# 1 Elements of Set Theory

## 1.1 Sets

Intuitively speaking, a "set" is a collection of objects.[1] The distinguishing feature of a "set" is that while it may contain numerous objects, it is nevertheless conceived as a *single* entity. In the words of Georg Cantor, the great founder of abstract set theory, "a set is a Many which allows itself to be thought of as a One." It is amazing how much follows from this simple idea.

The objects that a set $S$ contains are called the "elements" (or "members") of $S$. Clearly, to know $S$, it is necessary and sufficient to know all elements of $S$. The principal concept of set theory is, then, the relation of "being an element/member of." The universally accepted symbol for this relation is $\in$, that is, $x \in S$ (or $S \ni x$) means that "$x$ is an element of $S$" (also read "$x$ is a member of $S$," or "$x$ is contained in $S$," or "$x$ belongs to $S$," or "$x$ is in $S$," or "$S$ includes $x$," etc.). We often write $x, y \in S$ to denote that both $x \in S$ and $y \in S$ hold. For any natural number $m$, a statement like $x_1, ..., x_m \in S$ (or equivalently, $x_i \in S$, $i = 1, ..., m$) is understood analogously. If $x \in S$ is a false statement, then we write $x \notin S$, and read "$x$ is not an element of $S$."

If the sets $A$ and $B$ have exactly the same elements, that is, $x \in A$ iff $x \in B$, then we say that $A$ and $B$ are identical, and write $A = B$, otherwise we write $A \neq B$.[2] (So, for instance, $\{x, y\} = \{y, x\}$, $\{x, x\} = \{x\}$, and $\{\{x\}\} \neq \{x\}$.) If every member of $A$ is also a member of $B$, then we say that $A$ is a **subset** of $B$ (also read "$A$ is a set in $B$," or "$A$ is contained in $B$") and write $A \subseteq B$ (or $B \supseteq A$). Clearly, $A = B$ holds iff both $A \subseteq B$ and $B \subseteq A$ hold. If $A \subseteq B$ but $A \neq B$, then $A$ is said to be a **proper subset** of $B$, and we denote this situation by writing $A \subset B$ (or $B \supset A$).

For any set $S$ that contains finitely many elements (in which case we say $S$ is **finite**), we denote by $|S|$ the total number of elements that $S$ contains, and refer to this number as the **cardinality** of $S$. We say that $S$ is a **singleton** if $|S| = 1$. If $S$ contains infinitely many elements (in which case we say $S$ is **infinite**), then we write $|S| = \infty$. Obviously, we have $|A| \leq |B|$ whenever $A \subseteq B$, and if $A \subset B$ and $|A| < \infty$, then $|A| < |B|$.

We sometimes specify a set by enumerating its elements. For instance, $\{x, y, z\}$ is the set that consists of the objects $x$, $y$ and $z$. The contents of the sets $\{x_1, ..., x_m\}$ and $\{x_1, x_2, ...\}$ are similarly described. For example, the set $\mathbb{N}$ of positive integers can be written as $\{1, 2, ...\}$. Alternatively, one may describe a set $S$ as a collection of all objects $x$ that satisfy a given property $P$. If $P(x)$ stands for the (logical) statement "$x$ satisfies the property $P$," then we can write $S = \{x : P(x) \text{ is a true statement}\}$ or simply $S = \{x : P(x)\}$. If $A$ is a set and $B$ is the set that contains all elements $x$

[1] The notion of an "object" is left undefined, that is, it can be given any meaning. All we demand of our "objects" is that they be logically *distinguishable*. That is, if $x$ and $y$ are two objects, $x = y$ and $x \neq y$ cannot hold simultaneously, and that the statement "either $x = y$ or $x \neq y$" is a tautology.

[2] *Reminder.* iff = if and only if.

of $A$ such that $P(x)$ is true, we write $B = \{x \in A : P(x)\}$. For instance, where $\mathbb{R}$ is the set of all real numbers, the collection of all real numbers greater than or equal to 3 can be written as $\{x \in \mathbb{R} : x \geq 3\}$.

The symbol $\emptyset$ denotes the **empty set**, that is, the set that contains no elements (i.e. $|\emptyset| = 0$). Formally speaking, we can define $\emptyset$ as the set $\{x : x \neq x\}$; for this description entails that $x \in \emptyset$ is a false statement for any object $x$. Consequently, we write

$$\emptyset := \{x : x \neq x\},$$

meaning that the symbol on the left hand side is *defined* by that on the right hand side.[3] Clearly, we have $\emptyset \subseteq S$ for any set $S$, which, in particular, implies that $\emptyset$ is unique. (Why?) If $S \neq \emptyset$, we say that $S$ is **nonempty**. For instance, $\{\emptyset\}$ is a nonempty set. Indeed, $\{\emptyset\} \neq \emptyset$ – the former, after all, is a set of sets that contains the empty set, while $\emptyset$ contains nothing. (An empty box is not the same thing as nothing!)

We define the class of all subsets of a given set $S$ as

$$2^S := \{T : T \subseteq S\},$$

which is called the **power set** of $S$. (The choice of notation is motivated by the fact that the power set of a set that contains $m$ elements has exactly $2^m$ elements.) For instance, $2^\emptyset = \{\emptyset\}$, $2^{2^\emptyset} = \{\emptyset, \{\emptyset\}\}$, and $2^{2^{2^\emptyset}} = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\}$, and so on.

**Notation.** Throughout this text, the class of all nonempty *finite* subsets of any given set $S$ is denoted by $\mathcal{P}(S)$, that is,

$$\mathcal{P}(S) := \{T : T \subseteq S \text{ and } 0 < |T| < \infty\}.$$

Of course, if $S$ is finite, then $\mathcal{P}(S) = 2^S \backslash \{\emptyset\}$.

Given any two sets $A$ and $B$, by $A \cup B$ we mean the set $\{x : x \in A \text{ or } x \in B\}$ which is called the **union** of $A$ and $B$. The **intersection** of $A$ and $B$, denoted as $A \cap B$, is defined as the set $\{x : x \in A \text{ and } x \in B\}$. If $A \cap B = \emptyset$, we say that $A$ and $B$ are **disjoint**. Obviously, if $A \subseteq B$, then $A \cup B = B$ and $A \cap B = A$. In particular, $\emptyset \cup S = S$ and $\emptyset \cap S = \emptyset$ for any set $S$.

Taking unions and intersections are *commutative* operations in the sense that

$$A \cap B = B \cap A \quad \text{and} \quad A \cup B = B \cup A$$

for any sets $A$ and $B$. They are also *associative*, that is,

$$A \cap (B \cap C) = (A \cap B) \cap C \quad \text{and} \quad A \cup (B \cup C) = (A \cup B) \cup C,$$

[3] Recall my notational convention: For any symbols $\clubsuit$ and $\heartsuit$, either one of the expressions $\clubsuit := \heartsuit$ and $\heartsuit =: \clubsuit$ means that $\clubsuit$ is defined by $\heartsuit$.

and *distributive*, that is,

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \quad \text{and} \quad A \cup (B \cap C) = (A \cup B) \cap (A \cup C),$$

for any sets $A$, $B$ and $C$.

*Exercise 1.* Prove the commutative, associative and distributive laws of set theory stated above.

***Exercise 2.*** Given any two sets $A$ and $B$, by $A \backslash B$ – the **difference** between $A$ and $B$ – we mean the set $\{x : x \in A \text{ and } x \notin B\}$.
($a$) Show that $S \backslash \emptyset = S$, $S \backslash S = \emptyset$, and $\emptyset \backslash S = \emptyset$ for any set $S$.
($b$) Show that $A \backslash B = B \backslash A$ iff $A = B$ for any sets $A$ and $B$.
($c$) (*De Morgan Laws*) Prove: For any sets $A$, $B$ and $C$,

$$A \backslash (B \cup C) = (A \backslash B) \cap (A \backslash C) \quad \text{and} \quad A \backslash (B \cap C) = (A \backslash B) \cup (A \backslash C).$$

Throughout this text we use the terms "class" or "family" only to refer to a nonempty collection of sets. So if $\mathcal{A}$ is a class, we understand that $\mathcal{A} \neq \emptyset$ and that any member $A \in \mathcal{A}$ is a set (which may itself be a collection of sets). The union of all members of this class, denoted as $\bigcup \mathcal{A}$, or $\bigcup \{A : A \in \mathcal{A}\}$, or $\bigcup_{A \in \mathcal{A}} A$, is defined as the set $\{x : x \in A \text{ for some } A \in \mathcal{A}\}$. Similarly, the intersection of all sets in $\mathcal{A}$, denoted as $\bigcap \mathcal{A}$, or $\bigcap \{A : A \in \mathcal{A}\}$, or $\bigcap_{A \in \mathcal{A}} A$, is defined as the set $\{x : x \in A$ for each $A \in \mathcal{A}\}$.

A common way of specifying a class $\mathcal{A}$ of sets is by designating a set $I$ as a set of indices, and by defining $\mathcal{A} := \{A_i : i \in I\}$. In this case, $\bigcup \mathcal{A}$ may be denoted as $\bigcup_{i \in I} A_i$. If $I = \{k, k+1, ..., K\}$ for some integers $k$ and $K$ with $k < K$, then we often write $\bigcup_{i=k}^{K} A_i$ (or $A_k \cup \cdots \cup A_K$) for $\bigcup_{i \in I} A_i$. Similarly, if $I = \{k, k+1, ...\}$ for some integer $k$, then we may write $\bigcup_{i=k}^{\infty} A_i$ (or $A_k \cup A_{k+1} \cup \cdots$) for $\bigcup_{i \in I} A_i$. Furthermore, for brevity, we frequently denote $\bigcup_{i=1}^{K} A_i$ as $\bigcup^K A_i$, and $\bigcup_{i=1}^{\infty} A_i$ as $\bigcup^\infty A_i$, throughout the text. Similar notational conventions apply to intersections of sets as well.

**Warning.** The symbols $\bigcup \emptyset$ and $\bigcap \emptyset$ are left *undefined* (much the same way the symbol $0/0$ is undefined in number theory).

*Exercise 3.* Let $A$ be a set and $\mathcal{B}$ a class of sets. Prove that

$$A \cap \bigcup \mathcal{B} = \bigcup \{A \cap B : B \in \mathcal{B}\} \quad \text{and} \quad A \cup \bigcap \mathcal{B} = \bigcap \{A \cup B : B \in \mathcal{B}\},$$

while

$$A \backslash \bigcup \mathcal{B} = \bigcap \{A \backslash B : B \in \mathcal{B}\} \quad \text{and} \quad A \backslash \bigcap \mathcal{B} = \bigcup \{A \backslash B : B \in \mathcal{B}\}.$$

A word of caution may be in order before we proceed further. While duly intuitive, the "set theory" we outlined so far provides us with no demarcation criterion for identifying what exactly constitutes a "set." This may suggest that one is completely free in deeming any given collection of objects as a "set." But, in fact, this would be a pretty bad idea that would entail serious foundational difficulties. The best known example of such difficulties was given by Bertrand Russell in 1902 when he asked if the set of all objects that are not members of *themselves* is a set: Is $S := \{x : x \notin x\}$ a set?[4] There is nothing in our intuitive discussion above that forces us to conclude that $S$ is not a set; it is a collection of objects (sets in this case) which is considered as a single entity. But we cannot accept $S$ as a "set," for if we do, we have to be able to answer the question: Is $S \in S$? If the answer is yes, then $S \in S$, but this implies $S \notin S$ by definition of $S$. If the answer is no, then $S \notin S$, but this implies $S \in S$ by definition of $S$. That is, we have a contradictory state of affairs no matter what! This is the so-called *Russell's paradox* which started a severe foundational crisis for mathematics that eventually led to a complete axiomatization of set theory in the early twentieth century.[5]

Roughly speaking, this paradox would arise only if we allowed "unduly large" collections to be qualified as "sets." In particular, it will not cause any harm for the mathematical analysis that will concern us here, precisely because in all of our discussions, we will *fix* a universal set of objects, say $X$, and consider sets like $\{x \in X : P(x)\}$, where $P(x)$ is an unambiguous logical statement in terms of $x$. (We will also have occasion to work with sets of such sets, and sets of sets of such sets, and so on.) Once such a domain $X$ is fixed, Russell's paradox cannot arise. Why, you may ask, can't we have the same problem with the set $S := \{x \in X : x \notin x\}$? No, because now we can answer the question "Is $S \in S$?". The answer is no! The statement $S \in S$ is false, simply because $S \notin X$. (For, if $S \in X$ was the case, then we would end up with the contradiction $S \in S$ iff $S \notin S$.)

So when the context is clear (that is, when a universe of objects is fixed), and when we define our sets as just explained, Russell's paradox will not be a threat against the resulting set theory. But can there be any other paradoxes? Well, there is really not an easy answer to this. To even discuss the matter unambiguously, we must leave our intuitive understanding of the notion of "set," and address the problem through a completely axiomatic approach (in which we would leave the expression "$x \in S$" as undefined, and give meaning to it *only* through axioms). This is, of course, not

[4] While a bit unorthodox, $x \in x$ may well be a statement that is true for some objects. For instance, the collection of all sets that I have mentioned in my life, say $x$, is a set that I have just mentioned, so $x \in x$. But the collection of all cheesecakes I have eaten in my life, say $y$, is not a cheesecake, so $y \notin y$.

[5] Russell's paradox is a classical example of the dangers of using *self-referential* statements carelessly. Another example of this form is the ancient *paradox of the liar*: "Everything I say is false." This statement can be declared neither true nor false! To get a sense of some other kinds of paradoxes and the way axiomatic set theory avoids them, you might want to read the popular account of Rucker (1995).

at all the place to do this. Moreover, the "intuitive" set theory that we covered here is more than enough for the mathematical analysis to come. We thus leave this topic by referring the reader who wishes to get a broader introduction to abstract set theory to Chapter 1 of Schechter (1997) or Marek and Mycielski (2001); both of these expositions provide nice introductory overviews of axiomatic set theory. If you want to dig deeper, then try the first three chapters of Enderton (1977).

## 1.2 Relations

An **ordered pair** is an ordered list $(a, b)$ consisting of two objects $a$ and $b$. This list is *ordered* in the sense that, as a defining feature of the notion of an ordered pair, we assume the following: For any two ordered pairs $(a, b)$ and $(a', b')$, we have $(a, b) = (a', b')$ iff $a = a'$ and $b = b'$.[6]

The (**Cartesian**) **product** of two nonempty sets $A$ and $B$, denoted as $A \times B$, is defined as the set of all ordered pairs $(a, b)$ where $a$ comes from $A$ and $b$ comes from $B$. That is,

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

As a notational convention, we often write $A^2$ for $A \times A$. It is easily seen that taking the Cartesian product of two sets is not a commutative operation. Indeed, for any two *distinct* objects $a$ and $b$, we have $\{a\} \times \{b\} = \{(a, b)\} \neq \{(b, a)\} = \{b\} \times \{a\}$. Formally speaking, it is not associative either, for $(a, (b, c))$ is not the same thing as $((a, b), c)$. Yet there is a natural correspondence between the elements of $A \times (B \times C)$ and $(A \times B) \times C$, so one can really think of these two sets as the same, thereby rendering the status of the set $A \times B \times C$ unambiguous.[7] This prompts us to define an $n$-**vector** (for any natural number $n$) as a list $(a_1, ..., a_n)$ with the understanding that $(a_1, ..., a_n) = (a'_1, ..., a'_n)$ iff $a_i = a'_i$ for each $i = 1, ..., n$. The (**Cartesian**) **product** of $n$ sets $A_1, ..., A_n$, is then defined as

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

We often write $\mathsf{X}^n A_i$ to denote $A_1 \times \cdots \times A_n$, and refer to $\mathsf{X}^n A_i$ as the $n$-**fold product**

[6] This defines the notion of an ordered pair as a new "primitive" for our set theory, but in fact, this is not really necessary. One can define an ordered pair by using only the concept of "set" as $(a, b) := \{\{a\}, \{a, b\}\}$. With this definition, which is due to Kazimierz Kuratowski, one can *prove* that, for any two ordered pairs $(a, b)$ and $(a', b')$, we have $(a, b) = (a', b')$ iff $a = a'$ and $b = b'$. The "if" part of the claim is trivial. To prove the "only if" part, observe that $(a, b) = (a', b')$ entails that either $\{a\} = \{a'\}$ or $\{a\} = \{a', b'\}$. But the latter equality may hold only if $a = a' = b'$, so we have $a = a'$ in all contingencies. Therefore, $(a, b) = (a', b')$ entails that either $\{a, b\} = \{a\}$ or $\{a, b\} = \{a, b'\}$. The latter case is possible only if $b = b'$, while the former possibility arises only if $a = b$. But if $a = b$, then we have $\{\{a\}\} = (a, b) = (a, b') = \{\{a\}, \{a, b'\}\}$ which holds only if $\{a\} = \{a, b'\}$, that is, $b = a = b'$.

*Quiz.* (Wiener) Show that we would also have $(a, b) = (a', b')$ iff $a = a'$ and $b = b'$, if we instead defined $(a, b)$ as $\{\{\emptyset, \{a\}\}, \{\{b\}\}\}$.

[7] What is this "natural" correspondence?

of $A_1, ..., A_n$. If $A_i = S$ for each $n$, we then write $S^n$ for $A_1 \times \cdots \times A_n$, that is, $S^n := \mathsf{X}^n S$.

*Exercise 4.* For any sets $A$, $B$, and $C$, prove that

$$A \times (B \cap C) = (A \times B) \cap (A \times C) \quad \text{and} \quad A \times (B \cup C) = (A \times B) \cup (A \times C).$$

Let $X$ and $Y$ be two nonempty sets. A subset $R$ of $X \times Y$ is called a (*binary*) **relation from** $X$ **to** $Y$. If $X = Y$, that is, if $R$ is a relation from $X$ to $X$, we simply say that it is a **relation on** $X$. Put differently, $R$ is a relation on $X$ iff $R \subseteq X^2$. If $(x, y) \in R$, then we think of $R$ as associating the object $x$ with $y$, and if $\{(x, y), (y, x)\} \cap R = \emptyset$, we understand that there is no connection between $x$ and $y$ as envisaged by $R$. In concert with this interpretation, we adopt the convention of writing $xRy$ instead of $(x, y) \in R$ throughout this text.

DEFINITION. A relation $R$ on a nonempty set $X$ is said to be **reflexive** if $xRx$ for each $x \in X$, and **complete** if either $xRy$ or $yRx$ holds for each $x, y \in X$. It is said to be **symmetric** if, for any $x, y \in X$, $xRy$ implies $yRx$, and **antisymmetric** if, for any $x, y \in X$, $xRy$ and $yRx$ imply $x = y$. Finally, we say that $R$ is **transitive** if $xRy$ and $yRz$ imply $xRz$ for any $x, y, z \in X$.

The interpretations of these properties are straightforward, so we do not elaborate on them here. But note: While every complete relation is reflexive, there are no other logical implications between these properties.

*Exercise 5.* Let $X$ be a nonempty set, and $R$ a relation on $X$. The **inverse** of $R$ is defined as the relation $R^{-1} := \{(x, y) \in X^2 : yRx\}$.
($a$) If $R$ is symmetric, does $R^{-1}$ have to be also symmetric? Antisymmetric? Transitive?
($b$) Show that $R$ is symmetric iff $R = R^{-1}$.
($c$) If $R_1$ and $R_2$ are two relations on $X$, the **composition** of $R_1$ and $R_2$ is the relation $R_2 \circ R_1 := \{(x, y) \in X^2 : xR_1z \text{ and } zR_2y \text{ for some } z \in X\}$. Show that $R$ is transitive iff $R \circ R \subseteq R$.

*Exercise 6.* A relation $R$ on a nonempty set $X$ is called **circular** if $xRz$ and $zRy$ imply $yRx$ for any $x, y, z \in X$. Prove that $R$ is reflexive and circular iff it is reflexive, symmetric and transitive.

***Exercise 7.***[H] Let $R$ be a reflexive relation on a nonempty set $X$. The **asymmetric part** of $R$ is defined as the relation $P_R$ on $X$ as $xP_Ry$ iff $xRy$ but not $yRx$. The relation $I_R := R \backslash P_R$ on $X$ is then called the **symmetric part** of $R$.
(***a***) Show that $I_R$ is reflexive and symmetric.
(***b***) Show that $P_R$ is neither reflexive nor symmetric.
(***c***) Show that if $R$ is transitive, so are $P_R$ and $I_R$.

***Exercise 8***. Let $R$ be a relation on a nonempty set $X$. Let $R_0 = R$, and for each positive integer $m$, define the relation $R_m$ on $X$ by $xR_my$ iff there exist $z_1, ..., z_m \in X$ such that $xRz_1$, $z_1Rz_2$, ..., $z_{m-1}Rz_m$ and $z_mRy$. The relation $tr(R) := R_0 \cup R_1 \cup \cdots$ is called the **transitive closure** of $R$. Show that $tr(R)$ is transitive, and if $R'$ is a transitive relation with $R \subseteq R'$, then $tr(R) \subseteq R'$.

## 1.3 Equivalence Relations

In mathematical analysis, one often needs to "identify" two distinct objects when they possess a particular property of interest. Naturally, such an identification scheme should satisfy certain consistency conditions. For instance, if $x$ is identified with $y$, then $y$ must be identified with $x$. Similarly, if $x$ and $y$ are deemed identical, and so are $y$ and $z$, then $x$ and $z$ should be identified. Such considerations lead us to the notion of *equivalence relation*.

DEFINITION. A relation $\sim$ on a nonempty set $X$ is called an **equivalence relation** if it is reflexive, symmetric and transitive. For any $x \in X$, the **equivalence class** of $x$ relative to $\sim$ is defined as the set

$$[x]_\sim := \{y \in X : y \sim x\}.$$

The class of all equivalence classes relative to $\sim$, denoted as $X/_\sim$, is called the **quotient set** of $X$ relative to $\sim$, that is,

$$X/_\sim := \{[x]_\sim : x \in X\}.$$

Let $X$ denote the set of all people in the world. "Being a sibling of" is an equivalence relation on $X$ (provided that we adopt the convention of saying that any person is a sibling of himself). The equivalence class of a person relative to this relation is the set of all of his/her siblings. On the other hand, you would probably agree that "being in love with" is not an equivalence relation on $X$. Here are some more examples (that fit better with the "serious" tone of this course).

EXAMPLE 1. [**1**] For any nonempty set $X$, the **diagonal relation** $D_X := \{(x, x) : x \in X\}$ is the smallest equivalence relation that can be defined on $X$ (in the sense that if $R$ is any other equivalence relation on $X$, we have $D_X \subseteq R$). Clearly, $[x]_{D_X} = \{x\}$ for each $x \in X$.[8] At the other extreme is $X^2$ which is the largest equivalence relation that can be defined on $X$. We have $[x]_{X^2} = X$ for each $x \in X$.

[**2**] By Exercise 7, the symmetric part of any reflexive and transitive relation on a nonempty set is an equivalence relation.

[8] I say an equally suiting name for $D_X$ is the "equality relation." What do you think?

**[3]** Let $X := \{(a,b) : a,b \in \{1,2,...\}\}$, and define the relation $\sim$ on $X$ by $(a,b) \sim (c,d)$ iff $ad = bc$. It is readily verified that $\sim$ is an equivalence relation on $X$, and that $[(a,b)]_\sim = \left\{(c,d) \in X : \frac{c}{d} = \frac{a}{b}\right\}$ for each $(a,b) \in X$.

**[4]** Let $X := \{..., -1, 0, 1, ...\}$ and define the relation $\sim$ on $X$ by $x \sim y$ iff $\frac{1}{2}(x-y) \in X$. It is easily checked that $\sim$ is an equivalence relation on $X$. Moreover, for any integer $x$, we have $x \sim y$ iff $y = x - 2m$ for some $m \in X$, and hence $[x]_\sim$ equals the set of all even integers if $x$ is even, and that of all odd integers if $x$ is odd. $\square$

One typically uses an equivalence relation to simplify a situation in a way that all things that are indistinguishable from a particular perspective are put together in a set and treated as if they are a single entity. For instance, suppose that for some reason we are interested in the signs of people. Then, any two individuals who are of the same sign can be thought of as "identical," so instead of the set of all people in the world, we would rather work with the set of all Capricorns, all Virgos and so on. But the set of all Capricorns is of course none other than the equivalence class of any given Capricorn person relative to the equivalence relation of "being of the same sign." So when someone says "a Capricorn is ...," then one is really referring to a whole class of people. The equivalence relation of "being of the same sign" divides the world into twelve equivalence classes, and we can then talk "as if" there are only twelve individuals in our context of reference.

To take another example, ask yourself how you would define the set of positive rational numbers, given the set of natural numbers $\mathbb{N} := \{1,2,...\}$ and the operation of "multiplication." Well, you may say, a positive rational number is the ratio of two natural numbers. But wait, what is a "ratio"? Let us be a bit more careful about this. A better way of looking at things is to say that a positive rational number is an ordered pair $(a,b) \in \mathbb{N}^2$, although, in daily practice, we write $\frac{a}{b}$ instead of $(a,b)$. Yet, we don't want to say that each ordered pair in $\mathbb{N}^2$ is a distinct rational number. (We would like to think of $\frac{1}{2}$ and $\frac{2}{4}$ as the same number, for instance.) So we "identify" all those ordered pairs who we wish to associate with a single rational number by using the equivalence relation $\sim$ introduced in Example 1.[3], and then define a rational number simply as an equivalence class $[(a,b)]_\sim$. Of course, when we talk about rational numbers in daily practice, we simply talk of a fraction like $\frac{1}{2}$, not $[(1,2)]_\sim$, even though, formally speaking, what we really mean is $[(1,2)]_\sim$. The equality $\frac{1}{2} = \frac{2}{4}$ is obvious, precisely because the rational numbers are constructed as equivalence classes such that $(2,4) \in [(1,2)]_\sim$.

This discussion suggests that an equivalence relation can be used to decompose a grand set of interest into subsets such that the members of the same subset are thought of as "identical" while the members of distinct subsets are viewed as "distinct." Let us now formalize this intuition. By a **partition** of a nonempty set $X$, we mean a class of pairwise disjoint, nonempty subsets of $X$ whose union is $X$. That is, $\mathcal{A}$ is a partition of $X$ iff $\mathcal{A} \subseteq 2^X \backslash \{\emptyset\}$, $\bigcup \mathcal{A} = X$ and $A \cap B = \emptyset$ for every distinct $A$ and $B$ in

$\mathcal{A}$. The next result says that the set of equivalence classes induced by any equivalence relation on a set is a partition of that set.

**Proposition 1.** *For any equivalence relation $\sim$ on a nonempty set $X$, the quotient set $X/_\sim$ is a partition of $X$.*

*Proof.* Take any nonempty set $X$ and an equivalence relation $\sim$ on $X$. Since $\sim$ is reflexive, we have $x \in [x]_\sim$ for each $x \in X$. Thus any member of $X/_\sim$ is nonempty, and $\bigcup\{[x]_\sim : x \in X\} = X$. Now suppose that $[x]_\sim \cap [y]_\sim \neq \emptyset$ for some $x, y \in X$. We wish to show that $[x]_\sim = [y]_\sim$. Observe first that $[x]_\sim \cap [y]_\sim \neq \emptyset$ implies $x \sim y$. (Indeed, if $z \in [x]_\sim \cap [y]_\sim$, then $x \sim z$ and $z \sim y$ by symmetry of $\sim$, so we get $x \sim y$ by transitivity of $\sim$.) This implies that $[x]_\sim \subseteq [y]_\sim$, because if $w \in [x]_\sim$, then $w \sim x$ (by symmetry of $\sim$), and hence $w \sim y$ by transitivity of $\sim$. The converse containment is proved analogously. ∎

The following exercise shows that the converse of Proposition 1 also holds. Thus the notions of equivalence relation and partition are really two different ways of looking at the same thing.

> *Exercise 9.* Let $\mathcal{A}$ be a partition of a nonempty set $X$, and consider the relation $\sim$ on $X$ defined by $x \sim y$ iff $\{x, y\} \subseteq A$ for some $A \in \mathcal{A}$. Prove that $\sim$ is an equivalence relation on $X$.

## 1.4 Order Relations

Transitivity property is the defining feature of any *order* relation. Such relations are given various names depending on the properties they possess in addition to transitivity.

Definition. A relation $\succsim$ on a nonempty set $X$ is called a **preorder** on $X$ if it is transitive and reflexive. It is said to be a **partial order** on $X$ if it is an antisymmetric preorder on $X$. Finally, $\succsim$ is called a **linear order** on $X$ if it is a partial order on $X$ which is complete.

By a **preordered set** we mean a list $(X, \succsim)$ where $X$ is a nonempty set and $\succsim$ is a preorder on $X$. If $\succsim$ is a partial order on $X$, then $(X, \succsim)$ is called a **poset** (short for *partially ordered set*), and if $\succsim$ is a linear order on $X$, then $(X, \succsim)$ is called either a **chain** or a **loset** (short for *linearly ordered set*).

While a preordered set $(X, \succsim)$ is not a set, it is convenient to talk as if it is a set when referring to properties that apply only to $X$. For instance, by a "finite preordered set," we understand a preordered set $(X, \succsim)$ with $|X| < \infty$. Or, when we say that $Y$ is a subset of the preordered set $(X, \succsim)$, we mean simply that $Y \subseteq X$. A similar convention applies to posets and losets as well.

**Notation.** Let $(X, \succsim)$ be a preordered set. Unless otherwise is stated explicitly, we denote by $\succ$ the asymmetric part of $\succsim$, and by $\sim$ the symmetric part of $\succsim$ (Exercise 7).

The main distinction between a preorder and a partial order is that the former may have a large symmetric part, while the symmetric part of the latter must equal the diagonal relation. As we shall see, however, in most applications this distinction is immaterial.

EXAMPLE 2. **[1]** For any nonempty set $X$, the diagonal relation $D_X := \{(x, x) : x \in X\}$ is a partial order on $X$. In fact, this relation is the only partial order on $X$ which is also an equivalence relation. (Why?) The relation $X^2$ is, on the other hand, a complete preorder, which is not antisymmetric unless $X$ is a singleton.

**[2]** For any nonempty set $X$, the equality relation $=$ and the subsethood relation $\supseteq$ are partial orders on $2^X$. The equality relation is not linear, and $\supseteq$ is not linear unless $X$ is a singleton.

**[3]** $(\mathbb{R}^n, \geq)$ is a poset for any positive integer $n$, where $\geq$ is defined coordinatewise, that is, $(x_1, ..., x_n) \geq (y_1, ..., y_n)$ iff $x_i \geq y_i$ for each $i = 1, ..., n$. When we talk of $\mathbb{R}^n$ without specifying explicitly an alterative order, we always have in mind this partial order (which is sometimes called the **natural** (or **canonical**) **order** of $\mathbb{R}^n$). Of course, $(\mathbb{R}, \geq)$ is a loset.

**[4]** Take any positive integer $n$, and preordered sets $(X_i, \succsim_i)$, $i = 1, ..., n$. The **product** of the preordered sets $(X_i, \succsim_i)$, denoted as $\boxtimes^n (X_i, \succsim_i)$, is the preordered set $(X, \succsim)$ with $X := \mathsf{X}^n X_i$ and

$$(x_1, ..., x_n) \succsim (y_1, ..., y_n) \;\text{ iff }\; x_i \succsim_i y_i \text{ for all } i = 1, ..., n.$$

In particular, $(\mathbb{R}^n, \geq) = \boxtimes^n (\mathbb{R}, \geq)$. □

EXAMPLE 3. In individual choice theory, a **preference relation** $\succsim$ on a nonempty alternative set $X$ is defined as a preorder on $X$. Here the reflexivity is a trivial condition to require, and transitivity is viewed as a fundamental rationality postulate. (More on this in Section B.4.) The **strict preference relation** $\succ$ is defined as the asymmetric part of $\succsim$ (Exercise 7). This relation is transitive but not reflexive. The **indifference relation** $\sim$ is then defined as the symmetric part of $\succsim$, and is easily checked to be an equivalence relation on $X$. For any $x \in X$, the equivalence class $[x]_\sim$ is called in this context the **indifference class** of $x$, and is simply a generalization of the familiar concept of "*the indifference curve that passes through* $x$." In particular, Proposition 1 says that no two distinct indifference sets can have a point in common. (This is the gist of the fact that "distinct indifference curves cannot cross!")

In social choice theory, one often works with multiple (complete) preference relations on a given alternative set $X$. For instance, suppose that there are $n$ individuals in the population, and $\succsim_i$ stands for the preference relation of the $i$th individual. The **Pareto dominance** relation $\succsim$ on $X$ is defined as $x \succsim y$ iff $x \succsim_i y$ for each $i = 1, ..., n$. This relation is a preorder on $X$ in general, and a partial order on $X$ if each $\succsim_i$ is antisymmetric. □

Let $(X, \succsim)$ be a preordered set. By an **extension** of $\succsim$ we understand a preorder $\trianglerighteq$ on $X$ such that $\succsim \subseteq \trianglerighteq$ and $\succ \subseteq \triangleright$, where $\triangleright$ is the asymmetric part of $\trianglerighteq$. Intuitively speaking, an extension of a preorder is "more complete" than the original relation in the sense that it allows one to compare more elements, but it certainly agrees exactly with the original relation when the latter applies. If $\trianglerighteq$ is a partial order, then it is an extension of $\succsim$ iff $\succsim \subseteq \trianglerighteq$. (Why?)

A fundamental result of order theory says that every partial order can be extended to a linear order, that is, for every poset $(X, \succsim)$ there is a loset $(X, \trianglerighteq)$ with $\succsim \subseteq \trianglerighteq$.While it is possible to prove this by mathematical induction when $X$ is finite, the proof in the general case is built on a relatively advanced method which we will cover later in the course. Relegating its proof to that Section 1.7, we only state here the result for future reference.[9]

**Sziplrajn's Theorem.** *Every partial order on a nonempty set $X$ can be extended to a linear order on $X$.*

A natural question is if the same result holds for preorders as well. The answer is yes, and the proof follows easily from Sziplrajn's Theorem by means of a standard method.

**Corollary 1.** *Let $(X, \succsim)$ be a preordered set. There exists a complete preorder on $X$ that extends $\succsim$.*

*Proof.* Let $\sim$ denote the symmetric part of $\succsim$, which is an equivalence relation. Then $(X/_\sim, \succsim^*)$ is a poset where $\succsim^*$ is defined on $X/_\sim$ by

$$[x]_\sim \succsim^* [y]_\sim \quad \text{if and only if} \quad x \succsim y.$$

By Sziplrajn's Theorem, there exists a linear order $\trianglerighteq^*$ on $X/_\sim$ such that $\succsim^* \subseteq \trianglerighteq^*$. We define $\trianglerighteq$ on $X$ by

$$x \trianglerighteq y \quad \text{if and only if} \quad [x]_\sim \trianglerighteq^* [y]_\sim.$$

[9] For an extensive introduction to the theory of linear extensions of posets, see Bonnet and Pouzet (1982).

It is easily checked that $\trianglerighteq$ is a complete preorder on $X$ with $\succsim \subseteq \trianglerighteq$ and $\succ \subseteq \triangleright$, where $\succ$ and $\triangleright$ are the asymmetric parts of $\succsim$ and $\trianglerighteq$, respectively. ■

*Exercise 10.* Let $(X, \succsim)$ be a preordered set, and define $\mathcal{L}(\succsim)$ as the set of all complete preorders that extend $\succsim$. Prove that $\succsim = \bigcap \mathcal{L}(\succsim)$. (Where do you use Sziplrajn's Theorem in the argument?)

*Exercise 11.* Let $(X, \succsim)$ be a finite preordered set. Taking $\mathcal{L}(\succsim)$ as in the previous exercise, we define $dim(X, \succsim)$ as the smallest positive integer $k$ such that $\succsim = R_1 \cap \cdots \cap R_k$ for some $R_i \in \mathcal{L}(\succsim)$, $i = 1, ..., k$.
($a$) Show that $dim(X, \succsim) \leq |X^2|$.
($b$) What is $dim(X, D_X)$? What is $dim(X, X^2)$?
($c$) For any positive integer $n$, show that $dim(\boxtimes^n (X_i, \succsim_i)) = n$, where $(X_i, \succsim_i)$ is a loset with $|X_i| \geq 2$ for each $i = 1, ..., n$.
($d$) Prove or disprove: $dim(2^X, \supseteq) = |X|$.

DEFINITION. Let $(X, \succsim)$ be a preordered set and $\emptyset \neq Y \subseteq X$. An element $x$ of $Y$ is said to be **$\succsim$-maximal** in $Y$ if there is no $y \in Y$ with $y \succ x$, and **$\succsim$-minimal** in $Y$ if there is no $y \in Y$ with $x \succ y$. If $x \succsim y$ for all $y \in Y$, then $x$ is called the **$\succsim$-maximum** of $Y$, and if $y \succsim x$ for all $y \in Y$, then $x$ is called the **$\succsim$-minimum** of $Y$.

Obviously, for any preordered set $(X, \succsim)$, every $\succsim$-maximum of a nonempty subset of $X$ is $\succsim$-maximal in that set. Also note that if $(X, \succsim)$ is a poset, then there can be at most one $\succsim$-maximum of any $Y \in 2^X \backslash \{\emptyset\}$.

EXAMPLE 4. **[1]** Let $X$ be any nonempty set, and $\emptyset \neq Y \subseteq X$. Every element of $Y$ is both $D_X$-maximal and $D_X$-minimal in $Y$. Unless it is a singleton, $Y$ has neither a $D_X$-maximum nor a $D_X$-minimum element. On the other hand, every element of $Y$ is both $X^2$-maximum and $X^2$-minimum of $Y$.

**[2]** Given any nonempty set $X$, consider the poset $(2^X, \supseteq)$, and take any nonempty $\mathcal{A} \subseteq 2^X$. The class $\mathcal{A}$ has a $\supseteq$-maximum iff $\bigcup \mathcal{A} \in \mathcal{A}$, and it has a $\supseteq$-minimum iff $\bigcap \mathcal{A} \in \mathcal{A}$. In particular, the $\supseteq$-maximum of $2^X$ is $X$ and the $\supseteq$-minimum of $2^X$ is $\emptyset$.

**[3]** (*Choice Correspondences*) Given a preference relation $\succsim$ on an alternative set $X$ (Example 3) and a nonempty subset $S$ of $X$, we define the "set of choices from $S$" for an individual whose preference relation is $\succsim$ as the set of all $\succsim$-maximal elements in $S$. That is, denoting this set as $C_{\succsim}(S)$, we have

$$C_{\succsim}(S) := \{x \in S : y \succ x \text{ for no } y \in S\}.$$

Evidently, if $S$ is a finite set, then $C_{\succsim}(S)$ is nonempty. (Proof?) Moreover, if $S$ is finite and $\succsim$ is complete, then there exists at least one $\succsim$-maximum element in $S$. Finiteness requirement cannot be omitted in this statement, but as we shall see throughout this course, there are various ways in which it can be substantially weakened. □

*Exercise 12.* ($a$) Which subsets of the set of positive integers have a $\geq$-minimum? Which ones have a $\geq$-maximum?
($b$) If a set in a poset $(X, \succsim)$ has a unique $\succsim$-maximal element, does that element have to be a $\succsim$-maximum of the set?
($c$) Which subsets of a poset $(X, \succsim)$ possess an element which is *both* $\succsim$-maximum *and* $\succsim$-minimum?
($d$) Give an example of an infinite set in $\mathbb{R}^2$ which contains a unique $\geq$-maximal element that is also the unique $\geq$-minimal element of the set.

***Exercise 13.***[H] Let $\succsim$ be a complete relation on a nonempty set $X$, and $S$ a nonempty finite subset of $X$. Define

$$c_{\succsim}(S) := \{x \in S : x \succsim y \text{ for all } y \in S\}.$$

(***a***) Show that $c_{\succsim}(S) \neq \emptyset$ if $\succsim$ is transitive.
(***b***) We say that $\succsim$ is **acyclic** if there does not exist a positive integer $k$ such that $x_1, ..., x_k \in X$ and $x_1 \succ x_2 \succ \cdots \succ x_k \succ x_1$. Show that every transitive relation is acyclic, but not conversely.
(***c***) Show that $c_{\succsim}(S) \neq \emptyset$ if $\succsim$ is acyclic.
(***d***) Show that if $c_{\succsim}(T) \neq \emptyset$ for every finite $T \in 2^X \backslash \{\emptyset\}$, then $\succsim$ must be acyclic.

*Exercise 14.*[H] Let $(X, \succsim)$ be a poset, and take any $Y \in 2^X \backslash \{\emptyset\}$ which has a $\succsim$-maximal element, say $x^*$. Prove that $\succsim$ can be extended to a linear order $\trianglerighteq$ on $X$ such that $x_*$ is $\trianglerighteq$-maximal in $Y$.

***Exercise 15.*** Let $(X, \succsim)$ be a poset. For any $Y \subseteq X$, an element $x$ in $X$ is said to be an **$\succsim$-upper bound** for $Y$ if $x \succsim y$ for all $y \in Y$; a **$\succsim$-lower bound** for $Y$ is defined similarly. The **$\succsim$-supremum** of $Y$, denoted $\sup_{\succsim} Y$, is defined as the $\succsim$-minimum of the set of all $\succsim$-upper bounds for $Y$, that is, $\sup_{\succsim} Y$ is an $\succsim$-upper bound for $Y$ and has the property that $z \succsim \sup_{\succsim} Y$ for any $\succsim$-upper bound $z$ for $Y$. The **$\succsim$-infimum** of $Y$, denoted as $\inf_{\succsim} Y$, is defined analogously.
(***a***) Prove that there can be only one $\succsim$-supremum and only one $\succsim$-infimum of any subset of $X$.
(***b***) Show that $x \succsim y$ iff $\sup_{\succsim}\{x, y\} = x$ and $\inf_{\succsim}\{x, y\} = y$, for any $x, y \in X$.
(***c***) Show that if $\sup_{\succsim} X \in X$ (that is, if $\sup_{\succsim} X$ exists), then $\inf_{\succsim} \emptyset = \sup_{\succsim} X$.
(***d***) If $\succsim$ is the diagonal relation on $X$, and $x$ and $y$ are any two distinct members of $X$, does $\sup_{\succsim}\{x, y\}$ exist?
(***e***) If $X := \{x, y, z, w\}$ and $\succsim := \{(z, x), (z, y), (w, x), (w, y)\}$, does $\sup_{\succsim}\{x, y\}$ exist?

***Exercise 16.***[H] Let $(X, \succsim)$ be a poset. If $\sup_{\succsim}\{x, y\}$ and $\inf_{\succsim}\{x, y\}$ exist for all $x, y \in X$, then we say that $(X, \succsim)$ is a **lattice**. If $\sup_{\succsim} Y$ and $\inf_{\succsim} Y$ exist for all $Y \in 2^X$, then $(X, \succsim)$ is called a **complete lattice**.
(***a***) Show that every complete lattice has an upper and a lower bound.
(***b***) Show that if $X$ is finite and $(X, \succsim)$ is a lattice, then $(X, \succsim)$ is a complete lattice.

($\boldsymbol{c}$) Give an example of a lattice which is not complete.
($\boldsymbol{d}$) Prove that $(2^X, \supseteq)$ is a complete lattice.
($\boldsymbol{e}$) Let $\mathcal{X}$ be a nonempty subset of $2^X$ such that $X \in \mathcal{X}$ and $\bigcap \mathcal{A} \in \mathcal{X}$ for any (nonempty) class $\mathcal{A} \subseteq \mathcal{X}$. Prove that $(\mathcal{X}, \supseteq)$ is a complete lattice.

## 1.5 Functions

Intuitively, we think of a *function* as a rule that transforms the objects in a given set to those of another. While this is not a formal definition – what is a "rule"? – we may now use the notion of a binary relation to formalize the idea. Let $X$ and $Y$ be any two nonempty sets. By a **function** $f$ **that maps** $X$ **into** $Y$, denoted as $f : X \to Y$, we mean a relation $f \in X \times Y$ such that

(i) for every $x \in X$, there exists a $y \in Y$ such that $x\, f\, y$,
(ii) for every $y, z \in Y$ with $x\, f\, y$ and $x\, f\, z$, we have $y = z$.

Here $X$ is called the **domain** of $f$ and $Y$ the **codomain** of $f$. The **range** of $f$ is, on the other hand, defined as

$$f(X) := \{y \in Y : x\, f\, y \text{ for some } x \in X\}.$$

The set of all functions that map $X$ into $Y$ is denoted by $Y^X$. For instance, $\{0,1\}^X$ is the set of all functions on $X$ whose values are either 0 or 1, and $\mathbb{R}^{[0,1]}$ is the set of all real-valued functions on $[0,1]$. The notation $f \in Y^X$ will be used interchangeably with the expression $f : X \to Y$ throughout this course. Similarly, the term **map** is used interchangeably with the term "function."

While our definition of a function may look at first a bit strange, it is hardly anything other than a set-theoretic formulation of the concept we use in daily discourse. After all, we want a *function* $f$ that maps $X$ into $Y$ to assign each member of $X$ to a member of $Y$, right? Our definition says simply that one can think of $f$ simply as a set of ordered pairs, so "$(x, y) \in f$" means "$x$ is mapped to $y$ by $f$." Put differently, all that $f$ "does" is completely identified by the set $\{(x, f(x)) \in X \times Y : x \in X\}$, which is what $f$ "is." The familiar notation $f(x) = y$ (which we shall also adopt in the rest of the exposition) is then nothing but an alternative way of expressing $x\, f\, y$. When $f(x) = y$, we refer to $y$ as the **image** (or **value**) **of** $x$ **under** $f$. Condition (i) says that every element in the domain $X$ of $f$ has an image under $f$ in the codomain $Y$. In turn, condition (ii) states that no element in the domain of $f$ can have more than one image under $f$.

Some authors adhere to the intuitive definition of a function as a "rule" that transforms one set into another, and refer to the set of all ordered pairs $(x, f(x))$ as the **graph** of the function. Denoting this set by $Gr(f)$, then, we can write

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

According to the formal definition of a function, $f$ and $Gr(f)$ are the same thing. So long as we keep this connection in mind, there is no danger in thinking of a function as a "rule" in the intuitive way. In particular, we say that two functions $f$ and $g$ are equal if they have the same graph, or equivalently, if they have the same domain and codomain, and $f(x) = g(x)$ for all $x \in X$. In this case, we simply write $f = g$.

If its range equals its codomain, that is, if $f(X) = Y$, then one says that $f$ maps $X$ **onto** $Y$, and refers to it as a **surjection** (or as a **surjective** function/map). If $f$ maps distinct points in its domain to distinct points in its codomain, that is, if $x \neq y$ implies $f(x) \neq f(y)$ for all $x, y \in X$, then we say that $f$ is an **injection** (or a **one-to-one**, or an **injective** function/map). Finally, if $f$ is both injective and surjective, then it is called a **bijection** (or a **bijective** function/map). For instance, if $X := \{1, ..., 10\}$, then $f := \{(1,2), (2,3), ..., (10,1)\}$ is a bijection in $X^X$, while $g \in X^X$, defined as $g(x) := 3$ for all $x \in X$, is neither an injection nor a surjection. When considered as a map in $(\{0\} \cup X)^X$, $f$ is an injection but not a surjection.

**Warning.** Every injective function can be viewed as a bijection, provided that one views the codomain of the function as its range. Indeed, if $f : X \to Y$ is an injection, then the map $f : X \to Z$ is a bijection, where $Z := f(X)$. This is usually expressed as saying that $f : X \to f(X)$ is a bijection.

Before we consider some examples, let us note that a common way of defining a particular function in a given context is to describe the domain and codomain of that function, and the image of a generic point in the domain. So one would say something like "let $f : X \to Y$ be defined by $f(x) := ...$" or "consider the function $f \in Y^X$ defined by $f(x) := ...$". For example, by the function $f : \mathbb{R} \to \mathbb{R}_+$ defined by $f(x) := x^2$, we mean the surjection that transforms every real number $x$ to the nonnegative real number $x^2$. Since the domain of the function is understood from the expression $f : X \to Y$ (or $f \in Y^X$), it is redundant to add the phrase "for all $x \in X$" after the expression "$f(x) := ...$," although sometimes we may do so for clarity. Alternatively, when the codomain of the function is clear, a phrase like "the map $x \mapsto f(x)$ on $X$" is commonly used. For instance, one may refer to the quadratic function mentioned above unambiguously as "the map $t \mapsto t^2$ on $\mathbb{R}$."

EXAMPLE 5. In the following examples $X$ and $Y$ stand for arbitrary nonempty sets.

[**1**] A **constant function** is the one who assigns the same value to every element of its domain, that is, $f \in Y^X$ is constant iff there exists a $y \in Y$ such that $f(x) = y$ for all $x \in X$. (Formally speaking, this constant function is the set $X \times \{y\}$.) Obviously, $f(X) = \{y\}$ in this case, so a constant function is not surjective unless its codomain is a singleton, and it is not injective unless its domain is a singleton.

[**2**] A function whose domain and codomain are identical, that is, a function in $X^X$, is called a **self-map** on $X$. An important example of a self-map is the **identity function** on $X$. This function is denoted as $\mathrm{id}_X$, and it is defined as $\mathrm{id}_X(x) := x$ for

all $x \in X$. Obviously, $\mathrm{id}_X$ is a bijection, and formally speaking, it is none other than the diagonal relation $D_X$.

[**3**] Let $S \subseteq X$. The function that maps $X$ into $\{0,1\}$ such that every member of $S$ is assigned to 1 and all the other elements of $X$ are assigned to zero is called the **indicator function of** $S$ **in** $X$. This function is denoted as $\mathbf{1}_S$ (assuming that the domain $X$ is understood from the context). By definition, we have

$$\mathbf{1}_S(x) := \begin{cases} 1, & \text{if } x \in S \\ 0, & \text{if } x \in X \backslash S \end{cases}.$$

You can check that, for every $A, B \subseteq X$, we have $\mathbf{1}_{A\cup B} + \mathbf{1}_{A\cap B} = \mathbf{1}_A + \mathbf{1}_B$ and $\mathbf{1}_{A\cap B} = \mathbf{1}_A \mathbf{1}_B$. □

The following set of examples point to some commonly used methods of obtaining new functions from a given set of functions.

Example 6. In the following examples $X$, $Y$, $Z$, and $W$ stand for arbitrary nonempty sets.

[**1**] Let $Z \subseteq X \subseteq W$, and $f \in Y^X$. By the **restriction** of $f$ to $Z$, denoted as $f|_Z$, we mean the function $f|_Z \in Y^Z$ defined by $f|_Z(z) := f(z)$. By an **extension** of $f$ to $W$, on the other hand, we mean a function $f^* \in Y^W$ with $f^*|_X = f$, that is, $f^*(x) = f(x)$ for all $x \in X$. If $f$ is injective, so must $f|_Z$, but surjectivity of $f$ does not entail that of $f|_Z$. Of course, if $f$ is not injective, $f|_Z$ may still turn out to be injective (e.g. $x \mapsto x^2$ is not injective on $\mathbb{R}$, but it is so on $\mathbb{R}_+$).

[**2**] Sometimes it is possible to extend a given function by *combining* it with another function. For instance, we can combine any $f \in Y^X$ and $g \in W^Z$ to obtain the function $h : X \cup Z \to Y \cup W$ defined by

$$h(t) := \begin{cases} f(t), & \text{if } t \in X \\ g(t), & \text{if } t \in Z \end{cases},$$

provided that $X \cap Z = \emptyset$, or $X \cap Z \neq \emptyset$ and $f|_{X\cap Z} = g|_{X\cap Z}$. Note that this method of combining functions does *not* work if $f(t) \neq g(t)$ for some $t \in X \cap Z$. For, in that case $h$ would not be well-defined as a function. (What would be the image of $t$ under $h$?)

[**3**] A function $f \in X^{X\times Y}$ defined by $f(x, y) := x$ is called the **projection from** $X\times Y$ **onto** $X$.[10] (The projection from $X\times Y$ onto $Y$ is similarly defined.) Obviously, $f(X \times Y) = X$, that is, $f$ is necessarily surjective. It is not injective unless $Y$ is a singleton.

[10] Strictly speaking, I should write $f((x, y))$ instead of $f(x, y)$, but that's just splitting hairs.

[4] Given functions $f : X \to Z$ and $g : Z \to Y$, we define the **composition** of $f$ and $g$ as the function $g \circ f : X \to Y$ by $g \circ f\,(x) := g(f(x))$. (For easier reading, we often write $(g \circ f)(x)$ instead of $g \circ f\,(x)$.) This definition accords with the way we defined the composition of two relations (Exercise 5). Indeed, we have $(g \circ f)(x) = \{(x, y) : x\,f\,z \text{ and } z\,g\,y \text{ for some } z \in Z\}$.

Obviously, $\mathrm{id}_Z \circ f = f = f \circ \mathrm{id}_X$. Even when $X = Y = Z$, the operation of taking compositions is not commutative. For instance, if the self-maps $f$ and $g$ on $\mathbb{R}$ are defined by $f(x) := 2$ and $g(x) := x^2$, respectively, then $(g \circ f)(x) = 4$ and $(f \circ g)(x) = 2$ for any real number $x$. The composition operation is, however, associative, that is, $h \circ (g \circ f) = (h \circ g) \circ f$ for all $f \in Y^X$, $g \in Z^Y$ and $h \in W^Z$. $\square$

*Exercise 17.* Let $\sim$ be an equivalence relation on a nonempty set $X$. Show that the map $x \mapsto [x]_\sim$ on $X$ (called the **quotient map**) is a surjection on $X$ which is injective iff $\sim\, = D_X$.

*Exercise 18.*[H] (*A Factorization Theorem*) Let $X$ and $Y$ be two nonempty sets. Prove: For any function $f : X \to Y$, there exists a nonempty set $Z$, a surjection $g : X \to Z$ and an injection $h : Z \to Y$ such that $f = h \circ g$.

*Exercise 19.* Let $X$, $Y$ and $Z$ be nonempty sets, and consider any $f, g \in Y^X$ and $u, v \in Z^Y$. Prove:
($a$) If $f$ is surjective and $u \circ f = v \circ f$, then $u = v$;
($b$) If $u$ is injective and $u \circ f = u \circ g$, then $f = g$;
($c$) If $f$ and $u$ are injective (respectively, surjective), then so is $u \circ f$.

*Exercise 20.*[H] Show that there is no surjection of the form $f : X \to 2^X$ for any nonempty set $X$.

For any given nonempty sets $X$ and $Y$, the (*direct*) **image** of a set $A \subseteq X$ under $f \in Y^X$, denoted $f(A)$, is defined as the collection of all elements $y$ in $Y$ with $y = f(x)$ for some $x \in A$. That is,

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

The range of $f$ is thus the image of its entire domain: $f(X) = \{f(x) : x \in X\}$. (*Note.* If $f(A) = B$, then one says that "$f$ maps $A$ onto $B$.")

The **inverse image** of a set $B$ in $Y$, denoted as $f^{-1}(B)$, is defined as the set of all $x$ in $X$ whose images under $f$ belong to $B$, that is,

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

By convention, we write $f^{-1}(y)$ for $f^{-1}(\{y\})$, that is,

$$f^{-1}(y) := \{x \in X : f(x) = y\} \quad \text{for any } y \in Y.$$

Obviously, $f^{-1}(y)$ is a singleton for each $y \in Y$ iff $f$ is an injection. For instance, if $f$ stands for the map $t \mapsto t^2$ on $\mathbb{R}$, then $f^{-1}(1) = \{-1, 1\}$ whereas $f|_{\mathbb{R}_+}^{-1}(1) = \{1\}$.

The issue of whether or not one can express the image (or the inverse image) of a union/intersection of a collection of sets as the union/intersection of the images (inverse images) of each set in the collection arises quite often in mathematical analysis. The following exercise summarizes the situation in this regard.

***Exercise 21.*** Let $X$ and $Y$ be nonempty sets and $f \in Y^X$. Prove that, for any (nonempty) classes $\mathcal{A} \subseteq 2^X$ and $\mathcal{B} \subseteq 2^Y$, we have

$$f\left(\bigcup \mathcal{A}\right) = \bigcup\{f(A) : A \in \mathcal{A}\} \quad \text{and} \quad f\left(\bigcap \mathcal{A}\right) \subseteq \bigcap\{f(A) : A \in \mathcal{A}\},$$

whereas

$$f^{-1}\left(\bigcup \mathcal{B}\right) = \bigcup\{f^{-1}(B) : B \in \mathcal{B}\} \quad \text{and} \quad f^{-1}\left(\bigcap \mathcal{B}\right) = \bigcap\{f^{-1}(B) : B \in \mathcal{B}\}.$$

A general rule that surfaces from this exercise is that inverse images are quite well-behaved with respect to the operations of taking unions and intersections, while the same cannot be said for direct images in the case of taking intersections. Indeed, for any $f \in Y^X$, we have $f(A \cap B) \supseteq f(A) \cap f(B)$ for all $A, B \subseteq X$ if, *and only if*, $f$ is injective.[11] The "if" part of this assertion is trivial. The "only if" part follows from the observation that, if the claim was not true, then, for any distinct $x, y \in X$ with $f(x) = f(y)$, we would find $\emptyset = f(\emptyset) = f(\{x\} \cap \{y\}) = f(\{x\}) \cap f(\{y\}) = \{f(x)\}$, which is absurd.

Finally, we turn to the problem of *inverting* a function. For any function $f \in Y^X$, let us define the set

$$f^{-1} := \{(y, x) \in Y \times X : x \, f \, y\}$$

which is none other than the inverse of $f$ viewed as a relation (Exercise 5). This relation simply *reverses* the map $f$ in the sense that if $x$ is mapped to $y$ by $f$, then $f^{-1}$ maps $y$ back to $x$. Now $f^{-1}$ may or may not be a function. If it is, we say that $f$ is **invertible** and $f^{-1}$ is the **inverse** of $f$. For instance, $f : \mathbb{R} \to \mathbb{R}_+$ defined by $f(t) := t^2$ is not invertible (since $(1, 1) \in f^{-1}$ *and* $(1, -1) \in f^{-1}$, that is, 1 does not have a unique image under $f^{-1}$), whereas $f|_{\mathbb{R}_+}$ is invertible and $f|_{\mathbb{R}_+}^{-1}(t) = \sqrt{t}$ for all $t \in \mathbb{R}$.

The following result gives a simple characterization of invertible functions.

**Proposition 2.** *Let $X$ and $Y$ be two nonempty sets. A function $f \in Y^X$ is invertible if, and only if, it is a bijection.*

***Exercise 22.*** Prove Proposition 2.

[11] Of course, this does not mean that $f(A \cap B) = f(A) \cap f(B)$ can never hold for a function that is not one-to-one. It only means that, for any such function $f$, we can always find nonempty sets $A$ and $B$ in the domain of $f$ such that $f(A \cap B) \supseteq f(A) \cap f(B)$ is false.

By using the composition operation defined in Example 6.[4], we can give another useful characterization of invertible functions.

**Proposition 3.** *Let $X$ and $Y$ be two nonempty sets. A function $f \in Y^X$ is invertible if, and only if, there exists a function $g \in X^Y$ such that $g \circ f = \mathrm{id}_X$ and $f \circ g = \mathrm{id}_Y$.*

*Proof.* The "only if" part is readily obtained upon choosing $g := f^{-1}$. To prove the "if" part, suppose there exists a $g \in X^Y$ with $g \circ f = \mathrm{id}_X$ and $f \circ g = \mathrm{id}_Y$, and note that, by Proposition 2, it is enough to show that $f$ is a bijection. To verify the injectivity of $f$, pick any $x, y \in X$ with $f(x) = f(y)$, and observe that

$$x = \mathrm{id}_X(x) = (g \circ f)(x) = g(f(x)) = g(f(y)) = (g \circ f)(y) = \mathrm{id}_X(y) = y.$$

To see the surjectivity of $f$, take any $y \in Y$ and define $x := g(y)$. Then we have

$$f(x) = f(g(y)) = (f \circ g)(y) = \mathrm{id}_Y(y) = y,$$

which proves $Y \subseteq f(X)$. Since the converse containment is trivial, we are done. ■

## 1.6 Sequences, Vectors and Matrices

By a *sequence* in a given nonempty set $X$, we intuitively mean an ordered array of the form $(x_1, x_2, ...)$ where each term $x_i$ of the sequence is a member of $X$. (Throughout this text we denote such a sequence by $(x_m)$, but note that some books prefer instead the notation $(x_m)_{m=1}^{\infty}$.) As in the case of ordered pairs, one could introduce the notion of a sequence as a new object to our set theory, but again there is really no need to do so. Intuitively, we understand from the notation $(x_1, x_2, ...)$ that the $i$th term in the array is $x_i$. But then we can think of this array as a function that maps the set $\mathbb{N}$ of positive integers into $X$ in the sense that it tells us that "the $i$th term in the array is $x_i$" by mapping $i$ to $x_i$. With this definition, our intuitive understanding of the ordered array $(x_1, x_2, ...)$ is formally captured by the function $\{(i, x_i) : i = 1, 2...\} = f$. Thus, we define a **sequence** in a nonempty set $X$ as any function $f : \mathbb{N} \to X$, and *represent* this function as $(x_1, x_2, ...)$ where $x_i := f(i)$ for each $i \in \mathbb{N}$. Consequently, the set of all sequences in $X$ is equal to $X^{\mathbb{N}}$. As is common, however, we denote this set as $X^\infty$ throughout the text.

By a *subsequence* of a sequence $(x_m) \in X^\infty$, we mean a sequence that is made up of the terms of $(x_m)$ which appear in the subsequence in the same order they appear in $(x_m)$. That is, a subsequence of $(x_m)$ is of the form $(x_{m_1}, x_{m_2}, ...)$ where $(m_k)$ is a sequence in $\mathbb{N}$ such that $m_1 < m_2 < \cdots$. (We denote this subsequence as $(x_{m_k})$.) Once again, we use the notion of function to formalize this definition. Strictly speaking, a **subsequence** of a sequence $f \in X^{\mathbb{N}}$ is a function of the form $f \circ \sigma$ where $\sigma : \mathbb{N} \to \mathbb{N}$ is strictly increasing (that is, $\sigma(k) < \sigma(l)$ for any $k, l \in \mathbb{N}$ with $k < l$). We *represent* this function as the array $(x_{m_1}, x_{m_2}, ...)$ with the understanding that