CHAPTER 0

Notation and terminology

In this chapter we collect together the principal notations and definitions used in this book. It is an article of faith among present-day mathematicians that every mathematical structure can be described in set theoretic terms, and so, although we make no attempt to explain the set theory involved, all the definitions in this chapter are made in terms of sets. The rules for manipulating sets, and the definition of function (in Section 0–3), must be thoroughly understood before any progress can be made, but no great harm will be done if, at any rate on a first reading, the other definitions given here are treated in an informal and intuitive way, the reader returning to the formal definitions when the logical and aesthetic necessity for them has become apparent. In any case it is suggested that the reader reviews Sections 0–1, 0–2, 0–3, and 0–5 and then proceeds to Chapter 1, returning to this chapter when the need arises.

0–1 SETS

We take an entirely naive and intuitive view of set theory.† For our purposes a set is what is suggested by any of the terms "collection", "class", "aggregate", "family". The objects which make up a given set are called its *elements*, *points*, or *members*. A set is determined by its elements, and two sets are equal if, and only if, they have the same elements. The fact that x is a member of a set A is expressed by writing $x \in A$; if x is not a member of A we write $x \notin A$. A set may be specified by enumerating its elements, usually between braces, $\{\dots\}$. Thus the set of positive integers less than five is $\{1, 2, 3, 4\}$; the first three letters of the English alphabet form the set $\{a, b, c\}$. It should be realized that in this notation there is no significance in the order in which the elements are written, so that the sets $\{1, 2\}$ and $\{2, 1\}$ are the same set. Moreover the sets $\{1\}$ and $\{1, 1\}$ are the same because they have the same elements, but the sets $\{1\}$ and $\{\{1\}\}$ are not since they do not have the same elements, the first consisting of the number 1 and the second consisting of the set $\{1\}$, which is not equal to 1.

There is another very widely used notation for sets. The symbol $\{x: P(x)\}$, where $P(x)$ is a proposition about x, is the set of all x for which $P(x)$ is true. Thus $\{1, 2, 3, 4\} = \{x: x \text{ is a positive integer and } x < 5\}$. If A is a set, then

† An extremely readable account of the basic facts of set theoretic life will be found in Halmos [13].

$\{x: x \in A\}$ is a set and, in view of the definition of equality of sets, $A = \{x: x \in A\}$.

There is one set which has a symbol of its own. This is the empty set, denoted by $\varnothing$, which has no elements. In the notation above, $\varnothing = \{x: x \neq x\}$ or $\varnothing = \{x: x \text{ is a rational number such that } x^2 = 2\}$. The empty set is often awkward to handle, but its extreme usefulness will become apparent later.

Suppose that A and B are sets. If every member of A is a member of B we say that A is a *subset* of B, that B *contains* A or that A is *contained in* B, and in this case we write $A \subset B$ or $B \supset A$. If $A \subset B$ and $B \subset A$, then $A = B$; if $A \subset B$ but $A \neq B$, then A is a *proper subset* of B. If A is not a subset of B we write $A \not\subset B$. The empty set is a subset of every set, that is for any set A, $\varnothing \subset A$. This may be a little puzzling, and possibly the easiest way of seeing it is to observe that if A is a set then $x \notin A$ implies $x \notin \varnothing$ (because $\varnothing$ has no members), so that $\varnothing \subset A$ (because the proposition $A \subset B$ holds if, and only if, $x \notin B$ implies $x \notin A$).

0–2 ALGEBRA OF SETS

We now introduce some operations on sets. It is important to obtain facility in handling these operations, and many of the statements made in the text below should be regarded as exercises.

The *union* of two sets A and B is the set formed by putting together the elements of A and B in a single set (see Figure 0–1). Formally:†

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

It is easy to verify the following properties of the union operation.

$$A \cup B = B \cup A;$$

$$A \cup (B \cup C) = (A \cup B) \cup C;$$

$$A \cup A = A;$$

$$A \cup \varnothing = A.$$

As an example of the sort of argument involved we prove the following proposition:

"If A and B are sets, $A \subset B$ if, and only if, $A \cup B = B$."

If $A \subset B$ and $x \in A \cup B$, then either $x \in B$ or $x \in A$, so that $x \in B$ (since $A \subset B$). Hence $x \in A \cup B$ implies $x \in B$ and so $A \cup B \subset B$. Since it is clear that $B \subset A \cup B$, we have $A \cup B = B$. Conversely, suppose that $A \cup B = B$. Then if $x \in A$, $x \in A \cup B = B$ and so $x \in B$. Thus $x \in A$ implies $x \in B$, and we have $A \subset B$.

† The word "or" is used here, and elsewhere, in its logical sense; for example "$x \in A$ or $x \in B$" means that one of $x \in A$, $x \in B$ must be true, but the possibility that both are true is not excluded.

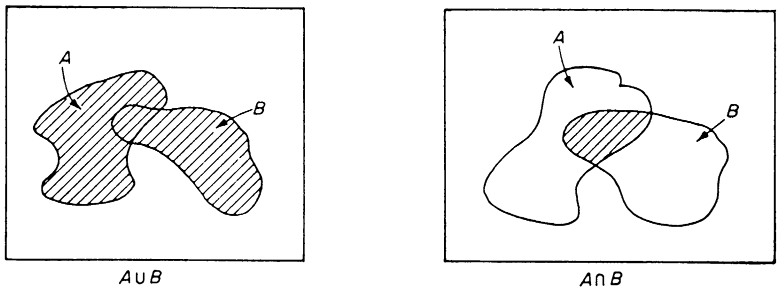

Figure 0–1

If A and B are sets, their *intersection* is the set of points which they have in common. Formally:

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

If A and B have no points in common, then $A \cap B = \varnothing$, and A and B are then said to be *disjoint*.

Exercises 0–1

1. $A \cap B = B \cap A$.
2. $A \cap (B \cap C) = (A \cap B) \cap C$.
3. $A \cap A = A$, $A \cap \varnothing = \varnothing$.
4. $A \subset B$ if, and only if, $A \cap B = A$.
5. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.
6. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

If A and B are sets, the *complement of B relative to A*, written $A \sim B$, is the set of all points which are in A and not in B (see Figure 9–2). Formally:

$$A \sim B = \{x : x \in A \text{ and } x \notin B\}.$$

If, in a given context, all sets are subsets of a given set X, then $X \sim A$ is called simply *the complement of A*.

Exercises 0–2

1. $A \sim (A \sim B) = A \cap B$.
2. $A \sim B = A$ if, and only if, $A \cap B = \varnothing$.
3. $A \sim (B \cup C) = (A \sim B) \cap (A \sim C)$.
4. $A \sim (B \cap C) = (A \sim B) \cup (A \sim C)$.
5. $(A \sim B) \cup B = A \cup B$.
6. $A \sim (B \sim C) = (A \sim B) \cup (A \cap C)$.
7. $(A \cup B) \sim C = (A \sim C) \cup (B \sim C)$.

8. Suppose that A and B are subsets of a set X. Show that $A \sim B = A \cap (X \sim B)$. Show also that $A \subset B$ if, and only if, one of the following conditions is satisfied: $X \sim B \subset X \sim A$, $A \cap (X \sim B) = \varnothing$, $(X \sim A) \cup B = X$.
9. The *symmetric difference* of two sets A and B, sometimes denoted by $A \triangle B$, is defined by
$$A \triangle B = (A \sim B) \cup (B \sim A).$$
Show that for any A, B, C
$$A \triangle A = \varnothing, \qquad A \triangle \varnothing = A, \qquad A \triangle B = B \triangle A,$$
$$A \triangle (B \triangle C) = (A \triangle B) \triangle C, \qquad A \cap (B \triangle C) = (A \cap B) \triangle (A \cap C).$$

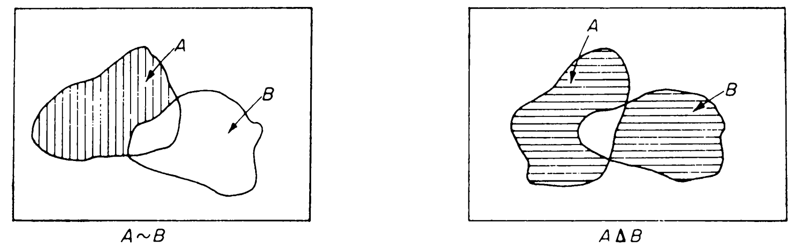

Figure 0–2

0–3 ORDERED PAIRS AND FUNCTIONS

The intuitive idea of an ordered pair is quite simple; an ordered pair (x, y) consists of two objects x and y of which one, x, is regarded as the first and the other, y, as the second. Two ordered pairs are equal if, and only if, they have the same first member and the same second member.

It is a fact of more than passing interest that it is possible to make an adequate definition of ordered pair in terms of sets. The discussion is rather technical and, since the intuitive description given above will be adequate for later use, it may be omitted at a first reading.

If x and y are any two objects, the *ordered pair* (x, y) is defined to be the set $\{\{x\}, \{x, y\}\}$; x is said to be the first member of the ordered pair and y the second. We have of course to prove that our definition has the properties that an ordered pair ought to have.

Theorem 0–1. If (x, y) and (s, t) are ordered pairs, then $(x, y) = (s, t)$ if, and only if, $x = s$ and $y = t$.

Proof. It is clear that if $x = s$ and $y = t$, then $(x, y) = (s, t)$.

Suppose conversely that $(x, y) = (s, t)$, that is that $\{\{x\}, \{x, y\}\} = \{\{s\}, \{s, t\}\}$. First observe that this implies $\{x\} \in \{\{s\}, \{s, t\}\}$ and hence that either $\{x\} = \{s\}$ or $\{x\} = \{s, t\}$, and in any case that $s \in \{x\}$. Consequently $x = s$.

We now have to prove that $y = t$. Suppose first that $x = y$. Then

$$\{\{x\}, \{x, t\}\} = \{\{s\}, \{s, t\}\} = \{\{x\}, \{x, y\}\} = \{\{x\}, \{x\}\} = \{\{x\}\},$$

and it follows that $\{x, t\} = \{x\}$, so that $x = t = y$. If $x \neq y$ then $\{x, y\} \neq \{x\}$, and the hypothesis $\{\{x\}, \{x, y\}\} = \{\{x\}, \{x, t\}\}$ implies $\{x, y\} = \{x, t\}$. Consequently, $y \in \{x, t\}$ and, since $y \neq x$, we must have $y = t$, and the proof is complete. ∎

We are now ready to make what is possibly the most important definition in mathematics—that of function.

In elementary analysis a function is usually thought of as being given by a formula such as

$$f(x) = x^2 + 1,$$

or

$$f(x) = \log(|\cos x| + 1).$$

In this case, given a value of the independent variable x, we can compute the value of the function. Sometimes, to cover more general situations, a function is described as a "rule" which enables us to associate with each member of one set a unique member of another set, the sets involved usually being sets of real numbers. The weakness of these descriptions of the idea of function lies in the vagueness of the words "formula" and "rule". In view of the frequency with which functions are used in analysis and elsewhere, it is desirable to have a definition of function formulated in terms of sets. Fortunately a simple definition of this kind is available; its construction is based upon the idea of representing a function graphically.

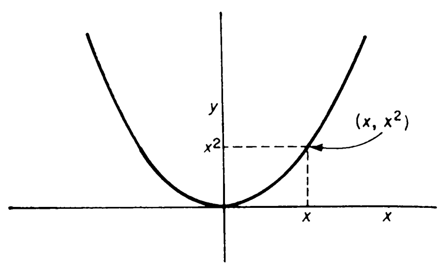

Figure 0–3. Graph of $f(x) = x^2$.

The function $f(x) = x^2$ $(x \in \mathbf{R})$, where $\mathbf{R}$ is the set of real numbers, is represented graphically in $\mathbf{R}^2$ by representing it as the set (Figure 0–3) $\{(x, x^2) : x \in \mathbf{R}\}$.

A moment's reflection shows us that the graph of $f(x) = x^2$ gives us as much information about the function as the formula used for computing it; for, to determine whether or not $y = f(x)$, we have only to decide whether or not the point (x, y) lies on the graph (or in the set $\{(x, x^2) : x \in \mathbf{R}\}$). We have not, of course, proved that the graph of a function can be identified with the function; such a statement is not susceptible of proof without a proper definition of function. What we have done is to suggest quite strongly that, if we were to make a definition of function in terms of the concept of graph (and hence in terms of sets), we should have an object which intuitively differs very little from what is commonly thought of as a function.

Accordingly we make the following definition. A *function* (*map, mapping, transformation*) f is a set of ordered pairs (that is a set, each member of which is an ordered pair) with the property that if $(x, z) \in f$ and $(x, y) \in f$ then $y = z$. If f is a function and $(x, y) \in f$, y is called the *value of f at x* and is often denoted by $f(x)$. The *domain of f*, denoted by dom (f), is the set $\{x : (x, y) \in f$ for some $y\}$ so that roughly speaking dom (f) is the set of points on which f is defined. The point of the condition imposed on the set f, that $(x, y) \in f$ and $(x, z) \in f$ implies $y = z$, is to ensure that f can have only one value at each member of its domain. In other words the only functions we shall consider are *one-valued*.

If f is a function, the *range of f*, denoted by ran (f), is the set $\{y : (x, y) \in f$ for some $x\}$; that is, the range of f is the set of values that f assumes.

The notation $f: X \to Y$ conveys the following information: f is a function with domain X and range a subset of Y; we sometimes say that f is on X *to* (or *into*) Y. If $f: X \to Y$ and ran $(f) = Y$, we say that f is a function on X *onto* Y.

It is sometimes convenient, particularly when discussing functions which are defined on a set of real numbers, to use a different notation. The phrase "$x \to f(x)$ $(x \in A)$" means that f is a function with domain A whose value at x $(\in A)$ is $f(x)$. Thus the squaring function $\{(x, x^2) : x \in \mathbf{R}\}$ may be written $x \to x^2$ $(x \in \mathbf{R})$.

It is important, particularly in formal arguments, to adhere strictly to the terms of the above definition. The tendency in elementary analysis to identify a function with the formula used to define it must be avoided. For example, the functions $f_1 = \{(x, x^2) : 0 \leqslant x \leqslant 1\}$ and $f_2 = \{(x, x^2) : 0 \leqslant x \leqslant 2\}$ have the sets $\{x : 0 \leqslant x \leqslant 1\}$, $\{x : 0 \leqslant x \leqslant 2\}$ as their respective domains, and so are not the same function despite the fact that each uses the same formula $(f(x) = x^2)$ for computing its value at a given point of the domain.

The distinction between f_1 and f_2 is sufficiently important to have generated a notation to describe it. Suppose that f is a function with domain X and that A is a subset of X; then *the restriction of f to A*, denoted by $f|_A$, is the function $\{(x, y) : x \in A$ and $(x, y) \in f\}$, that is, roughly speaking, the function whose domain is A but which in this domain takes the same values as f.† In the example above, if

$$A = \{x : 0 \leqslant x \leqslant 1\}$$

† This distinction is not just a logical nicety but a point of practical importance. Thus in Chapter 3 we shall encounter the function f defined on $X = \{x : 0 \leqslant x \leqslant 1\}$ by $f(x) = 1$ if x is rational, $f(x) = 0$ if x is irrational. If A is the set of rational numbers in x, then $f|_A$ (being constant) is continuous on A but f is not continuous at any point of A. Thus a restriction of f may behave quite differently from f.

then

$$f_1 = f_2|_A.$$

If f is a function and A is a subset of dom (f), the *image of A under f*, denoted by $f[A]$, is the set $\{f(x): x \in A\}$; evidently $f[A] = \operatorname{ran}(f|_A)$. Note also that if $x \in \operatorname{dom}(f)$, $f[\{x\}] \neq f(x)$ because $f[\{x\}]$ is the *set* $\{f(x)\}$. If $f: X \to Y$ and $B \subset Y$, the *inverse image* (or *counter-image*) *of B under f*, denoted by $f^{-1}[B]$, is the set $\{x: x \in X \text{ and } f(x) \in B\}$. This last definition does not require B to be a subset of ran (f); in fact, if B ran $(f) = \varnothing$, we shall have $f^{-1}[B] = \varnothing$. It should be observed that even if B is a singleton (that is, has just one element), $f^{-1}[B]$ may have more than one element. For example, if f is the function $x \to x^2$ $(x \in \mathbf{R})$ then $f^{-1}[\{1\}] = \{1, -1\}$.

One method of representing functions visually is shown in Figure 0-4.

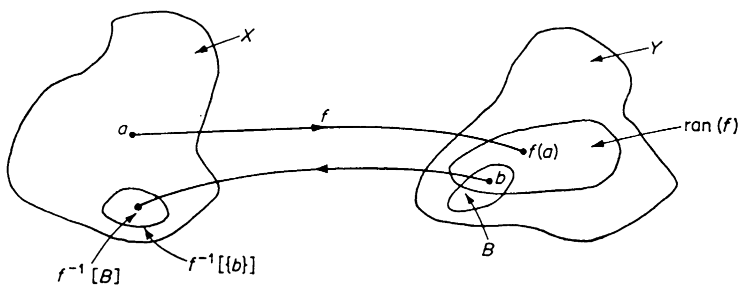

Figure 0-4. The function f: $X \to Y$.

Functions for which $f^{-1}[\{y\}]$ consists of a single point for every y in the range of f are sufficiently important to deserve a special name. A function f: $X \to Y$ is (1-1) (or a *one-to-one mapping*) if, whenever x and y are in dom (f), $f(x) = f(y)$ implies $x = y$. If f: $X \to Y$ is (1-1), the set $\{(y, x):(x, y) \in f\}$ defines a function, denoted by f^{-1}, with domain equal to the range of f and range equal to the domain of f. To see that f^{-1} is a function, suppose that $(x, z) \in f^{-1}$ and $(x, y) \in f^{-1}$; then $(z, x) \in f$ and $(y, x) \in f$ so that $f(z) = f(y)$ $(= x)$, and hence, since f is (1-1), $y = z$. If f: $X \to Y$ is (1-1), the function f^{-1}: ran $(f) \to X$ is called the *inverse of f*. It evidently has the properties:

$$f(f^{-1}(y)) = y \quad (y \in \operatorname{ran}(f)),$$

$$f^{-1}(f(x)) = x \quad (x \in \operatorname{dom}(f)).$$

There is a certain ambiguity in the notation we have just introduced; the symbol $f^{-1}[B]$ could stand for the inverse image of B under f or, if f^{-1} exists (that is, if f is (1-1)), for the image of B under f^{-1}. However, if f^{-1} exists, these two interpretations of $f^{-1}[B]$ give the same set so that no real confusion can arise. In any case this ambiguity will be removed when we discuss relations in general.

Exercises 0-3

If f: $X \to Y$ is a function, and A and B are subsets of Y, then:

1. $f^{-1}[A \sim B] = f^{-1}[A] \sim f^{-1}[B]$.
2. $f^{-1}[A \cup B] = f^{-1}[A] \cup f^{-1}[B]$.
3. $f^{-1}[A \cap B] = f^{-1}[A] \cap f^{-1}[B]$.

Suppose that f: $X \to Y$ is a function.

4. $f[f^{-1}[B]] \subset B$ for every $B \subset Y$; and $f[f^{-1}[B]] = B$ for every $B \subset Y$ if, and only if, f is onto Y.
5. $A \subset f^{-1}[f[A]]$ for every $A \subset X$ and $A = f^{-1}[f[A]]$ for every $A \subset X$ if, and only if, f is (1-1).
6. $f(A \cap B) = f(A) \cap f(B)$ for every $A \subset X$ and $B \subset X$ if, and only if, f is (1-1).
7. $Y \sim f[A] \subset f[X \sim A]$ if, and only if, f is onto.

For any set S let i_S: $S \to S$ be the mapping defined by $i_S(x) = x$ $(x \in S)$, called the *identity mapping*.

8. If f: $X \to X$ is (1-1) and onto, then $f \circ f^{-1} = i_X$ and $f^{-1} \circ f = i_X$. If g: $X \to X$ is such that $f \circ g = g \circ f = i_X$, then $g = f^{-1}$ (see Section 0-6). Give an example of two functions f, g on $\mathbf{R}$ to $\mathbf{R}$ such that $f \circ g = g \circ f$ but $f \neq g^{-1}$.
9. Give an example (i) of a function f: $X \to Y$ such that for some subsets $A \subset X$, $B \subset X$, $f[A \cap B] \neq f[A] \cap f[B]$; (ii) of a function f: $X \to Y$ such that for some subset $A \subset X$, $f[X \sim A] \neq f[X] \sim f[A]$.

0-4 INDEXED FAMILIES AND SEQUENCES

Suppose that f is a function with domain I. Then f is the set $\{(i, f(i)): i \in I\}$, where the ordered pair $(i, f(i))$ denotes that f has the value $f(i)$ at i. Now provided we know in advance that the object we are dealing with is a function, there is a certain amount of redundancy in writing $(i, f(i))$ to indicate that f has value $f(i)$ at i; this information is equally well conveyed by writing $f(i)$, and indeed this is the definition of the symbol $f(i)$. It has become customary to use this abbreviation systematically and to write $\{f(i)\}_{i \in I}$, or more frequently $\{f_i\}_{i \in I}$, for a function on I whose value at i is $f(i)$, or f_i.

We sometimes also write, for typographical convenience, $\{f_i\}$ $(i \in I)$ instead of $\{f_i\}_{i \in I}$ and sometimes altogether omit mention of the set I if the context makes it feasible. A function f written in the form $\{f_i\}_{i \in I}$ is sometimes referred to as an indexed family (or set), and the set I as the indexing set. It is important, when using the indexed family notation for a function, not to confuse the symbol $\{f_i\}$ $(i \in I)$ with $\{f_i: i \in I\}$; the first is a function, and the second is the range of the function.

In the case when I is an infinite subset of **P** (the set of non-negative integers)† the above notation has been in use for a very long time.

A *sequence* is a function whose domain is an infinite subset of **P**. If x is a sequence whose domain is D and whose value at $n \in D$ is x_n, then the symbol $\{x_n\}_{n \in D}$ stands for the function (sequence) x. For our purposes all sequences will have domain **P** or domain ω (the set of strictly positive integers), and we shall usually suppress mention of the domain and write $\{x_n\}$ for $\{x_n\}_{n \in D}$, relying on the context to make it clear which domain is intended. A sequence $\{x_n\}$ is *finite* if its range $\{x_n : n \in D\}$ is finite, otherwise it is *infinite*.

Suppose that $\{x_n\}$ is a sequence, then $\{y_n\}$ is a *subsequence* of $\{x_n\}$ if there is a function $f: \omega \to \omega$ with the properties: $f(k+1) \geqslant f(k)$ $(k \in \omega)$, $f[\omega]$ is infinite (or given $K \in \mathbf{R}$, $f(k) > K$ for some $k \in \omega$), and $y_n = x_{f(n)}$ $(n \in \omega)$. An ancient device used when discussing subsequences is to denote the function $k \to f(k)$ $(k \in \omega)$ described above by $k \to n_k$ $(k \in \omega)$. A subsequence of $\{x_n\}$ is then of the form $\{x_{n_k}\}$ where $n_{k+1} \geqslant n_k$ $(k \in \omega)$ and $\{n_k\}$ is infinite. As examples of subsequences we can cite the subsequences $\{x_{2n}\}$, $\{x_{n^2}\}$ of $\{x_n\}$.

The indexed family notation can be used to extend the definition of union and intersection of sets. Suppose that $\{A_i\}$ $(i \in I)$ is an indexed family of sets. Then

$$\bigcup \{A_i : i \in I\} = \{x : x \in A_i, \text{ for some } i \in I\}$$

$$\bigcap \{A_i : i \in I\} = \{x : x \in A_i, \text{ for every } i \in I\}.$$

Exercises 0–4

1. Let X be a set, and let $\{X_i\}$ $(i \in I)$ be an indexed family of subsets of X. Then

$$X \sim \bigcup \{X_i : i \in I\} = \bigcap \{X \sim X_i : i \in I\}$$

$$X \sim \bigcap \{X_i : i \in I\} = \bigcup \{X \sim X_i : i \in I\}.$$

2. If f is a function and $\{X_i\}$ $(i \in I)$ is an indexed family of sets, then

$$f^{-1}[\bigcup \{X_i : i \in I\}] = \bigcup \{f^{-1}[X_i] : i \in I\}$$

$$f^{-1}[\bigcap \{X_i : i \in I\}] = \bigcap \{f^{-1}[X_i] : i \in I\}.$$

0–5 CARTESIAN PRODUCTS

Suppose that A and B are sets. The *Cartesian product of A and B* denoted by $A \times B$ is the set of all ordered pairs (a, b) with a in A and b in B; that is,

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

† The relevant facts about the integers and finite and infinite sets appear in Section 1–2.

The Euclidean plane is the set of ordered pairs (x, y) with $x \in \mathbf{R}$ and $y \in \mathbf{R}$ so that $\mathbf{R} \times \mathbf{R}$ is the Euclidean plane; we denote it by $\mathbf{R}^2$. The Cartesian product $\{x : 0 \leqslant x \leqslant 1\} \times \{x : 1 \leqslant x \leqslant 2\}$ is the square in $\mathbf{R}^2$ bounded by the lines $x = 0$, $x = 1$, $y = 1$, $y = 2$.

There is an alternative description of Cartesian product which enables us to generalize the notion to more than two sets. Suppose that $I = \{a, b\}$ is any set with just two distinct elements ($\{1, 2\}$ for example). Let C be the set of all families $\{z_i\}$ $(i \in I)$ indexed by I which have the property $z_a \in A$ and $z_b \in B$; in other words, C is the family of all functions g on $\{a, b\}$ such that $g(a) \in A$ and $g(b) \in B$. If we define $f: C \to A \times B$ by $f(z) = (z_a, z_b)$ $(z \in C)$, then f is a (1–1) mapping of C onto $A \times B$ (the verification of this assertion is left as an exercise). This mapping enables us to identify the Cartesian product $X_1 \times X_2$ with the set of all families z indexed on a set $\{1, 2\}$ of two elements such that $z_1 \in X_1$ and $z_2 \in X_2$. It also suggests the following generalization. Let $\{X_i\}$ $(i \in I)$ be an indexed family of sets. The Cartesian product of the family, denoted by $\times \{X_i : i \in I\}$, is the set of all indexed families $\{x_i\}$ $(i \in I)$ which have the property $x_i \in X_i$ $(i \in I)$. Alternatively (and equivalently), $\times \{X_i : i \in I\}$ is the set of all functions g on I which have the property $g(i) \in X_i$ $(i \in I)$. However, the first description of the Cartesian product is probably the one with the most intuitive content.

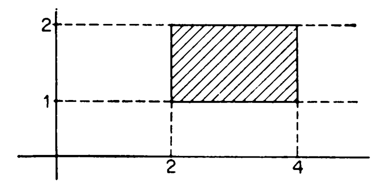

Figure 0–5. The Cartesian product $\{x : 2 \leqslant x \leqslant 4\} \times \{x : 1 \leqslant x \leqslant 2\}$.

When the index set is I finite consisting (say) of n elements, we take it to be the set of integers $\{r : r \in \omega,\ 1 \leqslant r \leqslant n\}$, and then the Cartesian product $\times \{X_i : i \in I\}$ (also written $X_1 \times X_2 \times \cdots \times X_n$) is the set of indexed families $\{x_i\}$ $(i \in I)$ with $x_i \in X_i$. In this case we write the family $\{x_i\}$ as $(x_1, x_2, \ldots, x_n)$ and call it an *ordered n-tuple*. We have seen above that when $n = 2$ this terminology is compatible with the earlier definition of ordered pair. In the case when $X_1 = X_2 = \cdots = X_n = X$ we write X^n for $X_1 \times X_2 \times \cdots \times X_n$, and $\mathbf{R}^n$ is called Euclidean space of n-dimensions.

Exercises 0–5

1. Let A, B, C, and D be sets. Then

$$(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$$

$$(A \times B) \sim (C \times D) = \{(A \sim C) \times (B \sim D)\} \cup \{(A \cap C) \times (B \sim D)\} \cup \{(A \sim C) \times (B \cap D)\}.$$

2. Let $X_1, X_2, \ldots, X_n$ be sets, and let $(x_1, \ldots, x_n)$, $(y_1, y_2, \ldots, y_n)$ be ordered n-tuples in (that is, members of) $X_1 \times X_2 \times \cdots \times X_n$. Show that $(x_1, x_2, \ldots, x_n) = (y_1, y_2, \ldots, y_n)$ if, and only if, $y_i = x_i$ $(i = 1, \ldots, n)$.

0–6 RELATIONS

A relation is usually thought of as an expression like "wife of" in which the related objects are husbands and wives. If y is a husband, then the set of all the wives of y is the set of all objects which bear this relation to y. As in the case of functions, the relation is adequately expressed by writing down the set of ordered pairs (x, y) such that y bears the relation to x. This leads us to the following definition. A *relation* is a set of ordered pairs. Thus the relation "son of" is the set of all ordered pairs (x, y) where y is a father and x is one of his sons. If A is a set, the relation $\in$ is $\{(a, A): a \in A\}$; the relation $\subset$ is $\{(B, A): B \subset A\}$. A function is a relation.

If R is a relation we write $(x, y) \in R$ or xRy interchangeably, and say that x is R-related to y if $(x, y) \in R$. If R is a relation, the *domain of* R is the set $\{x: (x, y) \in R \text{ for some } y\}$; the *range of* R is the set $\{y: (x, y) \in R \text{ for some } x\}$. Evidently R is a subset of (domain of R) $\times$ (range of R). If R is a relation, and A is a subset of the domain of R, then $R[A] = \{y: (x, y) \in R$ for some $x \in A\}$; the relation R^{-1}, the *relation inverse to* R, is $\{(x, y): (y, x) \in R\}$. It is a worthwhile exercise to verify that these notations agree with those already set out for functions.

If R and S are relations, their *composition* $R \circ S$ is the relation $\{(x, z):$ for some y, $(x, y) \in S$ and $(y, z) \in R\}$. We leave it as an exercise to verify that, if R and S are relations, then $(R \circ S)^{-1} = S^{-1} \circ R^{-1}$, and that, if R and S are functions, then $R \circ S$ is a function, and that $(R \circ S)(x) = R(S(x))$ $(x \in \text{domain } S)$.

If R is a relation with domain X and range a subset of X, R is called a relation *in* X. A relation in X is *reflexive* if $(x, x) \in R$ for every $x \in X$. The *identity relation in* X, denoted by Δ, is the relation $\{(x, x): x \in X\}$ so that a relation R in X is reflexive if $\Delta \subset R$. A relation R in X is *symmetric* if $(x, y) \in R$ implies $(y, x) \in R$, that is, if $R = R^{-1}$. A relation R in X is *transitive* if $(x, y) \in R$ and $(y, z) \in R$ imply $(x, z) \in R$; equivalently R is transitive if $R \circ R \subset R$.

A relation which is reflexive, transitive and symmetric is called an *equivalence relation*. If R is an equivalence relation in X and $x \in X$, the set $R[\{x\}] = \{y: (x, y) \in R\}$ is called the *equivalence class* of x. The set of equivalence classes generated by an equivalence relation R in X is denoted by X/R. An equivalence relation is completely determined by its equivalence classes as the following theorem shows.

Theorem 0–2. A relation R in X is an equivalence relation if, and only if, there is a family $\mathscr{A}$ of subsets of X such that

a) $X = \bigcup \{A: A \in \mathscr{A}\}$,

b) if $A \in \mathscr{A}$ and $B \in \mathscr{A}$, then either $A = B$ or $A \cap B = \varnothing$,

c) $R = \bigcup \{A \times A: A \in \mathscr{A}\}$, and in this case $\mathscr{A}$ is the family of equivalence classes of R.

Proof. Suppose that R is an equivalence relation on X. Let $\mathscr{A} = \{R[\{x\}]: x \in X\}$. Then since, for each $x \in X$, $x \in R[\{x\}] \subset X$, $X \subset \bigcup \{A: A \in \mathscr{A}\} \subset X$, and so (a) holds.

Suppose that $x \in X$, $y \in X$, and $R[\{x\}] \cap R[\{y\}] \neq \varnothing$. Then for some $t \in X$, $(x, t) \in R$ and $(y, t) \in R$. If $s \in R[\{x\}]$, then $(x, s) \in R$ so that, since R is symmetric, $(s, x) \in R$, and hence $(s, t) \in R \circ R$ and so, because R is transitive, $(s, t) \in R$. Hence, using symmetry again, $(t, s) \in R$ and hence $(y, s) \in R \circ R$. Thus, using transitivity again, $(y, s) \in R$ and so $s \in R[\{y\}]$. Hence $R[\{x\}] \subset R[\{y\}]$. We may prove similarly that $R[\{y\}] \subset R[\{x\}]$, and so (b) holds. Since $(x, y) \in R$ iff $y \in R[\{x\}]$, (c) is clear and the necessity of the conditions has been proved.

Now suppose that $\mathscr{A}$ is a family of subsets of X with the properties (a) and (b), and that R is the relation defined by (c). We have to show that R is an equivalence relation. First, because of (a), if $x \in X$, then $x \in A$ for some $A \in \mathscr{A}$, and so $(x, x) \in R$. Thus R is reflexive. Next, if $(x, y) \in R$ then, by (c), for some $A \in \mathscr{A}$, $x \in A$ and $y \in A$; thus $y \in A$ and $x \in A$ so, by (c), $(y, x) \in R$ and R is symmetric. Finally, if $(x, y) \in R$ and $(y, z) \in R$, then for some $A \in \mathscr{A}$ and some $B \in \mathscr{A}$, x and y belong to A and y and z belong to B. By (b), since $y \in A$ and $y \in B$, $A = B$ and so x and z belong to A. Hence $(x, z) \in R$ so that R is transitive. Hence R is an equivalence relation. Finally, it is clear that if $x \in X$, then there is just one $A \in \mathscr{A}$ such that $x \in A$ and that $R[\{x\}] = A$. This completes the proof of the theorem. ∎

Transitive relations have a notation and terminology of their own. A *partial order* in X is a transitive relation on X. It is usual to denote a partial order in X by the symbol $<$, to write $x < y$ for $(x, y) \in <$, and to say in this case that x is less than y. If $<$ is a partial order in X, we define the relation $\leqslant$ by $\leqslant \; = \; < \cup \Delta$; that is, $x \leqslant y$ if either $x < y$ or $x = y$. The reason for the qualification "partial" in partial order is that if $<$ is a partial order it may happen that $x \in X$, $y \in X$ and neither $x < y$ nor $y < x$. A *linear* (or *total*) *order* $<$ in X is a partial order with the property that for any $x \in X$, $y \in X$, either $x < y$, $x = y$, or $y < x$.

Exercises 0–6

1. Let X be a set and let $P(X)$ be the family of all subsets of X. Exhibit inclusion as a relation in $P(X)$. Is this relation transitive? symmetric? reflexive?

2. Let $f: X \to Y$ be a function. Define a relation R in X by xRy if, and only if, $f(x) = f(y)$. Show that R is an equivalence relation and determine its equivalence classes.

3. Find examples of relations which are (i) reflexive but not transitive or symmetric, (ii) symmetric but not transitive or reflexive, (iii) transitive but not reflexive or symmetric.

4. Find the domain and range of the relation
$$\{(x, y): x^2 + y^2 = 1,\ x \in \mathbf{R}, y \in \mathbf{R}\} \quad \text{in} \quad \mathbf{R}.$$
Is this relation a function?
5. Find the domain and range of the relation
$$\{(x, y): |x| + |y| = 1, x \in \mathbf{R}, y \in \mathbf{R}\} \quad \text{in} \quad \mathbf{R}.$$
Is this relation a function?
6. Find the domain and range of the relation
$$\{(x, y): x + y = 5, x \in \mathbf{R}, y \in \mathbf{R}\} \quad \text{in} \quad \mathbf{R}.$$
Is this relation a function?

0–7 ALGEBRAIC CONCEPTS

We summarize here, very briefly, some algebraic terminology. A proper discussion of those concepts, of which essential use is made later, will be found in Chapter 1. The others are only referred to occasionally, usually in the problems.

A *group* is a pair $(G, *)$ where G is a set and $*$ is a function on $G \times G$ to G (and we write $a * b$ for $*(a, b)$) having the properties

i) $a * (b * c) = (a * b) * c$ (for every a, b, c in G),

ii) there is a neutral element $e \in G$ such that $a * e = e * a = a$, for every $a \in G$,

iii) for each $a \in G$ there is an inverse element a^{-1} such that $a^{-1} * a = a * a^{-1} = e$.

If in addition

iv) $a * b = b * a$ (for every a and b in G),

the group $(G, *)$ is said to be *Abelian* or *commutative*.

A *field* is a triple $(F, +, *)$ such that F has at least two members, $(F, +)$ is an Abelian group, $(F \sim \{0\}, *)$ is an Abelian group, where 0 is the neutral element of $(F, +)$, and $a * (b + c) = (a * b) + (a * c)$ for every a, b, c in F.

An *ordered field* is a quadruple $(F, +, *, <)$ such that $(F, +, *)$ is a field, $<$ is a linear order on F, and $a < b$ implies $a + c < b + c$ for every $c \in F$ and $ac < bc$ for every $c \in F$ such that $c > 0$.

A *ring* is a triple $(R, +, \cdot)$ such that $(R, +)$ is an Abelian group and $\cdot$ is a function on $R \times R$ to R (with $a \cdot b$ written for $\cdot(a, b)$) such that
$$a \cdot (b \cdot c) = (a \cdot b) \cdot c \quad \text{for every } a, b, c \text{ in } R,$$
$$a \cdot (b + c) = (a \cdot b) + (a \cdot c) \quad \text{for every } a, b, c \text{ in } R,$$
$$(a + b) \cdot c = a \cdot c + b \cdot c \quad \text{for every } a, b, c \text{ in } R.$$

If, in addition, $a \cdot b = b \cdot a$ for every a, b in R, the ring is said to be *commutative*. If there is an element e in R, not equal to the neutral element of $(R, +)$, such that $e \cdot a = a \cdot e = a$ for every $a \in R$, then R is said to be a ring with *identity* and e is a (unique) identity for R.

A *linear* (or *vector*) *space* is a quadruple $(X, \oplus, \odot, \mathscr{F})$, where $(X, \oplus)$ is an Abelian group, $\mathscr{F}$ is a field (called the field of scalars) and $\odot$ is a function on $\mathscr{F} \times X$ to X with the properties
$$\alpha \odot (x \oplus y) = (\alpha \odot x) \oplus (\alpha \odot y),$$
$$(\alpha + \beta) \odot x = (\alpha \odot x) \oplus (\beta \odot x),$$
$$(\alpha \cdot \beta) \odot x = \alpha \odot (\beta \odot x),$$
$$1 \odot x = x,$$
for every α, β in $\mathscr{F}$ and x, y in X where $+$ and $\cdot$ are the operations in $\mathscr{F}$, and 1 is the neutral element in $(\mathscr{F}, \cdot)$.

An *algebra* is a structure consisting of a ring R, a field $\mathscr{F}$ such that R is a linear space over $\mathscr{F}$, and the scalar multiplication $\cdot$ related to the ring multiplication $\odot$ by
$$\alpha \cdot (x \odot y) = (\alpha \cdot x) \odot y = x \odot (\alpha \cdot y).$$

0–8 POINTWISE OPERATIONS ON FUNCTIONS

Suppose that $\mathscr{F}$ is a family of functions on a set X to a set Y (that is, all the functions in $\mathscr{F}$ have domain X and range in Y). If Y has some algebraic structure, for example if Y is a group, it is possible to introduce a corresponding structure into the family $\mathscr{F}$. To keep matters simple we shall suppose that Y is the field $\mathbf{R}$ of real numbers, but the method is obviously applicable to other structures.

Suppose that X is a set and that f and g are functions on X to $\mathbf{R}$. The *pointwise sum* of f and g, denoted by $f + g$, is the function with domain X which satisfies
$$(f + g)(x) = f(x) + g(x) \quad (x \in X).$$
The *pointwise product* of f and g, denoted by fg, is the function with domain X which satisfies
$$(fg)(x) = f(x)g(x) \quad (x \in X).$$
If $\lambda \in \mathbf{R}$, *the scalar product of* λ *with* f, denoted by λf, is the function with domain X which satisfies
$$(\lambda f)(x) = \lambda f(x) \quad (x \in X).$$
If, for every $x \in X$, $f(x) \neq 0$, the *pointwise reciprocal of* f denoted by $1/f$ is the function with domain X which satisfies
$$(1/f)(x) = \frac{1}{f(x)}.$$

It is an instructive exercise to show that if X is a set, then the family of all functions on X to $\mathbf{R}$ is, when endowed with the pointwise operations of addition, multiplication and multiplication by real scalars, an algebra over the field $\mathbf{R}$.

0–9 INTERVALS

An *interval* in $\mathbf{R}$ is a subset of $\mathbf{R}$ having one of the following forms:

$$
\begin{aligned}
(a, b) &= \{x : a < x < b\} \\
(a, \infty) &= \{x : x > a\} \\
(a, b] &= \{x : a < x \leqslant b\} \\
(-\infty, b) &= \{x : x < b\} \\
(-\infty, \infty) &= \mathbf{R} \\
(-\infty, b] &= \{x : x \leqslant b\} \\
[a, b) &= \{x : a \leqslant x < b\} \\
[a, \infty) &= \{x : x \geqslant a\} \\
[a, b] &= \{x : a \leqslant x \leqslant b\},
\end{aligned}
$$

where a and b are in $\mathbf{R}$ and $a < b$. We also regard $\{a\} = [a, a]$ and $\varnothing = (a, a)$ as intervals.

Intervals of the form (a, b), $(-\infty, \infty)$, (a, ∞), $(-\infty, a)$, (a, a) are called *open intervals* (because they are open subsets of $\mathbf{R}$). Intervals of the form $[a, b]$, $(-\infty, \infty)$, $[a, \infty)$, $(-\infty, a]$, $[a, a]$ are called *closed intervals*, and $[a, b]$ is sometimes called *compact*. The remaining types of intervals are called *half-open*.

We adopt, for the purposes of Chapter 5, the convention that

$$[b, a] = [a, b] \quad \text{if } a < b.$$