Chapter 2
Essential Dictionary I

In writing mathematics we use words and symbols to describe facts. We need to explain the meanings of words and symbols, and to state and prove the facts.

We'll be concerned with facts later. In this chapter and the next we list mathematical words with accompanying notation. This is our essential mathematical dictionary. It contains some 200 entries, organised around few fundamental terms: **set**, **function**, **sequence**, **equation**. As we introduce new words, we use them in short phrases and sentences.

Dictionaries are not meant to be read through, so don't be surprised if you find the exposition demanding. Take it in small doses. The last section of this chapter deals with advanced terminology and may be skipped on first reading.

2.1 Sets

A **set** is a collection of *well-defined*, *unordered*, *distinct* objects. (This is the so-called 'naive definition' of a set, due to Cantor.[1]) These objects are called the **elements** of a set, and a set is determined by its elements. We may write

The set of all odd integers
The set of vertices of a pentagon
The set of differentiable real functions

In simple cases, a set can be defined by listing its elements, separated by commas, enclosed within curly brackets. The expression

$$\{1, 2, 3\}$$

[1] Georg Cantor (German: 1845–1918).

F. Vivaldi, *Mathematical Writing*, Springer Undergraduate Mathematics Series,
DOI 10.1007/978-1-4471-6527-9_2

denotes the set whose elements are the integers 1, 2 and 3. Two sets are equal if they have the same elements:

$$\{1, 2, 3\} = \{3, 2, 1\}.$$

(By definition, the order in which the elements of a set are listed is irrelevant.)

It is customary to ignore repeated set elements: $\{2, 1, 3, 1, 3\} = \{2, 1, 3\}$. This convention, adopted by computer algebra systems, simplifies the definition of sets. If repeated elements are allowed and not collapsed, then we speak of a **multiset**: $\{2, 1, 3, 1, 3\}$. The **multiplicity** of an element of a multiset is the number of times the element occurs. Reference to multiplicity usually signals that there is a multiset in the background:

Every quadratic equation has two complex solutions, counting multiplicities.

Multisets are not as common as sets.

The set $\{\}$ with no elements is called the **empty set**, denoted by the symbol $\emptyset$. The empty set is distinct from 'nothing', it is more like an empty container. For example, the statements

This equation has no solutions.
The solution set of this equation is empty.

have the same meaning.

To assign a symbol to a mathematical object, we use an **assignment statement** (or **definition**), which has the following syntax:

$$A := \{1, 2, 3\}. \tag{2.1}$$

This expression assigns the symbolic name A to the set $\{1, 2, 3\}$, and now we may use the former in place of the latter. The symbol ':=' denotes the **assignment operator**. It reads '*becomes*', or '*is defined to be*', rather than '*is equal to*', to underline the difference between assignment and equality (in computer algebra, the symbols = and **:=** are not interchangeable at all!). So we can't write $\{1, 2, 3\} := A$, because the left operand of an assignment operator must be a symbol or a symbolic expression.

The right-hand side of an assignment statement such as (2.1) is a collection of symbols or words that pick out a unique thing, which logicians call the *definiens* (Latin for 'thing that defines'). The left-hand side is a symbol that will be used to stand for this unique thing, which is called the *definiendum* (Latin for 'thing to be defined'). These terms are rather heavy, but they are widely used [36, Chap. 8]. The definiendum may also be a symbolic expression—see below.

While it's very common to use the equal sign '=' also for an assignment, the specialised notation := improves clarity. There are other symbols for the assignment operator, namely

$$\stackrel{\text{def}}{=} \qquad \stackrel{\nabla}{=}, \tag{2.2}$$

which make an even stronger point.

To indicate that x is an element of a set A, we write

$$x \in A \qquad x \textit{ is an element of } A \qquad x \textit{ belongs to } A.$$

The symbol $\notin$ is used to negate membership. Thus

$$\{7, 5\} \in \{5, \{5, 7\}\} \qquad 7 \notin \{5, \{5, 7\}\}.$$

(Think about it.)

A **subset** B of a set A is a set whose elements all belong to A. We write

$$B \subset A \qquad B \textit{ is a subset of } A \qquad B \textit{ is contained in } A$$

and we use $\not\subset$ to negate set inclusion. For example

$$\{3, 1\} \subset \{1, 2, 3\} \qquad \emptyset \subset \{1\} \qquad \{2, 3\} \not\subset \{2, \{2, 3\}\}.$$

Every set has at least two subsets: itself and the empty set. Sometimes these are referred to as the **trivial** subsets. Every other subset—if any—is called a **proper subset**. Motivated by an analogy with $\leqslant$ and $<$, some authors write $\subseteq$ in place of $\subset$, reserving the latter for proper inclusion: $\mathbb{R} \subseteq \mathbb{R}$, $\mathbb{Q} \subset \mathbb{R}$. Proper inclusion is occasionally expressed with the pedantic notation $\subsetneq$.

The **cardinality** of a set is the number of its elements, denoted by the prefix #:

$$\#\{7, -1, 0\} = 3 \qquad \#A = n.$$

The absolute value symbol $|\cdot|$ is also used to denote cardinality: $|A| = n$. Common sense will tell when this choice is sensible. A set is **finite** if its cardinality is an integer, and **infinite** otherwise. To indicate that the set A is finite, without disclosing its cardinality, we write

$$\#A < \infty. \tag{2.3}$$

A more rigorous account of cardinality will be given in Sect. 2.3.3.

Next we consider the words associated with operations between sets. We write $A \cap B$ for the **intersection** of the sets A and B: this is the set comprising elements that belong to both A and B. If $A \cap B = \emptyset$, we say that A and B are **disjoint**, or have **empty intersection**. The sets $A_1, A_2, \ldots$ are **pairwise disjoint** if $A_i \cap A_j = \emptyset$ whenever $i \neq j$.

We write $A \cup B$ for the **union** of A and B, which is the set comprising elements that belong to A or to B (or to both A and B).

We write $A \setminus B$ for the (**set**) **difference** of A and B, which is the collection of the elements of A that do not belong to B. The **symmetric difference** of A and B, denoted by $A \,\Delta\, B$, is defined as

$$A \,\Delta\, B \stackrel{\text{def}}{=} (A \setminus B) \cup (B \setminus A).$$

The assignment operator '$\stackrel{\text{def}}{=}$' [cf. (2.2)] makes it clear that this is a definition. This notation establishes the meaning of $A \,\Delta\, B$, which is a symbolic expression rather than an individual symbol. The following examples illustrate the action of set operators:

$$\begin{aligned}
\{1, 2, 3\} \cap \{3, 4, 5\} &= \{3\} \\
\{1, 2, 3\} \cup \{3, 4, 5\} &= \{1, 2, 3, 4, 5\} \\
\{1, 2, 3\} \setminus \{3, 4, 5\} &= \{1, 2\} \\
\{1, 2, 3\} \,\Delta\, \{3, 4, 5\} &= \{1, 2, 4, 5\}.
\end{aligned}$$

The above **set operators** are **binary**; they have two sets as **operands**. The identities

$$A \cap B = B \cap A \qquad (A \cap B) \cap C = A \cap (B \cap C)$$

express the **commutative** and **associative** properties of the intersection operator. Union and symmetric difference enjoy the same properties, but set difference doesn't.

Let A be a subset of a set X. The **complement** of A (in X) is the set $X \setminus A$, denoted by A' or by A^c. The complement of a set is defined with respect to an **ambient set** X. Reference to the ambient set may be omitted if there is no ambiguity. So we write

The odd integers is the complement of the even integers

since it's clear that the ambient set is the integers.

With set operators we can construct new sets from old ones, although, in a sense, we are recycling things we already have. To create genuinely new sets, we introduce the notion of **ordered pair**. This is an expression of the type (a, b), with a and b arbitrary quantities. Ordered pairs are defined by the property

$$(a, b) = (c, d) \qquad \text{if} \qquad a = c \quad \text{and} \quad b = d. \tag{2.4}$$

The ordered pair (a, b) should not be confused with the set $\{a, b\}$, since for pairs order is essential and repetition is allowed. (Ordered pairs may be defined solely in terms of sets—see Exercise 2.14.) Let A and B be sets. We consider the set of all ordered pairs (a, b), with a in A and b in B. This set is called the **cartesian product** of A and B, and is written as

$$A \times B.$$

Note that A and B need not be distinct; one may write A^2 for $A \times A$, A^3 for $A \times A \times A$, etc. Because the cartesian product is **associative**, the product of more than two sets is defined unambiguously. Also note that the explicit presence of the multiplication operator '$\times$' is needed here, because the expression AB has a different meaning [see Eq. (2.21), Sect. 2.3].

2.1.1 Defining Sets

Defining a set by listing its elements is inadequate for all but the simplest situations. How do we define large or infinite sets? A simple device is to use the **ellipsis** '...', which indicates the deliberate omission of certain elements, the identity of which is made clear by the context. For example, the set $\mathbb{N}$ of **natural numbers** is defined as

$$\mathbb{N} := \{1, 2, 3, \ldots\}.$$

Here the ellipsis represents all the integers greater than 3. Some authors regard 0 as a natural number, so the definition

$$\mathbb{N} := \{0, 1, 2, 3, \ldots\}$$

is also found in the literature. Both definitions have merits and drawbacks; mathematicians occasionally argue about it, but this issue will never be resolved. So, when using the symbol $\mathbb{N}$, one may need to clarify which version of this set is employed.[2] The set of **integers**, denoted by $\mathbb{Z}$ (from the German *Zahlen*, meaning numbers), can also be defined using ellipses:

$$\mathbb{Z} := \{\ldots, -2, -1, 0, 1, 2, \ldots\} \quad \text{or} \quad \mathbb{Z} := \{0, \pm 1, \pm 2, \ldots\}.$$

To define general sets we need more powerful constructs. A **standard definition** of a set is an expression of the type

$$\{x : x \text{ has } \mathscr{P}\} \tag{2.5}$$

where $\mathscr{P}$ is some unambiguous property that things either have or don't have. This expression identifies the set of all objects x that have property $\mathscr{P}$. The colon ':' separates out the object's symbolic name from its defining properties. The vertical bar '|' or the semicolon ';' may be used for the same purpose.

Thus the empty set may be defined symbolically as

$$\emptyset \stackrel{\text{def}}{=} \{x \ : \ x \neq x\}. \tag{2.6}$$

The property $\mathscr{P}$ is 'x is not equal to x', which is not satisfied by any x. Likewise, the cartesian product $A \times B$ of two sets (see Sect. 2.1) may be specified as

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

The rule 'x has property $\mathscr{P}$' now reads: 'x is of the form (a, b) with $a \in A$ and $b \in B$'. The same set may be defined more concisely as

[2] Some authors denote the second version by the symbol $\mathbb{N}_0$.

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

This is a variant of the standard definition (2.5), where the type of object being considered (ordered pair) is specified at the outset. This form of standard definition can be very effective.

The set $\mathbb{Q}$ of **rational numbers**—ratios of integers with non-zero denominator—is defined as follows:

$$\mathbb{Q} := \Big\{ \frac{a}{b} \ : \quad a \in \mathbb{Z}, \quad b \in \mathbb{N}, \quad \gcd(a, b) = 1) \Big\}. \tag{2.7}$$

The property $\mathscr{P}$ is phrased in such a way as to avoid repetition of elements. This is the so-called **reduced form** of rational numbers. The rational numbers may also be defined abstractly, as infinite sets of equivalent fractions—see Sect. 4.6.

One might think that in the expression for a set we could choose any property $\mathscr{P}$. Unfortunately this doesn't work for a reason known as the *Russell-Zermelo paradox*[3] (1901). Consider the set definition

$$W := \{x \ : \ x \notin x\} \tag{2.8}$$

in which $\mathscr{P}$ is the property of being a set that is not a member of itself. The quantity

$$x = \{3, \{3, \{3, \{3\}\}\}\}$$

has property $\mathscr{P}$ and hence belongs to W, whereas

$$x = \{3, \{3, \{3, \{3, \ldots\}\}\}\} \qquad \text{or} \qquad x = \{3, x\}$$

does not have property $\mathscr{P}$ and hence does not belong to W. (In the above expression, the nested parentheses must match, so the notation $\{3, \{3, \{3, \{3, \ldots\}$ is incorrect.)

Given that W is a set of sets, we ask: does W belong to W? We see that if $W \in W$, then W has property $\mathscr{P}$, that is, $W \notin W$, and vice-versa. Impossible! Thus the standard definition (2.8), so deceptively similar to (2.6), does not actually define any set.

Fortunately, we can define a set in such a way that the definition guarantees the existence of the set. A **Zermelo definition** identifies a set W by describing it as

The set of members of X *that have property* $\mathscr{P}$

where the **ambient set** X is given beforehand, and $\mathscr{P}$ is a property that the members of X either have or do not have. In symbols, this is written as

[3] Bertrand Russell (British: 1872–1970); Ernst Zermelo (German: 1871–1953).

$$W := \{x \in X : x \text{ has } \mathscr{P}\}. \tag{2.9}$$

For example, the expression

The set of real numbers strictly between 0 and 1

is a Zermelo definition: the ambient set is the set of real numbers, and we form our set by choosing from it the elements that have the stated property.

Zermelo definitions work because it's a basic principle of mathematics (the so-called *subset axiom*) that for any set X of objects and any property $\mathscr{P}$, there is exactly one set consisting of the objects that are in X and have property $\mathscr{P}$. In Sect. 4.3 we shall see that the definiens of a Zermelo definition—a sentence with a variable x in it—is just a special type of function, called a **predicate**.

Both styles of definitions, standard and Zermelo, are widely used in mathematical writing.

2.1.2 Arithmetic

The notation for arithmetical operations is familiar and established. The **sum** and **difference** of two numbers x and y are always written $x + y$ and $x - y$, respectively. By contrast, their **product** may be written in several equivalent ways:

$$xy \qquad x \cdot y \qquad x \times y, \tag{2.10}$$

and so may their **quotient**:

$$\frac{x}{y} \qquad x/y \qquad x : y.$$

(The notation $x : y$ is used mostly in elementary texts.) Do not confuse the product dot '$\cdot$' with the **decimal point** '.', e.g., $3 \cdot 4 = 12$ and $3.4 = 17/5$.

The **reciprocal** of x, defined for $x \neq 0$, is also written in several ways:

$$\frac{1}{x} \qquad 1/x \qquad x^{-1}$$

while the **opposite** of x is $-x$.

The notation for exponentiation is $x^{\,y}$, where x is the **base** and y the **exponent**. Defining exponentiation for a general exponent is a delicate matter, as it requires the logarithmic and exponential functions. The case of a positive integer exponent is easier, because exponentiation reduces to repeated multiplication:

$$x^n \stackrel{\text{def}}{=} \underbrace{x \cdots x}_{n} \qquad n \geqslant 1.$$

The assignment operator $\stackrel{\text{def}}{=}$ [see (2.2)] indicates that this is a definition, giving meaning to the expression on the left. The use of the under-brace is necessary to specify the number of terms in the product, because all terms are identical. Also note the use of the **raised ellipsis** '$\cdots$' to represent repeated multiplication (or repeated applications of any operator), to be compared with the ordinary ellipsis '...', used for sets and sequences (see Sect. 3.1). Thus

$$\underbrace{x \cdots x}_{4} = x \cdot x \cdot x \cdot x \qquad \underbrace{x, \ldots, x}_{4} = x, x, x, x$$

whereas the notation $x \ldots x$ is incorrect.

In integer arithmetic, the symbol '|' is used for **divisibility**.

$$3|x \qquad 3 \textit{ divides } x \qquad x \textit{ is a multiple of } 3.$$

EXAMPLE. Turn symbols into words:

$$\{x \in \mathbb{Z} : x \geqslant 0,\ 2 \mid x\}.$$

BAD: The set of integers that are greater than or equal to zero, and such that 2 divides them. (*Robotic.*)

GOOD: The set of non-negative even integers.

A positive divisor of n, which is not 1 or n, is called a **proper divisor**, and a **prime** is an integer greater than 1 that has no proper divisors. The acronyms **gcd** and **lcm** are used for **greatest common divisor** and **least common multiple**. (The expression **highest common factor** (hcf)—a variant of gcd which is popular in schools—is seldom used in higher mathematics.) Two integers are **co-prime** (or **relatively prime**) if their greatest common divisor is 1. Some authors use (a, b) for $\gcd(a, b)$; this is to be avoided, since this notation is already overloaded.

The following concise notation represents certain infinite sets of integers (here k and m are any integers):

$$\begin{aligned} k + m\mathbb{Z} &\stackrel{\nabla}{=} \{x \in \mathbb{Z} : m|(x - k)\} \\ &= \{\ldots, k - 2m, k - m, k, k + m, k + 2m, \ldots\}. \end{aligned} \tag{2.11}$$

This definition gives meaning to the symbolic expression $k + m\mathbb{Z}$ on the left of the assignment operator, which otherwise would be meaningless (you can't form the sum or product of an integer and a set!). The two expressions on the right represent the same object. While any of the two would suffice, their combination adds clarity (we shall expand this idea in Sect. 6.4).

This notation is economical and effective:

$$\begin{array}{ll} x \in 1 + 2\mathbb{Z} & x \text{ is odd} \\ x \in m\mathbb{Z} & x \text{ is a multiple of } m \\ x \in n^2\mathbb{Z} \setminus 2\mathbb{Z} & x \text{ is an odd multiple of } n^2. \end{array}$$

This is a special case of a more general notation for sums and products of sets, to be developed in Sect. 2.3.

2.1.3 Sets of Numbers

The 'open face' symbols $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ were introduced in Sect. 2.1.1 to represent the natural numbers, the integers, and the rationals, respectively. Likewise, we denote by $\mathbb{R}$ the set of **real** numbers (its symbolic definition is left as Exercise 2.13), while the set of **complex** numbers is denoted by $\mathbb{C}$. The set $\mathbb{C}$ may be written as

$$\mathbb{C} \stackrel{\text{def}}{=} \left\{ x + \mathrm{i}y \,:\, \mathrm{i}^2 = -1, \;\; x, y \in \mathbb{R} \right\}.$$

The symbol i is called the **imaginary unit**, while x and y are, respectively, the **real part** $\mathrm{Re}(z)$ and the **imaginary part** $\mathrm{Im}(z)$ of the complex number $z = x + \mathrm{i}y$. The sets $\mathbb{R}$ and $\mathbb{C}$ are represented geometrically as the **real line** and the **complex plane** (or **Argand plane**), respectively. A plot of complex numbers in the Argand plane is called an **Argand diagram**. We have the chain of proper inclusions

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

We now construct new sets from the sets of numbers introduced above. An **interval** is a subset of $\mathbb{R}$ of the type

$$[a, b] := \{x \in \mathbb{R} \,:\, a \leqslant x \leqslant b\}$$

where a, b are real numbers, with $a < b$. This interval is **closed**, that is, it contains its end points. (A point is sometimes regarded as a degenerate closed interval, by allowing $a = b$ in the definition.) We also have **open** intervals

$$(a, b) := \{x \in \mathbb{R} \,:\, a < x < b\}$$

as well as **half-open** intervals

$$[a, b) \qquad (a, b].$$

The notational clash between an open interval $(a, b) \subset \mathbb{R}$ and an ordered pair $(a, b) \in \mathbb{R}^2$ is unfortunate but unavoidable, since both notations are firmly established. For (half) open intervals, there is the following alternative—and very logical—notation

$$]a, b[\qquad [a, b[\qquad]a, b],$$

which, for some reason, is not so common.

The interval with end-points $a = 0$ and $b = 1$ is the (open, closed, half-open) **unit interval**. A semi-infinite interval

$$\{x \in \mathbb{R} : a < x\} \qquad \{x \in \mathbb{R} : x \leqslant b\}$$

is called a **ray**. The rays consisting of all positive real or rational numbers are particularly important, and have a dedicated notation

$$\mathbb{R}^+ := \{x \in \mathbb{R}, \ \ x > 0\} \qquad \mathbb{Q}^+ := \{x \in \mathbb{Q}, \ \ x > 0\} \tag{2.12}$$

whereas $\mathbb{Z}^+$ is just $\mathbb{N}$.

Some authors extend the meaning of interval to include also rays and lines, and use expressions such as

$$(-\infty, \infty) \qquad [a, \infty) \qquad (-\infty, b]. \tag{2.13}$$

As infinity does not belong to the set of real numbers, the notation $[1, \infty]$ is incorrect.

A variant of (2.12) is used to denote non-zero real and rational numbers:

$$\mathbb{R}^* := \{x \in \mathbb{R}, \ \ x \neq 0\} \qquad \mathbb{Q}^* := \{x \in \mathbb{Q}, \ \ x \neq 0\}. \tag{2.14}$$

This notation is common but not universally recognised; before using these symbols, a clarifying comment may be appropriate (see Sect. 6.2).

The set $\mathbb{R}^2$ of all ordered pairs of real numbers is called the **cartesian plane**, which is the cartesian product of the real line with itself. If $(x, y) \in \mathbb{R}^2$, then the first component x is called the **abscissa** and the second component y the **ordinate**.

The set $\mathbb{Q}^2 \subset \mathbb{R}^2$, the collection of points of the plane having rational coordinates, is called the set of **rational points** in $\mathbb{R}^2$. The set $[0, 1]^2 \subset \mathbb{R}^2$ is called the **unit square**. In $\mathbb{R}^3$ we have the **unit cube** $[0, 1]^3$, and for $n > 3$ we have the **unit hypercube** $[0, 1]^n \subset \mathbb{R}^n$. The following subsets of the cartesian plane are related to the geometrical figure of the circle:

$$\begin{aligned} &\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\} && \textbf{unit circle} \\ &\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leqslant 1\} && \textbf{closed unit disc} \\ &\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\} && \textbf{open unit disc}. \end{aligned} \tag{2.15}$$

Thus the closed unit disc is the union of the open unit disc and the unit circle. The (unit) circle is denoted by the symbol $\mathbb{S}^1$.

For $n \geqslant 0$, the n-**dimensional unit sphere** $\mathbb{S}^n$ is defined as follows:

$$\mathbb{S}^n = \{(x_0, \ldots, x_n) \in \mathbb{R}^{n+1} \,:\, x_0^2 + \cdots + x_n^2 = 1\}$$

This Zermelo definition, to be compared with the Definition (2.15) of the unit circle $\mathbb{S}^1$, employs a combination of ordinary and raised ellipses. For $n = 0$, we have $\mathbb{S}^0 = \{-1, 1\}$.

2.1.4 Writing About Sets

The vocabulary on sets developed so far is sufficient for our purpose. We begin to use it in short phrases which define sets.

1. *The set of ordered pairs of complex numbers.*
2. *The set of rational points on the unit circle.*
3. *The set of prime numbers with fifty decimal digits.*
4. *The set of lines in the cartesian plane, passing through the origin.*

Note that we haven't used any symbols. The set in item 1 is $\mathbb{C}^2$. In item 2, among the infinitely many points of the unit circle, we consider those having rational co-ordinates. There is no difficulty in writing this set symbolically:

$$\{(x, y) \in \mathbb{Q}^2 \,:\, x^2 + y^2 = 1\}$$

although its properties are not obvious from the definition. This set is non-empty (the points $(0, \pm 1)$, $(\pm 1, 0)$ belong to it), but is it infinite? This example illustrates the power of a verbal definition. Item 3, which defines a subset of $\mathbb{N}$, makes an even stronger point. This set must be extremely large, but can we even show that it is non-empty? In item 4, each line counts as a single element, rather than an infinite collection of points (otherwise our set of lines would be the whole plane). The symbolic definition of this set is awkward; to simplify it we'll consider suitable **representations** of this set (Sect. 2.3.3).

It is possible to specify a *type* of set, without revealing its precise identity. In each of the following sets there is at least one unspecified quantity.

The set of fractions representing a rational number.
The set of divisors of an odd integer.
A proper infinite subset of the unit circle.
The cartesian product of two finite sets of complex numbers.
A finite set of consecutive integers.

Next we define sets in two ways, first with a combination of words and symbols, and then with words only. One should consider the relative merits of the two formulations.

Let $X = \{3\}$.
The set whose only element is the integer 3.
Let $X = \{m\}$, *with* $m \in \mathbb{Z}$.
A set whose only element is an integer.
Let $m \in \mathbb{Z}$, *and let* X *be a set such that* $m \in X$.
A set which contains a given integer.
Let X *be a set such that* $X \cap \mathbb{Z} \neq \emptyset$.
A set which contains at least one integer.
Let X *be a set such that* $\#(X \cap \mathbb{Z}) = 1$.
A set which contains precisely one integer.

In the first two examples the combination of 'let' and '=' replaces an assignment operator. An expression such as '*Let* $X \stackrel{\triangledown}{=} \{3\}$' would be overloaded.

The distinction between definite and indefinite articles is essential, the former describing a unique object, the latter an unspecified representative of a class of objects. In the following phrases, a change in one article, highlighted in boldface, has resulted in a logical mistake.

BAD: *A proper infinite subset of* **a** *unit circle.*
BAD: **A** *set whose only element is the integer 3.*
BAD: **The** *set whose only element is an integer.*
BAD: **The** *set which contains precisely one integer.*

As a final exercise, we express some geometric facts using set terminology.

The intersection of a line and a conic section has at most two points.
The set of rational points in any open interval is infinite.
A cylinder is the cartesian product of a segment and a circle.
The complement of the unit circle consists of two disjoint components.

The reader should re-visit familiar mathematics and describe it in the language of sets.

Exercise 2.1 For each of the following topics:

prime numbers, fractions, complex numbers,

(i) write five short sentences; (ii) ask five questions. The sentences should give a definition or state a fact; the questions should have mathematical significance, and preferably possess a certain degree of generality. **[**$\not\in$**]**[4]

Exercise 2.2 Define five interesting finite sets. **[**$\not\in$**]**

[4] The symbol **[**$\not\in$**]** indicates that the exercise must be completed without using any mathematical symbol.

Exercise 2.3 The following expressions define sets. Turn words into symbols, using standard or Zermelo definitions. (Represent geometrical objects, e.g., planar curves, by their cartesian equations.)

1. *The set of negative odd integers.*
2. *The set of natural numbers with three decimal digits.*
3. *The set of rational numbers which are the ratio of consecutive integers.*
4. *The set of rational points in the closed unit cube.*
5. *The complement of the open unit disc in the complex plane.*
6. *The set of vectors of unit length in three-dimensional euclidean space.*
7. *The set of circles in the plane, passing through the origin.*
8. *The set of hyperbolae in the plane, whose asymptotes are the coordinate axes.*
9. *The set of lines tangent to the unit circle.*

Exercise 2.4 The following expressions define sets. Turn symbols into words. [$\not\in$]

1. $\{x \in \mathbb{Q} \,:\, 0 < x < 1\}$
2. $\{1/(2n+1) \,:\, n \in \mathbb{Z}\}$
3. $\{m2^{-k} \,:\, m \in 1 + 2\mathbb{Z},\ \ k \in \mathbb{N}\}$
4. $\{x \in \mathbb{R} \smallsetminus \mathbb{Z} \,:\, x^2 \in \mathbb{Z}\}$
5. $\{z \in \mathbb{C} \smallsetminus \mathbb{R} \,:\, z^2 \in \mathbb{R}\}$
6. $\{z \in \mathbb{C} \,:\, |\mathrm{Re}(z)| + |\mathrm{Im}(z)| \leqslant 1\}$
7. $\{(m, n) \in \mathbb{Z}^2 \,:\, m \,|\, n\}$
8. $\{(x, y, z) \in \mathbb{R}^3 \,:\, x\, y\, z = 0\}$
9. $\{(x_1, \ldots, x_n) \in \mathbb{R}^n \,:\, \sum_k x_k = 0\}$
10. $\{x \in \mathbb{R} \,:\, \sin(2\,\pi\, x) = 0\}$
11. $\{(x, y) \in \mathbb{R}^2 \,:\, \sin(\pi\, x)\sin(\pi\, y) = 0\}$.

2.2 Functions

Functions are everywhere. Whenever a process transforms a mathematical object into another object, there is a function in the background. 'Function' is arguably the most used word in mathematics.

A **function** consists of two sets together with a rule[5] that assigns to *each* element of the first set a *unique* element of the second set. The first set is called the **domain** of the function and the second set is called the **co-domain**. A function whose domain is a set A may be called a function **over** A or a function **defined on** A. The terms **map** or **mapping** are synonymous with function. The term **operator** is used to describe certain types of functions (see below).

A function is usually denoted by a single letter or symbol, such as f. If x is an element of the domain of a function f, then the **value of** f **at** x, denoted by $f(x)$

[5] Below, we'll replace the term 'rule' with something more rigorous.

is the unique element of the co-domain that the rule defining f assigns to x. The notation

$$f : A \to B \qquad x \mapsto f(x) \tag{2.16}$$

indicates that f is a function with domain A and co-domain B that **maps** $x \in A$ **to** $f(x) \in B$. The symbol x is the **variable** or (the **argument**) of the function. The symbols $\to$ and $\mapsto$ have different meanings, and should not be confused. The function

$$I_A : A \to A \qquad x \mapsto x$$

is called the **identity** (**function**) on A. When explicit reference to the set A is unnecessary, the identity is also denoted by Id or $\mathbb{1}$.

In Definition (2.16) the symbols used for the function's name and variable are inessential; the two expressions

$$f : \mathbb{R} \setminus \{0\} \to \mathbb{R} \quad x \mapsto \frac{1}{x} \qquad x : \mathbb{R} \setminus \{0\} \to \mathbb{R} \quad f \mapsto \frac{1}{f}$$

define exactly the same function (even though the rightmost expression breaks just about every rule concerning mathematical notation—see Sect. 6.2).

Let us use the word 'function' in short expressions. These are function definitions:

1. *The integer function that squares its argument.*
2. *The function that returns 1 if its argument is rational, and 0 otherwise.*
3. *The function that counts the number of primes smaller than a given real number.*
4. *The function that gives the distance between two points on the unit circle, measured along the circumference.*

We surmise that the function in item 2 is defined over the real numbers. Item 3 is a much-studied function in number theory. The set of values assumed by the function in item 4 is the closed interval $[0, \pi]$.

Functions of several variables are defined over cartesian products of sets. For example, the function

$$f : \mathbb{Z} \times \mathbb{Z} \to \mathbb{N} \qquad (x, y) \mapsto \gcd(x, y)$$

depends on two integer arguments, and hence is defined over the cartesian product of two copies of the integers. This definition requires a value for $\gcd(0, 0)$, which normally is taken to be zero.

Let $f : A \to B$ be a function. The set

$$\{(x, f(x)) \in A \times B \, : \, x \in A\} \tag{2.17}$$

is called the **graph** of f. A function is completely specified by three sets: domain, co-domain and graph. We can now reformulate the definition of a function, replacing the imprecise term 'rule' with the precise term 'graph'. We write a formal definition.

DEFINITION. *A* **function** f *is a triple* (X, Y, G) *of non-empty sets. The sets* X *and* Y *are arbitrary, while* G *is a subset of* $X \times Y$ *with the property that for every* $x \in X$ *there is a unique pair* $(x, y) \in G$. *The quantity* y *is called the* **value of the function at** x, *denoted by* $f(x)$.

We see that, besides sets, the definition of a function requires the constructs of ordered pair and triple. It turns out that these quantities can be defined solely in terms of sets (see Exercise 2.14). So, to define functions, all we need are sets after all.

Given a function $f : A \to B$, and a subset $X \subset A$, the set

$$f(X) \stackrel{\text{def}}{=} \{f(x) \,:\, x \in X\} \tag{2.18}$$

is called the **image of** X **under** f. The assignment operator gives meaning to the symbolic expression $f(X)$, which otherwise would be meaningless, since we stipulated that the argument of a function is an element of the domain, not a subset of it. Thus $\sin(\mathbb{R})$ is the closed interval $[-1, 1]$.

Clearly, $f(X) \subset B$, and $f(A)$ is the smallest set that can serve as co-domain for f. The set $f(A)$ is often called the **image** or the **range** of the function f. This term is sometimes used to mean co-domain, which should be avoided. A **constant** is a function whose image consists of a single point.

The notation (2.18) is suggestive and widely used. However, in computer algebra, the quantities $f(x)$ and $f(X)$ (with x an element and X a subset of the domain, respectively) are written with a different syntax, e.g., `f(x)` and `map(f,X)` with Maple.

A function is said to be **injective** (or **one-to-one**) if distinct points of the domain map to distinct points of the co-domain. A function is **surjective** (or **onto**) if $f(A) = B$, that is, if the image coincides with the co-domain. A function that is both injective and surjective is said to be **bijective**.

For any non-empty subset X of the domain A, we define the **restriction of** f **to** X as

$$f|_X : X \to B \qquad\qquad x \mapsto f(x).$$

Given two functions $f : A \to B$ and $g : B \to C$, their **composition** is the function

$$g \circ f : A \to C \qquad x \mapsto g(f(x)). \tag{2.19}$$

The notation $g \circ f$ reminds us that f acts before g. The image $g(f(x))$ of x under $g \circ f$ is denoted by $(g \circ f)(x)$, where the parentheses isolate $g \circ f$ as the function's symbolic name. The hybrid notation $g \circ f(x)$ should be avoided.

If $f : A \to B$ is a bijective function, then the **inverse** of f is the function $f^{-1} : B \to A$ such that

$$f^{-1} \circ f = \mathbb{1}_A \qquad f \circ f^{-1} = \mathbb{1}_B$$

where $\mathbb{1}_{A,B}$ are the identities in the respective sets. A function is said to be **invertible** if its inverse exists. If $f : A \to B$ is injective, then we can always define the inverse of f by restricting its domain to $f(A)$ if necessary. In absence of injectivity, it may still be possible to construct an inverse by a suitable restriction of the function. Thus the arcsine may be defined by restricting the sine to the interval $[-\pi/2, \pi/2]$.

Let $f : A \to B$ be a function, and let C be a subset of B. The set of points

$$f^{-1}(C) \stackrel{\text{def}}{=} \{x \in A : f(x) \in C\} \tag{2.20}$$

is called the **inverse image** of the set C.

Since the definition of inverse image does not involve the inverse function, the inverse image exists even if the inverse function does not. These two concepts must be distinguished carefully. When the reciprocal of a function comes into play, things get very confusing, since we now have three unrelated objects represented by closely related notation:

$$f^{-1}(x) \qquad f^{-1}(\{x\}) \qquad f(x)^{-1}.$$

The first expression is well-defined if x belongs to the image of f and f is invertible there. In the second expression there is no condition on f, and x need only be an element of the co-domain. In the third expression the point x must belong to the domain of f, and $f(x)$ must be non-zero. Thus

$$\sin^{-1}(1) = \frac{\pi}{2} \qquad \sin^{-1}(\{1\}) = \frac{\pi}{2} + 2\pi\,\mathbb{Z} \qquad \sin(1)^{-1} = \csc(1).$$

In the first expression we tacitly assume that $\sin^{-1} = \arcsin : [-1, 1] \to [-\pi/2, \pi/2]$. In the third expression the symbol csc denotes the co-secant ($\csc(x) = 1/\sin(x)$), defined in the domain $\mathbb{R} \setminus \pi\,\mathbb{Z}$.

With a judicious use of definite and indefinite articles, we can specify a function's type without committing ourselves to a specific object.

1. *The inverse of a trigonometric function.*
2. *The composition of a function with itself.*
3. *An integer-valued bijective function.*
4. *A function which coincides with its own inverse.*

In item 2, we infer that the function maps its domain into itself. Functions of type 3 will be considered in the next section to define cardinality of sets. Functions of type 4 are called **involutions** (e.g., $x \mapsto -x$, over a suitable domain).

Writing about **real functions** is considered in Chap. 5.

Exercise 2.5 Turn symbols into words. [$\not\in$]

1. f^{-1}
2. $f^{-1}(0)$
3. $f(x^{-1})$
4. $f(0)^{-1}$
5. $f \circ f$
6. $(f \circ g)^{-1}$
7. $f|_{\mathbb{Z}}$
8. $f(\mathbb{R}^{+})$
9. $f(\mathbb{R} \setminus \mathbb{Q})$
10. $f(A) \cap f(B)$
11. $f(\mathbb{R}) \cap \mathbb{Q}$
12. $\mathbb{Z} \cap f^{-1}(\mathbb{Z})$.

Exercise 2.6 Explain clearly and plainly. [$\not\in$, 50][6]

1. How do I divide two fractions?
2. I have a positive integer. How do I check if it's prime?
3. I have a positive integer. How do I check if it's a cube?
4. I have a positive integer. How do I check if it's the sum of two squares?
5. I have a cartesian equation of a circle, and a point. How do I check if the point lies inside the circle?
6. I have two lines in three-dimensional space. How do I check if they intersect?
7. I have two vectors on the plane. How do I check if they are linearly independent?
8. I have a list of quadratic polynomial functions and I must select the functions that assume both positive and negative values. What shall I do?
9. I have two real functions. From a sketch, it seems that their graphs intersect, and do so at a right angle. How do I verify that this is indeed the case?
10. I have three points on the plane. How do I compute the centre of the circle passing through them?
11. I have four points on the plane. How do I check if the points are vertices of a square?

Exercise 2.7 I have two finite sets and a function between them. I am able to compute the value of the function at each point of the domain, and to count and compare the elements of these sets. I need explicit instructions for answering the following questions. [$\not\in$, 50]

1. How do I check that my function is surjective?
2. How do I check that my function is injective?

Exercise 2.8 Answer each question as clearly as you can.

1. Let A and B be sets. Why are the sets $A^2 \setminus B^2$ and $(A \setminus B)^2$ not necessarily equal? Under what conditions are they equal?
2. Let $f : X \to Y$ be a function, and let A be a subset of X. Why are the sets A and $f^{-1}(f(A))$ not necessarily equal? Under what conditions are they equal?
3. Let A and B be subsets of the domain of a function f. Why are $f(A) \cap f(B)$ and $f(A \cap B)$ not necessarily equal? Under what conditions are they equal?

[6] Each assignment should contain no mathematical symbols and approximately 50 words.

2.3 Some Advanced Terms

We develop some advanced terminology and notation on sets. The content of this section is not essential for the rest of the book.

2.3.1 Families of Sets

The **power set** $\mathbf{P}(A)$ of a set A is the set of all subsets of A. Thus if $A = \{1, 2, 3\}$, then

$$\mathbf{P}(A) = \{\{\}, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}\,.$$

If A has n elements, then $\mathbf{P}(A)$ has 2^n elements. Indeed to construct a subset of A we consider each element of A and we decide whether to include it or leave it out, giving n binary choices. (A rigorous proof requires induction, see Exercise 8.4.)

A **partition** of a set A is a collection of pairwise disjoint non-empty subsets of A whose union is A. So a partition of A is a subset of $\mathbf{P}(A)$. For instance, the set $\{\{2\}, \{1, 3\}\}$ is a partition of $\{1, 2, 3\}$, and the even and odd integers form a partition of $\mathbb{Z}$. A partition may be described as a decomposition of a set into **classes**.

We write some phrases using these terms. Consider carefully the distinction between definite and indefinite articles (see Sect. 2.1.4 and Exercise 2.9).

The power set of a finite set.
The power set of a power set.
A partition of a power set.
The set of all partitions of a set.
A set of partitions of the natural numbers.
A finite partition of an infinite set.

Now some sentences:

Let us partition our interval into finitely many equal sub-intervals.
The plane may be partitioned into concentric annuli.
There is no finite partition of a triangle into squares.

2.3.2 Sums and Products of Sets

Let X and Y be sets of numbers. The **algebraic sum** $X + Y$ and **product** XY (also known as **Minkowski**[7] **sum** (**product**)), are defined as follows:

$$X + Y \stackrel{\text{def}}{=} \{x + y : x \in X,\ y \in Y\} \qquad XY \stackrel{\text{def}}{=} \{xy : x \in X,\ y \in Y\} \tag{2.21}$$

[7] Hermann Minkowski (Polish: 1864–1909).

with the stipulation that repeated elements are to be ignored. For example, if $X = \{1, 3\}$ and $Y = \{2, 4\}$, then

$$X + Y = \{3, 5, 7\} \qquad XY = \{2, 4, 6, 12\}.$$

The expression 'sum of sets' is always understood as an algebraic sum. In the case of product, it is advisable to use the full expression to avoid confusion with the cartesian product.

If $X = \{x\}$ consists of a single element, then we use the shorthand notation $x + Y$ and xY in place of $\{x\} + Y$ and $\{x\}Y$, respectively (as we did in Sect. 2.1.2 for integers). For example

$$\frac{1}{2} + \mathbb{N} = \left\{\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \ldots\right\} \qquad \pi\,\mathbb{Z} = \{\ldots, -2\,\pi, \pi, 0, \pi, 2\,\pi, \ldots\}.$$

This notation leads to concise statements such as

$$m\mathbb{Z} + n\mathbb{Z} = \gcd(m, n)\mathbb{Z}$$

which combines algebraic sum and product of sets (see Exercise 7.5).

Elementary—but significant—applications of this construct are found in **modular arithmetic**. Let m be a positive integer. We say that two integers x and y are **congruent modulo** m if m divides $x - y$. This relation is denoted by[8]

$$x \equiv y \pmod{m}.$$

Thus

$$-3 \equiv 7 \pmod 5 \qquad 1 \not\equiv 12 \pmod 7.$$

The integer m is called the **modulus**. The set of integers congruent to a given integer is called a **congruence** (or **residue**) **class**. One verifies that the congruence class of k modulo m is the infinite set $k + m\mathbb{Z}$ given explicitly in (2.11). The congruence class of k modulo m is also denoted by $[k]_m$, $k \pmod m$, or, if the modulus is understood, by $[k]$ or $\overline{k}$.

The set of congruence classes modulo m is denoted by $\mathbb{Z}/m\mathbb{Z}$. If $m = p$ is a prime number, the notation $\mathbb{F}_p$ (meaning 'the field with p elements') may be used in place of $\mathbb{Z}/p\mathbb{Z}$. The set $\mathbb{Z}/m\mathbb{Z}$ contains m elements, which form a **partition** of $\mathbb{Z}$:

$$\mathbb{Z}/m\mathbb{Z} = \{m\mathbb{Z}, 1 + m\mathbb{Z}, 2 + m\mathbb{Z}, \ldots, (m - 1) + m\mathbb{Z}\}.$$

Variants of this notation are used extensively in algebra, where one defines the sum/product of more general sets, such as groups and rings.

[8] This notation is due to Carl Friedrich Gauss (German: 1777–1855).

2.3.3 Representations of Sets

Consider the following geometrical sets:

The set of triangles with a vertex at the origin.
The set of triples of mutually tangent circles.

These definitions are easy to grasp, but how are we meant to work with sets of this kind? Suppose that we require a data structure suitable for computer implementation. We must then identify each element of our set with one or more concrete objects, such as numbers or matrices. This identification gives a description of a set in terms of another set, hopefully easier to handle.

Two sets A and B are said to be **equivalent** (written $A \sim B$) if there is a **bi-unique correspondence** between the elements of A and the elements of B, namely, if there exists a bijective function $f : A \to B$. A set equivalent to $\{1, 2, \ldots n\}$ is said to have **cardinality** n, and a set equivalent to $\mathbb{N}$ is said to be **countable** or **countably infinite**. The set $\mathbb{Z}$ is countable, and so is $m\mathbb{Z}$, for any $m \in \mathbb{N}$. A set X is **uncountable** if it contains a countably infinite subset Y, but X is not equivalent to Y. The set $\mathbb{R}$ is uncountable. We see that characterising the cardinality of infinite sets requires a more sophisticated approach than mere 'counting'.

A **representation** of a set A is any set B which is equivalent to A. (This is the most general acceptation of the term representation; in algebra, representations are based on a more specialised form of equivalence.)

For instance, the open unit interval and the real line are equivalent, as established by the bijective function

$$f : \mathbb{R} \to (0, 1) \qquad x \mapsto \frac{1}{\pi} \arctan(x) + \frac{1}{2}. \tag{2.22}$$

Likewise, the exponential function establishes the equivalence $\mathbb{R} \sim \mathbb{R}^+$.

Let us consider representations of the set L of lines in the plane passing through a given point (a, b). The set L is uncountable. An element λ of L is an infinite subset of $\mathbb{R}^2$, which we write symbolically as

$$\lambda = \left\{(x, y) \in \mathbb{R}^2 \,:\, y = b + s(x - a)\right\}$$

where s is a real number representing the line's slope. The line $x = a$ is not of this form, and must be treated separately. Collecting all the lines together, we obtain a symbolic description of L:

$$L = \left\{\left\{(x, y) \in \mathbb{R}^2 \,:\, y = b + s(x - a)\right\} \,:\, s \in R\right\} \cup \left\{\left\{(a, y) \in \mathbb{R}^2 \,:\, y \in \mathbb{R}\right\}\right\}.$$

The simple verbal definition of L seems to have drowned in a sea of symbols!

We look for a set equivalent to L with a more legible structure. An obvious simplification results from representing L as a set of **cartesian equations**:

$$L \sim \{y = b + s(x - a) \,:\, s \in \mathbb{R}\} \cup \{x = a\}.$$

We have merely replaced the solution set of an equation with the equation itself (cf. Sect. 3.3). This identification provides the desired bi-unique correspondence between the two sets.

We can simplify further. Because a and b are fixed, there is no need to specify them explicitly; it suffices to give the (possibly infinite) value of the slope. Alternatively, we could identify a line by an angle θ between 0 and π, measured with respect to some reference axis passing through the point (a, b). Because the angles 0 and π correspond to the same line, only one of them is to be included, resulting in the half-open interval $[0, \pi)$. The equivalence between $\mathbb{R} \cup \{\infty\}$ and $[0, \pi)$ may be achieved with a transformation of the type (2.22), where the included end-point 0 corresponds to the point at infinity.

Finally, any half-open interval may be identified with the **circle** $\mathbb{S}^1$, by gluing together the end-points of the interval. In our case this is achieved with the function $\theta \mapsto (\cos(2\theta), \sin(2\theta))$. The essence of our set is now clear:

$$L \sim \mathbb{R} \cup \{\infty\} \sim [0, \pi) \sim \mathbb{S}^1.$$

Exercise 2.9 Consider the phrases displayed in Sect. 2.3.1. Provide an example of each object being defined.

Exercise 2.10 Why are the sets $\emptyset$ and $\{\emptyset\}$ distinct? What are the elements of the power set $\mathbf{P}(\mathbf{P}(\mathbf{P}(\emptyset)))$? Explain.

Exercise 2.11 Represent the algebraic sum of sets as a function.

Exercise 2.12 Consider the function that performs the prime factorization of a natural number greater than 1. What would you choose for co-domain? Explain, discussing possible representations.

Exercise 2.13 Represent the following sets:

1. the set of open segments in the plane; the subset of segments of unit length;
2. the set of triangles with one vertex at the origin;
3. the polynomial set $\mathbb{Z}[x]$, as a set of integer sequences;
4. the real numbers, as a set of integer sequences;
5. the set of all finite subsets of $\mathbb{N}$, as a set of rational numbers.

Exercise 2.14 Prove that the definition

$$(a, b) \stackrel{\text{def}}{=} \{\{a\}, \{a, b\}\}$$

satisfies (2.4). (This shows that an ordered pair can be defined in terms of a set, so there's no need to introduce a new object.) Hence define an ordered triple in terms of sets.

http://www.springer.com/978-1-4471-6526-2

Mathematical Writing
Vivaldi, F.
2014, XVII, 204 p. 16 illus. in color., Softcover
ISBN: 978-1-4471-6526-2