal{R}$ is not reflexive, not symmetric, and not transitive. As $\mathcal{R}$ is not reflexive, it is not a connex relation.
7. Let $A = \mathbb{R}$ and $\mathcal{R} = \left\{(x, y) \in \mathbb{R} \times \mathbb{R} : x^2 + y^2 = 1\right\}$. The relation $\mathcal{R}$ is not reflexive since $2 \in A$, but $(2, 2) \notin \mathcal{R}$ ( $2^2 + 2^2 = 8 \neq 1$). It is symmetric, for if $(x, y) \in \mathcal{R}$ then $x^2 + y^2 = 1$. i.e., $y^2 + x^2 = 1$ and $(y, x) \in \mathcal{R}$. The relation $\mathcal{R}$ is not antisymmetric as both $(1, 0), (0, 1) \in \mathcal{R}$ but $0 \neq 1$. Finally, it is not transitive as $(1, 0), (0, -1) \in \mathcal{R}$, but $(1, -1) \notin \mathcal{R}$. Clearly, it is not a connex relation (Fig. 1.6).

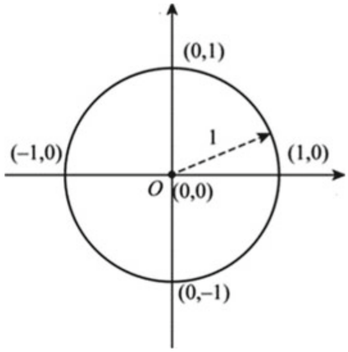



**Fig. 1.6** Graph of $\mathcal{R} = \{(x, y) \in \mathbb{R} \times \mathbb{R} : x^2 + y^2 = 1\}$

## 1.4 Types of Binary Relations on Sets

A relation can be classified depending on its properties. In this book, we will encounter specific types of relations, such as equivalence and order relations. Readers can refer to (Halmos, 2013) for more details regarding the types of relations.

**Definition 1.4.1** (*Types of relations*) Let $A$ be a set and $\mathcal{R}$ be a relation on $A$. The relation $\mathcal{R}$ is said to be

1. An equivalence relation if $\mathcal{R}$ is reflexive, symmetric, and transitive.
2. A partial order relation if $\mathcal{R}$ is reflexive, antisymmetric, and transitive.
3. A total order relation if $\mathcal{R}$ is antisymmetric, transitive, and a connex relation.

As any connex relation is reflexive, a total order relation must be a partial order relation. By order relation, we mean a partial order relation.

The identity relation (Example 1.3.7(4)) is the canonical example of an equivalence relation, where for any $a, b \in A$, $(a, b) \in \mathcal{R}$ if and only if $a = b$. The partial order relation generalizes the concept of ordering or arranging the elements of a set. For $a, b \in A$, the pair $(a, b)$ belongs to a partial order relation means that one of the elements precedes the other in the order. The word "partial" indicates that not every pair of elements in $A$ are related, i.e., if $a$ and $b$ are arbitrary elements in $A$, then the partial order relation does not require $(a, b)$ or $(b, a)$ to be $\mathcal{R}$. In contrast, due to the connexity property, the total order relation requires that either $(a, b) \in \mathcal{R}$ or $(b, a) \in \mathcal{R}$ for each $a, b \in A$. i.e., in a total order relation, any two elements in $A$ are comparable. The set endowed with a total order relation is called a chain. For order relations, the notation $\leq$ is usually used instead of $\mathcal{R}$, and the notation $a \leq b$ is used instead of $a\mathcal{R}b$.

***Example 1.4.2***

1. In Example 1.3.7,
   - The relation $\emptyset$ in (1) is an equivalence relation, a partial order, and a total order relation.

- The set $\emptyset$ in (2) and the relation in (3) are not reflexive; hence these relations are neither equivalence nor order relations.
- The relation $\Delta_A$ in (4) is an equivalence relation and a partial order relation. If $|A| \geq 2$, we pick $a, b \in A$ such that $a \neq b$. Since neither $(a, b)$ nor $(b, a)$ belongs to $\Delta_A$, then $\Delta_A$ cannot be a total order relation for any set that contains more than one element.
- The relation in (5) is an equivalence relation, it is not an order relation since it is not antisymmetric.
- The relation in (6) is neither an equivalence relation nor an order relation. This result remains true if $\mathbb{N}$ is replaced with $\mathbb{Z}$, $\mathbb{Q}$, or $\mathbb{R}$.
- The relation in (7) is not reflexive, so it is neither equivalence nor order relation.

2. Let $A = \mathbb{R}$, and $\mathcal{R}$ be the natural order on $\mathbb{R}$. i.e., $a\mathcal{R}b \Leftrightarrow a \leq b$ for all $a, b \in \mathbb{R}$. It is straightforward to verify that $\mathcal{R}$ is reflexive, antisymmetric, and transitive. The relation $\mathcal{R}$ is not symmetric. For any two distinct real numbers, one must be greater than the other. Hence, the natural order is a total (hence, a partial) order relation, but not an equivalence relation.

**Definition 1.4.3** Let $A$ be a nonempty set and $\leq$ be a partial order relation on $A$. For any nonempty subset $B$ of $A$,

1. an element $a \in A$ is called a lower bound of $B$ if $a \leq b$ for all $b \in B$,
2. an element $d \in A$ is called an upper bound of $B$ if $b \leq d$ for all $b \in B$,
3. the set $B$ is called bounded below (bounded above) if $B$ has a lower bound (an upper bound),
4. the set $B$ is called bounded if it is bounded below and above; otherwise, it is called unbounded.

**Definition 1.4.4** Let $A$ be a nonempty set and $\leq$ be a partial order relation on $A$. For any nonempty subset $B$ of $A$,

1. an element $c \in B$ is called minimal in $B$ if for all $b \in B$, $b \leq c \Rightarrow c = b$,
2. an element $c \in B$ is called the minimum of $B$ if $c \leq b$ for all $b \in B$,
3. an element $d \in B$ is called maximal in $B$ if for all $b \in B$, $d \leq b \Rightarrow d = b$,
4. an element $d \in B$ is called the maximum of $B$ if $b \leq d$ for all $b \in B$.

Note that if we read $c \leq b$ as $c$ is less than or equal to $b$, then

in the above definition, (1) states that $c \in B$ is minimal if $c$ is less than or equal to every element in $B$ that is comparable with $c$. Item (2) states that $c \in B$ is minimum if $c$ is less than or equal to every element in $B$. Thus, the minimum is always a minimal element, the maximal and maximum elements can be similarly distinguished, and the maximum element is always maximal. The converse is not true, as shown in the following example.

***Example 1.4.5*** Let $A = \{a, b, c\}$.

1. Let $\mathcal{R} = \{(a, a), (b, b), (c, c), (a, b), (a, c)\}$. The relation $\mathcal{R}$ is a partial order relation on $A$ (Check!). Clearly, $b$ is a maximal element in $A$, but is not maximum, since there exists an element $c \in A$ such that $c \not\leq b$. The element $c$ is also maximal.
2. Let $\mathcal{R} = \{(a, a), (b, b), (c, c), (b, c), (a, c)\}$. The relation $\mathcal{R}$ is a partial order relation on $A$ (Check!). The element $b$ is a minimal element in $A$ that is not minimum, since there exists an element $a \in A$ such that $a \not\leq c$. The element $a$ is also minimal.
3. Let $\mathcal{R} = \{(a, a), (b, b), (c, c), (a, b), (a, c), (b, c)\}$. It is easy to check that $c$ is a maximal and maximum element in $A$, and $a$ is minimal and minimum.

**Proposition 1.4.6** *Let $A$ be a nonempty set and $\leq$ be a partial order relation on $A$. If $B$ is a nonempty subset of $A$, then $B$ has at most one maximum and one minimum element. If $\leq$ is a total order relation on $B$, then a minimal (res. maximal) is the minimum (res. maximum) element in $B$.*

Next, we focus on equivalence relations. Recall the definition of a partition of a set given in Definition 1.1.16. We show that any equivalence relation results in a partition of the underlying set (Theorem 1.4.10 below). We begin with the following definition.

**Definition 1.4.7** (*Equivalence classes*) Let $A$ be a nonempty set and $\mathcal{R}$ be an equivalence relation on $A$. For all $a \in A$, the equivalence class of $a$, denoted by $[a]_{\mathcal{R}}$ or simply $[a]$, is the set of elements of $A$ that are related to $a$ via $\mathcal{R}$. That is,

$$\begin{aligned} [a]_{\mathcal{R}} &= \{b \in A : (a, b) \in \mathcal{R}\} = \{b \in A : (b, a) \in \mathcal{R}\} \\ &= \{b \in A : a\mathcal{R}b\} = \{b \in A : b\mathcal{R}a\}. \end{aligned}$$

Two elements of the set $A$ are called equivalent if and only if they belong to the same equivalence class, i.e., if and only if they are related by $\mathcal{R}$. The set of all equivalence classes for all elements in $A$ is called equivalence classes of $\mathcal{R}$.

***Example 1.4.8***

1. Let $A = \{1, 2, 3\}$ and $\mathcal{R}_1 = \{(1, 1), (2, 2), (3, 3), (1, 2), (2, 1)\}$. The relation $\mathcal{R}_1$ is an equivalence relation (Verify!). The equivalence classes are Fig 1.7 visualizes the equivalence class for each element in $A$.

$$[1] = \{1, 2\} = [2] \text{ and } [3] = \{3\}$$

2. Let $A = \{1, 2, 3, 4\}$ and $\Delta_A = \{(1, 1), (2, 2), (3, 3), (4, 4)\}$. As explained in Example 1.3.7, the relation $\Delta_A$ is an equivalence relation. The equivalence classes are

$$[1] = \{1\}, [2] = \{2\}, [3] = \{3\}, \text{ and } [4] = \{4\}.$$

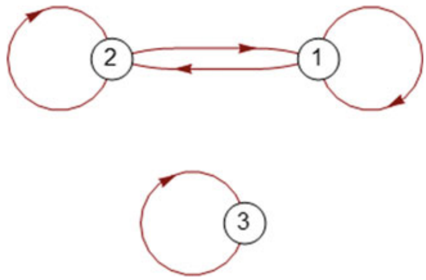



**Fig. 1.7** Equivalence classes of $\mathcal{R}_1$

In general, the equivalence class of an element in a set $A$, endowed with $\Delta_A$, consists of one element $[a] = \{a\}$, see also Fig. 1.8.

**Proposition 1.4.9** *Let $A$ be a nonempty set and $\mathcal{R}$ be an equivalence relation on $A$. For all $a, b \in A$, the following statements are satisfied:*

1. $a \in [a]$.
2. $b \in [a]$ if and only if $[a] = [b]$.
3. $[a] \cap [b] \neq \emptyset$ if and only if $[a] = [b]$.
4. If $(a, b) \notin \mathcal{R}$, then $[a] \cap [b] = \emptyset$.

Item (3) in the above proposition states that any two equivalence classes are either equal or disjoint. This is an important fact for proving the following theorem.

**Theorem 1.4.10** *Let $A$ be a nonempty set. The equivalence classes of any equivalence relation on $A$ form a partition of $A$.*

***Example 1.4.11*** Let $A = \mathbb{Z}$ and $\mathcal{R} = \left\{(a, b) \in \mathbb{Z}^2 : (a - b)/3 \in \mathbb{Z}\right\}$. The relation $\mathcal{R}$ is an equivalence relation on $\mathbb{Z}$ and satisfies the following properties:

- Reflexive: For all $a \in \mathbb{Z}$, $((a - a)/3) = 0 \in \mathbb{Z}$, which implies that $(a, a) \in \mathcal{R}$.
- Symmetric: Assume that $(a, b) \in \mathcal{R}$ for arbitrary integers $a$ and $b$. According to the definition of $\mathcal{R}$, $(a - b)/3 \in \mathbb{Z}$, which is equivalent to $(b - a)/3 = -(a - b)/3 \in \mathbb{Z}$. Hence, $(b, a) \in \mathcal{R}$.
- Transitive: Assume that $(a, b) \in \mathcal{R}$ and $(b, c) \in \mathcal{R}$ for arbitrary integers $a, b$, and $c$. According to the definition of $\mathcal{R}$, both $(a - b)/3$ and $(b - c)/3 \in \mathbb{Z}$. Consequently,

$$\frac{a - c}{3} = \frac{a - b}{3} + \frac{b - c}{3} \in \mathbb{Z}.$$

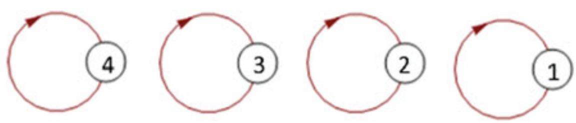



**Fig. 1.8** Equivalence classes of $\Delta_A$

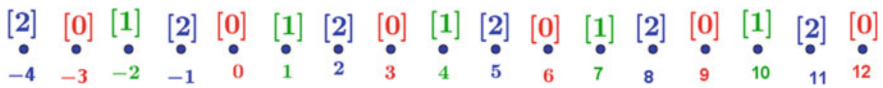



**Fig. 1.9** Equivalence classes of $\mathcal{R}$ on $A$

Therefore, the relation $\mathcal{R}$ is an equivalence relation. For any $a \in \mathbb{Z}$, the equivalence class of $a$ is

$$[a] = \{b \in \mathbb{Z} : (a, b) \in \mathcal{R}\} = \left\{b \in \mathbb{Z} : \frac{a-b}{3} \in \mathbb{Z}\right\} = \left\{b \in \mathbb{Z} : \exists\, k \in \mathbb{Z} \ni \frac{\mathrm{a}-\mathrm{b}}{3} = \mathrm{k}\right\}$$
$$= \{b \in \mathbb{Z} : \exists\, k \in \mathbb{Z} \ni \mathrm{b} = \mathrm{a} - 3\mathrm{k}, \mathrm{k} \in \mathbb{Z}\} = \mathrm{a} + \{3l : l \in \mathbb{Z}\} = \mathrm{a} + 3\mathbb{Z}.$$

For example,

$$[0] = 0 + 3\mathbb{Z} = \{\ldots - 6, -3, 0, 3, 6, \ldots\}, [1] = 1 + 3\mathbb{Z} = \{\ldots - 5, -2, 1, 4, 7, \ldots\}$$
$$[2] = 2 + 3\mathbb{Z} = \{\ldots - 4, -1, 2, 5, 8, \ldots\}, [3] = 3 + 3\mathbb{Z} = \{\ldots - 3, 0, 3, 6, 9, \ldots\} = [0]$$

By Proposition 1.4.9 (3),
$\ldots = [-6] = [-3] = [0] = [3] = [6] = \ldots =$ multiple of 3
$\ldots = [-5] = [-2] = [1] = [4] = [7] = \ldots =$ ( multiple of 3) + 1.
$\ldots = [-4] = [-1] = [2] = [5] = [8] = \ldots =$ ( multiple of 3) + 2 (Fig. 1.9).

To show that these classes are the only equivalent classes for the relation $\mathcal{R}$, let $m$ be an arbitrary element in $\mathbb{Z}$. Confine $m$ between two consecutive multiples of 3, i.e., find $n$ such that

$$3n \leq m < 3(n+1) = 3n + 3$$

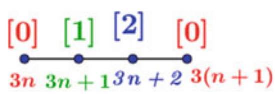
The possible values for the integer $m$ are

$m = 3n \in 3\mathbb{Z} = [0]$, $m = 3n + 1 \in 3\mathbb{Z} + 1 = [1]$, or $m = 3n + 2 \in \mathbb{Z} + 2 = [2]$.

This indicates that at most three equivalence classes exist, namely: [0], [1], [2], see Fig 1.10. According to Proposition 1.4.9 (4), as (0, 1), (0, 2), and (1, 2) are not elements of $\mathcal{R}$, the three equivalence classes are disjoint. The set of equivalence classes of $\mathcal{R}$ is $\{[0], [1], [2]\} = \{[6], [-8], [-4]\} = \ldots$ etc.

In the previous example, using an equivalence relation on $\mathbb{Z}$, the set $\mathbb{Z}$ was divided into three disjoints parts. By defining another similar equivalence relation on $\mathbb{Z}$, the set $\mathbb{Z}$ can be divided into another number of sets. For example, the relation $\mathcal{R} = \left\{(a, b) \in \mathbb{Z}^2 : (a-b)/9 \in \mathbb{Z}\right\}$ on $\mathbb{Z}$ divides $\mathbb{Z}$ into nine disjoint sets. In general, for any $n \in \mathbb{Z}^*$, the relation

**Fig. 1.10** Only three different equivalence classes of $\mathcal{R}$



$$\mathcal{R} = \left\{(a, b) \in \mathbb{Z}^2 : \frac{a-b}{n} \in \mathbb{Z}\right\}$$

divides $\mathbb{Z}$ into $n$ disjoint sets in the form $m + n\mathbb{Z}$, where $m = 0, 1, 2, \dots, n-1$ (Exercise 1.19).

***Example 1.4.12*** Let $\mathbb{Z}^*$ be the set of nonzero integers and

$$A = \mathbb{Z} \times \mathbb{Z}^* = \{(m, n) : m, n \in \mathbb{Z} \wedge n \neq 0\}.$$

On $A$, define the relation $\sim$ as $(m, n) \sim (r, s) \Leftrightarrow ms = nr$. The relation $\sim$ satisfies the following properties:

1. Reflexive: For an arbitrary element $(m, n) \in A$, the integers $m, n$ satisfy $mn = nm$, which implies that $(m, n) \sim (m, n)$.
2. Symmetric: Assume that $(m, n), (r, s) \in A$ are arbitrary elements such that $(m, n) \sim (r, s)$, then

$$(m, n) \sim (r, s) \Rightarrow ms = nr \Rightarrow rn = sm \Rightarrow (r, s) \sim (m, n).$$

3. Transitive: Suppose $(m, n), (r, s), (k, l)$ are arbitrary elements in $A$ such that $(m, n) \sim (r, s)$ and $(r, s) \sim (k, l)$. Therefore, $m, n, r, s, k, l$ are integers where $n, s, l \neq 0$ , and

$$(m, n) \sim (r, s) \text{ and } (r, s) \sim (k, l) \Rightarrow ms = nr \wedge rl = sk.$$

Multiplying both sides of $ms = nr$ by $l$ and both sides of $rl = sk$ by $n$ yields

$$msl = nrl \text{ and } nrl = nsk$$

which implies that $msl = nsk$. As $s \neq 0$, dividing both sides of $msl = nsk$ by $s$ yields $ml = nk$. Thus, $(m, n) \sim (k, l)$, and $\sim$ is transitive.

Thus, the relation $\sim$ is an equivalence relation. By definition, the equivalence class of $(m, n)$ is given by

$$[(m, n)] = \{(r, s) \in A : (m, n) \sim (r, s)\} = \{(r, s) \in A : ms = nr\}.$$

If $(m, n)$ is identified with the rational fraction $m/n$, then $[(m, n)]$ is identified with the set of all equivalent fractions to $m/n$. In fact, the set of rational fractions $\mathbb{Q}$ is defined as the set of equivalence classes of the relation $\sim$.

$$\mathbb{Q} = \{[(m, n)] : m, n \in \mathbb{Z} \wedge n \neq 0\} := \{m/n : m, n \in \mathbb{Z} \wedge n \neq 0\}.$$

The set $\mathbb{Q}^* = \mathbb{Q} \backslash [(0, n)]$ is identified with $\{m/n : m, n \in \mathbb{Z} \wedge m, n \neq 0\}$.

If $A$ is any nonempty set and $\mathcal{C}$ is any partition of $A$, a relation $\mathcal{R}$ on $A$ can be defined as follows:

$(a, b) \in \mathcal{R}$ if and only if $a, b$ belong to the same element (set) of $\mathcal{C}$.

It is straightforward to verify that the relation $\mathcal{R}$ is an equivalence relation. The equivalence class of any element $a$ in $A$ is $[a] = \{b \in A : (a, b) \in \mathcal{R}\} = \{b \in A : a, b \text{ belong to the same set in } \mathcal{C}\} = E$

where $E$ is the element in $\mathcal{C}$ containing $a$.

**Proposition 1.4.13** *Let $A$ be a nonempty set. Any partition $\mathcal{C}$ of $A$ defines an equivalence relation on $A$ whose equivalence classes are the elements of $\mathcal{C}$.*

**Corollary 1.4.14** *Let $A$ be a nonempty set. There exists a one to one corresponding between the set of equivalence relations on $A$ and the set partitions of $A$.*

Next, we provide examples of relations defined on the set of equivalence classes. If $A$ is a nonempty set, $\mathcal{R}$ is an equivalence relation on $A$, and $\mathcal{C} = \{[a] : a \in A\}$ is the set of equivalence classes of $\mathcal{R}$, then a relation $\mathcal{R}'$ can be defined from $\mathcal{C}$ to a given set $B$, i.e.,

$$\mathcal{R}' \subseteq \{([a], b) \in \mathcal{C} \times B : a \in A, b \in B\}$$

A relation on the set of equivalence classes $\mathcal{C}$ is in the form

$$\mathcal{R}' \subseteq \{([a], [b]) \in \mathcal{C} \times \mathcal{C} : a, b \in A\}$$

***Example 1.4.15***

1. Consider the relation in Example 1.4.11 and its equivalence classes $\mathcal{C} = \{[0], [1], [2]\}$. The set

$$\mathcal{R}' = \{([0], 1), ([5], 2), ([12], 3)\} \text{ and } \mathcal{R}'' = \{([0], 1), ([5], 2), ([4], 3)\}$$

are relations from $\mathcal{C}$ to the set $\{1, 2, 3\}$. The sets

$$\mathcal{S}' = \{([0], [1]), ([5], [-4]), ([12], [-7])\},$$
$$\mathcal{S}'' = \{([0], [1]), ([5], [-4]), ([12], [-8])\}$$

are relations on $\mathcal{C}$.

2. Consider the equivalence relation in Example 1.4.12 and its equivalence classes $\mathbb{Q} = \{[(m, n)] : m, n \in \mathbb{Z} \wedge n \neq 0\}$. The set $\mathcal{J} = \{([(m, n)], m) : m, n \in \mathbb{Z} \wedge n \neq 0\}$ forms a relation from $\mathbb{Q}$ to $\mathbb{Z}$. The set $\mathcal{T} = \mathbb{Q}^* \times \mathbb{Q}^* = \{([(m, n)], [(r, s)]) : m, n, r, s \in \mathbb{Z}^*\}$ is a relation on $\mathbb{Q}^*$.

## 1.5 Functions

In this section, a specific type of relations, called functions, is examined. The importance of studying functions springs from their presence and role in almost every branch of mathematics. The symbol $f$, the first letter of function, is used instead of $\mathcal{R}$ to denote a relation that is a function.

**Definition 1.5.1** Let $A$ and $B$ be any sets and $f \subseteq A \times B$ be a relation from $A$ to $B$. The relation $f$ is said to be a function on $A$ if for each element $a \in A$, there exists a unique element $b \in B$ such that $(a, b) \in f$, i.e., the conditional statement

$$a \in A \Rightarrow \exists\,!\, b \in B \text{ such that } (a, b) \in f$$

is true for all $a \in A$. The notation $\exists\,!$ indicates uniqueness.

The notation $f : A \to B$ denotes that $f$ is a function from $A$ to $B$, and the notation $b = f(a)$ is used instead of $(a, b) \in f$. Note that the uniqueness requirement in this definition is equivalent to the following statement:

$$a_1 = a_2 \Rightarrow f(a_1) = f(a_2) \text{ for all } a_1, a_2 \in A.$$

***Remark 1.5.2*** If $\mathcal{R}$ is an equivalence relation on $A$, and $f : \{[a] : a \in A\} \to B$ is a function on the set of all equivalence classes of $\mathcal{R}$, then uniqueness requirement means that the definition of $f$ does not depend on the representatives used for the equivalence class of $a$, i.e., if $a$ and $b$ are elements in $A$ such that $[a] = [b]$, then their images under $f$ must be equal. This is called well-defined property of $f$, i.e., a function $f$ from $\{[a] : a \in A\}$ to $B$ must satisfy that

$$a_1 \mathcal{R} a_2 \Rightarrow f([a_1]) = f([a_2]) \quad \forall a_1, a_2 \in A$$

In general, if the domain of $f$ involves equivalence classes (e.g., $f : \mathcal{A} \to B$, where $\mathcal{A}$ is a set defined using equivalence classes), then the well-defined property must be verified. For more explanation, see the Example 1.5.6 (4–9).

**Definition 1.5.3** Let $A$ and $B$ be any sets, $f : \mathcal{A} \to B$ be a function from $A$ to $B$, and $A_1$ be any subset of $A$. The restriction of $f$ on $A_1$, denoted by $f_{|A_1}$, is the map from $A_1$ to $B$ defined by $f_{|A_1}(a) = f(a)$ for all $a \in A_1$.

As for any other relation on sets (Definition 1.3.3), the following definition holds:

**Definition 1.5.4** Let $A$ and $B$ be any two sets and $f : A \to B$ be a function.

- The set $A$ is called the domain of $f$, denoted by $D(f)$, and $B$ is called the codomain of $f$.
- The element $b \in B$ such that $b = f(a)$ is called the image of $a$ under $f$ or the value of $a$.
- The set of all images of $A$ is $f(A) = \{f(a) : a \in A\}$, which is called the range of $f$, denoted by $Rang(f)$.
- If $b \in B$, then the preimage (inverse image) of $b$ under $f$ is the set

$$f^{-1}(b) = \{a \in A : b = f(a)\}.$$

- If $Y \subseteq B$, then the preimage (inverse image) of $Y$ under $f$ is the set

$$f^{-1}(Y) = \{a \in A : f(a) \in Y\}.$$

***Remark 1.5.5***

1. To calculate $f^{-1}(Y)$, it is easier to
   - calculate $f^{-1}(\{b\}) = \{x : f(x) = b\}$ for all $b \in Y$, then
   - use the equality $f^{-1}(Y) = \bigcup\limits_{b \in Y} f^{-1}(\{b\})$.
2. If a function is defined on a subset of $\mathbb{R}$ as an algebraic expression $f(x)$ in a variable $x$, the domain of $f$ is taken to be all possible values $x$, where the expression is valid. For example, the domain of $f(x) = \frac{x^2+1}{x^2-3x+2}$ is $\mathbb{R}\backslash\{1, 2\}$.

***Example 1.5.6*** Let $A = \{a, b, c, d\}$ and $B = \{1, 2, 3, 4\}$.

1. The relation $f = \{(a, 1), (b, 1), (c, 1), (d, 2)\}$ from $A$ to $B$ is a function whose domain is $A$, and the image of each element in $A$ is given as

$$f(a) = 1,\ f(b) = 1,\ f(c) = 1,\ \text{and } f(d) = 2.$$

2. The relation $g = \{(a, 1), (b, 1), (a, 2), (c, 3), (d, 1)\}$ from $A$ to $B$ is not a function, as both $(a, 1)$ and $(a, 2)$ are in $g$ but $1 \neq 2$.
3. The relation $h = \{(a, 2), (b, 3), (c, 4)\}$ from $A$ to $B$ is not a function, as it is not defined for every element in $A$.
4. The relation $\mathcal{R}'$ in Example 1.4.15 is not a function. It is not defined for every element in $\mathcal{C}$ ($\{[0], [5], [12]\} = \{[0], [2]\} \neq \mathcal{C}$), and it is not well-defined since $[0] = [12]$, but $1 \neq 3$.
5. The relation $\mathcal{R}''$ in Example 1.4.15 represents a function on $\mathcal{C}$; it is defined for every element in $\mathcal{C}$ ($\{[0], [5], [4]\} = \{[0], [1], [2]\}$). Since $[0] \neq [5]$, $[0] \neq [4]$ and $[4] \neq [5]$, there are no equal equivalence classes in the domain of $\mathcal{R}''$.
6. The relation $\mathcal{S}'$ in Example 1.4.15 is not a function. It is not defined for every element in $\mathcal{C}$, and it is not well-defined since $[0] = [12]$ but $[1] \neq [-7]$.
7. The relation $\mathcal{S}''$ in Example 1.4.15 is not a function. Note that [0] and [12] are the only equal equivalence classes in the domain of $\mathcal{S}''$, and as their images under $\mathcal{S}''$ ([1] and [−8], respectively) are equal, then $\mathcal{S}''$ is well-defined. However, it is not a function on $\mathcal{C}$ as it is not defined for every element in $\mathcal{C}$.
8. The relation $\mathcal{J}$ in Example 1.4.15 is defined for every element in $\mathbb{Q}$, but it is not well-defined since $[(2, 4)] = [(1, 2)]$, but their images under $\mathcal{J}$ are not equal.
9. The relation $\mathcal{T}$ in Example 1.4.15 is defined for every element in $\mathbb{Q}^*$, but it is not well-defined since every element in $\mathbb{Q}^*$ is an image of all other elements in $\mathbb{Q}^*$. For example, $[(1, 2)]$ has infinitely many different images under $\mathcal{T}$.

***Example 1.5.7*** Let $f : \{1, 2, 3, 4, 5, 6\} \to \{a, b, c, d\}$ be the function

$$f = \{(1, a), (2, b), (3, a), (4, b), (5, c), (6, d)\}$$

The images of the elements in $\{1, 2, 3, 4, 5, 6\}$ are given by

$$f(1) = a,\ f(2) = b,\ f(3) = a,\ f(4) = b,\ f(5) = c,\ f(6) = d$$

Therefore, $Rang(f) = \{a, b, c, d\}$. If $X = \{1, 3, 6\}$, then the image of $X$ is

$$f(X) = \{f(1), f(3), f(6)\} = \{a, d\}.$$

If $Y = \{a, b\}$, the preimage (inverse image) of $Y$ is

$$f^{-1}(Y) = f^{-1}(\{a\}) \cup f^{-1}(\{b\}) = \{1, 3\} \cup \{2, 4\} = \{1, 2, 3, 4\}.$$

***Example 1.5.8***

1. Consider the identity relation on a nonempty set $A$, defined by $\Delta_A = \{(a, a) : a \in A\}$. Clearly, for each $a \in A$ there exists a unique element $a$ in $A$ such that $\Delta_A(a) = a$. Therefore, $\Delta_A$ is a function on $A$. This function is called the identity function on $A$, and usually written as $I : A \to A$ such that $I(a) = a$.
2. Let $A$ be any set and $B \subseteq A$. The inclusion map of $B$, denoted by $\iota_B : B \to A$ and defined by $\iota_B(a) = a$ for all $a \in B$, is a function from $B$ to $A$. In fact, the map $\iota_B$ is the restriction of the identity function $\Delta_A$ on the subset $B$.
3. The relations

$$f = \left\{(x, y) \in \mathbb{R} \times \mathbb{R} : y = x^2\right\} \text{ and } h = \left\{(x, y) \in \mathbb{R} \times \mathbb{R} : y = x^2 - 5\right\}$$

   are functions on $\mathbb{R}$ (Check!). Geometrically, the function $f$ is the parabola $x^2$, and $h$ is the same parabola shifted 5 units downwards.
4. The relation $f = \left\{(x, y) \in \mathbb{R} \times \mathbb{R} : x^2 + y^2 = 1\right\}$ is not a function, as both $(0, 1)$ and $(0, -1)$ belong to $f$, and $1 \neq -1$. The equation $x^2 + y^2 = 1$ implies that $y = \pm\sqrt{1 - x^2}$. This means that every element $x$ such that $x \notin \{-1, 1\}$ is related to two elements in the codomain, which violates the uniqueness condition.
5. The relation $\left\{(x, y) \in \mathbb{R} \times [0, \infty) : x^2 + y^2 = 1\right\}$ retains the uniqueness condition in the definition of functions, but it is not defined for every element of $\mathbb{R}$. For example, $(3, y)$ does not belong to the relation for any $y$ in $(0, \infty)$, and thus, the relation is not a function.
6. The relation $\left\{(x, y) \in [-1, 1] \times [0, \infty) : x^2 + y^2 = 1\right\}$ defines a unique element $\sqrt{1 - x^2}$ in $[0, \infty)$ for every element in $[-1, 1]$. Therefore, this relation is a function that represents the top semicircle of the unit circle. Such a relation can be easily expressed as $f(x) = \sqrt{1 - x^2}$, where $-1 \leq x \leq 1$.
7. The relation $\{(x, y) \in \mathbb{R} \times \mathbb{R} : y = x + 1\}$ defines a unique element $x + 1$ in $\mathbb{R}$ for every element $x \in \mathbb{R}$. Therefore, this relation is a function on $\mathbb{R}$ that is written as $f(x) = x + 1$.

***Example 1.5.9*** Let $\mathbb{Q} = \{m/n : m, n \in \mathbb{Z} \wedge n \neq 0\}$ (Example 1.4.12). Define

$$\begin{array}{ll} f : \mathbb{Q} \times \mathbb{Q} \longrightarrow \mathbb{Q} & \text{and } g : \mathbb{Q} \times \mathbb{Q} \longrightarrow \mathbb{Q} \quad \text{where} \\ f(a/b,\ c/d\ ) = \frac{ad+cb}{bd} & g(a/b,\ c/d\ ) = \frac{ac}{bd} \end{array}$$

Both $f$ and $g$ are functions on $\mathbb{Q} \times \mathbb{Q}$. To show this, let $(a/b, c/d)$ be an arbitrary element in $\mathbb{Q} \times \mathbb{Q}$. Since both $b \neq 0$ and $d \neq 0$, then $bd \neq 0$. Therefore, both $f(a/b, c/d)$ and $g(a/b, c/d)$ belong to $\mathbb{Q}$. To verify the uniqueness requirement, assume that $(a/b, c/d)$ and $\left(a'/b', c'/d'\right)$ are two elements in $\mathbb{Q} \times \mathbb{Q}$ such that $(a/b, c/d) = \left(a'/b', c'/d'\right)$. It is necessary to show that $f(a/b, c/d) = f\left(a'/b', c'/d'\right)$ and $g(a/b, c/d) = g\left(a'/b', c'/d'\right)$. As $a/b = a'/b'$ and $c/d = c'/d'$, then $ab' = a'b$ and $cd' = c'd$ (Example 1.4.12). Thus,

$$\begin{aligned} (ad + cb)\left(b'd'\right) &= adb'd' + cbb'd' \\ &= \left(ab'\right)\left(dd'\right) + \left(cd'\right)\left(bb'\right) \\ &= \left(a'b\right)\left(dd'\right) + \left(c'd\right)\left(bb'\right) \\ &= \left(a'd'\right)(bd) + \left(c'b'\right)(bd) \\ &= (a'd' + c'b')(bd) \end{aligned}$$

According to the definition of the equivalence relation (Example 1.4.12),

$$\frac{ad + cb}{bd} = \frac{a'd' + c'b'}{bd}$$

Similarly,

$$ac\left(b'd'\right) = ab'\left(cd'\right) = a'b\left(c'd\right) = a'c'(bd)$$

which implies that

$$\frac{ac}{bd} = \frac{a'c'}{b'd'}$$

as required.

Functions can be expressed as equations, tables, or graphs. The functions in Example 1.5.8(3) and (6) are expressed as equations. If the domain of a function is a finite set, the function can be presented in tabular form. In this representation, the inputs and the outputs are listed in different columns. For example, the function in Example 1.5.6 (1) can be represented as in Table 1.1.

If $f : A \times B \to C$ is a function in which $A$ and $B$ are both finite, then the outputs of $f$ can be presented in a tabular form. If $A = \{a_1, a_2, \ldots, a_n\}$, $B = \{b_1, b_2, \ldots, b_m\}$, and $f : A \times B \to C$ is a function, then the outputs of $f$ can be listed as in Table 1.2.

**Table 1.1** Representation of the function in Example 1.5.6 (1)

| $x$ | $f(x)$ | $(x, f(x))$ |
|---|---|---|
| $a$ | 1 | $(a, 1)$ |
| $b$ | 1 | $(b, 1)$ |
| $c$ | 1 | $(c, 1)$ |
| $d$ | 2 | $(d, 2)$ |

**Table 1.2** Tabular representation of function on finite domain and codomain

| $f$ | $a_1$ | $a_2$ | $\cdots$ | $a_n$ |
|---|---|---|---|---|
| $b_1$ | $f(a_1, b_1)$ | $f(a_2, b_1)$ | $\cdots$ | $f(a_n, b_1)$ |
| $b_2$ | $f(a_1, b_2)$ | $f(a_2, b_2)$ | $\cdots$ | $f(a_n, b_2)$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $b_m$ | $f(a_1, b_m)$ | $f(a_2, b_2)$ | $\cdots$ | $f(a_n, b_m)$ |

If the domain is a subset of the real numbers $\mathbb{R}$, or a subset of the plane $\mathbb{R}^2$, then $f$ can be graphically visualized. The graph of a function $f$ from $A$ to $B$ is the set $\{(x, f(x)) : x \in A\}$. To learn more about graphing functions, the reader can refer to (Hungerford & Shaw, 2009).

**Definition 1.5.10** (Piecewise defined function) A function defined by multiple expressions is called a piecewise defined function, or simply a piecewise function. Each expression is applied to a certain part of the domain.

The following are examples of piecewise functions:

$$f(x) = \begin{cases} 2 & x \geq 7 \\ -x^2 + 1 & x < 7 \end{cases} \qquad g(x) = \begin{cases} x & x > 0 \\ 5 & x = 0 \\ -1 & x < 0 \end{cases} \qquad h(x) = |x| = \begin{cases} x & x \geq 0 \\ -x & x < 0 \end{cases}$$

**Definition 1.5.11** (*Composition of functions*) Let $A$, $B$, $H$ and $K$ be any sets. Let $f : A \to B$ and $g : H \to K$ be two functions such that $Rang(f) \subseteq D(g)$. The composition of $f$ and $g$ is the function.

$$g \circ f : A \longrightarrow K, \text{ where } g \circ f(x) = g(f(x)).$$

It is left to the reader (Exercise 1.21) to verify that a composition of two functions is a function.

***Example 1.5.12*** Let $f : \mathbb{R} \to \mathbb{R}$, where $f(x) = x - 1$, and let $g : \mathbb{R} \to \mathbb{R}$, where $g(x) = x^2$. The corresponding compositions are.

$$g \circ f(x) = g(f(x)) \underset{g(z)=z^2}{=} (f(x))^2 \underset{f(x)=x-1}{=} (x-1)^2 = x^2 - 2x + 1.$$

and

$$f \circ g(x) = f(g(x)) \underset{f(z)=z-1}{=} g(x) - 1 = x^2 - 1.$$

Clearly, $g \circ f$ and $f \circ g$. are not equal.

Note that for $g \circ f$ to be defined, the range of $f$ must be a subset of the domain of $g$. Otherwise, the composition cannot be defined. For example, if $f : \mathbb{R} \to \mathbb{R}$ is defined by $f(x) = x - 1$ and $g : (\mathbb{R}^+ \cup \{0\}) \to \mathbb{R}$ is defined by $g(x) = \sqrt{x}$, then the composition $g \circ f$ cannot be donfined at $x = 0$ (Verify!).

The following proposition can be obtained by applying the composition of two functions twice.

**Proposition 1.5.13** *Let $A$, $B$ and $C$ be any three sets. If $f : A \to B$, $g : B \to C$, and $h : C \to D$ are three functions such that $Rang(f) \subseteq D(g)$ and $Rang(g) \subseteq D(h)$, then*

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

**Definition 1.5.14** (Injective and surjective functions) Let $A$ and $B$ be two sets. Let $f : A \to B$ be a function. The map $f$ is called

1. Injective (one-to-one) if $x \neq y \Rightarrow f(x) \neq f(y)$ for all $x, y \in A$. Equivalently,

$$f(x) = f(y) \Rightarrow x = y \text{ for all } x, y \in A.$$

2. Surjective (onto) if for each $y \in B$, there exists an $x \in A$ such that $f(x) = y$.
3. Bijective if $f$ is both injective and surjective.

***Example 1.5.15***

1. The map $f : \mathbb{R} \to \mathbb{R}$, where $f(x) = x$, is an injective and a surjective map, so it is bijective.
2. The map $g : \mathbb{R} \to \mathbb{R}$, where $g(x) = x^2$, is neither injective nor surjective. However, we can obtain injective or surjective functions from $g$ by restricting the domain or codomain. For example,
   - The map $g_1 : \mathbb{R} \to \mathbb{R}^+$, where $g_1(x) = x^2$, is surjective but not injective.
   - The map $g_2 : \mathbb{R}^+ \to \mathbb{R}$, where $g_2(x) = x^2$, is injective but not surjective.
   - The map $g_3 : \mathbb{R}^+ \to \mathbb{R}^+$, where $g_3(x) = x^2$, is both injective and surjective. Therefore, this map is bijective.

For the proof of the following lemma, see Exercise 1.6.

**Lemma 1.5.16** *Let $A$ and $B$ be two finite sets such that $|A| = |B|$, and let $f : A \to B$ be a function. The map $f$ is injective if and only if it is surjective.*

***Example 1.5.17*** For $n \in \mathbb{N}$, let $s, t \in \{1, 2, \ldots, n\}$ and $\mathcal{R}_{s,t} : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ be defined as

$$\mathcal{R}_{s,t}(k) = \begin{cases} k & k \neq s \wedge k \neq t \\ t & k = s \\ s & k = t \end{cases}$$

Clearly, the map $\mathcal{R}_{\mathrm{s,t}}$ is defined for any element in $\{1, 2, \ldots, n\}$, and any element in the domain has a unique image under $\mathcal{R}_{\mathrm{s,t}}$. Hence, it is a function on $\{1, 2, \ldots, n\}$. To show that $\mathcal{R}_{\mathrm{s,t}}$ is an injective map, let $k, k'$ be elements in $\{1, 2, \ldots, n\}$ such that $\mathcal{R}_{s,t}(k) = \mathcal{R}_{s,t}(k')$.

- If $k \neq s \wedge k \neq t$, then $k = \mathcal{R}_{s,t}(k) = \mathcal{R}_{s,t}(k')$, which implies that $k = k'$.
- If $k = s$, then $t = \mathcal{R}_{s,t}(k) = \mathcal{R}_{s,t}(k')$, which implies that $k' = s = k$.
- If $k = t$, then $s = \mathcal{R}_{s,t}(k) = \mathcal{R}_{s,t}(k')$, which implies that $k' = t = k$.

In all cases $k' = k$. Thus, $\mathcal{R}_{s,t}$ is injective. By Lemma 1.5.16, the map $\mathcal{R}_{s,t}$ is also surjective.

The map $\mathcal{R}_{s,t}$ permutes the two elements $s$ and $t$.Consequently, it is known as a "transposition". The following proposition is proved in Chap 6.

**Proposition 1.5.18** *Any bijective map on* $\{1, 2, \ldots, n\}$ *is a composition of transpositions.*

**Definition 1.5.19** (*Invertible function*) Let $A$ and $B$ be two sets. A function $f : A \to B$ is said to be invertible if there exists a function $g : B \to A$ such that $g \circ f = \iota_A$ and $f \circ g = \iota_B$. The map $g$ is called the inverse of $f$ and is denoted by $f^{-1}$.

The conditions $g \circ f = \iota_A$ and $f \circ g = \iota_B$ means $(a, b) \in f \Leftrightarrow (b, a) \in g$. This relation can be intuitively expressed as follows: if $f$ connects $a$ to $b$, then $g$ returns $b$ to its preimage $a$, and vice versa. There are many examples of noninvertible functions. For example, the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$ is not invertible. If $f$ was invertible, then there would exist $g : \mathbb{R} \to \mathbb{R}$ such that

$$f \circ g(x) = (g(x))^2 = x$$

which implies that $g(x) = \pm\sqrt{x}$, i.e., $g$ takes either the value of $\sqrt{x}$ or $-\sqrt{x}$, and thus, is not a function. In addition, $g$ is not defined for the negative real numbers.

**Theorem 1.5.20** Let $f : A \to B$ be a function. The map $f$ is invertible if and only if it is a bijective map.

According to the Theorem above, to show that a map $f : A \to B$ is bijective, it suffices to show that there exists a map $g : B \to A$ such that $g \circ f = \iota_A$ and $f \circ g = \iota_B$ (Exercise 1.8).

***Example 1.5.21***

1. Let $f : \mathbb{R} \to \mathbb{R}$ be defined as $f(x) = 2x + 1$, then $f$ is a bijective map and is thus invertible. To find its inverse, let $y = f(x)$, and solve for $x$ as follows: $y = 2x + 1 \Leftrightarrow x = (y - 1)/2$. Interchanging $x$ and $y$ yields $y = (x - 1)/2$. Hence, the inverse of $f(x)$ is the map $g(x) = (x - 1)/2$. It can be verified that $f \circ g(x) = x = g \circ f(x)$. Therefore, $g$ is the inverse of $f$.
2. Let $f : \mathbb{R} \to [-1, 1]$ be defined as $f(x) = \cos x$. Since $\cos\left(\frac{\pi}{4}\right) = \cos\left(\frac{-\pi}{4}\right)$, then $f$ is not one to one, hence it is not invertible. However, if the domain of $\cos x$ is restricted to be $0 \le x \le \pi$, then $\cos x$ will be invertible, with the inverse $g(x) = \cos^{-1} x$. The domain of $g$ is $[-1, 1]$.

## 1.6 Matrices

This section focuses on matrices, a type of maps that can be presented using arrays. Many functions on matrices, such as matrix addition and matrix multiplication are defined and briefly discussed. For additional information and proofs, the reader can refer to (Burton, 2007) and (Hartman, 2011).

**Definition 1.6.1** Let $m, n \in \mathbb{N}$, and let $K$ be any set. A $K$-matrix of type $m \times n$ (read $m$ by $n$) is a map

$$A : \{1, 2, \dots, m\} \times \{1, 2, \dots, n\} \to K$$

that assigns an element $a_{ij}$ in $K$ for each $(i, j) \in \{1, 2, \dots, m\} \times \{1, 2, \dots, n\}$.

Such a map can be presented as a rectangular array with $m$ rows and $n$ columns in the following form:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

where $A((i, j)) = a_{ij} \in K$ for all $1 \le i \le m$ and $1 \le j \le n$.

If there is no risk of ambiguity, a $K$-matrix is simply referred to as a matrix. If a formula for the elements $a_{ij}$ is given, the matrix can be written as $\left(a_{ij}\right)_{\substack{1 \le i \le m \\ 1 \le j \le n}}$ or simply $\left(a_{ij}\right)$. Here, the elements $a_{ij}$ are called the entries of the matrix $A$, and the integers $m$ and $n$ are called the dimensions of the matrix $A$. A matrix that consists of $m$ rows and $n$ columns is called an $m \times n$ matrix. The set of all $m \times n$ matrices with entries in $K$ is denoted by $\mathcal{M}_{mn}(K)$. The set $\mathcal{M}_{nn}(K)$ can be abbreviated to $\mathcal{M}_n(K)$. A matrix in $\mathcal{M}_n(K)$ is called a square matrix of dimension $n$. The entries $a_{ii}$ in a square matrix are called diagonal entries.

***Example 1.6.2***

1. The set $\mathcal{M}_{3\times 2}(\mathbb{Z})$ consists of all $3 \times 2$ matrices with integers entries.
2. The set $\mathcal{M}_{5\times 1}(\mathbb{R})$ consists of all $5\times 1$ matrices with real entries, and such matrices have the form

$$\begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \\ a_{41} \\ a_{51} \end{pmatrix}$$

where $a_{i1}$ is an element in $\mathbb{R}$. This matrix can be considered an element in $\mathbb{R}^5$. In fact, there exists a bijection map between $\mathcal{M}_{5\times 1}(\mathbb{R})$ and $\mathbb{R}^5$. In general, an $n \times 1$ matrix consists of one column in the form

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{(n-1)1} \\ a_{n1} \end{pmatrix}$$

This matrix can be identified with an element in $\mathbb{R}^n$.
3. A $1 \times n$ matrix consisting of one row can be expressed as

$$\begin{pmatrix} a_{11} \ a_{12} \ \cdots \ a_{1(n-1)} \ a_{1n} \end{pmatrix}$$

and identified with an element in $\mathbb{R}^n$.
4. The set $\mathcal{M}_7(\mathbb{C})$ consists of all square matrices of dimension 7 with complex entries.
5. Let $X = \{a, b, c\}$. The set $\mathcal{M}_{3\times 2}(\mathcal{P}(X))$ consists of all $3 \times 2$ matrices whose entries are subsets of $X$. An example of a matrix in $\mathcal{M}_{3\times 2}(\mathcal{P}(X))$ is

$$\begin{pmatrix} \{a, b\} & \{b, c\} \\ \emptyset & \{b\} \\ \{a, c\} & \{a\} \end{pmatrix}$$

**Definition 1.6.3** Let $n \in \mathbb{N}$, and let $K \neq \emptyset$ be a subset of $\mathbb{C}$ such that $K$ contains 0 and 1. Let $\mathcal{M}_n(K)$ be the set of square matrices of dimension $n$ over $K$.

1. A matrix $\left(a_{ij}\right) \in \mathcal{M}_n(K)$ with entries $a_{ij} = 0$ whenever $i \neq j$ can be expressed as

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$$

This matrix is called a diagonal matrix. The set of all diagonal matrices in $\mathcal{M}_n(K)$ is denoted by $D(K)$. A diagonal matrix with all diagonal entries are equal to 1 is called a unit matrix, denoted by $I_n$.

$$I_n = \left(a_{ij}\right) \text{ where } a_{ij} = \begin{cases} 1 \, if \; i = j \\ 0 \, if \; i \neq j \end{cases}$$

A notable example of diagonal matrices is the subset $\{A \in \mathcal{M}_n(\mathbb{C}) : A = \lambda I_n\}$ in which all the diagonal elements are equal.

2. A matrix $\left(a_{ij}\right) \in \mathcal{M}_n(K)$ with entries $a_{ij} = 0$ whenever $i > j$ can be expressed as

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & a_{24} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & a_{34} & \cdots & a_{3n} \\ \vdots & \vdots & 0 & a_{44} & \cdots & a_{4n} \\ 0 & 0 & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

This matrix is called an upper triangular matrix. The set of all upper triangle matrices over $K$ is denoted by $U(K)$. i.e.,

$$U(K) = \left\{\left(a_{ij}\right) \in \mathcal{M}_n(K) : a_{ij} = 0 \quad \forall \; i > j\right\}$$

3. A matrix $\left(a_{ij}\right) \in \mathcal{M}_n(K)$ with all its entries $a_{ij} = 0$ whenever $i < j$ is in the form

$$\begin{pmatrix} a_{11} & 0 & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & 0 & 0 & \cdots & 0 \\ a_{31} & a_{32} & a_{33} & 0 & \cdots & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & 0 \\ a_{n1} & a_{n2} & a_{n3} & a_{n4} & \cdots & a_{nn} \end{pmatrix}.$$

This matrix is called a lower triangular matrix. The set of all lower triangle matrices over $K$ is denoted by $L(K)$. i.e.,

$$L(K) = \{(a_{ij}) \in \mathcal{M}_n(K) : a_{ij} = 0 \quad \forall\ i < j\}.$$

In the following, we define several algebraic operations on matrices. For the remainder of this chapter, the set $K$ denotes a nonempty subset of the complex numbers $\mathbb{C}$ such that.

1. $K$ contains 0 and 1.
2. $K$ is closed under the usual addition on $\mathbb{C}$, the usual multiplication on $\mathbb{C}$, and the conjugation.

**Definition 1.6.4** Let $m, n \in \mathbb{N}$. Let $\mathcal{M}_{mn}(K)$ be the set of $m \times n$ matrices with $K$-entries. Let $A = (a_{ij})$ and $B = (b_{ij})$ be two matrices in $\mathcal{M}_{mn}(K)$.

1. For any $c \in K$, the multiplication of the matrix $A$ by any constant $c$ is the matrix $cA$ obtained from $A$ by multiplying each entry by $c$, i.e., $cA = (ca_{ij})$.
2. The addition of $A$ and $B$ is the matrix $A + B$ obtained from $A$ and $B$ by adding the entries of $A$ and $B$ that have the same indices $i$ and $j$. That is,

$$A + B = (a_{ij} + b_{ij})$$

where $a_{ij} + b_{ij}$ is the usual sum of the complex numbers $a_{ij}$ and $b_{ij}$.

Note that matrices with two different dimensions cannot be added. For example, a $2 \times 3$ matrix and a $4 \times 7$ matrix cannot be added. Therefore, matrix addition is defined only for matrices with the same dimensions.

***Example 1.6.5***

1. The matrices

$$A = \begin{pmatrix} 2 & -3 & 7 \\ 3 & 2 & 4 \end{pmatrix},\ B = \begin{pmatrix} -3 & 5 & -5 \\ -1 & 2 & -2 \end{pmatrix}$$

are elements in $\mathcal{M}_{2\times 3}(\mathbb{Z})$. Multiplying these matrices by 2 and 0 respectively, yields

$$2A = 2\begin{pmatrix} 2 & -3 & 7 \\ 3 & 2 & 4 \end{pmatrix} = \begin{pmatrix} 4 & -6 & 14 \\ 6 & 4 & 8 \end{pmatrix},\ 0B = 0\begin{pmatrix} -3 & 5 & -5 \\ -1 & 2 & -2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The sum $A + B$ is $\begin{pmatrix} -1 & 2 & 2 \\ 2 & 4 & 2 \end{pmatrix}$.

2. The matrices

$$E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},\ E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},\ E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

are elements in $\mathcal{M}_{2\times 2}(\mathbb{C})$. Any matrix $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ in $\mathcal{M}_{2\times 2}(\mathbb{C})$ can be expressed using such matrices, as follows:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}E_{11} + a_{12}E_{12} + a_{21}E_{21} + a_{22}E_{22} = \sum_{i=1}^{2}\sum_{j=1}^{2} a_{ij}E_{ij}.$$

3. In general, a matrix $(a_{ij})$ in $\mathcal{M}_{mn}(\mathbb{C})$ can be expressed as

$$\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}E_{ij}$$

where $E_{ij} = (a_{kl})$ in $\mathcal{M}_{m\times n}(\mathbb{C})$ with $a_{kl} = \begin{cases} 1 \; if \; k = i \wedge l = j \\ 0 \quad \text{elsewhere} \end{cases}$

**Definition 1.6.6** Let $m, n, l \in \mathbb{N}$. Let $A = (a_{ij}) \in \mathcal{M}_{mn}(K)$ and $B = (b_{ij}) \in \mathcal{M}_{nl}(K)$. The product of $A$ and $B$ is the matrix $AB \in \mathcal{M}_{ml}(K)$, defined as $(c_{ij})$, where

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}.$$

In this definition, the entry $c_{ij}$ is formed using the $i$-th row in $A$ and the $j$-th column in $B$ entries. i.e.,

$$c_{ij} = \begin{pmatrix} a_{i1} & a_{i2} & \ldots & a_{in} \end{pmatrix} \cdot \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{rj} \end{pmatrix} = \sum_{k=1}^{n} a_{ik}b_{kj}.$$

Note that matrix multiplication can be performed if and only if the number of columns in the left matrix equals the number of rows in the right matrix. The multiplication of square matrices is only defined if the two matrices have the same dimension. If $A$ is an $m \times n$ matrix and $B$ is an $n \times l$ matrix, then their product $A \cdot B$ is an $m \times l$ matrix. We will also use $AB$ to denote the multiplication $A \cdot B$ of any two matrices $A$ and $B$.

***Example 1.6.7***

1. Let $A = \begin{pmatrix} 1 & -3 \\ 4 & 1 \\ 0 & 7 \end{pmatrix} \in \mathcal{M}_{3\times 2}(\mathbb{Z})$ and $B = \begin{pmatrix} 2 & 6 & 7 \\ -3 & 2 & 4 \end{pmatrix} \in \mathcal{M}_{2\times 3}(\mathbb{Z})$.

$$AB = \begin{pmatrix} 11 & 0 & -5 \\ 5 & 26 & 32 \\ -21 & 14 & 28 \end{pmatrix} \in \mathcal{M}_{3\times 3}(\mathbb{Z}) \text{ and } BA = \begin{pmatrix} 26 & 49 \\ 5 & 39 \end{pmatrix} \in \mathcal{M}_{2\times 2}(\mathbb{Z}).$$

2. Let $A = \begin{pmatrix} 1 & 5 \\ 6 & -1 \end{pmatrix} \in \mathcal{M}_{2\times 2}(\mathbb{Z})$ and $B = \begin{pmatrix} 6 & 1 & 0 \\ -2 & 3 & 3 \end{pmatrix} \in \mathcal{M}_{2\times 3}(\mathbb{Z})$.

   $AB = \begin{pmatrix} -4 & 16 & 15 \\ 38 & 3 & -3 \end{pmatrix} \in \mathcal{M}_{2\times 3}(\mathbb{Z})$ and $BA$ is not possible.

3. Let $E_{ij} \in \mathcal{M}_{mn}(\mathbb{C})$ and $E_{ks} \in \mathcal{M}_{nl}(\mathbb{C})$ be defined as in Example 1.6.5 (3). It is straightforward to show that $E_{ij}E_{ks} = \delta_{jk}E_{is}$, where $\delta_{jk} = \begin{cases} 1 & j = k \\ 0 & j \neq k \end{cases}$.

**Proposition 1.6.8** *Let $m, n \in \mathbb{N}$. The matrix addition $+ : \mathcal{M}_{mn}(K) \times \mathcal{M}_{mn}(K) \to \mathcal{M}_{mn}(K)$ and the matrix multiplication $\cdot : \mathcal{M}_n(K) \times \mathcal{M}_n(K) \to \mathcal{M}_n(K)$ are both functions.*

***Proof*** Let $(A, B) \in \mathcal{M}_{mn}(K) \times \mathcal{M}_{mn}(K)$ be an arbitrary element. As all elements in $\mathcal{M}_{mn}(K)$ are of the same type, then the addition of $A$ and $B$ is defined. To verify the uniqueness requirement, assume that $(A, B)$ and $(A', B')$ are elements in $\mathcal{M}_{mn}(K) \times \mathcal{M}_{mn}(K)$ such that $(A, B) = (A', B')$. According to the equality of the order pairs, $A = A'$ and $B = B'$, which implies that.

$$+(A, B) = A + B = A' + B' = +(A', B')$$

Thus, the matrix addition identifies a unique element in $\mathcal{M}_{mn}(K)$ for each pair of matrices in $\mathcal{M}_{mn}(K) \times \mathcal{M}_{mn}(K)$.

For the matrix multiplication, if $(A, B) \in \mathcal{M}_n(K) \times \mathcal{M}_n(K)$, then $A$ and $B$ are square matrices of the same dimension. The multiplication of $A$ and $B$ is defined and yields a square matrix of dimension $n$. To verify the uniqueness requirement, assume that $(A, B)$ and $(A', B')$ are elements in $\mathcal{M}_n(K) \times \mathcal{M}_n(K)$ such that $(A, B) = (A', B')$. According to the equality of the order pairs, $A = A'$ and $B = B'$, which implies that $a_{ik} = a'_{ik}$ and $b_{kj} = b'_{kj}$ for all $1 \leq i, j \leq n$, i.e.,

$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = \sum_{k=1}^{n} a'_{ik}b'_{kj}$ is the $ij$-entry in the multiplication $A'B'$.

Thus,

$$\cdot(A, B) = AB = A'B' = \cdot(A', B')$$

Therefore, the matrix multiplication identifies a unique element in $\mathcal{M}_n(K)$ for each pair of matrices in $\mathcal{M}_n(K) \times \mathcal{M}_n(K)$. ■

**Proposition 1.6.9** *Let $n \in \mathbb{N}$ and $\mathcal{M}_n(K)$ be the set of all square matrices of dimension $n$.*

1. *The multiplication of two diagonal matrices is a diagonal matrix. If $A$ and $B$ are diagonal matrices, then $AB = BA$.*
2. *The multiplication of two upper matrices is an upper matrix.*
3. *The multiplication of two lower matrices is a lower matrix.*

***Proof***

1. Assume that $A$ and $B$ are arbitrary matrices in $D(K)$. The multiplication $A \cdot B$ is $\left(c_{ij}\right)$ where

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

For $s \neq t$, $a_{st} = b_{st} = 0$, which implies that for $i \neq j$,

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} \underset{k \neq i \Rightarrow a_{ik}=0}{=} a_{ii} b_{ij} \underset{i \neq j \Rightarrow b_{ij}=0}{=} 0$$

i.e., $AB$ is a diagonal matrix. Similarly, $BA$ is a diagonal matrix. The diagonal entries in $AB$ are

$$c_{ii} = \sum_{k=1}^{n} a_{ik} b_{ki} \underset{k \neq i \Rightarrow a_{ik}=0}{=} a_{ii} b_{ii}$$

which are also the diagonal entries in $BA$ as

$$\sum_{k=1}^{n} b_{ik} a_{ki} \underset{k \neq i = b_{ik}=0}{=} b_{ii} a_{ii} = a_{ii} b_{ii}$$

2. Assume that $A$ and $B$ are arbitrary matrices in $U(K)$. The multiplication $AB$ is $(c_{ij})$ where

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

As for $s > t$, $a_{st} = b_{st} = 0$, then for $i > j$

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} \underset{k < i \Rightarrow a_{ik}=0}{=} \sum_{k=i}^{n} a_{ik} b_{kj} \underset{k \geq i > j \Rightarrow b_{kj}=0}{=} 0.$$

That is, $AB$ is an upper matrix.

3. Similar to (2). ■

Many functions that can be defined on matrices, several of which are introduced below.

**Definition 1.6.10** (*The conjugate map*) Let $m, n \in \mathbb{N}$ and $\mathcal{M}_{mn}(K)$ be the set of all $m \times n$ matrices over $K$. Define the map

$$\begin{aligned} \mathcal{M}_{mn}(K) &\longrightarrow \mathcal{M}_{mn}(K) \\ (a_{ij}) &\mapsto (\overline{a}_{ij}) \end{aligned}$$

where $\overline{a}_{ij}$ is the complex conjugate of $a_{ij}$. This map is called the conjugate map. The conjugate of a matrix $A$ is denoted by $\overline{A}$.

It is clear that for any matrix $A \in \mathcal{M}_{mn}(\mathbb{R})$, $A = \overline{A}$. The conjugate of a matrix over $\mathbb{C}$ can be computed using the conjugates of complex numbers as in the following example:

$$\text{If } A = \begin{pmatrix} 1+3i & 2 & 2-i \\ 5i & -1 & 6+7i \end{pmatrix} \in \mathcal{M}_{2\times 3}, \text{ then } \overline{A} = \begin{pmatrix} 1-3i & 2 & 2+i \\ -5i & -1 & 6-7i \end{pmatrix}$$

The complex conjugate of a complex number $x + iy$ is $x - iy$. The following propositions are straightforward.

**Proposition 1.6.11** *Let $m, n, l \in \mathbb{N}$. Let $\mathcal{M}_{mn}(K)$ be the set of all $m \times n$ matrices over $K$, and $\mathcal{M}_{nl}(K)$ be the set of all $n \times l$ matrices over $K$. For $A \in \mathcal{M}_{mn}(K)$ and $B \in \mathcal{M}_{nl}(K)$,*

$$\overline{\overline{A}} = A, \quad \text{and} \quad \overline{AB} = \overline{A}\,\overline{B}$$

**Proposition 1.6.12** *Let $m, n \in \mathbb{N}$ and $\mathcal{M}_{mn}(K)$ be the set of all $m \times n$ matrices over $K$. The conjugate map is a function on $\mathcal{M}_{mn}(K)$.*

The second example of a function on matrices is the transpose map, which exchanges the rows of a matrix with its columns.

**Definition 1.6.13** (The transpose map) Let $m, n \in \mathbb{N}$ and $\mathcal{M}_{mn}(K)$ be the set of all $m \times n$ matrices over $K$. The map

$$\begin{aligned} T : \mathcal{M}_{mn}(K) &\longrightarrow \mathcal{M}_{nm}(K) \\ (a_{ij}) &\mapsto (a_{ji}) \end{aligned}$$

is called the transpose map. The transpose of a matrix $A$ is denoted by $A^T$.

The reader can easily check that $(8)^T = (8)$, $(1\ 2)^T = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^T = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$, and

$$\begin{pmatrix} 1 & -3 \\ 4 & 1 \\ 0 & 7 \end{pmatrix}^T = \begin{pmatrix} 1 & 4 & 0 \\ -3 & 1 & 7 \end{pmatrix}$$

**Proposition 1.6.14** *Let $m, n \in \mathbb{N}$ and $\mathcal{M}_{mn}(K)$ be the set of all $m \times n$ matrices over $K$. The transpose map is a function from $\mathcal{M}_{mn}(K)$ to $\mathcal{M}_{nm}(K)$, which satisfies the following statements:*

1. $\left(A^T\right)^T = A$
2. *For all* $\alpha \in \mathbb{C}$, $(\alpha A)^T = \alpha A^T$
3. $(A + B)^T = A^T + B^T$

*for any* $A, B \in \mathcal{M}_{mn}(K)$.

**Proposition 1.6.15** *Let $m, n, l \in \mathbb{N}$. For any $A \in \mathcal{M}_{mn}(K)$ and $B \in \mathcal{M}_{nl}(K)$,*

$$(AB)^T = B^T A^T.$$

**Definition 1.6.16** (*The Hermitian conjugate map*) Let $m, n \in \mathbb{N}$ and $\mathcal{M}_{mn}(K)$ be the set of all $m \times n$ matrices over $K$. The map

$$* : \mathcal{M}_{mn}(K) \longrightarrow \mathcal{M}_{nm}(K)$$
$$\left(a_{ij}\right) \mapsto \left(\overline{a}_{ji}\right)$$

is called the Hermitian conjugate map. The entry $\overline{a}_{ji}$ is the element in $K$ obtained by taking the conjugate of the element $a_{ji}$ in the transpose matrix of $(a_{ij})$. The Hermitian conjugate of a matrix $A$ is denoted by $A^*$.

For any $n \in \mathbb{N}$, the Hermitian conjugate of $I_n$ is $I_n$. The Hermitian conjugate map is a composition of the conjugate and the transpose maps ($A^* = \overline{A}^T$ for any matrix $A$ in $\mathcal{M}_{mn}(K)$). Therefore, the Hermitian conjugate is a function. It is also known by conjugate transpose, or Hermitian transpose map. The following proposition is straightforward.

**Proposition 1.6.17** *Let $m, n, l \in \mathbb{N}$ and $\mathcal{M}_{mn}(K)$ be the set of all $m \times n$ matrices over $K$. For $A \in \mathcal{M}_{mn}(K)$ and $B \in \mathcal{M}_{nl}(K)$.*

$$\left(A^*\right)^* = A, \quad (AB)^* = B^* A^*.$$

The following functions are defined only on square matrices.

**Definition 1.6.18** (*The trace map*) Let $n \in \mathbb{N}$ and $\mathcal{M}_n(K)$ be the set of all $n \times n$ matrices over $K$. The trace map is defined as

$$\text{trace:}\ \mathcal{M}_n(K) \longrightarrow \mathbb{C}$$
$$A \mapsto \sum_{i=1}^{n} a_{ii}$$

**Proposition 1.6.19** *Let $n \in \mathbb{N}$ and $\mathcal{M}_n(K)$ be the set of all $n \times n$ matrices over $K$. The $trace : \mathcal{M}_n(K) \to \mathbb{C}$ is a function that satisfies the following statements:*

1. $\text{trace}(A^T) = \text{trace}(A)$
2. $\text{trace}(A + B) = \text{trace}(A) + \text{trace}(B)$
3. $\text{trace}(AB) = \text{trace}(BA)$

*for all $A, B \in \mathcal{M}_n(K)$.*

**Definition 1.6.20** (*The determinant map*) Let $n \in \mathbb{N}$ and $\mathcal{M}_n(K)$ be the set of all $n \times n$ matrices over $K$. Let $\det : \mathcal{M}_n(K) \to \mathbb{C}$ be the recursive map defined for square matrices as

1. $\det((c)) = c$(determinant of $1 \times 1$ matrix is the entry of such matrix).
2. For any $A \in \mathcal{M}_n(K)$, such that $n \geq 2$,

$$\det(A) = \sum_{k=1}^{n} (-1)^{i+k} a_{ik} \det(A_{ik})$$

where $i$ is any integer such that $1 \leq i \leq n$, and $A_{ik}$ is the matrix obtained from $A$ by deleting row $i$ and column $k$.

In some linear algebra books, the notation $|A|$ is used instead of $\det(A)$. Throughout this book, $\det(A)$ is used to denote the determinant of $A$

***Remarks 1.6.21*** In the abovementioned definition, a fixed row $i$ is chosen, and the entries of the row are used to compute the determinant of the matrix $A$. The determinant can also be defined using a fixed column $j$ and the entries of this column, as follows:

- $\det((c)) = c$ (determinant of $1 \times 1$ matrix is the entry of such matrix).
- For any $A \in \mathcal{M}_n(K)$ such that $n \geq 2$,

$$\det(A) = \sum_{k=1}^{n} (-1)^{j+k} a_{kj} \det(A_{kj})$$

where $j$ is any integer such that $1 \leq j \leq n$, and $A_{kj}$ is the matrix obtained from $A$ by deleting row $k$ and column $j$.

***Example 1.6.22***

1. Consider an arbitrary matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ in $\mathcal{M}_2(K)$. To find the determinant of $A$, either a row or a column must be selected. Using the first row, we have

$$\det(A) = \sum_{k=1}^{2} (-1)^{1+k} a_{1k} \det(A_{1k}) = (-1)^2 a \det((d)) + (-1)^3 b \det((c))$$

$$= ad - bc$$

For example,
$\det\left(\begin{pmatrix} 2 & 7 \\ -3 & 5 \end{pmatrix}\right) = 10 - (-21) = 31$ and

$$\det\left(\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\right) = \cos\theta^2 + \sin\theta^2 = 1 \text{ for any angle } \theta.$$

2. Consider an arbitrary matrix $A = \begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{pmatrix}$ in $\in \mathcal{M}_3(K)$. To compute the determinant of $A$, either a row or a column must be selected. Choosing the first column this time,

$$\det(A) = \sum_{k=1}^{n} (-1)^{1+k} a_{k1} \det(A_{k1})$$

i.e., $\det(A)$ is

$$(-1)^2 x_1 \det\left(\begin{pmatrix} y_2 & z_2 \\ y_3 & z_3 \end{pmatrix}\right) - x_2 \det\left(\begin{pmatrix} y_1 & z_1 \\ y_3 & z_3 \end{pmatrix}\right) + x_3 \det\left(\begin{pmatrix} y_1 & z_1 \\ y_2 & z_2 \end{pmatrix}\right)$$

$$= x_1(y_2z_3 - z_2y_3) - x_2(y_1z_3 - z_1y_3) + x_3(y_1z_2 - z_1y_2).$$

Notably, the choice of the column or row in calculating the determinant determines the ease of the calculation. For example, the second row is the optimal choice to calculate the determinant of

$$\begin{pmatrix} 1 & 5 & 1 \\ 2 & 0 & 0 \\ 3 & 2 & 7 \end{pmatrix}$$

because the second row contains many zero entries. Using the second row, we obtain

$$\det\begin{pmatrix} 1 & 5 & 1 \\ 2 & 0 & 0 \\ 3 & 2 & 7 \end{pmatrix} = -2\det\left(\begin{pmatrix} 5 & 1 \\ 2 & 7 \end{pmatrix}\right) + 0 + 0 = -2(35 - 2) = -66$$

**Proposition 1.6.23** *Let $n \in \mathbb{N}$ and $\mathcal{M}_n(K)$ be the set of all $n \times n$ matrices over $K$. Let $A, B \in \mathcal{M}_n(K)$ be arbitrary matrices. The map* $\det : \mathcal{M}_n(K) \to \mathbb{C}$ *is a function that satisfies*

1. $\det(A) = \det(A^T)$
2. $\det(kA) = k^n \det(A)$
3. $\det(A \cdot B) = \det(A) \cdot \det(B)$.

The abovementioned functions are used to create special subsets of $\mathcal{M}_n(K)$. These subsets have important applications in algebra and other mathematical fields. We briefly mention them below.

**Definition 1.6.24** Let $n \in \mathbb{N}$ and $\mathcal{M}_n(K)$ be the set of all $n \times n$ matrices over $K$.

1. A matrix $A \in \mathcal{M}_n(K)$ is called an orthogonal matrix if $A \cdot A^T = A^T \cdot A = I_n$. The subset of all orthogonal matrices is denoted by $O(n, K)$, or simply, $O(n)$. i.e.,
$$O(n) = \left\{ A \in \mathcal{M}_n(K) : A \cdot A^T = A^T \cdot A = I_n \right\}.$$
2. A matrix $A \in \mathcal{M}_n(K)$ is called a unitary matrix if $A \cdot A^* = A^* \cdot A = I_n$. The subset of all unitary matrices is denoted by $U(n, K)$, or simply $U(n)$. i.e.,
$$U(n) = \left\{ A \in \mathcal{M}_n(K) : A \cdot A^* = A^* \cdot A = I_n \right\}.$$

**Notation 1.6.25** Let $n \in \mathbb{N}$ and $\mathcal{M}_n(K)$ be the set of all $n \times n$ matrices over $K$.

1. The set of all matrices in $\mathcal{M}_n(K)$ whose determinant is not zero is denoted by $GL(n, K)$ or $GL_n(K)$. i.e.,
$$GL_n(K) = \{A \in \mathcal{M}_n(K) : \det(A) \neq 0\}.$$
2. The set of all matrices in $\mathcal{M}_n(K)$ whose determinant equals 1 is denoted by $SL(n, K)$ or $SL_n(K)$. i.e.,
$$SL_n(K) = \{A \in \mathcal{M}_n(K) : \det(A) = 1\}.$$

In the following, we discuss the invertibility of a square matrix. The invertibility of a matrix is not defined if the matrix is not a square matrix.

**Definition 1.6.26** (Invertible matrices) Let $n \in \mathbb{N}$ and $\mathcal{M}_n(K)$ be the set of all $n \times n$ matrices over $K$. Let $A$ be a matrix in $\mathcal{M}_n(K)$. The matrix $A$ is said to be invertible if there exists a matrix $B \in \mathcal{M}_n(K)$ such that $AB = BA = I_n$. The matrix $B$ (if it exists) is called the inverse of $A$ and is denoted by $A^{-1}$.

Not every matrix in $\mathcal{M}_n(K)$ is invertible. For example, the zero matrix $0_n = (0)$ belongs to $\mathcal{M}_n(K)$, but no matrix in $\mathcal{M}_n(K)$ satisfies $0_n B = B\, 0_n = I_n$. Therefore,
$$\begin{aligned} \mathcal{M}_n(K) &\longrightarrow \mathcal{M}_n(K) \\ A &\mapsto A^{-1} \end{aligned}$$

is not a function on $\mathcal{M}_n(K)$. To define a formula for the inverse of an invertible matrix, the following definition is needed.

**Definition 1.6.27** Let $n \in \mathbb{N}$ and $\mathcal{M}_n(K)$ be the set of all $n \times n$ matrices over $K$. The adjugate matrix of a matrix $A$ in $\mathcal{M}_n(K)$ is the matrix computed from $A$ as follows:

$$\mathrm{Adj}(A) = \left((-1)^{i+j} A_{ij}\right)^T$$

where $A_{ij}$ is the determinant of the matrix obtained from $A$ by deleting the $i$ th row and the $j$ th columnn.

**Proposition 1.6.28** *Let $n \in \mathbb{N}$ and $\mathcal{M}_n(K)$ be the set of all $n \times n$ matrices over $K$. A matrix $A$ in $\mathcal{M}_n(K)$ is invertible if and only if $det(A) \neq 0$. In this case,*

$$A^{-1} = \frac{1}{\det(A)}\mathrm{Adj}(A)$$

*where $Adj(A)$ is the adjugate matrix of A.*

**Proposition 1.6.29** *Let $n \in \mathbb{N}$ and $\mathcal{M}_n(K)$ be the set of all $n \times n$ matrices over $K$. If $A$ and $B$ are invertible matrices in $\mathcal{M}_n(K)$, then $(AB)^{-1} = B^{-1}A^{-1}$.*

## 1.7 Geometric Transformations and Symmetries in the Plane

The rotation and reflection are two important basic transformations that operate on a plane (generally in $\mathbb{R}^n$). This section will study $R_\theta$, the rotation around the origin with angle $\theta$, and $l_\theta$, the reflection of a line passing through the origin, inclined at an angle $\theta$ from the $x$-axis. Any other rotation or reflection of a line (in the plane) can be defined using either $R_\theta$ or $l_\theta$. Readers can refer to (Boyd & Vandenberghe, 2018) for more details regarding geometric transformations and vectors in the plane. As mentioned in Example 1.6.2, there exists a correspondence between the points in $\mathbb{R}^n$ and the $n \times 1$ matrices with real entries. By restricting the study to $\mathbb{R}^2$, this correspondence can be expressed in the following lemma.

**Lemma 1.7.1** *The map.*

$$g : \mathbb{R}^2 \longrightarrow \left\{\begin{pmatrix} x \\ y \end{pmatrix} : x, y \in \mathbb{R}\right\}$$
$$(x, y) \mapsto \begin{pmatrix} x \\ y \end{pmatrix}$$

*is a bijection map identifying the points in $\mathbb{R}^2$ with the matrices in $\mathcal{M}_{2\times1}(\mathbb{R})$.*

Using this lemma, any function on the plane can be defined directly on the matrices $\mathcal{M}_{2\times 1}(\mathbb{R})$.

**Reminder 1.7.2** Any point $P$ in the plane can be represented using at least two coordinate systems of the plane:

1. The Cartesian coordinates of $P = (x, y)$, where $x$ and $y$ are given by the projections of the point $P$ on the $x$-axis and $y$-axis, respectively.
2. The polar coordinates of $P = (r, \psi)$, where $r$ is the distance from $P$ to the origin of the plane, and $\psi$ is the angle that the line $\overrightarrow{OP}$ makes with $x$-axis.

The relations between the two representations are given by the following equations

$$x = r\cos\psi,\ y = r\sin\psi,\ r^2 = x^2 + y^2,\ y = x\tan\psi$$

Recall that a rotation around the origin with an angle $\theta$ in the plane changes the polar coordinates of a point from $(r, \psi)$ to $(r, \psi + \theta)$. According to the next proposition, a rotation by $\theta$ around the origin is represented by a matrix multiplication, which will be denoted by $R_\theta$. The matrix $R_\theta$ is called the rotation matrix by $\theta$.

**Proposition 1.7.3** *The rotation of the point $(x, y)$ in the plane around the origin with an angle $\theta$ is equivalent to the function.*

$$\begin{aligned} f : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ \begin{pmatrix} x \\ y \end{pmatrix} &\mapsto R_\theta \cdot \begin{pmatrix} x \\ y \end{pmatrix} \end{aligned}$$

*where* $R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$.

***Proof*** Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be the rotation by an angle $\theta$ around the origin. The rotation $f$ moves the point $(x, y) = (rcos\psi, rsin\psi)$ to $\left(x', y'\right) = (r\cos(\psi + \theta), r\sin(\psi + \theta))$. Thus, the coordinates after the rotation are

$$x' = r\cos(\psi + \theta) = (r\cos\psi)\cos\theta - (r\sin\psi)\sin\theta = x\cos\theta - y\sin\theta$$

and

$$y' = r\sin(\psi + \theta) = (r\cos\psi)\sin\theta + (r\sin\psi)\cos\theta = x\sin\theta + y\cos\theta.$$

These equations can be expressed as

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

The following proposition matches the expected geometrical fact that “rotation by an angle $\theta$ followed by rotation by an angle β is equivalent to a rotation by $\theta+\beta$”. ■

**Proposition 1.7.4** *Let $\theta$ and $\beta$ be any two angles. Then,*

$$R_{\theta+\beta} = R_\theta R_\beta = R_\beta R_\theta.$$

***Proof*** According to the identities for the trigonometric functions,

$$R_{\theta+\beta} = \begin{pmatrix} \cos(\theta+\beta) & -\sin(\theta+\beta) \\ \sin(\theta+\beta) & \cos(\theta+\beta) \end{pmatrix}$$

$$= \begin{pmatrix} \cos\theta\cos\beta - \sin\theta\sin\beta & -(\sin\theta\cos\beta + \cos\theta\sin\beta) \\ \sin\theta\cos\beta + \cos\theta\sin\beta & \cos\theta\cos\beta - \sin\theta\sin\beta \end{pmatrix}$$

$$= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix} = R_\theta R_\beta.$$

Similarly, $R_{\beta+\theta} = R_\beta R_\theta$. The result follows as $\theta+\beta = \beta+\theta$. ■

***Example 1.7.5***

1. According to Proposition 1.7.3,
   - The rotation matrix by 0 is $R_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2$.
   - The rotation matrix by $\frac{\pi}{2}$ is $R_{\pi/2} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.
   - The rotation matrix by $\pi$ is $R_\pi = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I_2$.

     $R_n$ can also be computed using Proposition 1.7.4, to obtain the same answer.
   - Proposition 1.7.4 can be used to obtain
     - $R_{3\pi/2} = R_{\pi/2}R_\pi = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$,
     - $R_{2\pi} = R_{\pi/2}R_{3\pi/2} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.
2. Similarly, $R_{\pi/3} = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}$, $R_{2\pi/3} = R_{\pi/3}R_{\pi/3} = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$
   - $R_\pi = R_{\pi/3}R_{2\pi/3} = -I_2$,,
   - $R_{4\pi/3} = R_{\pi/3}R_\pi = -R_{\pi/3}$,,
   - $R_{5\pi/3} = R_{\pi/3}R_{4\pi/3} = -R_{2\pi/3}$, and
   - $R_{2\pi} = R_{\pi/3}R_{5\pi/3} = R_0$.

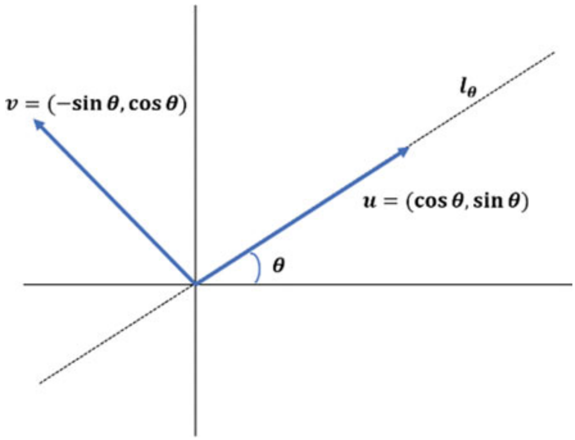



**Fig. 1.11** Rotating $u$ by $\pi/2$

To find a matrix that represents a reflection around a line $l_\theta$ passing the origin and making an angle $\theta$ with the $x$-axis, we need an expression for the unit vector in the direction of $l_\theta$. Recall that if $l_\theta$ makes an angle $\theta$ with the $x$-axis, then the coordinates for the unit vector $u$ on $l_\theta$ are $(\cos\theta, \sin\theta)$, which can be identified with $u = \begin{pmatrix}\cos\theta \\ \sin\theta\end{pmatrix}$. The unit vector $v$ that is perpendicular to $l_\theta$ can be obtained by rotating $u$ with $\frac{\pi}{2}$ see Fig. 1.11. According to Proposition 1.7.3,

$$v = R_{\pi/2}\begin{pmatrix}\cos\theta \\ \sin\theta\end{pmatrix} = \begin{pmatrix}-\sin\theta \\ \cos\theta\end{pmatrix}.$$

**Proposition 1.7.6** *Let $l_\theta$ be the straight line that passes the origin and makes an angle $\theta$ with the $x$-axis. The reflection of the point $(x, y)$ around $l_\theta$ in the plane is equivalent to the function.*

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}$$
$$\begin{pmatrix}x \\ y\end{pmatrix} \mapsto l_\theta \cdot \begin{pmatrix}x \\ y\end{pmatrix}$$

where $l_\theta = \begin{pmatrix}\cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta\end{pmatrix}$.

***Proof*** Let $w = \begin{pmatrix}x \\ y\end{pmatrix}$ be a point in the plane, then $w$ can be expressed as a line vector that starts from the origin and passes the point $w$. Such a vector is the sum of two vectors $w = w_\parallel + w_\perp$ where $w_\parallel$ is in the direction of $l_\theta$ and $w_\perp$ is perpendicular direction on $l_\theta$ (Fig. 1.12). i.e.,

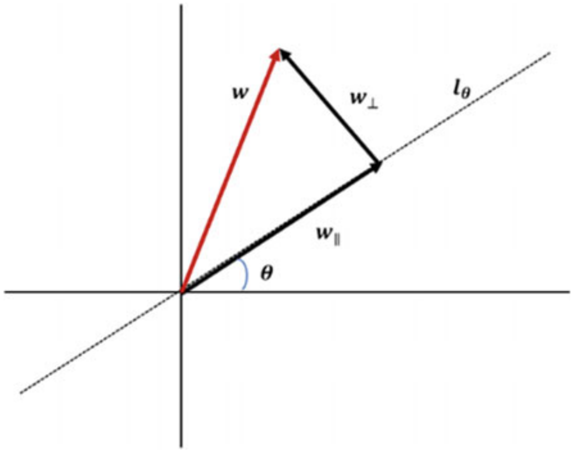



**Fig. 1.12** Vector $w = w_{\parallel} + w_{\perp}$

- $w_{\parallel} = \langle w, u \rangle u = (x\cos\theta + y\sin\theta)\begin{pmatrix}\cos\theta \\ \sin\theta\end{pmatrix} = \begin{pmatrix}x\cos^2\theta + y\sin\theta\cos\theta \\ x\sin\theta\cos\theta + y\sin^2\theta\end{pmatrix}$.
- $w_{\perp} = \langle w, v \rangle v = (-x\sin\theta + y\cos\theta)\begin{pmatrix}-\sin\theta \\ \cos\theta\end{pmatrix} = \begin{pmatrix}x\sin^2\theta - y\sin\theta\cos\theta \\ -x\sin\theta\cos\theta + y\cos^2\theta\end{pmatrix}$ where $\langle w, u \rangle$ and $\langle w, v \rangle$ are the inner products of $w$ with $u$ and $v$, respectively (Boyd & Vandenberghe, 2018).

The reflection of $w$ around $l_\theta$ is the point $w' = w_{\parallel} - w_{\perp}$ (Fig.1.13). Therefore,

$$w' = w_{\parallel} - w_{\perp}$$

$$= \begin{pmatrix}x\cos^2\theta + y\sin\theta\cos\theta \\ x\sin\theta\cos\theta + y\sin^2\theta\end{pmatrix} - \begin{pmatrix}x\sin^2\theta - y\sin\theta\cos\theta \\ -x\sin\theta\cos\theta + y\cos^2\theta\end{pmatrix}$$

$$= \begin{pmatrix}x\left(\cos^2\theta - \sin^2\theta\right) + 2y\sin\theta\cos\theta \\ 2x\sin\theta\cos\theta + y\left(\sin^2\theta - \cos^2\theta\right)\end{pmatrix}$$

$$= \begin{pmatrix}x\cos 2\theta + y\sin 2\theta \\ x\sin 2\theta - y\cos 2\theta\end{pmatrix} = \begin{pmatrix}\cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta\end{pmatrix}\begin{pmatrix}x \\ y\end{pmatrix}.$$

The above proposition shows that the reflection about a straight line that passes through the origin and makes the angle $\theta$ with the $x$-axis can be represented by a matrix multiplication. This matrix is called the reflection matrix about $l_\theta$ and is denoted by $l_\theta$. Applying the reflection about the line $l_\theta$ twice returns each point to itself. That is, ${l_\theta}^2 = I_2$

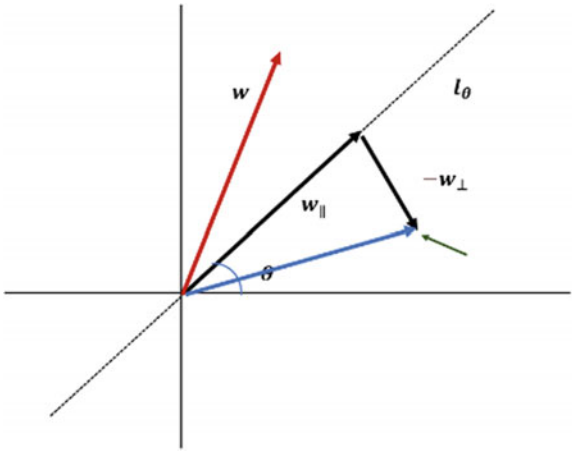



**Fig. 1.13** Reflection of $w$ around $l_\theta$

**Proposition 1.7.7** *For angles $\theta$ and $\beta$ the following identities hold:*

1. $R_{\theta+\beta} = R_\theta R_\beta = R_\beta R_\theta$.
2. $l_\theta l_\theta = R_0$.
3. $l_\theta l_\beta = R_{2(\theta-\beta)}$.
4. $R_\theta l_\beta = l_{\beta+\frac{\theta}{2}}$.
5. $l_\beta R_\theta = l_{\beta-\frac{\theta}{2}}$.
6. $\left(l_\beta R_\theta\right)^2 = \left(R_\theta l_\beta\right)^2 = R_0$

***Proof*** The first identity is the result in Proposition 1.7.4. Items (2) and (3) can be easily verified using Proposition 1.7.3, and 1.7.6, matrix multiplication, and trigonometric identities. To show (4), we first replace $\theta$ by $\frac{\theta}{2}+\beta$ in (3) to obtain $l_{\frac{\theta}{2}+\beta} l_\beta = R_\theta$. Using this identity and the identity in (2), we get

$$R_\theta l_\beta = \left(l_{\frac{\theta}{2}+\beta} l_\beta\right) l_\beta = l_{\frac{\theta}{2}+\beta}$$

To obtain the identity in (5), we compute $l_\beta l_{\beta-\frac{\theta}{2}}$ using the identity in (3) to get

$$l_\beta l_{\beta-\frac{\theta}{2}} = R_\theta$$

Using this identity and the identity in (2), we get

$$l_\beta R_\theta = l_\beta \left(l_\beta l_{\beta-\frac{\theta}{2}}\right) = l_{\beta-\frac{\theta}{2}}$$

The last equality follows directly from (2), (4), and (5). ■

The geometric interpretation for the relation in (3) is that a reflection about $l_\theta$ followed by a reflection about $l_\beta$ is equivalent to a rotation with an angle that is double the angle between $l_\theta$ and $l_\beta$. This aspect can be explained as follows.

It is straightforward to show, using Proposition 1.7.6, that a refection about $l_\theta$ maps a point $(r, \alpha)$ to the point $(r, 2\theta - \alpha)$. Therefore, applying another reflection $l_\theta$ leads to

$$(r, \alpha)l_\beta \to (r, 2\beta - \alpha)l_\theta \to (r, 2\theta - (2\beta - \alpha)) = (r, 2(\theta - \beta) + \alpha)$$

.i.e., the composition of $l_\theta$ and $l_\beta$ is a rotation by angle $2(\theta - \beta)$.

***Example 1.7.8*** By using the abovementioned notation,

$$l_0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \text{ and } l_{\pi/4} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\left(l_{\pi/4}\right)^2 = l_{\pi/4}l_{\pi/4} = I_2, \quad l_{\pi/4}R_{\pi/2} = l_0, \text{ and } l_{\pi/3} = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}.$$

From now on, a rotation by angle $\theta$ is identified with its rotation matrix $R_\theta$. A reflection about the line that passes through the origin and makes an angle $\theta$ with $x$-axis is identified with its reflection matrix $l_\theta$. Therefore, $R_\theta$ (similarly, $l_\theta$) represents both the matrix and the symmetry represented by the matrix.

Let $n \in \mathbb{N}$ such that $n \geq 3$. Consider a regular (sides of equal lengths and equal interior angles) $n$-polygon. There exist $2n$ types of symmetries for such a polygon: rotations about the center by an angle moving each vertex to the next vertex, reflections about the lines that pass the center of the polygon and the vertices, reflections about the lines that pass the center of the polygon and divide opposite sides of the polygon into equal halves, and their compositions. For example, see Fig. 1.14. The following steps can be implemented to identify such symmetries.

1. Select one of the vertices, and number it as vertex 1.
2. Identify the center of the polygon with the origin of the plane such that the line passes the center and vertex 1 lies on the $x$-axis.

As the polygon is regular, the required symmetries are

- the rotations by the angles $0, \frac{2\pi}{n}, \frac{4\pi}{n}, \ldots, \frac{2(n-1)\pi}{n}$,
- the reflections about lines passing the origin and making angles $0, \frac{\pi}{n}, \frac{2\pi}{n}, \frac{3\pi}{n}, \ldots, \frac{(n-1)\pi}{n}$ with the $x$-axis, and
- any compositions of these.

According to a Propositions 1.7.3 and 1.7.6, these symmetries can be represented by the matrices

$$R_0, R_{\frac{2\pi}{n}}, R_{\frac{4\pi}{n}}, \ldots, R_{\frac{2(n-1)\pi}{n}}, l_0, l_{\frac{\pi}{n}}, \ldots, l_{\frac{(n-1)\pi}{n}}$$

and any matrix resulting by their multiplications. Note that $l_\theta$ and $l_{\pi+\theta}$ represent the same line of symmetry. This can be checked easily using the result of Proposition 1.7.6.

***Example 1.7.9*** (The symmetries of a Square) Consider a square with the center at the origin and one of its vertices on the $x$-axis. The symmetries of the square are represented by $R_0 = I_2$, $R_{\pi/2}$, $R_\pi$, $R_{3\pi/2}$, $l_0$, $l_{\pi/4}$,$l_{\pi/2}$,$l_{3\pi/4}$ and the products of any of these matrices. The equations in Proposition 1.7.7 easily shows that these are all the different symmetries of the square. The following table is obtained using Proposition 1.7.7. Note that $l_0 = l_\pi$ as they represent the same line. Similarly, $l_{\pi/4} = l_{5\pi/4}$, $l_{\pi/2} = l_{3\pi/2}$ and $l_{3\pi/4} = l_{7\pi/4}$ (see Fig. 1.15 and Table 1.3).

The next example pertains to a polygon with an odd number of vertices.

***Example 1.7.10*** (The symmetries of a Pentagon) Consider a pentagon with the center at the origin, and one of its vertices on the $x$-axis. The symmetries of the pentagon are represented by

$$R_0 = I_2,\ R_{2\pi/5},\ R_{4\pi/5},\ R_{6\pi/5},\ R_{8\pi/5},\ l_0,\ l_{\pi/5},\ l_{2\pi/5},\ l_{3\pi/5},\ l_{4\pi/5}$$

and their products. The equations in Proposition 1.7.7 shows that these are all the different symmetries of the regular pentagon (Fig. 1.16).

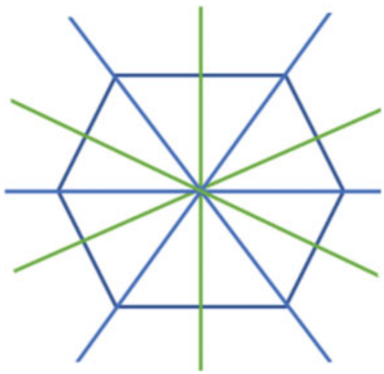


**Fig. 1.14** Regular 6-polygon

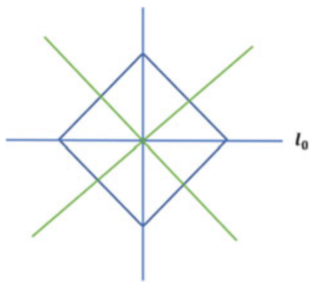
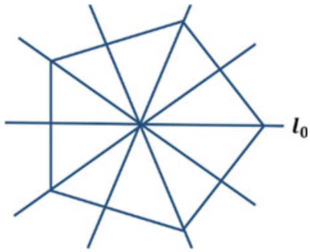



**Fig. 1.15** Regular 4-polygon

**Table 1.3** Composition of the symmetries of a square

| $\cdot$ | $R_0$ | $R_{\pi/2}$ | $R_\pi$ | $R_{3\pi/2}$ | $l_0$ | $l_{\pi/4}$ | $l_{\pi/2}$ | $l_{3\pi/4}$ |
|---|---|---|---|---|---|---|---|---|
| $R_0$ | $R_0$ | $R_{\pi/2}$ | $R_\pi$ | $R_{3\pi/2}$ | $l_0$ | $l_{\pi/4}$ | $l_{\pi/2}$ | $l_{3\pi/4}$ |
| $R_{\pi/2}$ | $R_{\pi/2}$ | $R_\pi$ | $R_{3\pi/2}$ | $R_0$ | $l_{\pi/4}$ | $l_{\pi/2}$ | $l_{3\pi/4}$ | $l_0$ |
| $R_\pi$ | $R_\pi$ | $R_{3\pi/2}$ | $R_0$ | $R_{\pi/2}$ | $l_{\pi/2}$ | $l_{3\pi/4}$ | $l_0$ | $l_{\pi/4}$ |
| $R_{3\pi/2}$ | $R_{3\pi/2}$ | $R_0$ | $R_{\pi/2}$ | $R_\pi$ | $l_{3\pi/4}$ | $l_0$ | $l_{\pi/4}$ | $l_{\pi/2}$ |
| $l_0$ | $l_0$ | $l_{3\pi/4}$ | $l_{\pi/2}$ | $l_{\pi/4}$ | $R_0$ | $R_{3\pi/2}$ | $R_\pi$ | $R_{\pi/2}$ |
| $l_{\pi/4}$ | $l_{\pi/4}$ | $l_0$ | $l_{3\pi/4}$ | $l_{\pi/2}$ | $R_{\pi/2}$ | $R_0$ | $R_{3\pi/2}$ | $R_\pi$ |
| $l_{\pi/2}$ | $l_{\pi/2}$ | $l_{\pi/4}$ | $l_0$ | $l_{3\pi/4}$ | $R_\pi$ | $R_{\pi/2}$ | $R_0$ | $R_{3\pi/2}$ |
| $l_{3\pi/4}$ | $l_{3\pi/4}$ | $l_{\pi/2}$ | $l_{\pi/4}$ | $l_0$ | $R_{3\pi/2}$ | $R_\pi$ | $R_{\pi/2}$ | $R_0$ |

**Fig. 1.16** Regular 5-polygon




**Summary 1.7.11** Let $n \in \mathbb{N}$ such that $n \geq 3$. The symmetries of the regular $n$-polygon are

- $n$ rotations, each of which shifts each vertex to the next vertex position,
- $n$ reflections, each of which pertain to the line passing the center and making an angle $\frac{\pi k}{n}$, where $k = 0, 1, \ldots, n-1$,
- any compositions of these entities.

Using Proposition 1.7.7 and $= l_\theta l_{\pi+\theta}$, one can easily show that the product of any two of these matrices

$$R_0, R_{\frac{2\pi}{n}}, \ldots, R_{\frac{2(n-1)\pi}{n}}, l_0, l_{\frac{\pi}{n}}, l_{\frac{2\pi}{n}} \cdots, l_{\frac{(n-1)\pi}{n}}$$

is again one of the matrices listed above.

**Corollary 1.7.12** *Let $n \in \mathbb{N}$ such that $n \geq 3$. The set.*

$$\left\{R_0, R_{\frac{2\pi}{n}}, \ldots, R^n_{\frac{2(n-1)\pi}{n}}, l_0, l_{\frac{\pi}{n}}, l_{\frac{2\pi}{n}} \ldots, l_{\frac{(n-1)\pi}{n}}\right\}$$

contains all the different symmetries of the regular $n$-polygon.

The following picture shows the effects of all the possible symmetries of an octagon (Fig. 1.17).

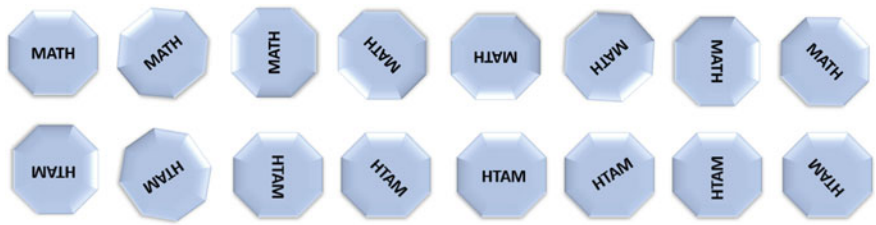



**Fig. 1.17** Symmetries of regular 8-polygon

## Exercises

### Solved Exercises

1.1 Show that any nonempty finite subset of $\mathbb{Z}$ has unique minimum and maximum elements.

**Solution:** Let $n \in \mathbb{N}$ and $A = \{a_1, a_2, \ldots, a_n\}$ be a nonempty finite subset of $\mathbb{Z}$. Using mathematical induction on $n$, we show that $A$ has minimum and maximum elements. i.e., we show that there exist $x, y \in A$ such that $x \le a \le y$ for all $a \in A$.

Base step: if $n = 1$, then $A = \{a_1\}$ for some integer $a_1 \in \mathbb{Z}$. As $a_1 \le a_1 \le a_1$, then by letting $x = y = a_1$, the statement is true for $n = 1$.

Inductive step: assume that the statement is true for $n$. That is, any subset of $\mathbb{Z}$ that contains $n$ elements must have minimum and maximum elements. Let $A = \{a_1, a_2, \ldots, a_n, a_{n+1}\}$ be a subset of $\mathbb{Z}$ with $n + 1$ elements. Let $B = A \backslash \{a_{n+1}\} = \{a_1, a_2, \ldots, a_n\}$ be a subset that only contains $n$ elements. According to the induction hypothesis, $B$ has minimum and maximum elements. i.e., there exist $x, y \in B$ such that $x \le a_i \le y$ for all $a_i \in B \subseteq A$. Three possibilities can be listed for $a_{n+1}$:

$$a_{n+1} \le x, x \le a_{n+1} \le y, \text{ or } y \le a_{n+1}$$

- If $a_{n+1} \le x$, then $a_{n+1} \le x \le a_i \le y$ for all $a_i \in A$, $1 \le i \le n$. Thus, $a_{n+1}$ is a minimum element of $A$ and $y$ is a maximum element.
- If $x \le a_{n+1} \le y$, then $x \le a_i \le y$ for all $a_i \in A$, $1 \le i \le n + 1$. Thus, $x$ is a minimum element of $A$ and $y$ is a maximum of $A$.
- If $y \le a_{n+1}$ then $x \le a_i \le y \le a_{n+1}$ for all $a_i \in A$, $1 \le i \le n$. Thus, $x$ is a minimum element of $A$ while $a_{n+1}$ is a maximum element of $A$.

In all three cases, $A$ has minimum and maximum elements. Therefore, according to the principle of mathematical induction, the statement is true for any $n \in \mathbb{N}$. The uniqueness follows as the relation $\le$ is a total order relation on $\mathbb{Z}$.

1.2 Let $A$ be any set. Consider the identity relation on $A$ that is defined in Example 1.3.7 (4). Show that

i Any subset of $\Delta_A$ is a transitive relation on $A$.

ii A relation $\mathcal{R}$ on $A$ is both symmetric and antisymmetric if and only if $\mathcal{R}$ is a subset of the identity relation $\Delta_A$.

**Solution**

i. Let $\mathcal{R}$ be any subset of $\Delta_A$ and $(a, b), (b, c) \in \mathcal{R}$, then $a = b$ and $b = c$. Therefore, $(a, c) = (a, b) \in \mathcal{R}$, and $\mathcal{R}$ is transitive.
ii. Assume that $\mathcal{R}$ is a relation on $A$ such that $\mathcal{R}$ is both symmetric and antisymmetric. Let $(a, b)$ be an arbitrary element in $\mathcal{R}$. Since $\mathcal{R}$ is symmetric, the ordered pair $(b, a)$ must also be in $\mathcal{R}$. However, since $\mathcal{R}$ is antisymmetric, then $a = b$, and thus, $\mathcal{R} \subseteq \Delta_A$. For the other direction, if $\mathcal{R} \subseteq \Delta_A$, any element in $\mathcal{R}$ is in the form $(a, a)$ for some $a \in A$. That is, $\mathcal{R} = \{(a, a) \in A \times A : a \in B\}$ for some $B \subseteq A$. i.e., $\mathcal{R} = \Delta_B$ for some subset $B$. By Example 1.3.7 (4), $\mathcal{R}$ is both symmetric and antisymmetric.

1.3 Let $\mathcal{R}$ be a relation on $\mathbb{Z}$ defined as $a\mathcal{R}b$ if and only if $a - b$ is divisible by 2. Determine whether $\mathcal{R}$ is reflexive, symmetric, antisymmetric, and/or transitive. What type of relation is $\mathcal{R}$?

**Solution:**

$\mathcal{R}$ is reflexive: since $a - a = 0$ and $0 = 0 \cdot 2$ is divisible by 2, and $a\mathcal{R}a$.

$\mathcal{R}$ is symmetric: if $a\mathcal{R}b$, then $a - b$ is divisible by 2. i.e., there exists $k \in \mathbb{Z}$ such that $a - b = 2k$. This implies that $b - a = 2(-k)$ is divisible by 2. i.e., $b\mathcal{R}a$.

$\mathcal{R}$ is not antisymmetric: $2\mathcal{R}4$ and $4\mathcal{R}2$ (Check!), but $2 \neq 4$.

$\mathcal{R}$ is transitive: if $a\mathcal{R}b$ and $b\mathcal{R}c$, then there exist $k, h \in \mathbb{Z}$ such that

$$a - b = 2k \text{ and } b - c = 2h.$$

Therefore,
$a - c = (a - b) + (b - c) = 2k + 2h = 2(k + h)$ is divisible by 2
i.e., $a\mathcal{R}c$, and thus, $\mathcal{R}$ is transitive.

1.4 Consider the set of positive integers $\mathbb{N}$, and let $A = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$. Define

$$\mathcal{R} = \left\{(x + 2, x + 3) \in \mathbb{N}^2 : x \in A\right\}.$$

Here $\mathcal{R}$ is a relation on $\mathbb{N}$. Determine the domain and the range of $\mathcal{R}$. What is the domain and range of $\mathcal{R}$ if $A$ is replaced by $\mathbb{N}$?

**Solution:** The relation $\mathcal{R}$ can be expressed as follows:

$$\mathcal{R} = \{(2, 3), (3, 4), (4, 5), (5, 6), (6, 7), (7, 8), (8, 9), (9, 10), (10, 11), (11, 12), (12, 13)\}$$

Therefore,

$$D(\mathcal{R}) = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$
$$Rang(\mathcal{R}) = \{3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13\}$$

If $A$ is replaced by $\mathbb{N}$, then the domain and the range would be the following infinite sets

$$D(\mathcal{R}) = \{3, 4, 5, 6, 7, \ldots\}, \quad Rang(\mathcal{R}) = \{4, 5, 6, 7, 8, \ldots\}$$

1.5 Let $A = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\} \times \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$. Let

$$\mathcal{R} = \left\{((x, y), (u, v)) \in A^2 : 3x + y \leq 3u + v\right\}$$

be a relation on $A$. Determine whether $\mathcal{R}$ is an equivalence, a partial order, or a total order relation.

Solve the question with $\mathcal{R} = \left\{((x, y), (u, v)) \in A^2 : 11x + y \leq 11u + v\right\}$ as well.

**Solution**: The properties of $\mathcal{R}$ can be checked as follows:

Reflexivity: since $3x + y \leq 3x + y$, we have $((x, y), (x, y)) \in \mathcal{R}$ for each $(x, y) \in A$. i.e., $\mathcal{R}$ is reflexive.

Symmetry: since $((0, 0), (1, 2)) \in \mathcal{R}$ and $((1, 2), (0, 0)) \notin \mathcal{R}$, the relation $\mathcal{R}$ is not symmetric.

Antisymmetry: since both $((0, 4), (1, 1))$ and $((1, 1), (0, 4))$ are elements in $\mathcal{R}$ and $(0, 4) \neq (1, 1)$, $\mathcal{R}$ is not antisymmetric.

Transitivity: If both $((x, y), (u, v))$ and $((u, v), (z, w))$ are elements in $\mathcal{R}$, then

$3x + y \leq 3u + v$ and $3u + v \leq 3z + w$.

This implies that $3x + y \leq 3z + w$ and $((u, v), (z, w)) \in \mathcal{R}$. So, $\mathcal{R}$ is transitive.

Therefore, $\mathcal{R}$ is not an equivalence or an order relation.

If $\mathcal{R} = \left\{((x, y), (u, v)) \in A^2 : 11x + y \leq 11u + v\right\}$, the abovementioned reasons can be used to show that $\mathcal{R}$ is reflexive, not symmetric and transitive. $\mathcal{R}$ is antisymmetric because if $((x, y), (u, v))$ and $((u, v), (x, y))$ are both elements in $\mathcal{R}$, then $11x + y = 11u + v$ implying that $v - y = 11(x - u)$ is a multiple of 11. That is, there exists $q \in \mathbb{Z}$ such that $v - y = 11q$. Since both $y$ and $v$ belong to $A$, we have $|v - y| \leq 10$. So, zero is the only possible value of $q$. i.e., $v = y$, and $x = u$. Therefore, $(x, y) = (u, v)$, and $\mathcal{R}$ is antisymmetric. Hence, $\mathcal{R}$ is a partial order relation. To show that $\mathcal{R}$ is a total order relation, assume that $(x, y)$ and $(u, v)$ are two elements in $A$. Since $11x + y$ and $11u + v$ are elements in $\mathbb{N}$, they are comparable. i.e., either $11x + y \leq 11u + v$ or $11u + v \leq 11x + y$. In general, $\mathcal{R} = \left\{((x, y), (u, v)) \in A^2 : kx + y \leq ku + v\right\}$ is a total order relation whenever $k > 10$.

1.6 Let $A$ and $B$ Be two finite sets such that $|A| = |B|$ and $f : A \to B$ be a function. Show that the map $f$ is injective if and only if $f$ is surjective.

**Solution:** Assume that $|A| = |B| = n$. Let $B = \{b_1, \cdots, b_n\}$, where $b_1, \cdots, b_n$ are distinct elements in $B$. For all $1 \leq i \leq n$, let $A_i = \{a \in A : f(a) = b_i\}$ and $k_i = |A_i|$, the number of elements in $A_i$. The sets

$\{A_i, 1 \le i \le n\}$ form a partition of $A$ (Check!). Therefore,

$$n = |A| = k_1 + \cdots + k_n$$

Note that,

- the function $f$ is surjective if and only if $A_i \neq \varnothing$ for all $1 \le i \le n$, that is, if and only if $k_i \ge 1$ for all $1 \le i \le n$.
- the function $f$ is injective if and only if for all $1 \le i \le n$, the set $A_i$ contains at most one element. That is, if and only if $k_i \le 1$ for all $1 \le i \le n$.

If $f$ is injective, then $k_i \le 1$ for all $1 \le i \le n$, so $k_i \in \{0, 1\}$. Since the sum $k_1 + \cdots + k_n = n$, then $\mathrm{k_i} = 1$ for all $1 \le i \le n$. Therefore, $f$ is surjective.

For the other direction, suppose $f$ is surjective, then $k_i \ge 1$ for all $1 \le i \le n$. Let $l_i = k_i - 1 \ge 0$. Here,

$$n = k_1 + \cdots + k_n = (1 + l_1) + \cdots + (1 + l_n) = \underbrace{1 + \cdots + 1}_{n \text{ times}} + l_1 + \cdots + l_n.$$

Therefore, $n = n + (l_1 + \cdots + l_n)$ and $l_1 + \cdots + l_n = 0$. Since $l_i \ge 0$ for all $1 \le i \le n$, then $l_i = 0$ for all $1 \le i \le n$, and $k_i = 1$ for all $1 \le i \le n$. Hence, $f$ is injective.

1.7 Consider the following relations:

1. $\mathcal{R} = A \times B$ where $A = \emptyset$ and $B$ is any nonempty set.
2. $\mathcal{S} = A \times B$ where $B = \emptyset$ and $A$ is any nonempty set.
3. $f = \{(x, y) \in \mathbb{R} \times \mathbb{R} : y^2 - x^2 = 1\}$.
4. $g = \{(x, y) \in \mathbb{R} \times \mathbb{R} : x^2 - y^2 = 1\}$.

Determine whether each relation is a function.

**Solution:**

1. If $A = \emptyset$, then $\mathcal{R} = A \times B = \emptyset$ for any nonempty set $B$. Since there are no elements in $A$, then the conditional statement

$$(a \in A \Rightarrow \exists\, !\, b \in B \text{ such that } (a, b) \in \mathcal{R})$$

is true. Therefore, $\mathcal{R}$ is a function.

2. Assume that $B = \varnothing$, and $A$ is a nonempty subset. If $\mathcal{S}$ is a function, then for each $a \in A$ there exists $b \in B$ such that $(a, b) \in \mathcal{S}$, which contradicts that $B$ is empty, So, $\mathcal{S}$ is not a function.
3. For each $x \in \mathbb{R}$, there exist two ordered pairs in $f$. Namely, $(x, \sqrt{1 + x^2})$ and $(x, -\sqrt{1 + x^2})$ preventing $f$ from being a function (Fig. 1.18). If we restrict the codomain to only the nonnegative real numbers, and define $f$ as

$$f = \{(x, y) \in \mathbb{R} \times (\mathbb{R}^+ \cup \{0\}) : y^2 - x^2 = 1\}$$

then $f$ would assign only one image for each $x$ in $\mathbb{R}$, and thus, it is a function on $\mathbb{R}$.

4. The relation $g = \{(x, y) \in \mathbb{R} \times \mathbb{R} : x^2 - y^2 = 1\}$ is not a function because the ordered pairs $\left(x, \sqrt{x^2 - 1}\right)$ and $(x, -\sqrt{x^2 - 1})$ are both in $g$. Even if we restrict the codomain to the nonnegative real numbers, the relation would not be a function on $\mathbb{R}$, because not every element in the domain has an image. If $x$ is any real number such that $|x| < 1$, then there is no $y$ such that $(x, y) \in g$ since that would give $1 + y^2 = x^2 < 1$, which implies that $y^2 < 0$. Thus, for $x$ with $|x| < 1$, no image exists under $g$ and $g$ is not a function (Fig. 1.19).

1.8 Let $f : A \to B$ Be a Function. Prove that the Map $f$ is Invertible if and Only if It is a Bijective Map.

**Solution:**

Assume that $f$ is invertible and $g : B \to A$ is the inverse function of $f$. If $f(a) = f(b)$, applying $g$ on both sides of the equation yields

$$a = g(f(a)) = g(f(b)) = b$$

which means that $f$ is injective. Let $b \in B$ be an arbitrary element, and $a = g(b)$. Then

$$f(a) = f(g(b)) = f \circ g(b) = \iota_B(b) = b.$$

So, $f$ is surjective. Therefore, $f$ is bijective.

For the other direction, assume that $f$ is bijective, and let

$$g = \{(b, a) : (a, b) \in f\} \subseteq B \times A.$$

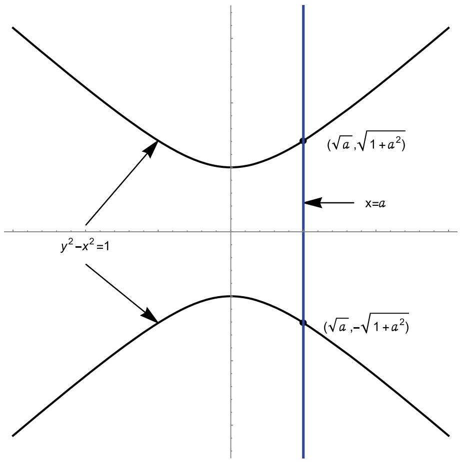



**Fig. 1.18** Graph of $f = \{(x, y) \in \mathbb{R} \times \mathbb{R} : y^2 - x^2 = 1\}$

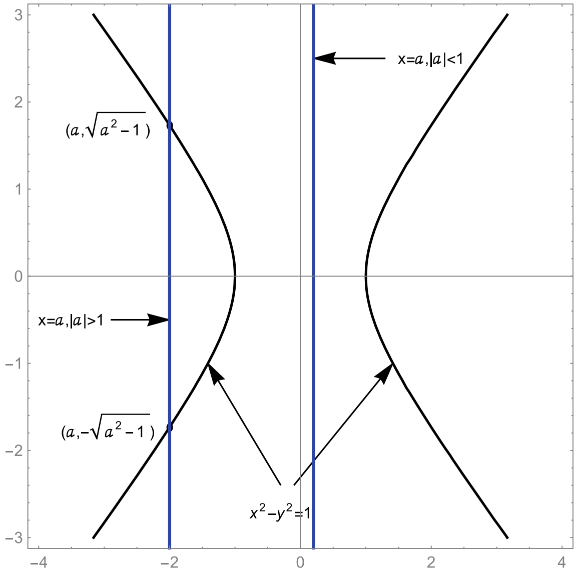



**Fig. 1.19** Graph of $g = \{(x, y) \in \mathbb{R} \times \mathbb{R} : x^2 - y^2 = 1\}$

We show that $g : B \to A$ is a function such that $g \circ f = \iota_A$ and $f \circ g = \iota_B$. Let $b \in B$ be an arbitrary element. Since $f$ is surjective (onto), there exists $a \in A$ such that $f(a) = b$. i.e., $(a, b) \in f$, so $(b, a) \in g$. That is, $g$ is defined for each element in $B$.

To show the uniqueness of images of elements in $B$ under $g$, let $(b, a_1)$ and $(b, a_2)$ be two elements in $g$, so $(a_1, b)$ and $(a_2, b)$ belong to $f$. Since $f$ is injective (one-to-one), then $a_1 = a_2$, as required. We still need to show that for all $a \in A$, $g \circ f(a) = a$ and for all $b \in B$, $f \circ g(b) = b$. Let $a \in A$, since $f$ is defined for all elements of $A$, then there exists a unique element $b \in B$ such that $(a, b) \in f$ i.e., $(b, a) \in g$. Therefore,

$$g \circ f(a) = g(f(a)) = g(b) = a.$$

Since $a$ is an arbitrary element in $A$, then $g \circ f = \iota_A$. Similarly, according to the surjectivity of $f$, for each $b \in B$ there exists $a \in A$ such that $(a, b) \in f$. This implies that $(b, a) \in g$. i.e.,

$$f \circ g(b) = f(g(b)) = f(a) = b.$$

Since $b$ is an arbitrary element in $B$, then $f \circ g = \iota_B$.

1.9 Let $n \in \mathbb{N}$, $K$ be any subset of $\mathbb{C}$, and $A \in \mathcal{M}_n(K)$ be an invertible matrix. Show that

$$\det\left(A^{-1}\right) = \frac{1}{\det(A)}$$

and $\det(A) = \det\left(A^{-1}\right)$ if and only if $\det(A) \in \{1, -1\}$.

**Solution:** If $A \in \mathcal{M}_n(K)$ is an invertible matrix, then there exists $A^{-1} \in \mathcal{M}_n(K)$ such that

$$AA^{-1} = I_n$$

Therefore,

$$\det(A)\det\left(A^{-1}\right) = \det\left(AA^{-1}\right) = \det(I_n) = 1$$

The result now follows. Moreover, we have

$$\det(A) = \det\left(A^{-1}\right) \Leftrightarrow \det(A) = \frac{1}{\det(A)} \Leftrightarrow (\det(A))^2 = 1 \Leftrightarrow \det(A) \in \{1, -1\}$$

1.10 Compute all possible symmetries of a regular triangle and list their multiplications.

**Solution:** According to Corollary 1.7.12, the set.

$$\left\{R_0, R_{2\pi/3}, R_{4\pi/3}, l_o, l_{\pi/3}, l_{2\pi/3}\right\}$$

contains all the symmetries of the regular triangle. Using the relations in Proposition 1.7.7, and,

$$l_\theta = l_{\pi+\theta}$$

we obtain the following Table 1.4 that contains all their compositions.

## Unsolved Exercises

**Table 1.4** Compositions of the symmetries of a regular triangle

| · | $R_0$ | $R_{2\pi/3}$ | $R_{4\pi/3}$ | $l_0$ | $l_{\pi/3}$ | $l_{2\pi/3}$ |
|---|---|---|---|---|---|---|
| $R_0$ | $R_0$ | $R_{2\pi/3}$ | $R_{4\pi/3}$ | $l_0$ | $l_{\pi/3}$ | $l_{2\pi/3}$ |
| $R_{2\pi/3}$ | $R_{2\pi/3}$ | $R_{4\pi/3}$ | $R_0$ | $l_{\pi/3}$ | $l_{2\pi/3}$ | $l_0$ |
| $R_{4\pi/3}$ | $R_{4\pi/3}$ | $R_0$ | $R_{2\pi/3}$ | $l_{2\pi/3}$ | $l_0$ | $l_{\pi/3}$ |
| $l_0$ | $l_0$ | $l_{2\pi/3}$ | $l_{2\pi/3}$ | $R_0$ | $R_{4\pi/3}$ | $R_{2\pi/3}$ |
| $l_{\pi/3}$ | $l_{\pi/3}$ | $l_0$ | $l_{2\pi/3}$ | $R_{2\pi/3}$ | $R_0$ | $R_{4\pi/3}$ |
| $l_{2\pi/3}$ | $l_{2\pi/3}$ | $l_{\pi/3}$ | $l_0$ | $R_{4\pi/3}$ | $R_{2\pi/3}$ | $R_0$ |

1.11 Let $\Delta$ denote the symmetric difference defined in Definition 1.1.8. Show that $(A\Delta B)\Delta C = A\Delta(B\Delta C)$ for any sets $A$, $B$ and $C$.

1.12 Let $r$ be a real number such that $0 < r < 1$. Show that

$$r^{-1} + 1 + r + r^2 + \cdots + r^n = \frac{r^{n+2} - 1}{r(r-1)}$$

for every integer $n$ such that $n \geq -1$.

1.13 Show that $2^{n+2} + 3^{2n+1}$ is divisible by 7 for every nonnegative integer $n$.

1.14 Show that if $x \in \mathbb{R}$, $x > -1$, then $(1+x)^n \geq 1 + nx$ for every integer $n$ such that $n \geq 0$.

1.15 For each of the following relations

1. $\mathcal{R} = \{(x, y) \in \mathbb{Z}^2 : x + y < 5\}$,
2. $\mathcal{T} = \{(x, y) \in \mathbb{N}^2 : x + y > 1\}$, and
3. $\mathcal{V} = \{(x, y) \in \mathbb{Z}^2 : x + y \text{ is even}\}$,

determine whether the relation is reflexive, symmetric, antisymmetric, or transitive.

1.16 Let $A$ be a nonempty set. For any relations $\mathcal{R}$ and $\mathcal{T}$ on $A$, the composition relation is defined as

$$\mathcal{R}\circ\mathcal{T} = \{(a, c) \in A^2 : \exists\, b \in A, (a, b) \in \mathcal{T} \wedge (b, c) \in \mathcal{R}\}$$

Show that if $\mathcal{R}$ and $\mathcal{T}$ are equivalence relations on $A$, then $\mathcal{R} \circ \mathcal{T}$ is an equivalence relation on $A$ if and only if $\mathcal{R} \circ \mathcal{T} = \mathcal{T} \circ \mathcal{R}$.

1.17 Let $A = \mathbb{Z}\backslash\{0\}$ and $\mathcal{R} = \{(x, y) \in A^2 : xy > 0\}$ be a relation on $A$. Determine whether $\mathcal{R}$ is an equivalence relation.

1.18 Let $A = \mathbb{Z}\backslash\{0\}$ and $\mathcal{R} = \{(x, y) \in A^2 : x|y\}$ be a relation on $A$. Show that $\mathcal{R}$ is a partial order relation. Is $\mathcal{R}$ a total order relation?

1.19 Let $n \in \mathbb{N}$ and $\mathcal{R} = \{(a, b) \in \mathbb{Z}^2 : \frac{b-a}{n} \in \mathbb{Z}\} = \{(a, b) \in \mathbb{Z}^2 : n|(b-a)\}$.

a. Show that $\mathcal{R}$ is an equivalence relation on $\mathbb{Z}$.

b Show that the equivalence classes of $\mathcal{R}$ can be expressed as $m + n\mathbb{Z}$, where $m = 0, 1, 2, \ldots, n-1$.

1.20 Consider the set of integers $\mathbb{Z}$. Let $f : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ and $g : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ be the maps defined by

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

Show that $f$ and $g$ define functions on $\mathbb{Z}$.

1.21 Show that the composition of two functions is a function, and the composition of two bijective maps is a bijective map. Show that the inverse of a bijective map is bijective.

1.22 Determine whether the functions

a. $f : \mathbb{R} \to \mathbb{R}$, where $f(x) = x^2 - 5$
b. $g : \mathbb{R}\backslash\{0\} \to \mathbb{R}$, where $g(x) = \frac{1}{x} + 2$

are injective or surjective. Find the domains and ranges of these functions.

1.23 Let $A$ and $B$ Be Two Sets and $f:A\to B$ Be a Function. Show that

$$\iota_B \circ f = f \circ \iota_A = f$$

where $i_A$ and $i_B$ are the inclusion functions of A and B respectively.

1.24 Let $A$ and $B$ be two sets and $f : A \to B$ be a Function. Show that

- $f$ is injective if and only if for any $H \subseteq A$, $f^{-1}(f(H)) = H$.
- $f$ is surjective if and only if for any $K \subseteq B$, $f\big(f^{-1}(K)\big) = K$.

where $f(H)$ denotes the image of $H$ under $f$, and $f^{-1}(K)$ is the preimage of $K$.

1.25 Let $f : \mathbb{C} \to \mathbb{C}$ be the map takes $z = x+iy$ to its complex conjugate $\overline{z} = x-iy$. Show that the map $f$ is a bijective function, and for all $z, z_1, z_2 \in \mathbb{C}$

$$\bar{\bar{z}} = z, \quad \text{and} \quad \overline{z_1 z_2} = \bar{z_1}\ \bar{z_2}$$

1.26 Show that the multiplication of diagonal matrices is commutative.

1.27 Show that if $A$ is an upper or a lower tringle matrix, then the determinant of $A$ is the product of its diagonal entries. i.e.,

$$det(A) = \prod_{i=1}^{n} a_i = a_{11}\ a_{22} \ldots\ldots a_{nn}.$$

1.28 List all possible symmetries for a regular octagon and the compositions of any two of these symmetries.

## References

Boyd, S., & Vandenberghe, L. (2018). *Introduction to applied linear algebra vectors, matrices, and least squares*. Cambridge University Press.
Burton, D. M. (2007). *Elementary number theory*. MacGraw-Hill.
Halmos, P. (2013). *Naive set theory*. Springer.
Hammack, R. (2013). *Book of Proof.* Creative Commons.
Hartman, G. (2011). *Fundamentals of Matrix Algebra.* Creative Commons.
Hungerford, T., & Shaw, D. (2009). *Contemporary precalculus: A graphing approach* (5th ed.). Thomson Higher Education.
Printer, C. C. (2014). *A book of set theory*. Dover Publications INC.