1

INTRODUCTION

1.1. The nature of algebra

Algebra is first a generalisation of arithmetic. Instead of dealing with particular numbers we use letters to denote arbitrary ones and work with these according to the usual rules of arithmetic. It is concerned with those properties and processes that are common to all numbers and not those, such as primeness, peculiar to certain numbers or integers. The interest of algebra lies in the processes we use and their consequences, some of the chief fields of the work being the use of the four rules of addition, subtraction, multiplication and division and problems involving them, such as simplification of expressions, the solution and investigation of equations, the study of polynomials and other functions and their graphs, and the investigation of inequalities.

Elementary algebra uses letters to stand for numbers of various types: fractions, real numbers and later complex numbers. Most of the work is similar in all these cases, the processes and rules being almost identical, and the algebra is not basically concerned with the particular set of numbers in question, but rather with the methods and rules for combining them.

Algebra is essentially a finite process. We often include under the heading of 'algebra' such topics as convergence of series and the study of transcendental functions such as the exponential and logarithmic functions, but these properly belong to analysis, which is concerned with limiting processes and the infinite and infinitesimal.

1.2. Abstract algebra

As we have indicated, algebra is concerned basically with the processes and rules of combination of numbers, rather than with the numbers themselves. This was not realised by the early workers, but in the early part of the nineteenth century mathematicians gradually came to understand that the same or similar processes to those used in elementary algebra could be applied to many objects or sets other than numbers, and modern abstract algebra came into being.

Modern algebra then is concerned with sets of objects and possible ways of combining the elements of the set. Many rules of combination have been investigated and others are still being studied: the most fruitful are those with similar properties to the ordinary four rules of elementary work, though applied much more generally than to numbers alone.

Sets, together with rules of combining their elements, form *algebraic structures*. Much of the interest is synthetic, i.e. is concerned with the shape of the structure as a whole, but the analytical aspect of investigation of the elements themselves is important in some cases.

In this book we deal first with various special structures, emphasising those aspects that are of general application, and later, particularly in volume 2, investigate more general structures. Thus the reader is led gradually into the purely abstract work, and even there many concrete examples are given.

1.3. The axiomatic approach to mathematics

Modern pure mathematics is almost entirely axiomatic in approach. This is a fairly recent development: for most of its history mathematics has been ostensibly based on the natural or everyday world. Thus, whatever Euclid himself believed his system of geometry to be, it has usually been taken to be a description of physical space, while the nature of real numbers was held to be self-evident by most people. Such foundations gradually proved unsatisfactory. The assumptions behind Euclid are difficult to state clearly and since the formulation of the theory of relativity it has been found that the natural world does not obey them anyway. Analysis became rigorous only about 200 years ago, and the nature of irrational numbers was not described until 1872, while the nature of number itself is still being discussed.

The intuitive ideas behind mathematics are thus not secure, and the study of the foundations belongs more properly to philosophy: mathematics itself is concerned with the deductions obtained from the basic ideas. It is therefore more satisfactory to lay down certain initial axioms or postulates and to deduce from them according to the accepted laws of logic (which are themselves the subject of study by philosophers). Mathematics is not concerned with the question whether the axioms are 'true' or not; all it can say is that given certain assumptions then other results and consequences follow logically. Theoretically any set of axioms may be chosen provided it is self-consistent, but obviously the work will be unfruitful unless the choice is a careful one, and natural phenomena and the traditional fields of study can lead us to suitable sets of axioms. Thus Euclid may be put on a proper footing by assuming certain axioms, while the choice of similar sets will lead to the various non-Euclidean geometries. The fewer the axioms the greater generality the resulting system possesses, but the fewer the results that may be deduced.

Thus modern mathematics lays down postulates and deduces from them. This can be a surprisingly fruitful process, both in practical terms (since many practical systems will obey the axioms chosen) and in aesthetic ones. It enables us to study systems that seem at first sight impossible but which often turn out to be extremely useful. For example, the study of space of four dimensions would seem useless at first sight, but it is vital to relativity theory and also in electromagnetism.

Abstract algebra lays down postulates for combining elements of sets and studies their consequences. For numbers these are the fundamental laws of addition and multiplication (the Commutative, Associative and Distributive Laws), while the choice of some only of these leads to the study of more general structures. A surprising amount of work may be done with very few axioms in this subject (group theory has only three basic laws but research is still very active in the subject).

1.4. Logic in mathematics

Mathematics uses the laws of logic and we do not attempt to lay down what these are or to study them. There are, however, a few important logical ideas which are often not understood properly by the mathematical student but which are vital to much of his work, especially that which is concerned with the consequences of axioms rather than the techniques of manipulation. We explain some of these here.

Equality and identity

In elementary work the equals sign is usually used to indicate that two expressions have the same value, e.g. $x+2 = 4$ means that $x+2$ and 4 have the same value, for some particular value of x, which often has to be found. If two expressions have the same value for all values of the variable concerned we usually use the identity sign; thus $(x+2)^2 \equiv x^2+4x+4$. The distinction is often blurred in practice. The equals sign is also used as a special case of an inequality, thus we use '$\leqslant$' and '$\geqslant$', and both being true implies equality.

We will use '$a = b$' to mean that a and b *are the same*. This implies nothing about inequalities and may be used whatever type of element we are dealing with. Thus for numbers '$x = y$' means that x and y are the same number, while if we are dealing with polynomials '$P(x) = Q(x)$' means that $P(x)$ and $Q(x)$ are the same polynomial, according to our definition of sameness of polynomials (they are of the same degree and have all coefficients the same). If A and B are sets '$A = B$' means that they are the same set, not merely that they are of equal size. Other uses of the symbol (for example, for isomorphic groups) will be given as we require them.

Implication

If a statement A leads logically to another statement B we say that 'A implies B' and write $A \Rightarrow B$. For example: R is a square $\Rightarrow$ R is a rectangle, or $x = y \Rightarrow x^2 = y^2$, or the positive integer n is even $\Rightarrow$ n can be divided by 2. If $A \Rightarrow B$ then B is implied by A and we sometimes write $B \Leftarrow A$. If it does not follow logically that B is true when A is, we write $A \not\Rightarrow B$ or $B \not\Leftarrow A$.

In the first two examples above we have $A \Rightarrow B$ but $B \not\Rightarrow A$, while in the third $B \Rightarrow A$ also. In this case we say that 'A implies and is implied by B' and write $A \Leftrightarrow B$. In such a case the statements are logically equivalent and the argument in which B follows A is reversible. This is not the case in the other two examples. Again the distinction is often not realised, as when solving an equation we square both sides. This is not reversible and so, although any solution must be given by our method, it does not follow that a solution we obtain is a solution of the original equation, and each must be checked.

If A and B must be either true or false statements, so that if we call the negation of A 'not A' then either A or not A is true, $A \Rightarrow B$ is logically equivalent to 'not $B \Rightarrow$ not A'.

The sign $\Rightarrow$ must not be confused with $\rightarrow$, nor $\Leftrightarrow$ with $\leftrightarrow$, the latter being used for various correspondences. Thus in inversion we could write $P \rightarrow P'$ or $P \leftrightarrow P'$ where P and P' are inverse points.

Necessary and sufficient conditions

A is a *necessary* condition for B means that $B \Rightarrow A$: if B is true then A must be.

A is a *sufficient* condition for B means that $A \Rightarrow B$: if A is true then B must be.

Thus a necessary condition for an integer greater than 2 to be prime is that it is odd, but this is not sufficient. A sufficient condition for a figure to be a rectangle is that it is a square, but this is not necessary. However, the necessary and sufficient condition for two triangles to have their sides in proportion is that they are equiangular.

We often have a set of necessary and sufficient conditions. (A is a *necessary and sufficient* condition for B means that $A \Leftrightarrow B$.) Thus necessary and sufficient conditions for a figure to be a square are that it is a rectangle and also a rhombus.

If and only if

This gives another way of thinking about the ideas of implication. If $A \Rightarrow B$ we say that B is true *if* A is true, while if $B \Rightarrow A$ we say that B is true *only if* A is. Thus 'B true if A is' means that A is a sufficient condition for B and vice versa, while 'B true only if A is' means that A is a necessary condition for B. If $A \Leftrightarrow B$ we say that B is true *if and only if* A is true.

Counter examples

If we have a theorem which seems likely to be true but which may not be, usually the best way of proving it false is to find an example where it is not true. Such an example is called a counter example, the German 'Gegenbeispiel' being sometimes used. A famous instance is Fermat's theorem on binary powers, which states that the number $2^{2^n}+1$ is prime for all n. Although Fermat believed in the truth of this theorem he could not prove it, and in 1732 Euler discovered that if $n = 5$ the number is composite: this counter example of course immediately disproves the theorem.

As a further example, it is easily proved that if Σu_r is convergent then $u_r \to 0$, and we may think that the converse is true, but $u_r = 1/r$ gives a counter example.

If after due consideration we cannot find a counter example we may reasonably suppose that the theorem is true, but this of course is not proved, and it may well be false in some obscure cases. Goldbach's conjecture, that every even integer may be expressed as the sum of two primes, has never been proved, but no counter example has been discovered and most mathematicians believe in the truth of the conjecture.

Reductio ad absurdum

A common way of proving a theorem is to assume that it is false and then to show that this leads to a logical contradiction or to an obviously false result. For example, to prove that there is no greatest prime we assume that there is and let the greatest prime be n. Then $n!+1$ (where $n!$ means the product $n(n-1)(n-2)\ldots2.1$) either is a prime greater than n or has a prime factor greater than n (all integers from 2 to n are factors of $n!$ and so cannot be factors of $n!+1$), and in either case we have a contradiction of our supposition. This method is particularly common in proving a converse: the theorem that if the opposite angles of a quadrilateral are supplementary then the quadrilateral is cyclic is usually proved in this way, assuming the basic theorem that the opposite angles of a cyclic quadrilateral are supplementary.

1.5. Historical summary

The first notable algebraists were the Arabs. The Egyptians, Greeks and Hindus had all done a little work in this subject, but the Arabs were the first to concentrate on it, mainly in connection with astronomy, and progressed so far as the solution of cubic equations. The word 'algebra' is a corruption of the Arabic 'al-jebr', meaning the transposing of negative terms in an equation to the other side.

At the time of the Renaissance, algebra became one of the main fields of mathematical study. Cubics were solved for the general case by Tartaglia (about 1499–1557) and Cardan (1501–76) and quartics by Ferrari (1522–65). Vieta (1540–1603) introduced letters to stand for unknown quantities, while the symbols + and − appear first in a book printed in 1489, and the exponential notation for powers was introduced by Descartes (1596–1650).

Newton (1642–1727) worked on the theory of equations and the binomial theorem, and about this time negatives came to be accepted as proper numbers. Complex numbers were also used but were imperfectly understood until later. Argand's famous paper on the geometrical interpretation of imaginary quantities was published in 1806, while Gauss finally put complex numbers on an equal footing with the real numbers in 1831. Gauss gave the first fully satisfactory proof of the 'Fundamental Theorem of Algebra' that a polynomial equation of the nth degree has exactly n roots, his first proof being discovered in 1797.

Determinants were studied by Wronski (1778–1853), Cauchy (1789–1857) and Jacobi (1804–51) among others, while matrices were introduced at about the same period, much of the work being by Hamilton (1805–65) and Cayley (1821–95). Hamilton also invented quaternions, the first non-commutative system to be studied intensively, which were superseded for practical purposes by matrices and tensors. The theory of invariants and linear transformations, connected with matrix theory and leading to modern linear algebra, was developed by Cayley and Sylvester (1814–97) and by Hermite (1822–1901).

A milestone in the development of modern abstract ideas was

Boole's publication of *The Laws of Thought* in 1854. This applied mathematics to logic and marked the first real break from traditional ideas, based on practical ideas of number and space. Boole's work showed that algebra is not necessarily concerned with numbers but that the processes may be used much more generally.

The beginnings of the ideas of group theory lay in the solubility of equations. Quartics had been solved by Ferrari in the sixteenth century but nobody had been able to give a solution for the general quintic, and this was finally proved impossible by Abel in 1826. Galois simplified his solution in about 1830 and discovered a great deal about groups in connection with the solution of equations, besides investigating invariant subgroups and the theory of fields. He was the first to use the word 'group' in the modern sense. The theory was elaborated by Lagrange, Cayley and particularly Cauchy (in about 1844–46). At this early period groups were thought of in terms of permutations or substitutions, or sometimes in connection with residues (Euler) and number theory. Definitions of abstract groups were given by Kronecker in 1870 and later simplified. Other notable early workers in group theory were Jordan (composition series and conditions for groups to be soluble), Sylow (1832–1918) (subgroups), Sophus Lie (1842–99) (topological groups) and Klein (groups of the regular polyhedra).

Topology, originally known as 'analysis situs', was studied by Euler and others, but was only gradually recognised as a separate subject, distinct from geometry.

In the present century the growth of abstract ideas has been rapid. Many algebraic systems have been studied and research is still active both in the 'pure' algebra of groups, rings and fields and more recently in algebra applied to topological structures.

2

SETS

2.1. The idea of a set

A set is merely a collection of objects. Although the idea is basically simple, and indeed hardly seems to need stating, it is the most important concept in mathematics. (The latter has even been described as being the study of various aspects of set theory!) The reason is that mathematics is essentially a process of abstraction—we select certain properties of the objects with which we are working and apply the laws of logic to deduce further properties. We cannot do this without putting some restriction on our objects: we deal, in a certain piece of work, only with objects which are in a given set.

It may not at first sight appear obvious that elementary mathematics restricts itself in this way. We tend to think of arithmetic as applying to everything, but of course this is a false idea. Arithmetic in the first instance is concerned merely with the properties of numbers—we do it in the *set* of numbers. Even here we work in different sets at different stages. At first we restrict ourselves to the set of positive whole numbers, which we later extend to include fractions, then negatives, and finally we do our arithmetic within the set of all real numbers. When we apply our arithmetic to problems we extend the sets that we use to include, for instance, all objects which have a monetary value, or all baths with two taps (in the famous calculations of this type). Notice, however, that the properties which we abstract from these sets are precisely those which are possessed by ordinary numbers. There is no new mathematics involved, and so these sets of practical objects have little purely mathematical interest.

When we start algebra we still, in the elementary stages, keep within the set of numbers. We let x 'stand for' any number. We very easily lose sight of the basic set, and later we start using our letters to stand for any *complex* number, thus extending our real number domain of operations to the complex domain. Remarkably our algebra is nearly the same, and the purpose of this book is to show how we can still keep the same algebra, or at any rate some of it, when working in sets other than the real or complex numbers.

The study of geometry leads us into new sets. We work, in Euclid, within any one of the set of planes, we deal with the set of points in that plane, and with the sets of lines and triangles. Our results apply to the sets of objects which satisfy certain postulates.

In advanced mathematics more varied sets are encountered. Differentiation can be performed only within the set of *differentiable* functions (i.e. functions which possess a derivative at a certain point, a property which by no means all functions have). There is a separate set of *integrable* functions. We are interested in the set of convergent series—we cannot speak of the sum of a series that isn't convergent. The method of induction applies only to the set of integers. We make excursions into the set of vectors.

The above examples are of mathematical sets, but of course the idea may be applied to any collection of objects. The objects are called *elements*. They may be of any type, and even of varied types. The set may consist of a finite number of elements, or of infinitely many. We may not even know how many. So long as we can say of any object that it is either an element or is not, then we have defined a set. Thus we may consider the set of all mammals who have been parents of live-born young that have ever lived. It would be difficult to give an estimate of the size of this set but, given any object, it is possible to tell whether or not it is in the set. (At least it is possible in theory, provided we have an exact definition of mammal.) No object that is not a mammal need be considered and every mammal either has been a parent or has not. Notice that the definition is precise.

The above example has a simple definition, but this need not be so. Any selection of, say, ten thousand insects forms a set, and there may be no obvious connection between the elements in this case. We may even take ten thousand insects and one jam-jar and thus form another set. The elements may even be