1

BASIC SET THEORY AND ANALYSIS

1. Sets and functions

The great German mathematician G. Cantor (1845–1918) is usually regarded as the creator of the theory of sets. As our starting point for this book we shall take Cantor's definition of a set: 'A set is any collection of definite, distinguishable objects of our thought, to be conceived as a whole.' The objects mentioned in the definition are called the elements or members of the set. Usually we denote sets by capital letters and elements by lower case letters. If X is a set then we write $x \in X$ to mean that x is an element of X. When an object x is not an element of a set X, we write $x \notin X$.

In what follows we shall take for granted the following sets, which occur throughout mathematics:

$N = \{1, 2, 3, \ldots\}$, the set of all positive integers,
$Z = \{0, 1, -1, \ldots\}$, the set of all integers,
Q, the set of all rational numbers,
R, the set of all real numbers,
C, the set of all complex numbers.

The notation arises as follows: N for natural numbers, Z for Zahlen (German for integers), Q for quotient. The notation R and C seems to require no explanation.

Usually, in a given discussion, we take a fixed set and everything is carried out with reference to it alone. In such a case the fixed set is called the universe of discourse. For example, in number theory the universe of discourse is Z. Within a universe of discourse X a common way of generating a set is to take an object in X of a certain type and then to consider the set of all such objects. For example, having defined an object in Z called a prime number we may then consider the set of all prime numbers.

In a work of the present nature we are primarily concerned with the manipulation of sets, rather than with their deeper properties. To this end we now introduce notation and definitions, and observe some simple results.

First, there is no way, in general, of explicitly writing down all the elements of a set. For example, it is in the nature of the positive integers N

that they cannot all be explicitly exhibited. We have to be content to write $N=\{1,2,3,\ldots\}$; the three dots leaving much to the imagination. Generally, we use the curly bracket notation for sets either writing down the first few elements and then some dots which we agree is to tell us that the law of formation of the elements is well known or obvious, or we put in the law of formation. For example, $\{x \mid x \in N \text{ and } x > 8\}$ is read as 'the set of all x such that x is a positive integer and x is greater than 8'. The vertical bar following x is read as 'such that'. Thus we could write this last set as $\{9, 10, 11, \ldots\}$. Again, $\{x \mid x \in R \text{ and } x > 0\}$ denotes the set of all strictly positive real numbers. In this case it is not possible to write down the elements explicitly, or even in such a way as to indicate the law of formation, such is the nature of the real numbers. In fact it will be seen later that the set $\{x \mid x \in R \text{ and } x > 0\}$ is uncountable, so that the elements cannot even be exhibited as an infinite sequence $x_1, x_2, x_3, \ldots$. We remark that the order of the elements in a set is generally irrelevant. For example, N is the same set as $\{2, 1, 4, 3, 6, 5, \ldots\}$.

If A, B are sets then the notation $A \subset B$ means that every element of A is also an element of B. If $A \subset B$ then we say that A is a subset of B, B is a superset of A, A is included in B and also B includes A. The notation $B \supset A$ is regarded as equivalent to $A \subset B$. We define $A = B$ if and only if $A \subset B$ and $B \subset A$. Also, we say A is a proper subset of B if and only if $A \subset B$ but $A \neq B$. For example, the set of odd integers is a proper subset of Z. We remark that some writers use the notation $A \subseteq B$, which allows equality, and reserve $A \subset B$ for proper subsets. On occasion we shall also say that '$A \subset B$, strictly', meaning that A is a proper subset of B.

Two simple properties of the set inclusion $\subset$ are:

(i) $A \subset A$,

(ii) $A \subset B$ and $B \subset C$ imply $A \subset C$.

If A is a given set let us consider that subset of A defined as $\{x \in A \mid x \neq x\}$. This set has no elements and is known as the *empty set*. It is denoted by $\varnothing$ and has the property that $\varnothing \subset A$ for every set A. Each set $A \neq \varnothing$ has at least two distinct subsets, A and $\varnothing$. If A has only these two subsets then A must be a one element set, $A = \{a\}$, say, where a is the sole element of A. Note that $\varnothing$ has no elements but that the one element set $\{\varnothing\}$ is not empty.

Unions and intersections of sets

Given sets A, B we may form two new sets from them:

$$A \cup B = \{x \mid x \text{ belongs to at least one of } A \text{ and } B\},$$

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

We call $A \cup B$ the *union* and $A \cap B$ the *intersection*, of A and B. For example, $\{1,2,3\} \cup \{1,4,3\} = \{1,2,3,4\}$; $\{2,3\} \cap \{1,3,2\} = \{2,3\}$. It is trivial that $A \cap B \subset A \subset A \cup B$ for any sets A and B. If $A \cap B = \varnothing$, then we say that A and B are disjoint.

We shall often want to form the union or intersection of a whole class (or collection) of sets. Let $\mathscr{S}$ be a class of sets A. Then we define

$$\bigcup \{A \mid A \in \mathscr{S}\} = \{x \mid x \in A \text{ for at least one } A \in \mathscr{S}\},$$

$$\bigcap \{A \mid A \in \mathscr{S}\} = \{x \mid x \in A \text{ for all } A \in \mathscr{S}\}.$$

Sometimes we write $\bigcup A_\alpha$, $\bigcap A_\alpha$, where we think of α as running through some indexing set. If α runs through N we usually write

$$\bigcup \{A_n \mid n \in N\} = \bigcup_{n=1}^{\infty} A_n,$$

and similarly for $\bigcap_{n=1}^{\infty} A_n$. The '$\infty$' in this notation is conventional, but superfluous, not to say confusing. It is emphasized that A_∞ is not in the collection $\{A_n \mid n \in N\}$. Observe also that no limiting process is involved in the above. Thus, for example, to say that $x \in \bigcup_{n=1}^{\infty} A_n$, is to say that there is a positive integer p such that $x \in A_p$.

Example 1. Let A_n be the interval $[0, 1 + 1/n)$ on the real line, i.e. $A_n = \{x \in R \mid 0 \leqslant x < 1 + 1/n\}$, $n = 1, 2, \ldots$. Then

$$\bigcap_{n=1}^{\infty} A_n = [0, 1] = \{x \in R \mid 0 \leqslant x \leqslant 1\}.$$

To show this we first prove $[0, 1] \subset \bigcap A_n$ and then prove $\bigcap A_n \subset [0, 1]$. Now $x \in [0, 1]$ implies $0 \leqslant x \leqslant 1 < 1 + 1/n$, for all $n \in N$, i.e. $x \in A_n$ for all $n \in N$, i.e. $x \in \bigcap A_n$. Conversely, $x \in \bigcap A_n$ implies $0 \leqslant x < 1 + 1/n$, for all $n \in N$, whence $0 \leqslant x \leqslant 1$ (either letting $n \to \infty$, or supposing $x > 1$ and obtaining a contradiction to $x < 1 + 1/n$ for all $n \in N$).

Cover for a set

Let $\mathscr{S}$ be a class of sets A. Then the class $\mathscr{S}$ is called a cover for a set X if and only if

$$X \subset \bigcup \{A \mid A \in \mathscr{S}\}.$$

Any subclass of $\mathscr{S}$ which also covers X is called a subcover of $\mathscr{S}$.

The notion of 'open' cover will be employed in chapter 2, in connection with compact sets. The 'open' here refers to the fact that the sets of the cover are open sets, in the sense of topology. For the moment we shall be content with a very simple example on covers.

Example 2. (i) Let I_n be the open interval

$$(n,n+1) = \{x \in R \mid n < x < n+1\}$$

on the real line. Then the class $\{I_n \mid n \in Z\}$ is not a cover for R, for no integer belongs to $\bigcup \{I_n \mid n \in Z\}$.

(ii) If $J_n = \{x \in R \mid n \leqslant x < n+1\} = [n, n+1)$, then the class $\{J_n \mid n \in Z\}$ is a cover for R.

(iii) Let $S[a,r] = \{z \in C \mid |z-a| \leqslant r\}$, where $a \in C$ and $r > 0$. Thus $S[a,r]$ is the closed disc of centre a and radius r in the complex plane. It is clear that the class $\{S[m+in, 1] \mid m, n \in Z\}$ is a cover for C.

Complementation

If X is our universe of discourse and $A, B \subset X$ then we define

$$A \sim B = \{x \in X \mid x \in A, x \notin B\}.$$

We call $A \sim B$ the *complement of B with respect to A*. By $\sim A$ we mean $X \sim A$, and we call $\sim A$ the complement of A. It is clear that $A \sim B = A \cap (\sim B)$, $\sim(\sim A) = A$, and that $A \subset B$ is equivalent to $\sim B \subset \sim A$.

The two following results concerning complementation are known as De Morgan's laws:

$$\sim \bigcup A_\alpha = \bigcap (\sim A_\alpha); \quad \sim \bigcap A_\alpha = \bigcup (\sim A_\alpha).$$

To prove the first of these, for example, we merely note that $x \in \sim A_\alpha$ for all α is equivalent to $x \notin A_\alpha$ for any α.

Some other properties of union and intersection which are easy to show are

(i) $\bigcap A_\alpha \subset A_\alpha \subset \bigcup A_\alpha$, for any α,
(ii) $A \cup (\bigcap A_\alpha) = \bigcap (A \cup A_\alpha)$,
(iii) $A \cap (\bigcup A_\alpha) = \bigcup (A \cap A_\alpha)$.

Ordered pair

Let x, y be any objects. Then the ordered pair (x, y) is defined as the set $\{\{x\}, \{x, y\}\}$. It is easy to check the fundamental property of ordered *pairs:* $(x, y) = (u, v)$ *if and only if* $x = u$ *and* $y = v$. *More generally we may define in a similar way an ordered n-tuple* $(x_1, \ldots, x_n)$ *with the property* $(x_1, \ldots, x_n) = (y_1, \ldots, y_n)$ *if and only if* $x_1 = y_1, \ldots, x_n = y_n$.

Relation

A relation ρ *is defined to be a set of ordered pairs. For example,* $\rho = \{(1, 2), (a, b)\}$ *is a relation.*

Equivalent notation for $(x, y) \in \rho$ is $x\rho y$. Thus in our example we might write $1\rho 2$ instead of $(1, 2) \in \rho$.

An important type of relation is the

Cartesian product

Let X, Y *be given sets. Then*

$$X \times Y = \{(x, y) \mid x \in X \text{ and } y \in Y\}$$

is called the Cartesian product of X *and* Y.

Another important relation is the

Equivalence relation

Let X *be a given set. A relation* ρ *is called an equivalence relation on* X *if and only if it is* (i) *reflexive, i.e.* $x\rho x$ *for all* $x \in X$, (ii) *symmetric, i.e.* $x\rho y$ *implies* $y\rho x$, (iii) *transitive, i.e.* $x\rho y$ *and* $y\rho z$ *imply* $x\rho z$. *It is usual to denote an equivalence relation by* $\sim$ *rather than* ρ. *There is little danger of confusion with complementation.*

Example 3. (i) Equality is obviously an equivalence relation on any set.

(ii) Let $X = \{(x, y) \mid x, y \in N\}$. Define $(x, y) \sim (u, v)$ to mean $xv = yu$. Then $\sim$ is an equivalence relation on X. For example, let us check transitivity: $(x, y) \sim (u, v)$ and $(u, v) \sim (z, w)$ imply $xv = yu$, $uw = vz$, whence $xvuw = yuvz$, and so $xw = yz$, i.e. $(x, y) \sim (z, w)$.

(iii) Define $x \sim y$ to mean $x - y$ is divisible by 2. It is easy to check that $\sim$ is an equivalence relation on Z.

Let us return to general relations. The *domain* of a relation is the set of all first co-ordinates of its members. The *range* is the set of all second co-ordinates. If $\sim$ is an equivalence relation on X, then we define $E_x =$

$\{y \in X \mid y \sim x\}$ and call E_x the equivalence class containing the element x. For example, in example 3 (iii) we have

$$E_0 = \{0, \pm 2, \pm 4, \ldots\}.$$

In general, it is easy to check that $E_x = E_y$ if and only if $x \sim y$, and that $E_x \cap E_y = \varnothing$ if $E_x \neq E_y$. It is thus evident that $\{E_x \mid x \in X\}$ forms a partition of X, i.e. X is the union of the disjoint classes E_x. For example, in example 3 (iii) we have

$$Z = \{0, \pm 2, \pm 4, \ldots\} \cup \{\pm 1, \pm 3, \pm 5, \ldots\}.$$

Probably the most significant type of relation that occurs in mathematics is that which is called a function. The following definition of a function may seem rather strange to those who are used to books of analysis which extensively employ functions but never actually define them.

Function

A function f is defined to be a relation, such that if $(x, y) \in f$ and $(x, z) \in f$ then $y = z$. Four other terms for function are map, mapping, operator, and transformation.

Our concept of a function as a certain set of ordered pairs is what some would call the graph of a function, since they define a function as a 'rule' or some such. On occasion we shall use the term 'graph of a function', when this seems more expressive. However, to us, a function and its graph are exactly the same thing.

Example 4. (i) $\{(1,2), (2,2)\}$ and $\{(z, z+1) \mid z \in C\}$ are functions.

(ii) $\{(1,2), (1,4)\}$ and $\{(x^2, x) \mid x \in R\}$ are not functions. For example $(1, 1)$ and $(1, -1)$ are in the second set.

(iii) $\{(x^2, x) \mid x \in R^+\}$ is a function. Here $R^+ = \{x \in R \mid x > 0\}$. In this case, if the first co-ordinates are equal, $x^2 = y^2$, then $(x - y)(x + y) = 0$, so $x = y$, i.e. the second co-ordinates are equal.

If f is a function and $(x, y) \in f$ then we write $y = f(x)$, which is the conventional notation for y as a function of x. We say that y is the value of f at x, or that y is the image of x under f.

The notation

$$f : X \to Y$$

is now widely used in mathematics. It is interpreted as 'f is a function from the set X into the set Y'. The meaning of $f : X \to Y$ is that X is the domain of f and that the range of f is a subset of Y, not necessarily the whole of Y.

If $f: X \to Y$ and $A \subset X$, then the function $g: A \to Y$, defined by $g(a) = f(a)$, for $a \in A$, is called the *restriction* of f to A.

Example 5. (i) Define f by $f(x) = e^x$, for $x \in R$, i.e. the domain of f is R and $f = \{(x, e^x) \mid x \in R\}$. The range of f is in fact R^+, as is well known. We may write, with increasing accuracy, $f: R \to R$, and $f: R \to R^+$.

(ii) Define f by $f(z) = |z|$, for $z \in C$. Then, with increasing precision, we have $f: C \to C$, $f: C \to R$, and $f: C \to \{x \in R \mid x \geqslant 0\}$.

Bijective maps

Let $f: X \to Y$. Then f is called injective if $f(x_1) = f(x_2)$ implies $x_1 = x_2$, for every $x_1, x_2 \in X$. If the range of f is the whole of Y, then f is called surjective. A mapping which is both injective and surjective is called bijective.

The terms 'one–one', 'onto' and 'one–one correspondence' are sometimes used instead of 'injective', 'surjective' and 'bijective mapping' respectively.

Example 6. $f: R \to \{x \in R \mid x \geqslant 0\}$, defined by $f(x) = x^2$, is surjective but not injective. The same prescription for f, but with $f: R^+ \to R^+$, is bijective.

Inverse function

Let $f: X \to Y$ be bijective. Since f is surjective, if $y \in Y$ then there exists $x \in X$ such that $y = f(x)$. This x is unique, since f is injective. Hence there is an inverse function $g: Y \to X$ such that $g(f(x)) = x$, for all $x \in X$, and $f(g(y)) = y$ for all $y \in Y$. It is usual to write $g = f^{-1}$.

Example 7. $f: R \to R^+$, defined by $f(x) = e^x$, is bijective. The inverse $f^{-1}: R^+ \to R$ is denoted by log.

Equivalent sets

Sets X, Y are called equivalent (written $X \sim Y$) if and only if there exists a bijective map $f: X \to Y$. We note that $\sim$ is an equivalence relation. For *example, $X \sim X$ since the map $f: X \to X$, given by $f(x) = x$, is a suitable bijection.*

Example 8. (i) $\{1,2,3\} \sim \{1,3,5\}$; $\{2,4,6,\ldots\} \sim N$ and $N \sim Z$.

(ii) $\{1,2\}$ is not equivalent to $\{1,2,3\}$.

(iii) The interval $(-1,1)$ in R is equivalent to R. A suitable bijection is $f(x) = \tan(\pi x/2)$.

Countable set

A set is called countable if and only if it is equivalent to N or to a subset of N. Otherwise it is called uncountable. In case the set is equivalent to $\{1,2,\ldots,n\}$ it is called finite, with n elements.

Examples of countable sets are N, Z and Q. The set R of real numbers is uncountable (see exercise 12). It is clear that any subset of a countable set is countable. Also, if A_n is countable for $n = 1,2,\ldots$, then $\bigcup\{A_n \mid n \in N\}$ is countable, i.e. a countable union of countable sets is countable. To 'count' the elements of $\bigcup A_n$ we proceed as follows. Let $A_n = \{a_{n1}, a_{n2}, \ldots\}$, $n = 1,2,\ldots$. Write

$$\{a_{11}, a_{12}, a_{21}, a_{13}, a_{22}, a_{31}, a_{14}, a_{23}, \ldots\}.$$

Then this set is countable and is the union of the A_n, provided we do not allow repetitions. For example, if

$$A_1 = \{2,3,4,5,6,\ldots\} \quad \text{and} \quad A_2 = \{2,4,6,8,10,\ldots\}$$

then our scheme gives $A_1 \cup A_2 = \{2, 3, 4, 5, 6, 8, 7, 10, \ldots\}$.

Example 9. Let $A_n = \{m/n \mid m \in Z\}$, $n = 1,2,\ldots$. Each A_n is countable, since Z is countable, and $\bigcup A_n = Q$. Hence Q is countable. Thus Q is equivalent to N, even though N is a very 'thin' subset of Q.

Image and inverse image

Let $f: X \to Y$, and let $A \subset X$, $B \subset Y$. Then the image of the set A under the map f is defined to be $f(A) = \{f(x) \mid x \in A\}$. The inverse image of B is defined to be $f^{-1}(B) = \{x \mid x \in X$ and $f(x) \in B\}$. Note that $f^{-1}(B)$ is a subset of X; it is not necessary that f should be bijective in order that we may write $f^{-1}(B)$.

The basic properties of image and inverse image are now given.

Theorem 1. *Let $f: X \to Y$, and suppose $\{A_\alpha\}$ is a class of subsets of X and $\{B_\alpha\}$ is a class of subsets of Y. Then*

(i) $A_\alpha \subset A_\beta$ *implies* $f(A_\alpha) \subset f(A_\beta)$;
(ii) $B_\alpha \subset B_\beta$ *implies* $f^{-1}(B_\alpha) \subset f^{-1}(B_\beta)$;
(iii) $f(\bigcup A_\alpha) = \bigcup f(A_\alpha)$;
(iv) $f(\bigcap A_\alpha) \subset \bigcap f(A_\alpha)$;
(v) $f^{-1}(\bigcup B_\alpha) = \bigcup f^{-1}(B_\alpha)$;
(vi) $f^{-1}(\bigcap B_\alpha) = \bigcap f^{-1}(B_\alpha)$;
(vii) $A_\alpha \subset f^{-1}(f(A_\alpha))$;
(viii) $f(f^{-1}(B_\alpha)) \subset B_\alpha$.

Proof. As a sample we prove (iii). Let $y \in f(\bigcup A_\alpha)$. Then $y = f(x)$, for some $x \in \bigcup A_\alpha$, i.e. $y = f(x)$ for $x \in A_\alpha$, some α. Thus $y \in f(A_\alpha) \subset \bigcup f(A_\alpha)$. Conversely, if $y \in \bigcup f(A_\alpha)$ then $y \in f(A_{\alpha'})$ for some α', so $y = f(x)$, where $x \in A_{\alpha'} \subset \bigcup A_\alpha$. Hence $y \in f(\bigcup A_\alpha)$. This proves (iii). The other results are left as exercises. One should observe the difference between (iv) and (vi).

Composition of functions

Let $f: X \to Y$ and $g: Y \to Z$, where f, g are any functions and X, Y, Z are any sets. Then we define the composite (or product) function $gf: X \to Z$, by $(gf)(x) = g(f(x))$, for each $x \in X$.

A composite function is obviously what is called, in elementary calculus, a function of a function.

Partial-order relation

Let X be a set. Then a partial-order relation on X, denoted in general by $\prec$, is a relation which is: (1) *reflexive (i.e. $x \prec x$ for each $x \in X$);* (2) *transitive (i.e. $x \prec y$ and $y \prec z$ imply $x \prec z$);* (3) *antisymmetric (i.e. $x \prec y$ and $y \prec x$ imply $x = y$).*

A total-order relation, $\prec$, is a partial-order relation with the extra property that for any x, y, either $x \prec y$ or $y \prec x$.

A partially ordered set is just a pair $(X, \prec)$, consisting of a set X and a partial order on it. Similarly for a totally ordered set.

Example 10. (i) Let $\mathscr{A}$ be any collection of sets. Then the set inclusion $\subset$ is a partial order on $\mathscr{A}$. It is not in general a total order on $\mathscr{A}$.

(ii) The 'natural' total order on R, the real numbers, is of course $\leqslant$.

(iii) Define $\prec$ on N by saying that $x \prec y$ if and only if x is an integer multiple of y. Then $\prec$ is obviously a partial ordering of N. It is not total. For example, 3 is not an integer multiple of 2, nor is 2 an integer multiple of 3.

Let $(X, \prec)$ be a partially ordered set. An element $a \in X$ is called maximal if and only if $a \prec x$ for $x \in X$ implies $x = a$. An element $b \in X$ is called minimal if and only if $x \prec b$ for $x \in X$ implies $x = b$.

Suppose $A \subset X$. An element $M \in X$ is called an upper bound for A if and only if $x \prec M$ for all $x \in A$. Similarly, $m \in X$ is a lower bound for A if $m \prec x$ for all $x \in A$.

Example 11. Let X be a non-empty set and let $P(X)$ denote the class of all subsets of X (one calls $P(X)$ the power set of X). Then $P(X)$ is partially ordered by set inclusion. If $\mathscr{A}$ is any class of subsets of X, then $\bigcup\{A \mid A \in \mathscr{A}\}$ is an upper bound for $\mathscr{A}$ and $\bigcap\{A \mid A \in \mathscr{A}\}$ is a lower bound for $\mathscr{A}$.

In order to prove various important results in several branches of mathematics it has been found necessary to invoke a fundamental axiom of set theory known as Zorn's lemma. We state this now and make some comments on it after.

Zorn's lemma

Let X be a partially ordered set with the property that every totally ordered subset has an upper bound. Then X contains a maximal element.

There are just two places in this book where we need to use this axiom. In chapter 3, section 2, we employ it to prove that every linear space has a Hamel base. Zorn's lemma is again used in the proof of the important Hahn–Banach extension theorem (chapter 4, section 5).

The intuitive meaning of Zorn's lemma does not seem to be immediately apparent, and the reader is asked to accept it as an *axiom* of set theory, which is vital if certain results of an existential type are desired.

We remark that it has been proved that Zorn's lemma is equivalent to two other axioms of set theory: (1) the axiom of choice, (2) the well-ordering principle. For our purposes, Zorn's lemma is, for technical reasons, the best axiom to adopt. Those interested in the equivalence of the three axioms should consult P. C. Suppes' book, which is listed in the bibliography.