1 Some mathematical language

Connectives

Mathematics is concerned with statements or propositions of a particular kind, namely those which are composed of mathematical signs and objects. Such statements are often called **relations**.

From these statements we can build up others by means of certain logical signs known as **connectives**. The basic connectives which we use to construct compound statements are

not	denoted by	$\sim$ ($\neg$),
or	denoted by	$\vee$,
and	denoted by	$\wedge$ (&),
if...then	denoted by	$\Rightarrow$ ($\rightarrow$, $\supset$),
if and only if	denoted by	$\Leftrightarrow$ ($\leftrightarrow$, $\equiv$).

These connectives are used in everyday speech, but often in an ambiguous and confusing manner. It is necessary then to fix their meaning more precisely and this is done by means of **truth tables** which provide information regarding the truth or falsehood of a compound statement in terms of the assumed truth or falsehood of its constituent parts. Thus the operation of **negation** (not) is defined by the table

X	$(\sim X)$
T	F
F	T

which indicates that if the statement X is true then its negation, $(\sim X)$, is false, and if the statement is false then its negation is true.

Conjunction (and), **disjunction** (or), **logical implication** (if... then), and **logical equivalence** (if and only if) are similarly defined by

X	Y	$(X \wedge Y)$	$(X \vee Y)$	$(X \Rightarrow Y)$	$(X \Leftrightarrow Y)$
T	T	T	T	T	T
T	F	F	T	F	F
F	T	F	T	T	F
F	F	F	F	T	T

Note. (i) In this interpretation of 'or' we say that $(X \vee Y)$ is true when X is true *and* Y is true. This is sometimes known as the '**inclusive or**', and, it will be noted, is not always the everyday usage of 'or'.

(ii) In the statement $(X \Rightarrow Y)$ (sometimes known as a **conditional** statement), X is known as the **antecedent** or **hypothesis** and Y as the **consequent** or **conclusion.**

Just as in everyday speech it is possible to express the same argument in different ways, it can happen that two mathematical statements 'say the same thing'. We say that two compound statements are **equivalent** when they have identical truth tables. Thus, for example, $(X \Rightarrow Y)$ and $((\sim X) \vee Y)$ are equivalent since we have

X	Y	$(\sim X)$	$X \Rightarrow Y$	$((\sim X) \vee Y)$
T	T	F	T	T
T	F	F	F	F
F	T	T	T	T
F	F	T	T	T

Note. (i) It follows that $\Rightarrow$ can be expressed in terms of $\sim$ and $\vee$, as indeed can all the other connectives. (The connectives can also be expressed in terms of $\sim$ and $\wedge$.)
More basic still are the **Sheffer stroke** $|$ and the connective $\downarrow$ defined by

X	Y	$(X \mid Y)$	$(X \downarrow Y)$
T	T	F	F
T	F	T	F
F	T	T	F
F	F	T	T

since all the other connectives can be defined in terms of $|$ (or $\downarrow$) alone.

(ii) The fact that two statements are equivalent is often used to simplify the proofs of theorems. Thus, to prove that A is true, we show that B, a statement equivalent to A, is true.

Axioms

Certain statements (relations) play a particularly important role in any mathematical system. These are relations which are stated explicitly, once and for all, and which are then known as the **axioms** of the system. Relative to the system, all axioms are said to be **true.**

In addition to this type of axiom there is a second type, the **logical axioms**, which tell us how the connectives are to be manipulated – these are the rules of reasoning.

We now say that a statement (relation) in a particular system is **true** (i.e. true relative to the system), if it can be obtained by repeated application of the rules:

(*a*) every relation obtained by applying an axiom is true,

(*b*) if R and S are relations such that $(R \Rightarrow S)$ is true and if, in addition, the relation R is true, then the relation S is true.

A relation R is **false** if $(\sim R)$ is true.

Note. (i) (*b*) is known in logic as the rule of **modus ponens.**

(ii) Given a relation $(R \Rightarrow S)$ we define its **converse** to be $(S \Rightarrow R)$, its **inverse** to be $((\sim R) \Rightarrow (\sim S))$, and its **contrapositive** to be $((\sim S) \Rightarrow (\sim R))$. A frequently used rule of reasoning is that

$$((R \Rightarrow S) \Leftrightarrow ((\sim S) \Rightarrow (\sim R))) \quad \text{is true,}$$

i.e. that a conditional relation and its contrapositive are equivalent (and, hence, that the converse relation is equivalent to the inverse relation).

(iii) It is not the case that a relation must be true or false. There are relations which are **undecidable** in the sense that they can be neither proved nor disproved. (Here a mathematical system differs from the propositional calculus of logic.)

Variables and quantifiers

The relations which we consider in mathematics frequently contain letters representing indeterminate mathematical objects which we refer to as **variables.** If R contains the variable x, we can replace x wherever it occurs in R by a mathematical object A, e.g. a number or a set. When we do this the resulting assembly of letters and signs is again a relation and is known as the relation obtained by **substituting** A for x in R. A is said to **satisfy** the relation R if the relation so obtained is true (thus 3 satisfies the relation $x^2+2 = 11$).

To indicate that x occurs in the relation R we write $R\{x\}$ and we denote the relation obtained by substituting A for x by $R\{A\}$. A relation of the form $R\{x\}$ is often known as a **predicate.**

Two questions which are frequently asked of mathematical relations containing a variable are: 'Is the relation true *for every A* which we care to substitute for x?' (in practice it would be assumed that A would belong to a particular class of objects, for otherwise the relation obtained by substituting for x would frequently be meaningless, e.g. $(\text{triangle})^2+2 = 11$) and 'Is there *some A* (belonging to a particular class) for which the relation is true?' These two notions lead to the introduction of the two **quantifiers** $\forall$ and $\exists$.

In mathematics we use as a shorthand for the phrases

for all, for each, for every, for any, given any

the sign $\forall$ which is known as the **universal quantifier**.

Thus we write

$$(\forall x) \quad (x \text{ is a real number} \Rightarrow (x+2)^2 = x^2+4x+4)$$

to indicate that the relation

$$(x+2)^2 = x^2+4x+4$$

is satisfied by every real number x.

Similarly, we express the fact that *there exists* an object which satisfies the relation $x^2+2 = 11$ by writing

$$(\exists x) \quad (x^2+2 = 11),$$

where $\exists$, the **existential quantifier**, is interpreted as shorthand for '*there exists...such that*' or '*for some*'.

Note. (i) The usage of $\forall$ and $\exists$ as explained above corresponds to the everyday usage of the logical connectives. Their usage in logic, from whence the symbols are borrowed, is, of course, more circumscribed.

(ii) It is often convenient to work with restricted notions of quantification and to write

$$(\forall x)\ ((x+2)^2 = x^2+4x+4)$$

rather than $(\forall x)$ (x is a real number $\Rightarrow (x+2)^2 = x^2+4x+4$) when it is clearly understood from the context that x is a real number.

(iii) Some authors choose to define $\forall$ in terms of $\exists$ and $\sim$ since $(\forall x)\ (R\{x\})$ and $\sim ((\exists x)\ (\sim R\{x\}))$ are equivalent statements.

(iv) A useful formula which enables one to write down the negation of a relation containing two or more quantifiers relating to different variables is $(\sim ((\forall x)\ (\exists y)\ R\{x, y\}))$ is equivalent to $((\exists x)\ (\forall y)\ (\sim R\{x, y\}))$.

(v) The symbol $\exists!x$ is often used to denote 'there exists a unique $x, \ldots$'.

When a variable occurs in a relation which is quantified, then we say that x is a **bound variable** in that relation, e.g. x is bound in the relation $(\exists x)\ (x^2+2 = 11)$. If no quantifier appears in the relation then x is said to be a **free variable** or an **unknown** and accordingly we speak of $R\{x\}$ as defined above as a **predicate with free variable** x.

A relation containing only bound variables or no variables at all is said to be **closed** whilst a relation containing a free variable is said to be **open**. Closed statements have the property of being true, false or undecidable.

Examples. (i) $(\exists x)\ (x^2+2 = 11)$ is true;

(ii) $(\exists x)\ (x^2 = -1)$ is false (assuming that we are again using quantification restricted to the real numbers; if we do not make this assumption then the relation is true);

(iii) (x is a prime) is an *open* relation which is neither true nor false in itself but only becomes so when a particular object is substituted for x. Those objects of a given set which satisfy an open statement form the **solution set** of the statement relative to that set. Thus the *solution set* of $x^2 = 2$ relative to the real numbers is the set $\{+\sqrt{2}, -\sqrt{2}\}$. Relative to the rational numbers (p. 24), the *solution set* of $x^2 = 2$ is empty (p. 9).

Equality

The mathematical sign of **equality** is introduced in order to form relations of the type $a = b$ which we take intuitively to mean that the objects a and b are identical.

The rules which we require for manipulating equalities are:

(*a*) $(\forall x)\,(x = x)$,
(*b*) $(\forall x, y)\,((x = y) \Leftrightarrow (y = x))$,
(*c*) $(\forall x, y, z)\,(((x = y) \wedge (y = z)) \Rightarrow (x = z))$,
(*d*) if $u = v$ and $R\{u\}$ and $R\{v\}$ are obtained by replacing the letter x in $R\{x\}$ by u and v respectively, then

$$R\{u\} \Leftrightarrow R\{v\}.$$

Axiom systems

An abstract deductive science is constructed by selecting an **axiom system**, i.e. certain undefined terms (e.g. 'set' in the Zermelo–Fraenkel Axiom system (p. 213)) and a number of axioms (relations of the specific 'once-and-for-all' type (p. 2)) containing them. From these axioms and using the rules of reasoning derived from the logical axioms, one obtains further relations which are *true* in the sense described on p. 3 and which are known as **theorems** or **lemmas** (the term used depends upon the extent to which mathematicians will wish to refer to them). The question of the 'truth' of the axioms is irrelevant within the framework of the axiom system. If, however, we can assign meanings to the undefined terms in such a way that the axioms can be judged 'true', then we say that we have a **model** or an **interpretation** of the abstract axiom system and the system is said to be **satisfiable.**

The undefined terms appearing in an axiom system will, in general, be of two kinds, namely, the *universal terms* such as 'set' and '$\in$' (p. 8), and the *technical terms* peculiar to that system (sometimes known as *primitive terms*) such as 'point' and 'line'.

Since our aim is to model rational thought, the most fundamental question to be asked about an axiom system is 'Does the system imply any contradictory theorems?' (i.e. is there a relation R such that both R and $\sim R$ can be deduced from the axioms?).

It can be shown that if a relation R exists such that both R and $\sim R$ can be deduced from the axioms, then any relation Q can be so deduced.

An axiom system Σ which does not imply contradictory statements is said to be **consistent** (see p. 23 for an example).

The usefulness of this definition is limited by our ability to recognise contradictions. It is not possible, for example, to show that the axioms of mathematics are consistent, yet all mathematicians proceed in the belief that they are. This belief is made explicit, for example, in the use of proofs based on **reductio ad absurdum**, i.e. proofs in which we infer the falsity of a relation P by adjoining it to our axiom system and then showing that in the new system there is some proposition Q such that both Q and $\sim Q$ are deducible.

Since there is, in general, no test for consistency we must be satisfied with what is termed **relative consistency**, obtained by showing that our axiom system Σ can be embedded in a second system Σ' (in whose consistency one is willing to believe). We take, therefore, as a 'working definition' of consistency: an axiom system is consistent if it is satisfiable.

If Σ is an axiom system and A is one of its axioms, then A is said to be **independent** in Σ, or an **independent axiom** of Σ, if it cannot be derived from other axioms in the system. This will be the case if both Σ and the axiom system obtained from Σ by replacing A by its negation (denoted by $(\Sigma - A) + (\sim A)$) are satisfiable (see p. 23 for an example).

An axiom system Σ is said to be **independent** if all the axioms of Σ are independent in Σ.

An axiom system Σ is said to be logically **complete** if it is impossible to add a new independent axiom to Σ which is consistent with Σ and which contains no new undefined terms, i.e. if there is no Σ-*statement* A such that A is an independent axiom of the system $\Sigma + A$.

An axiom system Σ is said to be **categorical** if every two models of Σ are isomorphic with respect to Σ, i.e. if there exists a one-to-one correspondence between the elements and relations of one model and those of the other such that whenever a given relation holds between two elements of one model, the corresponding relation holds between the corresponding two elements of the other.

Note. Categoricity implies logical completeness and so provides a means whereby the latter can be tested.

Example. The axioms for a group (see (*a*), (*b*), (*c*), p. 25) are not *categorical*, since it is not true that all groups are isomorphic, indeed it is possible to add further *independent* axioms to the system, for example, (*d*) the operation $*$ is commutative. We note though that if the axioms on p. 25 were changed so as to read '...consisting of a set G containing 7 elements and a binary operation $*$ ...', then the axiom system would be *categorical* since all models of the system, i.e., 'groups of order 7', are isomorphic.

2 Sets and functions

Sets

A set is a totality of certain definite, distinguishable objects of our intuition or thought – called the *elements* of the set.

This classic definition of a **set** was given by Georg Cantor in 1874. Such attempts to give elementary definitions of a set are, however, doomed to failure, their being in the main based on the use of undefined synonyms, such as 'collection', and leading to logical inconsistencies (see *Russell paradox* p. 210). For this reason, mathematicians now regard the notion of a set as an undefined, primitive concept (see *Zermelo–Fraenkel Axioms* p. 213).

Along with the idea of a 'set' we introduce the further basic mathematical sign of **membership** ($\in$).

$a \in A$ is then read as 'a is an **element** (**member**) of the set A' or 'a **belongs to** A'.

We write $a \notin A$ to signify that a is *not* an element of A.

Two sets, A and B, are equal (written $A = B$) if and only if they have the same elements, i.e.

$$\forall x,\ x \in A \Leftrightarrow x \in B.$$

B is said to be a **subset** of A (or is **included** in A) when every element of B is an element of A, i.e.

$$\forall x,\ x \in B \Rightarrow x \in A.$$

We denote that B is a subset of A by writing $B \subset A$.

If $B \subset A$ and $B \neq A$, we say that B is a **proper** subset of A.

Note. Some authors denote inclusion by $\subseteq$ and reserve the use of the sign $\subset$ to cases where the subset is proper.

If X is a set, then we denote the subset, A, of those elements $x \in X$ which satisfy a particular relation $R\{x\}$ by

$$A = \{x \mid x \in X,\ R\{x\}\} \quad \text{or, more concisely,} \quad A = \{x \in X \mid R\{x\}\}.$$

(It is an axiom that A is a set, see Zermelo–Fraenkel Axiom V, p. 214.)

Alternative notations are

$$A = \{x : x \in X,\ R\{x\}\}, \quad A = \{x \in X : R\{x\}\}.$$

If B is a subset of A, then we define the **complement** of B in A to be

$$\{x \in A \mid x \notin B\}.$$

(It follows that the complement will again be a set.)

The complement of B in A is denoted by $A - B$, $A \backslash B$, $\mathbf{C}_A B$ or, when there is no ambiguity concerning the set A, by B' or $B^{\sim}$.

Note. Many authors, particularly when writing on Boolean algebra and measure theory, write $A - B$ to denote $\{x \in A \mid x \notin B\}$ but do not imply thereby that $B \subset A$ (see the definition of *difference* below).

The complement in A of the set A itself is called the **empty** subset of A. This set has no elements and is independent of A, i.e. $A - A = B - B$. It is called the **empty (null, void) set** and is denoted by $\varnothing$.

$\varnothing$ is a subset of every set X.

The set containing the single element x is denoted by $\{x\}$. A set having exactly one element is called a **singleton**.

Similarly, the set with elements $a, b, c, \ldots, z$ is denoted by $\{a, b, c, \ldots, z\}$.

Given a set A we denote the set whose elements are the subsets of A by $\mathscr{P}(A)$ or 2^A (see Zermelo–Fraenkel Axiom IV, p. 213). $\mathscr{P}(A)$ is called the **power set** or **set of subsets** of A.

Let A and B be two subsets of some set X.

(*a*) The **intersection** of A and B, written $A \cap B$ or AB, is the set of all those elements which belong to both A and B, i.e.

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

A and B are said to be **disjoint** if they have no elements in common, i.e., if

$$A \cap B = \varnothing.$$

(*b*) The **union** of A and B, written $A \cup B$ or $A + B$, is the set of all those elements which belong to A or to B (or to both), i.e.

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

The definitions of *intersection* and *union* can be extended to cover the intersection or union of a *family* of sets.

If I is any set, then we say that $\{A_i\}_{i \in I}$ is a **family of sets indexed by** I, the **index set**, if, corresponding to each $i \in I$, there is an associated set which we denote by A_i.

The **intersection** of a non-empty family of sets $\mathscr{A} = \{A_i\}_{i \in I}$ (i.e. $I \neq \varnothing$) is defined to be the set comprising those elements x which satisfy $x \in A_i$ for all $i \in I$. This intersection is denoted by

$$\bigcap_{i \in I} A_i \text{ or } \bigcap_{A \in \mathscr{A}} A.$$

Similarly, the **union** of the family of sets $\mathscr{A} = \{A_i\}_{i \in I}$ is defined to be the set of all those elements which satisfy $x \in A_i$ for some $i \in I$. This union is denoted by

$$\bigcup_{i \in I} A_i \text{ or } \bigcup_{A \in \mathscr{A}} A.$$

Note. (i) The use here of Zermelo–Fraenkel Axiom III (p. 213).

(ii) The fact that the use of i is not significant; any other letter may be used in its place (see also p. 134). For this reason we refer to i as being a *dummy suffix*.

(*c*) The **difference** of A and B is the set

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

If $B \subset A$, the difference is said to be **proper**. If $B \not\subset A$, then the difference is sometimes referred to as the **relative complement** of B in A. As noted above, the difference is often denoted by $A - B$.

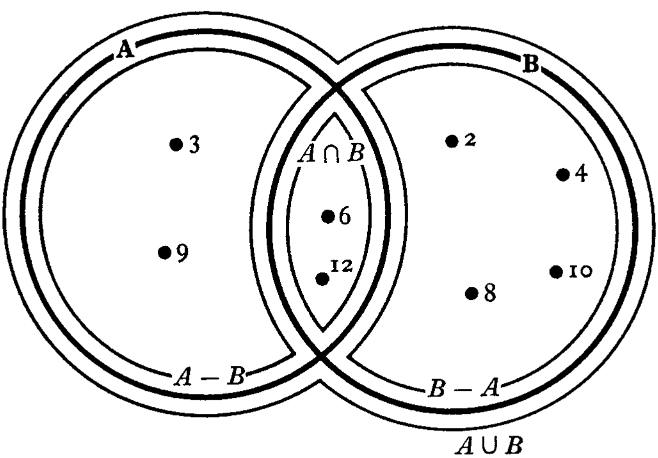
(*d*) The **symmetric difference** of A and B, $A \triangle B$, is

$$A \triangle B = (A \cap \mathbf{C}_X B) \cup (\mathbf{C}_X A \cap B),$$

or, in terms of *relative complements*,

$$A \triangle B = (A - B) \cup (B - A).$$

Example. Let $A = \{3, 6, 9, 12\}$ and $B = \{2, 4, 6, 8, 10, 12\}$.



Then $A \cap B = \{6, 12\}$ – the *intersection* of A and B,

$A \cup B = \{2, 3, 4, 6, 8, 9, 10, 12\}$ – the *union* of A and B,

$(A \cup B) - A = \{2, 4, 8, 10\}$ – the *complement* of A in $A \cup B$,