Chapter 1

Notation; Sets and Functions

It is sometimes convenient to use logical notation. We list the notation we use.

$P \implies Q$	for	"if P, then Q", or "Q whenever P", or "P only if Q";
$P \iff Q$	for	"P if and only if Q", that is to say P and Q are logically equivalent;
$P :\iff Q$	for	"P is defined to be equivalent to Q";
$\forall$	for	"For every . . . ";
$\exists$	for	"There is at least one . . . ";
$\exists!$	for	"There is a unique . . . ", or "There is one and only one . . .".

1.1 Sets

The mathematics we study in this course can be expressed entirely in terms of *sets* and *functions* between sets.

While the notion of sets and functions are presumed to be familiar from high school, we present a summary of the set-theoretical concepts and definitions used in this course and use the occasion to summarise notational conventions we use.

A *set* is almost any reasonable collection of things. We shall not even attempt a more formal definition in this course. The things in the collection are called the *elements* of the set in question. We write

$$x \in A$$

to denote that x is an element of the set A and

$$x \notin A$$

to denote that x is not an element of the set A.

Note that we do not exclude the possibility that x be a set in its own right, except that x cannot be A: **We explicitly exclude** $A \in A$.

Two sets are considered to be the same when they comprise precisely the same elements, in other words, when every element of the first set is also an element of the second and vice versa. Formally,

Definition 1.1. Given two sets A and B, $A = B$ if and only if x is an element of A when and only when x is an element of B.

We say that A is a *subset* of B if and only if $x \in B$ whenever $x \in A$. We write

$$A \subseteq B$$

whenever this is the case.

B is called a *proper subset* of A if and only if B is a subset of A, but $B \neq A$. In such a case we write

$$B \subset A.$$

Using our notational conventions, given two sets A and B,

$$A = B :\Longleftrightarrow \big((x \in A) \Leftrightarrow (x \in B)\big).$$
$$A \subseteq B :\Longleftrightarrow \big((x \in A) \Rightarrow (x \in B)\big).$$

We see that $A = B$ if and only if $A \subseteq B$ and $B \subseteq A$.

When we wish to describe a set, we can do so by *listing* all of its elements. Thus, if the set A has precisely a, b and c as its elements, then we write

$$A = \{a, b, c\}.$$

Example 1.2. By Definition 1.1, $\{a, b\}$, $\{a, b, b, b\}$ and $\{a, a, a, a, a, a, b\}$ are all the same set.

Another way of describing a set is by prescribing a number of *conditions* for membership of the set. In this case we write

$$A = \{x \mid P(x), Q(x), \ldots\}$$

to denote that the set in question consists of all those x for which $P(x), Q(x), \ldots$ all hold.

There are important operations on sets.

Definition 1.3. The *union* of the sets A and B is the set of all those objects which are in one, or other (or both). It is denoted by

$$A \cup B.$$

Using the notation above,

$$A \cup B := \{x \mid x \in A \text{ or } x \in B\}.$$

[Here, := has been used to signify that the expression on the left hand side is defined to be equal to the expression on the right hand side.]

Definition 1.4. The *intersection* of the sets A and B is the of all those objects which are elements of both. It is denoted by

$$A \cap B.$$

In other words,

$$A \cap B := \{x \mid x \in A \text{ and } x \in B\}.$$

Definition 1.5. Those elements of A that are not also elements of B form a set in their own right, which we denote by

$$A \setminus B,$$

so that

$$A \setminus B := \{x \in A \mid x \notin B\}.$$

Definition 1.6. Given sets A, B, their *(Cartesian) product* is the set of all ordered pairs, with the first member of each pair an element of A, and the second an element of B. It is denoted by

$$A \times B,$$

so that

$$A \times B := \{(x, y) \mid a \in A, y \in B\}.$$

We can extend unions, intersections and cartesian products to larger collections of sets than merely pairs.

Definition 1.7. An *indexed family* of sets, with *indexing set* Λ consists of a collection of sets, containing one set, say A_λ, for each element λ of the indexing set Λ. We write $\{A_\lambda \mid \lambda \in \Lambda\}$.

Definition 1.8. Given an indexed family of sets, say

$$\{A_\lambda \mid \lambda \in \Lambda\},$$

then their *union*, *intersection* and *Cartesian) product* are the sets defined respectively by

$$\bigcup_{\lambda \in \Lambda} A_\lambda := \{x \mid x \in A_\lambda \text{ for at least one } \lambda \in \Lambda\}$$
$$\bigcap_{\lambda \in \Lambda} A_\lambda := \{x \mid x \in A_\lambda \text{ for every } \lambda \in \Lambda\}$$
$$\prod_{\lambda \in \Lambda} A_\lambda := \{(x_\lambda)_{\lambda \in \Lambda} \mid x_\lambda \in A_\lambda \text{ for all } \lambda \in \Lambda\}.$$

Here $(x_\lambda)_{\lambda \in \Lambda}$ denotes a *generalised sequence*, namely, an ordered choice of elements x_λ, one for each $\lambda \in \Lambda$. Ordered pairs arise when $\Lambda = \{1, 2\}$ and sequences when $\Lambda = \mathbb{N}$.

A number of sets occur with such frequency that special notation has been introduced for them. These include the sets $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ consisting respectively of all *natural numbers*, all *integers*, all *rational numbers*, all *real numbers* and all *complex numbers*.

Explicitly,

$$\begin{aligned}
\mathbb{N} : &= \{0, 1, 2, 3, \ldots\} \\
\mathbb{Z} : &= \{\ldots - 3, -2 - 1, 0, 1, 2, 3, \ldots\} \\
\mathbb{Q} : &= \{x \in \mathbb{R} \mid x = \frac{p}{q} \text{ for some } p, q \in \mathbb{Z}, \text{ with } q \neq 0\} \\
&= \{x \in \mathbb{R} \mid x = \frac{p}{q} \text{ with } p \in \mathbb{Z} \text{ and } q \in \mathbb{N} \setminus \{0\}\,\}
\end{aligned}$$

Observe that

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

We write $\emptyset$ for the *empty set*, which is the (unique!) set with no elements. Note that it is a subset of every set, that is, if X is any set, then $\emptyset \subseteq X$.

1.2 Functions

To compare sets, we have the notion of a *function* or *map* or *mapping*.

Definition 1.9. A *function*, *map*, or *mapping* consists of three separate data, namely,

(i) a *domain* that is, a set on which the function is defined,

(ii) a *co-domain*, that is, a set in which the function takes its values, and

(iii) the assignment to each element of the domain of definition of a uniquely determined element from the set in which the function takes it values.

This is conveniently depicted diagrammatically by

$$f\colon X \longrightarrow Y,$$

or

$$X \xrightarrow{\quad f \quad} Y.$$

Here X is the domain of definition, Y is the set in which the function takes its values and f is the name of the function.

We emphasise that the function f need not be given in terms of a mathematical formula.

We write $X = \mathrm{dom}(f)$ and $Y = \mathrm{codom}(f)$ to indicate that X is the domain and Y the co-domain of f.

We often denote the function by f alone, but only when there is no danger of confusion. If we wish to express explicitly that the function, $f : X \longrightarrow Y$, assigns the element $y \in Y$ to the element $x \in X$, then we write $f : x \longmapsto y$ or, equivalently, $y = f(x)$, a form undoubtedly familiar to the reader. Sometimes the two parts are combined as

$$f\colon X \longrightarrow Y, \quad x \longmapsto y$$

or as

$$\begin{aligned} f\colon X &\longrightarrow Y \\ x &\longmapsto y. \end{aligned}$$

Definition 1.10. If f assigns $y \in Y$ to $x \in X$, then we say that y is the *image of x under f* or just the *image of x*.

Two functions f and g are *equal*, that is $f = g$ if and only if

(i) $\mathrm{dom}(f) = \mathrm{dom}(g)$

(ii) $\mathrm{codom}(f) = \mathrm{codom}(g)$

(iii) $f(x) = g(x)$ for every $x \in \mathrm{dom}(f)$.

In other words, to be the same, two functions must share both domain and co-domain as well as agreeing everywhere.

Observation 1.11. It is essential to note that the function f need not be given in terms of a mathematical formula.

Example 1.12. If X is the set of all human beings and Y is the set of all male human beings, we have a function

$$f\colon X \longrightarrow Y, \quad x \longmapsto y,$$

where f maps x to y if and only if y is the father of x.

Moreover, the same function maybe given by different formulæ.

Example 1.13. An example familiar from trigonometry is that the function

$$\mathbb{R} \longrightarrow \mathbb{R}, \quad x \longmapsto 1,$$

expressed by the formula $f(x) = 1$ is the same function as

$$\mathbb{R} \longrightarrow \mathbb{R}, \quad x \longmapsto \cos^2(x) + \sin^2(x),$$

expressed by the formula $f(x) = \cos^2(x) + \sin^2(x)$.

Furthermore, there are distinct functions whose domains agree, which agree at every point (and therefore have the same range). Thus the only difference between them is that they have different co-domains: *They only differ in the values they do* **not** *take!* At this stage, it may seem peculiarly pedantic to distinguish such functions, but there are important algebraic and geometric examples, whose detailed study lies beyond the scope of these notes.

Finally, we can define functions *piecewise*, so that its values are determined differently in different parts of its domain.

Lemma 1.14. *Given functions $g : A \to Y$ and $h : B \to Y$, with $g(x) = h(x)$ whenever $x \in A \cap B$, there is a unique function $f : A \cup B \to Y$ such that $f(a) = g(a)$ for all $a \in A$ and $f(b) = h(b)$ for all $b \in B$.*

Proof. Put $X := A \cup B$ and define f by

$$f\colon X \longrightarrow Y, \quad x \longmapsto \begin{cases} g(x) & \text{if } x \in A \\ h(x) & \text{if } x \in B \end{cases}$$

This definition is forced by the requirement that $f(a) = g(a)$ for $a \in A$ and $f(b) = h(b)$ for $b \in B$. This means that the only possible definition of f is the one we have just given. In other words, there cannot be more than one function meeting our requirements.

So the only question remaining is whether there is any such function at all, or, equivalently, whether our f is, in fact, a function.

(i) Since $X = A \cup B$ is the union of two sets, it is, itself, a set.

(ii) Y is, by hyhpothesis, also a set.

(iii) Take $x \in X$. Since $X = A \cup B$, either $x \in A$ or $x \in B$ (or possibly both).

If $x \in A$, then f assigns $g(x) \in Y$ to x, and $g(x)$ is uniquely determined, since $g\colon A \to Y$ is a function.

If $x \in B$, then f assigns $h(x) \in Y$ to x, and $h(x)$ is uniquely determined, since $h\colon B \to Y$ is a function.

Hence, $f\colon X \to Y$ is a function unless it happens to assign two different elements of Y to some element of X. This can only occur when $x \in A \cap B$, for then f assigns both $g(x)$ and $h(x)$ to x. But, by assumption, $g(x) = h(x)$ for all $x \in A \cap B$, so that $f\colon X \to Y$ is, indeed, a function.

□

Observation 1.15. In Lemma 1.14, the fact that $X = A \cup B$ ensures that there cannot be more than one function meeting our requirements, and the fact that g and h agree on $A \cap B$ ensure that there must be at least one such function.

Example 1.16. Consider the definition

$$| \ | \colon \mathbb{R} \longrightarrow \mathbb{R}, \quad x \longmapsto \begin{cases} -x & \text{for } x \leq 0 \\ x & \text{for } x \geq 0 \end{cases}$$

To see that $| \ |$ is a function, we define $\mathbb{R}_0^- := \{x \in \mathbb{R} \mid x \leq 0\}$ and $\mathbb{R}_0^+ := \{x \in \mathbb{R} \mid x \geq 0\}$. Then

(i) $g\colon \mathbb{R}_0^- \to \mathbb{R},\ x \mapsto -x$ and $h\colon \mathbb{R}_0^+ \to \mathbb{R},\ x \mapsto x$ are functions;

(ii) $\mathbb{R} = \mathbb{R}_0^- \cup \mathbb{R}_0^+$;

(iii) $\mathbb{R} = \mathbb{R}_0^- \cap \mathbb{R}_0^+ = \{0\}$ and $g(0) = -0 = 0 = h(0)$.

Hence, by Lemma 1.14, $| \ |$ is a function.

We adhere to the practice of specifying functions in formally correct manner, in order that it become ordinary matter of course for the reader to do so as well.

A function, $f\colon X \to Y$, can be represented by means of its *graph*.

Definition 1.17. The *graph*, $Gr(f)$, of the function $f : X \to Y$ is

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

This representation should be familiar from calculus.

Definition 1.18. The *range* or *image* of the function $f\colon X \to Y$ is the subset $\mathrm{im}(f)$ of Y defined by

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

Notice that $\mathrm{im}(f) \subseteq \mathrm{codom}(f)$ always holds, with equality holding only sometimes. For example if $f : \mathbb{R} \longrightarrow \mathbb{R}$ is defined by $f(x) := 1$ for every $x \in \mathbb{R}$, then $\mathrm{im}(f) = \{1\} \neq \mathbb{R} = \mathrm{codom}(f)$.

Definition 1.19. Given a function $f : X \longrightarrow Y$ and subsets A of X and B of Y, we define

$$\begin{aligned} f(A) &:= \{y \in Y \mid y = f(x) \text{ for some } x \in A\} \\ &= \{f(x) \mid x \in A\} \\ f^{-1}(B) &:= \{x \in X \mid f(x) \in B\}. \end{aligned}$$

Then $f(A)$ is called the *image of A under f* and $f^{-1}(B)$ is called the *inverse image of B under f*, or the *pre-image of B under f*.

Definition 1.20. The *identity function*, on the set X, denoted id_X, is the function

$$id_X : X \longrightarrow X,\ x \longmapsto x,$$

which we may also write as $id_X(x) = x$.

Notice that both the domain and co-domain must be precisely X for this definition to specify the identity function.

Definition 1.21. If X is a subset of Y, then the *inclusion map,* is

$$i_X^Y : X \longrightarrow Y,\ x \longmapsto x.$$

It is often denoted simply by i when the context makes the domain and co-domain clear.

If we use the equational notation, we again have $i(x) = x$. But here, the x on the left of the equality sign is viewed as an element of the set X, whereas on the right hand side it is viewed as an element of the set Y.

Definition 1.22. Given a function $f : X \longrightarrow Y$ and a subset A of X we can use f to define a function $f\mid_A : A \longrightarrow Y$, called the *restriction of* f *to* A by means of

$$f\mid_A (x) = f(x) \quad \text{for every } x \in A.$$

Note that unless $A = X$, this is *not* the same function as f, even though the two functions agree everywhere they are both defined.

Functions can sometimes be *composed*.

Definition 1.23. Given functions $f : X \to Y$ and $g : Y \to Z$ their *composition*, written $g \circ f$, is the function defined by

$$g \circ f : X \longrightarrow Z,\ x \longmapsto g\big(f(x)\big),$$

It is easy to see that

$$\begin{aligned} \operatorname{dom}(g \circ f) &= \operatorname{dom}(f) \\ \operatorname{codom}(g \circ f) &= \operatorname{codom}(g) \\ \operatorname{im}(g \circ f) &\subseteq \operatorname{im}(g). \end{aligned}$$

Equality need not hold in the last of these statements. To see this consider the functions $f : \mathbb{R} \longrightarrow \mathbb{R},\ x \longmapsto 1$ and $g : \mathbb{R} \longrightarrow \mathbb{R},\ y \longmapsto y$. Clearly $\operatorname{im}(g \circ f) = \{1\} \neq \mathbb{R} = \operatorname{im}(g)$.

Given $A \subseteq X$ and a function $f : X \longrightarrow A$, the restriction $f\mid_A : A \longrightarrow Y$ is, in fact, the composition of f and the inclusion of A into X:

$$f\mid_A = f \circ i_A^X.$$

The composition of functions is *associative.*

Lemma 1.24. *Take functions* $g\colon W \to X, f\colon X \to Y$ *and* $e\colon Y \to Z$*. Then the compositions* $(e \circ f) \circ g\colon W \to Z$ *and* $e \circ (f \circ g) : W \to Z$ *are the same function.*

Proof. $\operatorname{dom}\big((e \circ f) \circ g\big) = \operatorname{dom}\big(g\big) = \operatorname{dom}\big(f \circ g\big) = \operatorname{dom}\big(e \circ (f \circ g)\big)$, and, similarly, $\operatorname{codom}\big((e \circ f) \circ g\big) = \operatorname{codom}(e \circ f) = \operatorname{codom}(e) = \operatorname{codom}\big(e \circ (f \circ g)\big)$, it only remains to show that the two functions agree on their common domain, W. Given $w \in W$,

$$\begin{aligned} \big((f \circ g) \circ h\big)(w) &:= (f \circ g)\big(h(w)\big) \\ &:= f\Big(g\big(h(w)\big)\Big) \\ &=: f\big((g \circ h)(w)\big) \\ &=: \big(f \circ (g \circ h)\big)(w) \end{aligned}$$

∎

Lemma 1.25. *Let $f : X \to Y$ be a function, then $id_Y \circ f = f$ and $f \circ id_X = f$.*

Proof. Take $x \in X$. Then

$$\begin{aligned}(id_Y \circ f)(x) &:= id_Y\big(f(x)\big) := f(x)\\ (f \circ id_X)(x) &:= f\big(id_X(x)\big) := f(x)\end{aligned}$$

□

We say that the identity functions act as *neutral elements* with respect to composition.

Sometimes the effect of one function can be "undone" by another. If the first assigns y to x, the second allows us to determine x from y. Composition of functions and hte identity functions allow us to formulate this precisely.

Definition 1.26. The function $f : X \longrightarrow Y$ is said to be *invertible* if there is a function $g : Y \longrightarrow X$ such that $f \circ g = id_Y$ and $g \circ f = id_X$, that is to say if $f\big(g(y)\big) = y$ for every $y \in Y$ and $g\big(f(x)\big) = x$ for every $x \in X$. In this case g is said to be the *inverse* of f.

Note that in our definition of invertibility of the function $f : X \longrightarrow Y$, the function $g : Y \longrightarrow X$ needed to satisfy two conditions. We consider these separately, and introduce terminology for this purpose.

Definition 1.27. The function $g\colon Y \to X$ is a *left inverse* of $f\colon X \to Y$ if and only if $g \circ f = id_X$ and the function $h\colon Y \to X$ is a *right inverse* of $f\colon X \to Y$ if and only if $f \circ h = id_Y$.

Theorem 1.28. *If $f\colon X \to Y$ has both a left and a right inverse, then these must be the same, and hence f is invertible with a uniquely determined inverse.*

Proof. If $e\colon Y \to X$ is left inverse to $f\colon X \to Y$ and $g\colon Y \to X$ is right inverse, then

$$\begin{aligned} e &= e \circ id_Y && \text{by Lemma 1.25}\\ &= e \circ (f \circ g) && \text{as } g \text{ is right inverse to } f\\ &= (e \circ) \circ g && \text{by Lemma 1.24}\\ &= id_X \circ g && \text{as } e \text{ is left inverse to } f\\ &= g && \text{by Lemma 1.25}\end{aligned}$$

□

Observation 1.29. The fact that a function, $f\colon X \to Y$ cannot have more than one inverse justifies the notation f^{-1} usually used to denote function $Y \to X$ inverse to f, for it is uniquely determined by f whenever f is invertible.

To decide whether the function $f\colon X \to Y$ has an inverse does not seem to be an easy task at first glance. If we blindly follow our definition, we would need to try all possible functions from Y to X and see which, if any, satisfy the conditions in the definition. It would be preferable to be able to determine from *intrinsic* properties of f — that is, properties of f alone, without reference to other functions — whether it admits an inverse. Such an intrinsic criterion is available, as we show. To do so, we introduce some important properties of functions.

Definition 1.30. The function $f\colon X \to Y$ is said to be

(i) *1–1* or *injective* or *mono* if and only if it follows from $f(x) = f(x')$ that $x = x'$;

(ii) *onto* or *surjective* or *epi* if and only if given any $y \in Y$ there is an $x \in X$ with $f(x) = y$ — in other words $\operatorname{im}(f) = \operatorname{codom}(f)$;

(iii) *1–1 and onto* or *bijective* or *iso* if and only if it is both 1–1 and onto.

Thus a function is injective if and only if it distinguishes different elements of its domain: different elements of its domain are mapped to different elements of its co-domain.

Similarly, a function is surjective if and only if its range coincides with its co-domain.

Example 1.31. We write $\mathbb{R}_0^+$ for $\{x \in \mathbb{R} \mid x \geq 0\}$.

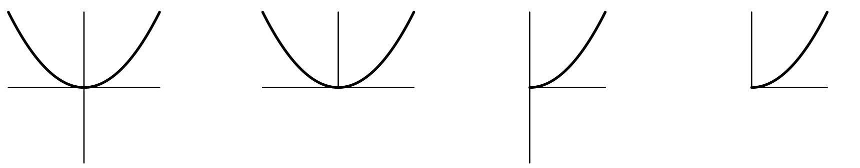
(i) $f : \mathbb{R} \longrightarrow \mathbb{R},\ x \longmapsto x^2$ is neither injective nor surjective, as $f(1) = f(-1)$ and there is no $x \in \mathbb{R}$ with $f(x) = -4$.

(ii) $g : \mathbb{R} \longrightarrow \mathbb{R}_0^+,\ x \longmapsto x^2$ is not injective, but it is surjective, as $f(1) = f(-1)$ and every non-negative real number can be written as the square of a real number.

(iii) $h : \mathbb{R}_0^+ \longrightarrow \mathbb{R},\ x \longmapsto x^2$ is injective, but it not surjective, as $f(x) = f(x')$ if and only if $x^2 = {x'}^2$ if and only if $x' = \pm x$ if and only if $x' = x$ as, by definition, $x, x' \geq 0$. On the other hand, there is no $x \in \mathbb{R}_0^+$ with $f(x) = -4$.

(iv) $k : \mathbb{R}_0^+ \longrightarrow \mathbb{R}_0^+,\ x \longmapsto x^2$ is both injective, and surjective, as should be clear from parts (ii) and (iii).

The differences between these functions is illustrated by their respective graphs



Graph of f Graph of g Graph of h Graph of k

Observation 1.32. The notions of injectivity, surjectivity and bijectivity can also be expressed in terms of equations.

Take sets X and Y, and suppose we have a relation between elements of X and elements of Y, which we express by writing

$$y = f(x)$$

whenever $y \in Y$ is related to $x \in X$.

Then f is a function if and only if for each $x \in X$, the equation $y = f(x)$ has one and only one solution $y \in Y$.

If we restrict attention to relations which are functions, then f is injective if and only if for each $y \in Y$, the equation $y = f(x)$ has *at most one* solution $x \in X$, and it is surjective if and only if for each $y \in Y$, the equation $y = f(x)$ has *at least one* solution $x \in X$.

The formulation in terms of equations suggests that a function has an inverse if and only if it is bijetive (1–1 and onto). This is indeed the case.

Theorem 1.33. *Take a non-empty set X and a function $f : X \longrightarrow Y$ has*

(i) a left inverse if and only if it is injective (or 1–1),

(ii) a right inverse if and only if it is surjective (or onto) and

(iii) an inverse if and only if it is bijective.

Proof. (i) We first suppose that $f : X \to Y$ has a left inverse, say, $g : Y \to X$.

To see that f must then be injective (that is 1–1), suppose that $f(x) = f(x')$. Then

$$x = id_X(x) = (g \circ f)(x) = g\big(f(x)\big) = g\big(f(x')\big) = (g \circ f)(x') = id_X(x') = x',$$

as required.

For the converse, suppose that $f : X \to Y$ is injective. Choose $x_0 \in X$ and define $g : Y \to X$ by

$$g(y) := \begin{cases} x & \text{if } \ y = f(x) \\ x_0 & \text{otherwise} \end{cases}$$

We must first show that g so defined is, in fact a function. For this, g must assign to each $y \in Y$ a uniquely determined $x \in X$. It follows from the definition of g, the only possible obstruction is that g might assign more than one element of X to some element y of Y. From the definition of g, this could only happen when $y \in \mathrm{im}(f)$. In turn, this is only possible when $y = f(x) = f(x')$. But then $x = x'$, since f is injective.

That $g \circ f = id_X$ follows from the definition of g.

(ii) We first suppose that $f : X \to Y$ has a right inverse, say, $g : Y \to X$.

To see that f must be surjective, take $y \in Y$ and put $x := g(y)$. Then

$$f(x) = f\big(g(y)\big) = (f \circ g)(y) = id_Y(y) = y,$$

showing that f is, indeed surjective.

For the converse, suppose that $f : X \to Y$ is surjective. Define $g : Y \to X$ by choosing for each $y \in Y$ an element, say x_y, of X with $f(x_y = x$. This is always possible is because f is surjective.

This g is obviously a function.

Now take $y \in Y$. Then $(f \circ g)(y) = f\big(g(y)\big) = f(x_y) = y$, by the definition of g. Thus $f \circ g = id_Y$.

(iii) This follows from Theorem (1.28) and parts (i) and (ii) here. □

Example 1.34. The last theorem illustrates one of the ways in which two functions can have the same domain and agree everywhere without being the same function.

Let X be a non-empty proper subset of the set Y, so that $X \subset Y$. Then the two functions

$$\begin{aligned} id_X &: X \to X, \quad x \longmapsto x \\ i_X^Y &: X \to Y \quad x \longmapsto x \end{aligned}$$

share a common domain and agree at every point, so that they have the same range: X. But they cannot be the same function. For whereas id_X is invertible — it is its own inverse — Theorem 1.33 tells us that i_X^Y cannot be invertible, since it fails to be surjective, whence it has no right inverse.

Functions can be depicted using diagrams. We say that the diagram

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[dr, "g \circ f"'] & Y \arrow[d, "g"] \\ & Z \end{tikzcd}$$

commutes whenever $h = g \circ f$, in other words, when the composition $g \circ f$ coincides with h.

Similarly, the diagram

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d, "j"'] & B \arrow[d, "g"] \\ C \arrow[r, "k"'] & D \end{tikzcd}$$

commutes when $k \circ j = g \circ f$, in other words, when the compositions $g \circ f$ and $k \circ j$ coincide. We can express the fact that the composition of functions is associative — that is, $(h \circ g) \circ f = h \circ (g \circ h)$ for all functions $f : W \longrightarrow X$, $g : X \longrightarrow Y$ and $h : Y \longrightarrow Z$ — by stating that the diagram

$$\begin{tikzcd} W \arrow[r, "f"] \arrow[d, "g \circ f"'] & X \arrow[d, "h \circ g"] \\ Y \arrow[r, "h"'] & Z \end{tikzcd}$$

commutes, or, equivalently, that the diagram

$$\begin{tikzcd} W \arrow[r, "f"] \arrow[dr, "g \circ f"'] & X \arrow[d, "g"] \arrow[dr, "h \circ g"] & \\ & Y \arrow[r, "h"'] & Z \end{tikzcd}$$

commutes.

1.3 Partitions and Equivalence Relations

For many purposes, two distinct objects can be indistinguishable, or their differences are irrelevant: they are equivalent. We next define this notion formally.

A *relation* between the elements of the set X and those of Y can be represented by the subset of $X \times Y$ comprising those pairs (x, y) $(x \in X, y \in Y)$ such that x stands in the relation R to y. We often write xRy to denote this.

An example is provided by the telephone book. Here we regard X as the set of all subscribers, and Y as all telephone numbers.

If Y happens to coincide with X, we speak of a *binary relation on* X.

Definition 1.35. An *equivalence relation* on X, $\sim$, is a binary relation on X, which is reflexive, symmetric and transitive. That it to say, for all $x, y, z \in X$, we have

Reflexiveness $x \sim x$

Symmetry $x \sim y$ if and only if $y \sim x$.

Transitivity If $x \sim y$ and $y \sim z$, then $x \sim z$.

Given an equivalence relation $\sim$ on X, and $x \in X$, we define

$$[x] := \{t \in X \mid x \sim t\},$$

and call it the *equivalence class* of x. We call any element z of $[x]$ a *representative* of $[x]$.

Finally, we let ${}^{X}\!/\!_{\sim}$ denote the set of all such equivalence classes, so that

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

We then have a function, the *natural map* or the *quotient map*

$$ \eta : X \longrightarrow {}^{X}\!/\!_{\sim}, \quad x \longmapsto [x]. $$

The above construction enjoys a *universal property*:

Theorem 1.36. *Let $\sim$ be an equvalence relation on the set X. Given any set Y and any function $f : X \longrightarrow Y$ with the property that $f(x) = f(x')$ whenever $x \sim x'$, there is a uniquely determined function $\hat{f} : {}^{X}\!/\!_{\sim} \longrightarrow Y$ such that $f = \hat{f} \circ \eta$, that is, $\hat{f}([x]) = f(x)$ for all $[x] \in {}^{X}\!/\!_{\sim}$.*

Proof. The proof is left as an exercise. $\square$

This theorem can be summarised by the commutative diagram

$$\begin{tikzcd}
X \arrow[r, "f"] \arrow[d, "\eta"'] & Y \\
X/\sim \arrow[ur, dashed, "\exists!\, \tilde{f}"'] &
\end{tikzcd}$$

Example 1.37. Given the function $f \circ X \to Y$ define the relation $\sim$ on X by

$$ x \sim u \qquad \text{if and only if} \qquad f(x) = f(u) $$

It is easy to verify directly that $\sim$ is an equivalence relation.

We may identify each equivalence class $[x]$ with the element $f(x)$ of Y, for these uniquely determine each other. This has the effect of identifying ${}^{X}\!/\!_{\sim}$ with $\{y \in Y \mid y = f(x) \text{ for some } x \in X\}$, that is, the range of f, $\operatorname{im}(f)$. The natural projection $\eta\colon X \to {}^{X}\!/\!_{\sim}$ then induces the function

$$ \eta_f \colon X \longrightarrow \operatorname{im}(f), \quad x \longmapsto f(x) $$

If we apply the universal property of the quotient construction to the function $f\colon X \to Y$, we obtain a uniquely determined function $\tilde{f}\colon {}^{X}\!/\!_{\sim} \to Y$ with $f = \tilde{f} \circ \eta$.

Using the identifications introduced, this becomes a unquely determined function $\tilde{f}^{\sharp}\colon \operatorname{im}(f) \to Y$ with $f = \tilde{f}^{\sharp} \circ \eta_f$.

As the inclusion function $i^{Y}_{\operatorname{im}(f)}\colon \operatorname{im}(f) \to Y$, shares this property, we have $\tilde{f}^{\sharp} = i^{Y}_{\operatorname{im}(f)}$.

$$\begin{tikzcd}
X \arrow[r, "f"] \arrow[d, "\eta_f"'] & Y \\
\operatorname{im}(f) \arrow[ur, "i^{Y}_{\mathrm{im(f)}}"'] &
\end{tikzcd}$$

Plainly, η_f is surjective (epi) and $i^{Y}_{\operatorname{im}(f)}$ is injective (mono). So we have shown that every function $f\colon X \to Y$ can be expressed as a mono (injective function) following an epi (surjective function). We summarise this in our next theorem, whose statement requires

Definition 1.38. A *mono-epi factorisation* of the function $f\colon X \to Y$ consists of a mono (injective function), $m\colon W \to Y$, and an epi (surjective function), $e\colon W \to W$ with $f = m \circ e$

Theorem 1.39. *Every function has a mon-epi factorisation.*

Another important notion is that of a *partition* of a set.

Definition 1.40. A *partition* of the set X is a collection of disjoint non-empty subsets of X, $\{X_\lambda \mid \lambda \in \Lambda\}$, whose union is X. Thus $\{X_\lambda \mid \lambda \in \Lambda\}$ is a partition of X if and only if

1. $\emptyset \subset X_\lambda \subseteq X$ for each $\lambda \in \Lambda$
2. $X_\lambda \cap X_\mu = \emptyset$ whenever $\lambda \neq \mu$
3. $X = \bigcup_{\lambda \in \Lambda} X_\lambda$

These two notions, however different they may appear, are intimately related. Indeed, they are two sides of the same coin, as the next theorem shows.

Theorem 1.41. *Every equivalence relation on the set X determines a unique partition of X, and conversely.*

Proof. We outline a proof, leaving the details as an exercise for the reader.

Given the equivalence relation $\sim$ on X, the equivalence classes form a partition of X, that is every $x \in X$ belongs to some equivalence class, and if $[x] \cap [x'] = \emptyset$, then $[x] = [x']$.

If $\{X_\lambda \mid \lambda \in \Lambda\}$ is a partition of X, then

$$x \sim x' \quad \text{if and only if} \quad x, x' \in X_\lambda \quad \text{for some } \lambda \in \Lambda$$

defines an equivalence relation on X

Now show that if we start with an equivalence relation, construct the associated partition, then the associated equivalence relation is the original one.

Finally, show that if we start with a partition, define the associated equivalence relation, then the assoicated partition is the original one. □

1.4 Exercises

Exercise 1.1. Given the function $f : X \longrightarrow Y$ and subsets A of X and B of Y, prove the following statements.

(i) $A \subseteq f^{-1}(f(A))$.

(ii) $f\left(f^{-1}(B)\right) \subseteq B$.

(iii) In general, equality need not hold in either (i) or (ii).

(iv) $G \subseteq f^{-1}(f(G))$ for every subset G of X if and only if f is injective (1–1).

(v) $f\left(f^{-1}(H)\right) \subseteq H$ for every subset H of Y if and only if f is surjective (onto).

Exercise 1.2. Take functions $f : X \longrightarrow Y$ and $g :\longrightarrow Z$. Prove the following statements.

(a) If f and g are both injective, then so is $g \circ f$.

(b) If f and g are both surjective, then so is $g \circ f$.

(c) If $g \circ f$ is injective, then so is f, but not necessarily g.

(d) If $g \circ f$ is surjective, then so is g, but not necessarily f.

(e) If f and g are bijective, so is $g \circ f$.

(f) If $g \circ f$ is bijective, then neither f nor g need be bijective.

Exercise 1.3. Let A, B, C and D be sets. Determine the relationships between

(i) $(A \times C) \cap (B \times D)$ and $(A \cap B) \times (C \cap D)$;

(ii) $(A \times C) \cup (B \times D)$ and $(A \cup B) \times (C \cup D)$.

Exercise 1.4. Given the function $f : X \longrightarrow Y$ and subsets G, H of Y, prove the following statements.

(i) $f^{-1}(G \cap H) = f^{-1}(G) \cap f^{-1}(H)$.

(ii) $f^{-1}(G \cup H) = f^{-1}(G) \cup f^{-1}(H)$.

(iii) $f^{-1}(G \setminus H) = f^{-1}(G) \setminus f^{-1}(H)$.

(iv) $f^{-1}(Y \setminus G) = X \setminus f^{-1}(G)$.

Exercise 1.5. Given the function $f : X \longrightarrow Y$ and subsets A, B of X, find the relationship between the following pairs of subsets of Y.

(i) $f(A \cap B)$ and $f(A) \cap f(B)$.

(ii) $f(A \cup B)$ and $f(A) \cup f(B)$.

(iii) $f(A \setminus B)$ and $f(A) \setminus f(B)$.

(iv) $f(X \setminus A)$ and $Y \setminus f(A)$.

Exercise 1.6. Let $\sim$ be an equvalence relation on the set X.

Prove that if Y is any set and if $f \colon X \longrightarrow Y$ is any function with the property that $f(x) = f(x')$ whenever $x \sim x'$, then there is a uniquely determined function $\hat{f} \colon X/_{\sim} \longrightarrow Y$ such that $f = \hat{f} \circ \eta$, that is, $\hat{f}([x]) = f(x)$ for all $[x] \in X/_{\sim}$.

$$\begin{tikzcd}
X \arrow[r, "f"] \arrow[d, "\eta"'] & Y \\
X/\sim \arrow[ur, dashed, "\exists!\, \tilde{f}"'] &
\end{tikzcd}$$