# SETS AND FUNCTIONS

## 1. Introduction of Sets

The concept of **sets** is the foundation of modern mathematics and definitions of most fundamental mathematics terms are based on sets. **Sets** are *well-defined* collection of objects. Here, “well-defined” means that there is a definite method to determine whether an object belongs to the set. For example, we may define $A$ to be the set of prime numbers. Recall that a positive integer is called a prime if it has exactly two positive divisors, 2, 3, 5, 7 are members of $A$ but 1, 4 are not. Mathematically, we write

$$A = \{x \mid x \text{ is a prime number}\}.$$

Generally, we write $\{x \mid P(x)\}$ to mean “the set of $x$ such that the statement $P(x)$ is true”.

- Here, the letter $x$ may be replaced by other symbols. For instance, $\{y \mid P(y)\}$ is just the same as $\{x \mid P(x)\}$.

- Sometimes the word “family” or “collection” is used in place of the word “set” to give a clear picture. For example, we usually say “the family of sets” instead of “the set of sets” when we refer to a set with its elements actually are sets themselves.

We may also define a set by listing out all its “objects”, like $B = \{1, 2, 3\}$. Notice that it is the same as $\{x \mid x \text{ is a positive integer not exceeding } 3\}$.

If $x$ is an object of the set $A$, we say $x$ is an **element** of $A$, and write $x \in A$. Otherwise, $x \notin A$. Two sets $A$ and $B$ are said to be equal, or $A = B$, if they contain the same elements, ie. $x \in A \Leftrightarrow x \in B$. The (unique) set with no element is called an **empty set**, denoted by $\varnothing$. If a set has one and only one element, it is called a **singleton**.

**Illustration.** Let $X = \{a, b, c, d\}$, then $a \in X$, $e \notin X$.

The symbols “$\forall$” and “$\exists$” means “for all” and “there exists” respectively. We say “$\forall x \in A$, $P(x)$” if every element $x$ of $A$ satisfies $P(x)$ and “$\exists x \in A$, $P(x)$” if at least one of the elements of A satisfies $P(x)$.

**Illustration.** Let $Y = \{1, 3, 5, 7, 9\}$, then $\forall x \in Y$, $x$ is odd. Also, $\exists x \in Y$ such that $x$ is a composite.

For any sets $A$ and $B$, if $\forall x \in A$, $x \in B$, i.e. $B$ contains all the elements of $A$, then $A$ is said to be a **subset** of $B$, denoted by $A \subseteq B$. Furthermore, if there is an element of $B$ not in $A$, $A$ is called a **proper subset** of $B$. Clearly, $A = B$ if and only if $A \subseteq B$ and $B \subseteq A$. This property is useful in proving that two sets are equal.

In convenience, we usually denote the set of natural numbers by $\mathbb{N}$, the set of integers by $\mathbb{Z}$, the set of rational numbers by $\mathbb{Q}$, the set of real numbers by $\mathbb{R}$, and the set of complex numbers by $\mathbb{C}$. Clearly, $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$. The positive-element subset of $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ are $\mathbb{Z}^+$, $\mathbb{Q}^+$, $\mathbb{R}^+$ respectively.

- There is still no standard agreement on the set of natural numbers. Some regard it as the set of positive integers, while the other regard it as the set of non-negative integers. For the rest of the text, we adopt the first definition.

- We may denote the subset of a set $A$ whose elements satisfied the statement $P$ by $\{x \in A \mid P(x)\}$.

**Example 1.1.**

Prove that {x | x is a triangle with every interior angle not smaller than $60^\circ$ } is equal to the set of equilateral triangles.

**Solution.**

Let $A$ be the set of triangles with every interior angle not smaller than $60^\circ$, $B$ be the set of equilateral triangles. Since the three interior angles of any equilateral triangle is $60^\circ$, every element of B satisfies the statement "a triangle with every interior angle not smaller than $60^\circ$". Therefore, $\forall x \in B$, $x \in A$, i.e. $B \subseteq A$.

On the other hand, suppose $x \in A$, Let $\theta_1$, $\theta_2$, $\theta_3$ be the three interior angles, with $60^\circ \le \theta_1 \le \theta_2 \le \theta_3$. Since $60^\circ \le \theta_3 = 180^\circ - \theta_1 - \theta_2 \le 180^\circ - 60^\circ - 60^\circ = 60^\circ$, we get $\theta_1 = \theta_2 = \theta_3 = 60^\circ$. Therefore $x \in B$. Hence $A \subseteq B$. The proof is completed.

## 2. Operations on Sets

For any two sets *A* and *B*, we define their **union**, $A\cup B$, by $A\cup B=\{x\mid x\in A \text{ or } x\in B\}$. Their **intersection**, $A\cap B$, is defined by $A\cap B=\{x\mid x\in A \text{ and } x\in B\}$. If $A\cap B=\varnothing$, then *A* and *B* are said to be **disjoint**.

**Illustration.** Let A = {1, 2, 3, 4}, B = {2, 4, 6}. Then $A\cup B$ = {1, 2, 3, 4, 6}, and $A\cap B$ = {2, 4}.

**Illustration.** The set of even numbers and the set of odd numbers are disjoint.

**Theorem 2.1.**
Let *A*, *B*, *C* be sets, then

1. $A\cup A=A$; (idempotent)
2. $A\cap A=A$; (idempotent)
3. $(A\cup B)\cup C=A\cup(B\cup C)$; (associative)
4. $(A\cap B)\cap C=A\cap(B\cap C)$; (associative)
5. $A\cup B=B\cup A$; (commutative)
6. $A\cap B=B\cap A$; (commutative)
7. $(A\cup B)\cap C=(A\cap C)\cup(B\cap C)$; (distributive)
8. $(A\cap B)\cup C=(A\cup C)\cap(B\cup C)$. (distributive)

**Proof.** (1)-(6) are rather trivial, and (7), (8) are similar, so only (7) will be proved here.

For (7), $\forall x\in(A\cup B)\cap C$, $x\in(A\cup B)$ and $x\in C$, so $x\in A\cap C$ or $x\in B\cap C$, which means $x\in(A\cap C)\cup(B\cap C)$. In other words, $(A\cup B)\cap C\subseteq(A\cap C)\cup(B\cap C)$.

Next, $\forall x\in(A\cap C)\cup(B\cap C)$, $x\in A\cap C$ or $x\in B\cap C$. Hence, $x\in C$, and either $x\in A$ or $x\in B$ is true. It follows that $x\in(A\cup B)\cap C$. This proves $(A\cap C)\cup(B\cap C)\subseteq(A\cup B)\cap C$. Combining the two, we get $(A\cup B)\cap C=(A\cap C)\cup(B\cap C)$.

Q.E.D.

As a result of the associative laws (3) and (4) in theorem 2.1, we can simply write $A\cup B\cup C$ and $A\cap B\cap C$ to represent the sets in (3) and (4) respectively. In the same manner, we can denote the union of the sets $A_1$, $A_2$,…, $A_n$ by $A_1\cup A_2\cup\ldots\cup A_n$, or simply $\bigcup_{i=1}^{n}A_i$, with

$$\bigcup_{i=1}^{n} A_i = \{x \mid x \text{ is an element of at least one } A_i\}.$$

The intersection of them, denoted by $A_1 \cap A_2 \cap ... \cap A_n$, or simply $\bigcap_{i=1}^{n} A_i$, is

$$\bigcap_{i=1}^{n} A_i = \{x \mid x \in A_i \text{ for all } i = 1, 2, \ldots, n\}.$$

Here are some more properties about intersection and union of sets. Although trivial, they may be useful. The readers may try to verify them.

**Theorem 2.2.**

Let *A*, *B*, *C* be sets, then

1. $\varnothing \subseteq A$;
2. $A \cup \varnothing = A$;
3. $A \cap \varnothing = \varnothing$;
4. $A \subseteq A \cup B$;
5. $B \subseteq A \cup B$;
6. $A \cap B \subseteq A$;
7. $A \cap B \subseteq B$;
8. $A \subseteq B, B \subseteq C \Rightarrow A \subseteq C$.

Let *A*, *B* be sets. The **complement** of *B* relative to *A*, denoted by $A \setminus B$, or $A - B$, is the set $\{x \mid x \in A \text{ and } x \notin B\}$.

**Illustration.** $\mathbb{Z} \setminus \mathbb{Z}^+$ is the set of non-positive integer, $\mathbb{R} \setminus \mathbb{Q}$ is the set of irrational number.

In any discussion involving sets, we call the set consisting of all elements under discussion to be the **universal set**, and is usually denoted by *U*. If such a set exists and $A \subset U$, we may denote the set $U \setminus A$ by $A^C$. In using this notation, we must make sure that the set *U* is clear and would not cause confusion to readers.

**Theorem 2.3.**

Let *A*, *B* be sets, and *U* is the universal set. Then

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

It is easy to prove theorem 2.3, so we shall not give a proof.

**Theorem 2.4. (De Morgan's theorem)**
Let $A$, $B$, $C$ be sets. Then
1. $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$;
2. $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$.

**Proof.** (1) is true because

$$\begin{aligned} x \in A \setminus (B \cup C) &\Leftrightarrow x \in A \text{ and } x \notin B \cup C \\ &\Leftrightarrow x \in A \text{ and } (x \notin B \text{ and } x \notin C) \\ &\Leftrightarrow x \in A \setminus B \text{ and } x \in A \setminus C \\ &\Leftrightarrow x \in (A \setminus B) \cap (A \setminus C) \end{aligned}$$

Similarly, for (2), it is because

$$\begin{aligned} x \in A \setminus (B \cap C) &\Leftrightarrow x \in A \text{ and } x \notin B \cap C \\ &\Leftrightarrow x \in A \text{ and } (x \notin B \text{ or } x \notin C) \\ &\Leftrightarrow (x \in A \text{ and } x \notin B) \text{ or } (x \in A \text{ and } x \notin C) \\ &\Leftrightarrow x \in A \setminus B \text{ or } x \in A \setminus C \\ &\Leftrightarrow x \in (A \setminus B) \cup (A \setminus C) \end{aligned}$$

Q.E.D.

- From theorem 2.4, we immediately get $(X \cup Y)^C = X^C \cap Y^C$ and $(X \cap Y)^C = X^C \cup Y^C$, special cases of them.

**Definition 2.1. (Intervals)**
Let $a, b \in \mathbb{R}$, with $a < b$. Then we define
1. $(a,b) = \{x \in \mathbb{R} \mid a < x < b\}$;
2. $[a,b] = \{x \in \mathbb{R} \mid a \le x \le b\}$;
3. $(a,b] = \{x \in \mathbb{R} \mid a < x \le b\}$;
4. $[a,b) = \{x \in \mathbb{R} \mid a \le x < b\}$.

($a$, $b$) is called an **open interval**, while [$a$, $b$] is called a **closed interval**. The other two are called **half-open** (or **half-closed**). $b - a$ is called the **length** of the interval. If there exists a positive integer $M$ such that for any $x$ in the interval, $|x| \le M$, the interval is said to be **bounded**. Otherwise, it is **unbounded**.

We usually use the symbols "$\infty$" and "$-\infty$" to mean **positive infinity** and **negative infinity** respectively. They may be considered as "numbers" beyond the real number system such that for any real number *x*, $-\infty < x < \infty$. Notice that neither of them belongs to $\mathbb{R}$. We may replace *a* by $-\infty$ and *b* by $\infty$ in open interval or half open intervals in definition 2.1 to represent unbounded intervals. For instances, $(-\infty,\infty) = \mathbb{R}$, $(0,\infty) = \mathbb{R}^+$.

- It is incorrect to write $(1,\infty]$, $[-\infty,0)$ to denote intervals in $\mathbb{R}$ since $\infty, -\infty \notin \mathbb{R}$.

**Definition 2.2. (Power Set)**
Let *A* be a set, then we define the **power set** of *A* to be the set of all subsets of *A*, denoted by *P*(*A*). Symbolically, we write $P(A) = \{x \mid x \subseteq A\}$.

- Note that $\varnothing \subseteq A$ for any set *A*. Indeed, $P(\varnothing) = \{\varnothing\}$ is not empty.

- In general, $A \subseteq P(A)$ is *false*, except when $A = \varnothing$.

**Illustration.** Let $A = \{0, 1, a\}$, then $P(A) = \{\varnothing, \{0\}, \{1\}, \{a\}, \{0, 1\}, \{0, a\}, \{1, a\}, A\}$

For any sets *A* and *B*, we define their **Cartesian product** $A \times B$ to be the set of all *ordered* pairs (*a*, *b*) with $a \in A$, $b \in B$. Shortly, $A \times B = \{(a,b) \mid a \in A \text{ and } b \in B\}$. In general, we may extend the definition to *n* arbitrary sets $A_1$, $A_2$, … , $A_n$, with

$$A_1 \times A_2 \times \ldots \times A_n = \{(a_1, a_2, \ldots, a_n) \mid a_i \in A_i \text{ for } i = 1, 2, \ldots, n\}.$$

If $A_1 = A_2 = \ldots = A_n = A$, we can simply write $A^n$ instead.

**Illustration.** Let $A = \{1, 2, 3\}$, $B = \{\mathrm{a}, \mathrm{b}\}$. Then $A \times B = \{(1,\mathrm{a}), (1,\mathrm{b}), (2,\mathrm{a}), (2,\mathrm{b}), (3,\mathrm{a}), (3,\mathrm{b})\}$

**Illustration.** (1, 2, 3) and (2, 1, 3) are two distinct elements of $\mathbb{N}^3$.

## 3. Functions

**Definition 3.1. (Function)**
Let *A*, *B* be non-empty sets. Then a **function** (mapping) *f* from *A* to *B*, denoted by $f : A \to B$, is a rule that associates *each* element *a* in *A* a *unique* element *b* in *B*. We say that *f* maps *a* into *b*, and write *f*(*a*) = *b*.
Here, *A* is called the **domain**, and *B* is called the **codomain** of *f*. *b* is called the **image** of *a*, and *a* is a **pre-image** of *b* under *f*.

- If the codomain is $\mathbb{R}$ or $\mathbb{C}$, $f$ is said to be a *real-valued function* or a *complex-valued function* respectively.

**Illustration.** Let $f:\mathbb{N}\to\mathbb{Z}$, defined by $f(x)=(-1)^x x$. It is well-defined since for every $x\in\mathbb{N}$, there exists an unique element $y\in\mathbb{Z}$ such that $y=(-1)^x x$. For example, $f(3)=(-1)^3\cdot 3=-3$.

**Illustration.** If we let $g:\mathbb{N}\to\mathbb{N}$, but with $g(x)=x/3$, then $g$ is not well-defined. For instance, 1 is an element of the domain, $\mathbb{N}$, but there does not exist an element $y$ from the codomain, also $\mathbb{N}$, such that $y=1/3$.

Sometimes, a function would be defined according to definition 3.2.

**Definition 3.2. (Function, in terms of sets)**
Let $A$, $B$ be non-empty sets. Then a function $f$ is a subset of $A\times B$, which satisfies
1. $\forall a\in A$, $\exists b\in B$ such that $(a,b)\in f$;
2. If $(a,b),(a,b')\in f$, then $b=b'$.

We write $f(a)=b$ if $(a,b)\in f$.

Readers should have already noted that the two conditions in definition 3.2 is actually corresponding to the words “each” and “unique” in the phase “associates each element $a$ in $A$ a unique element $b$ in $B$” in definition 3.1 respectively. Indeed, the difference in the two definitions is just a matter of presentation only.

**Definition 3.3. (Equality of functions)**
Let $f:A\to B$, $g:A\to B$ be functions, then we say $f$ equal to $g$, denoted by $f=g$, if $f(x)=g(x)$ for all $x$ in $A$.

**Definition 3.4. (Direct Image, Inverse Image and Range of a function)**
Let $f:A\to B$ be a function, then for any $X\subseteq A$, we define the **direct image** of $X$ under $f$, $f(X)$, to be

$$f(X)=\{b\in B\,|\,b=f(x)\text{ for some }x\in X\}.$$

For any $Y\subseteq B$, we define the **inverse image** of $Y$ under $f$, $f^{-1}(Y)$, to be

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

In particular, we define the **range** of $f$, denoted by $R(f)$, to be $f(A)$. That is

$$R(f)=\{b\in B\,|\,b=f(a)\text{ for some }a\in A\}.$$

**Example 3.1.**

Determine whether the following functions are well-defined. Give the range of them if so.

1. $f:\mathbb{N}\to\mathbb{N}$, with $f(n)$ defined to be the $n$-th prime number.
2. $g:\mathbb{R}^2\to\mathbb{R}$, with $g(x,y)=x^y$.
3. $h:\mathbb{R}\to\mathbb{Z}$, with $h(x)$ defined to be the largest integer not exceeding $x$.
4. $p:\mathbb{Q}\to\mathbb{Z}$, with $p(\frac{m}{n})=n$.

**Solution.**

Since there exists infinitely many prime numbers, and each prime number is a natural number, $f$ is well-defined, with $R(f)$ being the set of all prime numbers. The function $g$ is not well-defined because $g$ is not defined at (0, 0). Since for any real number $x$, there exists a unique integer $m$ such that $m\le x<m+1$, $h$ is well-defined. Indeed, $h(x)=m$, and $R(h)=\mathbb{Z}$. Finally $p$ is not well-defined because the definition is not clear. For example, $\frac{1}{2}=\frac{2}{4}$, but we can't decide whether $p(\frac{1}{2})=2$ or $p(\frac{1}{2})=p(\frac{2}{4})=4$.

**Theorem 3.1.**
Let $f:A\to B$ be a function, $A_1,A_2\subseteq A$, $B_1,B_2\subseteq B$, we have
1. If $A_1\subseteq A_2$, then $f(A_1)\subseteq f(A_2)$;
2. $f(A_1\cup A_2)=f(A_1)\cup f(A_2)$;
3. $f(A_1\cap A_2)\subseteq f(A_1)\cap f(A_2)$;
4. $A_1\subseteq f^{-1}(f(A_1))$;
5. If $B_1\subseteq B_2$, then $f^{-1}(B_1)\subseteq f^{-1}(B_2)$;
6. $f^{-1}(B_1\cup B_2)=f^{-1}(B_1)\cup f^{-1}(B_2)$;
7. $f^{-1}(B_1\cap B_2)=f^{-1}(B_1)\cap f^{-1}(B_2)$;
8. $f(f^{-1}(B_1))\subseteq B_1$.

**Proof.** We will provide proofs for (1), (3), (8).

For (1), since $\forall y\in f(A_1)$, $\exists x\in A_1\subseteq A_2$ such that $f(x)=y$, $y=f(x)\in f(A_2)$ also. (3) is true because $\forall y\in f(A_1\cap A_2)$, $\exists x\in A_1\cap A_2$ such that $f(x)=y$. As $x\in A_1$, $x\in A_2$, $y=f(x)\subseteq f(A_1)\cap f(A_2)$. For the last one, $\forall y\in f(f^{-1}(B_1))$, $\exists x\in f^{-1}(B_1)$ such that $y=f(x)$. Note that by definition of inverse, $f(x)\in B_1$, so $y\in B_1$. This establishes (8).

Q.E.D.

**Definition 3.5. (Composition of Functions)**
Let $f: A \to B$, $g: C \to D$ be functions. If $B \subset C$, we define the **composite function** $g \circ f: A \to D$ by

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

**Illustration.** Let $f: \mathbb{C} \to \mathbb{R}$, $g: \mathbb{R} \to \mathbb{R}$ be functions defined by $f(x) = |x|$, $g(x) = x^2$, then $g \circ f(x) = g(f(x)) = g(|x|) = |x|^2 = x \cdot \overline{x}$.

- In general, $g \circ f \neq f \circ g$.

**Theorem 3.2.**
Let $f: A \to B, g: B \to C, h: C \to D$ be functions. Then

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

**Proof.** Exercise.

- As a result of theorem 3.2, we can compute the composite function $f_1 \circ (f_2 \circ (...(f_{n-1} \circ f_n)))$ in any order, and so we can simply write it as $f_1 \circ f_2 \circ ... f_{n-1} \circ f_n$

**Theorem 3.3.**
Let $f: A \to B, g: B \to C$ be functions. If $X \subseteq C$,

$$(g \circ f)^{-1}(X) = f^{-1}(g^{-1}(X)).$$

**Proof.** Exercise.

**Definition 3.6. (Injective, Surjective and Bijective functions)**
Let $f: A \to B$ be a function.
If for any $b_1$, $b_2$ in *A*, $f(b_1) = f(b_2) \Rightarrow b_1 = b_2$, *f* is said to be **injective**.
If for any *b* in *B*, there exists *a* in *A* such that $f(a) = b$, *f* is said to be **surjective**.
If a function is both injective and surjective, it is said to be **bijective**.

- From the definition 3.6, we can disprove *f* being injective by giving an example of distinct $a_1$, $a_2$ in *A* such that $f(a_1) = f(a_2)$. If we can find a *b* in *B* such that for any *a* in *A*, $f(a) \neq b$, we can disprove *f* is surjective.

- If $f : A \to B$ is injective, then the function $g : A \to f(A)$ defined by $g(x) = f(x)$ is obviously bijective.

- Sometimes we say *f* is **one-one** if *f* is injective, **onto** if *f* is surjective and **one-one onto** if *f* is bijective. Such a function is called a **injection**, a **surjection** and **bijection** respectively.

**Example 3.2.**

Determine which of the following functions are injective and surjective.

1. $f : \mathbb{Q} \to \mathbb{Q}$, defined by $f(x) = x/2$.
2. $g : \mathbb{R} \to \mathbb{R}$, defined by $g(x) = x^2$.
3. $h : \mathbb{R}^2 \to \mathbb{C}$, defined by $h(x, y) = x + yi$.

**Solution.**

For (1), $f(x_1) = f(x_2) \Rightarrow x_1/2 = x_2/2 \Rightarrow x_1 = x_2$. Also, for any rational number *x*, 2*x* is also a rational number, with $f(2x) = x$. Thus, *f* is both injective and surjective.

For (2), $g(1) = g(-1) = 1$. Also, there is no real number *x* such that $g(x) = -1 \in \mathbb{R}$. Hence, *g* is neither injective nor surjective.

For (3), $h(x_1, y_1) = h(x_2, y_2) \Rightarrow x_1 + y_1 i = x_2 + y_2 i \Rightarrow x_1 = x_2, y_1 = y_2$.Furthermore, $\forall z \in \mathbb{C}$, we have $h(\mathrm{Re}(z), \mathrm{Im}(z)) = z$. Hence, *h* is both injective and surjective.

**Example 3.3.**

Prove that for any non-empty set *A*, there does not exist a bijective function from *A* to *P*(*A*), where *P*(*A*) is the power set of *A*.

**Solution.**

We will prove it by contradiction. Suppose such a function, says *f*, exists. We define a subset *B* of *A* by $B = \{x \in A \mid x \notin f(x)\}$. Since f is surjective, there exists $a \in A$ such that $f(a) = B$. We will prove such *a* does not exist by proving neither "$a \in B$" nor "$a \notin B$" is possible.

If $a \in B$, by the definition of *B*, we have $a \notin f(a) = B$, a contradiction. Conversely, if $a \notin B$, then $a \notin f(a)$. Thus, $a \in B$ by the definition of *B*, again a contradiction.

In example 3.2, if we confine the domain and codomain to the set of all non-negative real numbers, *g* will be bijective. Very often, we can get some useful property if we restrict our attention to a subset of the domain. This motivates the definition 3.7.

**Definition 3.7. (Restriction of Functions)**
Let $f: A \to B$ be a function, $S$ be a subset of $A$. Then the function $f_S: S \to B$, defined by

$$f_s(x) = f(x) \quad \forall x \in S,$$

is called the *restriction* of $f$ to $S$.

- For the sake of simplicity, we may use $f$ instead of $f_S$ for a restriction when it would not cause confusion to the readers.

**Illustration.** The trigonometric functions sin and cos from $\mathbb{R} \to \mathbb{R}$ are not injective. However, the restrictions of them to $[-\pi/2, \pi/2]$ and $[0, \pi]$ respectively are injective. Similarly, the restriction of tan to $(-\pi/2, \pi/2)$ is injective.

**Definition 3.8. (Identity Functions)**
We define the **identity function** $i_A: A \to A$ by

$$i_A(x) = x, \text{ for all } x \text{ in } A.$$

As discussed before, a well-defined function $f: A \to B$ must satisfy two conditions. First, each element $a$ in $A$ is associated to an element $b$ in $B$. Second, this element $b$ is unique. In view of this, if $f$ is surjective, we can consider each element $b$ in $B$ is associated to an element $a$ in $A$ where $f(a) = b$. Furthermore, if $f$ is injective, this element $a$ is unique. Hence, if $f$ is bijective, we may define a function in the other "direction", with the role of domain and codomain interchanged.

**Definition 3.9. (Inverse Functions)**
Let $f: A \to B$ be a bijective function. We define $f^{-1}: B \to A$ by $f^{-1}(b) = a,$ where $a$ is the unique pre-image of $b$.
The function $f^{-1}$ is called the **inverse function** of $f$.

- Clearly, $f^{-1} \circ f = i_A$, $f \circ f^{-1} = i_B$ in definition 3.9. In fact, if there is a function $g: B \to A$ such that $g \circ f = i_A$ and $f \circ g = i_B$, we can conclude $f$ is bijective, with $g = f^{-1}$ conversely.

- The notation of inverse function is very similar to that of inverse image. The only difference is that the "input" of inverse function is an element of codomain of the original function, and that of inverse image is a subset of the codomain. Sometimes, people may even use $f^{-1}(b)$ to represent $f^{-1}(\{b\})$ for a non-bijective function *f*. We should use this only when the meaning is clear from the text.

**Theorem 3.4.**

Let $f: A \to B$, $g: B \to C$ be functions. Then we have

1. If *f*, *g* are injective, $g \circ f$ is injective;
2. If *f*, *g* are surjective, $g \circ f$ is surjective;
3. If *f*, *g* are bijective, $g \circ f$ is bijective, with $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

**Proof.** Exercise.

With bijection defined, we can introduce the concept of cardinal number of a set.

**Definition 3.10. (Cardinal number)**

For any sets *A*, *B*, we say *A* and *B* have the same **cardinal number**, and write *A*~*B* if there is a bijection from *A* to *B*. Furthermore,

1. *A* is said to be **finite** if $A = \varnothing$ or $A \sim J_n$ for some positive integer *n*. In the second case, we may say the cardinal number of *A* is *n*, or simply *A* has *n* elements;
2. *A* is said to be **infinite** if *A* is not finite;
3. *A* is said to be **countable** if *A* is finite or $A \sim \mathbb{N}$;
4. *A* is said to be **uncountable** if *A* is not countable,

where $J_n = \{1, 2, \ldots n\}$ for any positive integer *n*.

- If $A \sim J_m$, $A \sim J_n$, then $m = n$. Hence the cardinal number of a finite set is *unique*.

- In fact, the sets $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, and any infinite subset of them have the same cardinal number, and thus are countable. Examples of uncountable sets are $\mathbb{R}$, $\mathbb{C}$ and $P(\mathbb{N})$.

**Example 3.4.**

Prove that $\mathbb{N} \sim \mathbb{Z}$.

**Solution.**

Define $f: \mathbb{N} \to \mathbb{Z}$ by

$$f(n)=\begin{cases} n/2, & n \text{ is even} \\ -(n-1)/2, & n \text{ is odd} \end{cases}$$

Since $f$ is bijective, $\mathbb{N} \sim \mathbb{Z}$.

**Exercise**

1. Prove that the cardinal number of a finite set is unique.

2. Give an example such that $f: A \to B$, $X, Y \subseteq A$ and $f(X) \cap f(Y) \neq f(X \cap Y)$.

3. Let $f:[0,1] \to (0,1)$ be a function, defined by

$$f(x)=\begin{cases} 1/(n+2), & x=1/n, n \in \mathbb{N} \\ 1/2, & x=0 \\ x, & \text{otherwise} \end{cases}$$

   Show that $f$ is bijective. Hence, $[0,1] \sim (0,1)$.

4. If definition 3.2 is used for function, give the corresponding definition for inverse in a similar way.

5. Give an example of f, g, from $\mathbb{R}$ to $\mathbb{R}$, such that $f \circ g = g \circ f$.

6. Prove theorem 3.2.

7. Prove theorem 3.3.

8. Prove theorem 3.4.

9. Show that for any distinct real number $a$, $b$, $(a,b) \sim [a,b] \sim \mathbb{R}$. (Actually, all non-degenerate intervals, i.e. intervals contains more than one element, have the same cardinality.)

# 4. Further Discussion on Real-valued Functions

In this section, we will introduce several types of real-valued functions and give some of their special properties.

We first introduce some arithmetic operations on functions.

**Definition 4.1. (Arithmetic Operations on Functions)**
For any real number *c*, real-valued functions *f* and *g* with the same domain, we define

1. $cf(x) = c \cdot f(x)$;
2. $(f+g)(x) = f(x)+g(x)$;
3. $(f-g)(x) = f(x)-g(x)$;
4. $(fg)(x) = f(x)\cdot g(x)$.

If $g(x) \neq 0$ for all *x* in the domain, we define

5. $(f/g)(x) = f(x)/g(x)$,

where these real-valued functions $cf$ $f+g$, $f-g$, $fg$, $f/g$ all have the same domain as *f* and *g* does.

**Definition 4.2. (Bounded Functions)**
For any real-valued function *f* defined on *A*, f is said to be **bounded** if there exists an real number *M* such that

$$|f(x)| \leq M \text{ , for all } x \text{ in } A.$$

We say *f* is **bounded above** if for all *x* in *A*, $f(x) \leq M_1$ for some real number $M_1$, and is **bounded below** if for all *x* in *A*, $f(x) \geq M_2$ for some real number $M_2$. $M_1$, $M_2$ are called **upper bound** and **lower bound** respectively.

- A bounded function is both bounded above and bounded below.

- The sum, difference and product of finitely many bounded functions are bounded.

**Illustration.** $f:\mathbb{R}\to\mathbb{R}$ defined by $f(x)=2^x$. $f$ is bounded below by 0. However, $f$ is not bounded above, so is not bounded. However *f* is bounded in any bounded intervals.

**Definition 4.3. (Even, Odd Functions)**
For any real-valued function *f* defined on a **symmetric set** *A* (i.e. $x \in A \Leftrightarrow -x \in A$) in $\mathbb{R}$, *f* is said to be **even** if

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

*f* is said to be **odd** if

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

- An odd function defined at 0 must attain zero there.

**Illustration.** The trigonometric function cos is an even function, and sin is an odd function.

The graph of an even function is symmetric about the *y*-axis.

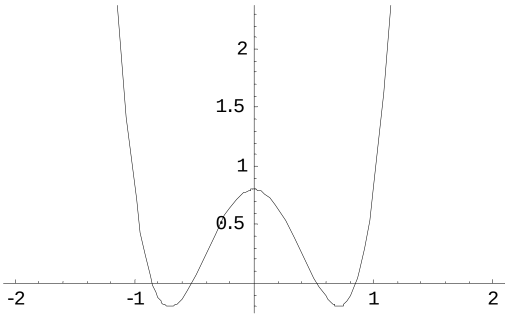


Figure 1: An even function

The left half side of the graph of an odd function looks like to be formed by rotating the right half side of it by $180^\circ$ about the origin.

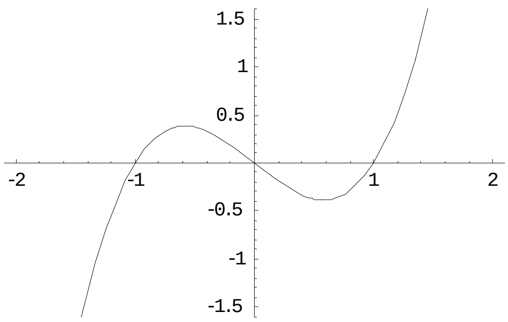


Figure 2: An odd function

**Theorem 4.1.**
Let *p*, *q* be odd functions, *r*, *s* be even functions defined on the same domain, *c* be a constant. Then
1. $-p$, $-q$ are odd, $-r$, $-s$ are even;
2. $cp$ is odd, $cr$ is even;
3. $p \pm q$ is odd, $r \pm s$ is even;
4. $pq$ and $rs$ are even, $pr$ is odd;
5. $1/p$ is odd, $1/r$ is even if they are defined.

**Proof.** Exercise.

**Definition 4.4. (Monotonic Functions)**
Let $f : A \to B$ where $A, B \subseteq \mathbb{R}$. We say that *f* is **increasing** if
$$f(x) \le f(y), \quad \forall x, y \in A \text{ with } x < y.$$
If $f(x) \le f(y)$ is replaced by $f(x) < f(y)$, *f* is said to be **strictly increasing** instead.
**Decreasing** and **strictly decreasing** functions are defined in similar way. All these functions are said to be **monotonic**.

**Illustration.** $x^3$ is strictly increasing on $\mathbb{R}$, while $1/x$ is strictly decreasing on $(0,\infty)$ The trigonometric function cos is an even function, and sin is an odd function.

If f is strictly increasing (decreasing), f must be injective. Hence, the function $g : A \to f(A)$ defined by $g(x) = f(x)$ is bijective and inverse exists. Actually the inverse function preserves this property.

**Theorem 4.2.**
If $A, B \subseteq \mathbb{R}$, $f : A \to B$ is strictly increasing (decreasing), and $g : A \to f(A)$ is defined by $g(x) = f(x)$, then g is bijective, and both *g* and $g^{-1}$ is strictly increasing (decreasing).

**Proof.** We have already shown that *g* is bijective, and *g* being strictly increasing is trivial. It remains to show $g^{-1}$ being strictly increasing. For all $x < y$ in *f*(*A*), there exist *a*, *b* such that $f(a) = x$, $f(b) = y$. $a \ge b$ would lead to a contradiction that $x = f(a) \ge f(b) = y$. Hence $g(x) = a < b = g(y)$.

Q.E.D.

**Definition 4.5. (Periodic Functions)**
If $f:\mathbb{R}\to\mathbb{R}$ satisfies $f(x+p)=f(x)$ for all real number *x*, *f* is said to be a periodic function, and *p* is called a period of the *f*.

- Clearly if *p* is a period of *f*, *mp* is also a period for any positive integer *m*.

**Illustration.** The trigonometric functions cos and sin are periodic functions with $2\pi$ being a common period of them.

**Exercise**

1. Prove theorem 4.1.

2. If $f:\mathbb{R}\to\mathbb{R}$ is both even and odd, prove that $f\equiv 0$.

## 5. Relations

We know that we can classify the set of integers into two subsets by their parities, namely odd numbers and even numbers. We call the collection of these two subsets to be a **partition** of integers. More generally, given an arbitrary set, we can define a partition by:

**Definition 5.1. (Partition)**
Given a non-empty set *A*, a partition of *A* is a collection of *disjoint* subsets of *A* whose *union is A*.

**Illustration.** $\{\mathbb{Z}^+,\mathbb{Z}^-,\{0\}\}$ is a partition of $\mathbb{Z}$.

**Illustration.** For each positive integer *q*, define $A_q=\left\{\frac{p}{q}\in\mathbb{Q}\mid p\in\mathbb{N},\gcd(p,q)=1\right\}$. Then $\{A_n\mid n\in\mathbb{N}\}$ is a partition of the set of all positive rational numbers.

With a partition of *A* defined, we see every element $a\in A$ belongs to one and only one subset in the partition. Therefore, we may consider a partition as a classification of elements of a set. Normally, we would classify the elements of a set, or define a partition on it in the sense that the elements of the same subset should share some common properties. A nice way to express this idea mathematically is by the concept of **equivalence relation**. Before explaining what it is, first one has to introduce what a **relation** on a set means.

**Definition 5.2. (Relation)**
A **relation** *R* on a non-empty set *A* is a subset of $A \times A$. For any elements $x, y \in A$, we say *x* is related to *y*, or $xRy$ if $(x, y) \in R$ and $x \not R y$ if $(x, y) \notin R$.

- It may be inconvenient to define a relation as a subset of the Cartesian product. Mostly we would just express it by a statement instead. For example, we define a relation ~ on $\mathbb{Z}$ by $x \sim y$ if $x - y = 2k$ for some integer *k*. It is equivalent to define ~ to be the set $\{(x, y\} \in \mathbb{Z}^2 \mid x - y = 2k, k \in \mathbb{Z}\}$.

- Notice that the order is important in general. That means $xRy$ does not necessarily imply $yRx$.

**Illustration.** Define a relation *R* on $\mathbb{R}$ by $xRy$ if $x - y \in \mathbb{N}$. Then we have $3R1$, but $1 \not R 3$.

As said before, usually we would define relations in a meaningful manner for some purpose. Two particularly important relations will be discussed below.

**Definition 5.3. (Equivalence Relation)**
A relation *R* on a non-empty set *A* is called an equivalence relation if
1. $xRx$ for all *x* in *A*. (Reflexive)
2. $xRy \Rightarrow yRx$. (Symmetric)
3. $xRy, yRz \Rightarrow xRz$ (Transitive)

**Example 5.1.**

Verify that the relation ~ on $\mathbb{R}$, defined by $x \sim y$ if $x - y \in \mathbb{Z}$, is an equivalence relation. How about if we change the condition to $x - y \in \mathbb{N}$

**Solution.**

To prove it, just notice that $\forall x, y, z \in \mathbb{R}$, $x - x = 0 \in \mathbb{Z}$, $x - y \in \mathbb{Z} \Rightarrow y - x = -(x - y) \in \mathbb{Z}$ and $x - y, y - z \in \mathbb{Z} \Rightarrow x - z = (x - y) + (y - z) \in \mathbb{Z}$. Hence the three properties in definition 5.3 hold. As a result, ~ is an equivalence relation. However, if the condition is changed to $x-y\in\mathbb{N}$, then the symmetric property would not hold because $2\sim 1$ but $1\nsim 2$ and it cannot be an equivalence relation.

**Example 5.2.**

Define a relation "~" on $\mathbb{N}^2$ with $(a_1,a_2)\sim(b_1,b_2)$ if $a_1+b_2=a_2+b_1$. Show that it is an equivalence relation.

**Solution.**

Since $a_1+a_2=a_2+a_1$ for any $a_1,a_2\in\mathbb{N}$, $(a_1,a_2)\sim(a_1,a_2)$ and ~ is reflexive. Also, if $(a_1,a_2)\sim(b_1,b_2)$, then $a_1+b_2=a_2+b_1 \Rightarrow b_1+a_2=b_2+a_1 \Rightarrow (b_1,b_2)\sim(a_1,a_2)$. Hence it is symmetric. Finally, if $(a_1,a_2)\sim(b_1,b_2)$, $(b_1,b_2)\sim(c_1,c_2)$, then $a_1+b_2=a_2+b_1$, $b_1+c_2=b_2+c_1$. Therefore, $(a_1+c_2)+(b_1+b_2)=(a_1+b_2)+(b_1+c_2)=(a_2+b_1)+(b_2+c_1)=(a_2+c_1)+(b_1+b_2)$. By cancellation law of addition of natural numbers, we get $a_1+c_2=a_2+c_1$. So ~ is transitive and thus is an equivalence relation.

For each $x\in[0,1)$, let $A_x=\{x+n\,|\,n\in\mathbb{Z}\}$ }. It is clear $\{A_x\,|\,x\in[0,1)\}$ form a partition of $\mathbb{R}$. Also, for any real numbers *a* and *b*, we have $a\sim b$, where ~ is the equivalence relation defined in example 5.1, if and only if *a*, *b* are in the same $A_x$ for some $x\in[0,1)$. The reason that equivalence relation is so important is that it induces a partition on a set. Indeed, there is a one-one correspondence between the two.

**Theorem 5.1.**
Given an equivalence relation *R* on a non-empty set *A*. For each $a\in A$, define $\langle a\rangle=\{x\in A\,|\,xRa\}$. Then $\{\langle a\rangle\,|\,a\in A\}$ forms a partition of *A*. We say that the partition is induced by *R*.
Conversely, if a partition of *A* is given, the relation on *A* defined by $x\sim y$ if *x*, *y* belong to the same subset of the partition is an equivalence relation.
Here, $\langle a\rangle=\{x\in A\,|\,xRa\}$ is called the **equivalence class** of *a*, and the set $\{\langle a\rangle\,|\,a\in A\}$ is called the **quotient set of *A* with respect to *R***, denoted by $A/R$.

**Proof.** We will prove the first half and leave the proof of the second half to the readers as exercise. By the reflexive property, $\forall a\in A, a\in\langle a\rangle\in\{\langle a\rangle\,|\,a\in A\}$. Hence, the union of the elements of collection is *A*. To show its elements are disjoint is equivalent to showing if $a,b\in A$, either $\langle a\rangle=\langle b\rangle$ or $\langle a\rangle\cap\langle b\rangle=\varnothing$.

Suppose $\langle a\rangle\cap\langle b\rangle\neq\varnothing$, then $\exists c\in A$ such that $c\in\langle a\rangle, c\in\langle b\rangle$ or $cRa, cRb$. By symmetric property, $aRc$. Now $\forall x\in\langle a\rangle$, by transitive property, $xRa, aRc, cRb\Rightarrow xRb\Rightarrow x\in\langle b\rangle$, that is, $\langle a\rangle\subseteq\langle b\rangle$. Similarly, $\langle b\rangle\subseteq\langle a\rangle$. Hence $\langle a\rangle=\langle b\rangle$ and we get the conclusion.

Q.E.D.

- Equivalence relation is called **congruence relation** sometimes, and one may use $a\equiv b$ to represent $a\sim b$ in this case. Symbols like $\dot{a}$ and $\overline{a}$ are commonly used to denote $\langle a\rangle$. Any element in $\langle a\rangle$ is called a **representative** of $\langle a\rangle$.

Since ~ in example 5.2 is an equivalence relation, it induces a partition of $\mathbb{N}^2$. It is an important and interesting example because each equivalence class $\langle(a,b)\rangle$ formed can be considered to represent the integer $a-b$. It may seem strange, but it is exactly how the number system is built up. After defining natural number and its addition and multiplication in our mathematical system (which is based on sets actually), we will try to extend it to integers, rational numbers, real numbers and then complex numbers stepwisely.

Readers should be careful that we cannot define subtraction in natural number since $1-2$ should not be something in $\mathbb{N}$. As a result, we have to avoid using subtraction in constructing integers from natural numbers, as we did in example 5.2. With the idea that $\langle(a,b)\rangle$ represent $a-b$ in mind, it is not difficult to expect we should define "addition" and "multiplication" on these equivalence class by

$$\langle(a_1,b_1)\rangle+\langle(a_2,b_2)\rangle=\langle(a_1+a_2,b_1+b_2)\rangle$$
$$\langle(a_1,b_1)\rangle\cdot\langle(a_2,b_2)\rangle=\langle(a_1a_2+b_1b_2,a_1b_2+a_2b_1)\rangle$$

One important point to notice is that, in defining operations on equivalence classes by its representatives, we must ensure that the outcome is *independent of the choice of the representative* in the classes. Take the addition above as an example. For the operation to be *well-defined*, one should check that

$$(a_1,b_1)\sim(a_3,b_3),(a_2,b_2)\sim(a_4,b_4)\Rightarrow\langle(a_1,b_1)\rangle+\langle(a_2,b_2)\rangle=\langle(a_3,b_3)\rangle+\langle(a_4,b_4)\rangle$$

It is true, since

$$\begin{aligned}&a_1+b_3=a_3+b_1, a_2+b_4=a_4+b_2\\&\Rightarrow(a_1+a_2)+(b_3+b_4)=(a_3+a_4)+(b_1+b_2)\\&\Rightarrow\langle(a_1+a_2,b_1+b_2)\rangle=\langle(a_3+a_4,b_3+b_4)\rangle\end{aligned}$$

The checking on multiplication is left as an exercise.

Equivalence relation can be viewed as a generalization of "equal". The triangles with side length 3, 4, 5 and 6, 8, 10 are not equal, but is related under the equivalence relation of similarity. If we just want to focus on properties which are invariant under similarity, like ratio of sides, we may simply consider the two triangles as the same thing. Equivalence relation appears in nearly every branch of mathematics and is a powerful tool.

**Exercise**

1. Recall that two triangles are similar if they have same interior angles, prove that the similarity of triangles is indeed an equivalence relation.

2. Complete the proof of theorem 5.1.

3. Check that the multiplication defined by
$$\langle(a_1,b_1)\rangle\cdot\langle(a_2,b_2)\rangle=\langle(a_1a_2+b_1b_2,a_1b_2+a_2b_1)\rangle$$
on the quotient set $\mathbb{N}^2/\sim$ in example 5.2 is a well-defined operation. You may use cancellation law of addition for natural numbers (i.e. $a+b=a+c\Rightarrow b=c$), but subtraction should never appear in your working.

## 6. Solutions to Selected Exercise

Functions

2. Let $f(x)=x^2$, $A=[-1,0]$, $B=[0,1]$, then $f(A\cap B)=\{0\}\neq[0,1]=f(A)\cap f(B)$.

4. If $f$ is bijective, the inverse function $f^{-1}$ of $f$ is defined to be the subset of $B\times A$, with $f^{-1}=\{(b,a)\mid(a,b)\in f\}$.

6. Theorem 3.2 is true because for all $x$ in $A$, we have
$$\begin{aligned}h\circ(g\circ f)(x)&=h(g\circ f(x))\\&=h(g(f(x)))\\&=h\circ g(f(x))\\&=(h\circ g)\circ f(x)\end{aligned}$$

9. Let $f:[0,1]\to[a,b]$, $g:(0,1)\to(a,b)$, $h:(0,1)\to\mathbb{R}$ defined by $f(x)=a+x(b-a)$,

$g(x) = a + x(b - a)$ , $h(x) = \tan[\pi(x - 0.5)]$ . Clearly, all the three functions are bijective. Therefore, by theorem 3.4 and exercise 3.3, we have $[a,b] \sim [0,1] \sim (a,b) \sim (0,1) \sim \mathbb{R}$ .