1

SOME FUNDAMENTAL IDEAS AND NOTATIONS

1. 'Knowns' and 'unknowns'; constants and variables. When a problem is stated in the language of algebra, there are usually certain numbers regarded as given and others whose values have to be found. The former are *constants* of the problem, and may be ordinary arithmetical numbers, or algebraic numbers denoted by symbols $a, b, c, d, \ldots$. The latter are often, but by no means always, denoted by letters like x, y, z from the end of the alphabet. In this context they are *unknowns*; but they are also called *variables* since they are, at any rate to begin with, capable of assuming varying values until finally pinned down by the conditions of the problem.

Illustration 1. *To determine two numbers whose sum is* $3a$ *and the sum of whose squares is* $5a^2$.

The statement of the problem involves the two arithmetical constants 3, 5 and the algebraic constant a. To solve it, assume that the 'unknowns' have values x, y. The first condition imposes the restriction

$$x+y = 3a,$$

and the second the restriction

$$x^2+y^2 = 5a^2.$$

If the first equation is considered in isolation, then x, y can take an infinite variety of values, for example $0, 3a$; $a, 2a$; $2a, a$; $3a, 0$; $4a, -a$; $5a, -2a$; and so on. Thus x, y appear as genuine variables. But when the second equation is superposed, the values of x, y are restricted very completely. For, by the first equation,

$$x = 3a-y,$$

so that, by the second equation,

$$(3a-y)^2+y^2 = 5a^2,$$

or

$$2y^2-6ay+4a^2 = 0,$$

or

$$2(y-a)(y-2a) = 0.$$

Hence *either* $y = a$, so that $x = 3a-a = 2a$,

or $y = 2a$, so that $x = 3a-2a = a$.

The solutions of the two equations are thus *either* $x = 2a, y = a$, *or* $x = a$, $y = 2a$.

What appeared as variables were in fact constants all the time; but the conception of them as variables *in the first instance* was essential to the solution of the problem.

Note for the more experienced reader. This is a convenient point for a short logical note which the beginner may prefer to postpone. What we have actually proved is that, *if the two equations*

$$x+y = 3a,$$
$$x^2+y^2 = 5a^2$$

have any solutions at all, then they are either $x = 2a$, $y = a$ or $x = a$, $y = 2a$. It is, logically, not yet established that the equations are in fact soluble, and the argument ought to be completed by substituting the available values in the given equation, when all is well.

In this case, substitution shows at once that both solutions are possible, and the problem is then solved absolutely. There are, however, examples in which this is not so. Suppose, say, that x is required to satisfy the equation

$$1+\sqrt{x} = \sqrt{(5-x)},$$

where *positive* square roots are to be taken. Since the two sides are equal, so also are their squares, so that

$$1+2\sqrt{x}+x = 5-x,$$

or

$$\sqrt{x} = 2-x.$$

Once again, the squares are equal, so that

$$x = 4-4x+x^2,$$

or

$$x^2-5x+4 = 0,$$

or

$$(x-1)(x-4) = 0.$$

Hence x is *either* 1 *or* 4. But verification in the original equation shows that $x = 1$ is a solution whereas $x = 4$ is not.

In practice, the final verification is often (indeed, usually) omitted, but care must be taken in doubtful cases, especially when square roots are involved. See the report *The Teaching of Algebra in Sixth Forms* issued by the Mathematical Association, 1957, where chapter II gives an excellent account of the problem.

Warning Example. Criticize the following solution of the equation

$$x+1 = 5:$$

'Square each side. Hence $(x+1)^2 = 25$, so that $x^2+2x-24 = 0$, or $(x-4)(x+6) = 0$. Thus $x = 4$ or $x = -6$.'

2. Functions. An isolated expression like

$$3^3+4^3+5^3$$

has the surprising property that its value is 6^3, but, in the context of 'constants' and 'variables', has little further interest. On the other hand, the expression

$$(x-1)^3+x^3+(x+1)^3$$

has greater possibilities, for it can take an endless succession of values for varying values of x. When $x = 0, 1, 2, 3, 4$, for example, the values are 0, 9, 36, 99, 216. It is a 'living' expression, completely responsive to the changes which occur in the variable x.

An expression whose value depends in this way on a variable x is said to be a *function* of x.

The generality of a function, as compared with a fixed expression, often serves to reveal unsuspected properties common to groups of numbers. Thus, by direct multiplication,

$$\begin{aligned}&(x-1)^3+x^3+(x+1)^3\\&\quad = (x^3-3x^2+3x-1)+x^3+(x^3+3x^2+3x+1)\\&\quad = 3x(x^2+2).\end{aligned}$$

But the right-hand side has $3x$ as a factor. Hence *the sum of the cubes of three consecutive integers is exactly divisible by three times the middle integer.*

For example,
$$6^3+7^3+8^3 = 1071 = 21\times 51.$$

A function of a variable x may be conveniently denoted for reference by a single letter such as f; for instance,

$$f = (x-1)^3+x^3+(x+1)^3.$$

The dependence on x may be emphasized by the more extended notation

$$f(x),$$

read as 'f of x'. Other symbols, like

$$g(x),\quad h(x),\quad F(x),\quad U(x)$$

are also used.

The symbol $f(4)$ is used to denote *the value of $f(x)$ when x has the value* 4. Thus, if $f(x)$ is x^3-60, then

$$f(4) = (4)^3-60 = 4.$$

In the same way, when $f(x)$ is x^3-60,

$$\begin{aligned}f(-1) &= (-1)^3-60 = -61,\\ f(0) &= (0)^3-60 = -60,\\ f(3) &= (3)^3-60 = -33.\end{aligned}$$

Note, too, that, if $f(x)$ is x^3-60, then

$$\begin{aligned}f(x^2) &= (x^2)^3-60 = x^6-60,\\ f(\sin x) &= (\sin x)^3-60 = \sin^3 x-60.\end{aligned}$$

These and similar adaptations of the notation for a function will be incorporated without special reference.

3. Identities and inequalities; notation. The symbol $\equiv$ is used to denote *definition* or *identity*. Thus the statement

$$f(x) \equiv (x-1)^3+x^3+(x+1)^3$$

reads: *$f(x)$ is defined to be* $(x-1)^3+x^3+(x+1)^3$ for all values of x. Again, the statement

$$(x^2-1)^2+(2x)^2 \equiv (x^2+1)^2$$

reads: *the function* $(x^2-1)^2+(2x)^2$ *is identically equal to the function* $(x^2+1)^2$, *the relation being true for all values of* x; for example, the values $x = 2, 4, 6$ give the formulae ('Pythagoras' triangles)

$$3^2+4^2 = 5^2, \quad 15^2+8^2 = 17^2, \quad 35^2+12^2 = 37^2.$$

In both cases, the emphasis is on 'all values of x'.

Few writers are as correct as they ought to be in this matter, and this book is unlikely to be found entirely consistent. Strictly speaking, a statement

$$ax+b \equiv 3x+2$$

ought to mean that the two sides are equal for all values of x, so that $a = 3$, $b = 2$; and a statement

$$ax+b = 3x+2$$

ought to mean that there is a value of x such that

$$(a-3)x = 2-b,$$

the value being

$$x = \frac{2-b}{a-3}.$$

The first statement is usually given correctly; but in the second the sign of equality $=$ is sometimes used when the sign of identity or definition $\equiv$ is really meant. In practice, the context indicates which interpretation is intended.

The symbol $\neq$ is used to denote *inequality*. Thus the statement

$$x^2-3x+2 \neq 0$$

asserts that x must not have either of the values, 1, 2.

The symbol $>$ is used for *greater than*, and the symbol $\geqslant$ for *greater than or equal to*. Thus the statement

$$x^2 > 9 \quad \text{if} \quad x > 3$$

asserts that x^2 is greater than 9 whenever x is greater than 3; and the statement

$$x^2-2x+8 \equiv (x-1)^2+7 \geqslant 7$$

asserts that (since the square $(x-1)^2$ is always greater than or equal to zero) the function x^2-2x+8 is always greater than or equal to 7 in value.

The symbol $<$ is used for *less than*, and the symbol $\leqslant$ for *less than or equal to*.

A statement like $$1 \leqslant x < 5$$
means that x lies between 1 and 5, and further that x may be equal to 1.

One other symbol may be given at this point. The inclusion of a number between vertical lines $$|x|$$
means that the *numerical value* of x is to be taken.

Thus $$|-5| = 5, \quad |-\tfrac{1}{2}| = \tfrac{1}{2}, \quad |3| = 3.$$

With this usage, the inequality
$$-2 < x < 2$$
may be written $$|x| < 2,$$
and the inequality $$-3 \leqslant x \leqslant 3$$
may be written $$|x| \leqslant 3.$$

4. Zero. The 'number' zero has one characteristic property that must be clearly grasped: *if a product has value zero, then at least one of its factors must be zero*. This is the basis of the well-known routine used to complete the solution of a quadratic equation
$$(x-1)(x-2) = 0.$$

'Hence *either* $x-1 = 0$, *or* $x-2 = 0$;

that is, *either* $x = 1$, *or* $x = 2$.'

Observe, too, that division by zero is a meaningless operation; a symbol like $$\tfrac{3}{0}$$
has no meaning. Again, *a relation*
$$a.0 = b$$
cannot hold unless $b = 0$, *in which case it is automatically true for all values of* a.

5. Polynomials. Let x be a variable, and $a, b, c, d, \ldots$ a succession of constants. The expressions
$$ax+b,$$
$$ax^2+bx+c,$$
$$ax^3+bx^2+cx+d,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$

are called *polynomials* in x; a polynomial is formed as a sum of terms, each of which is a constant multiple of a power of x. The highest power of x is called the *degree* of the polynomial. Polynomials of degree 1, 2, 3, 4, 5, 6 are called linear, quadratic, cubic, quartic, quintic, sextic. The three polynomials at the top of this paragraph are linear, quadratic, cubic.

A polynomial is a simple example of a function of x. Its dependence on x may be exhibited to the eye by drawing the *graph* of the function. For instance, the polynomial x^2-2x+3 may be exhibited by the graph

$$y = x^2-2x+3$$

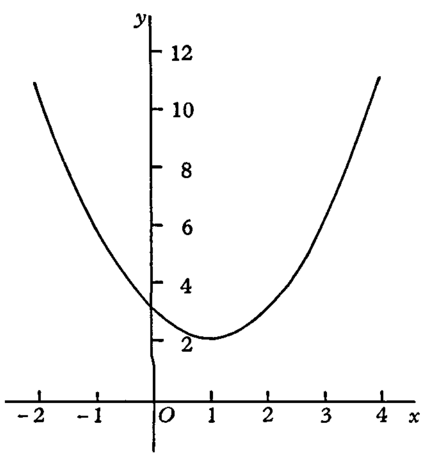

Fig. 1

shown in the diagram (fig. 1); the table, for values of x in integers between -2, 4, is

x	-2	-1	0	1	2	3	4
y	11	6	3	2	3	6	11

Some properties of the function

$$f(x) \equiv x^2-2x+3$$

may be deduced from the graph. For example,

(i) $f(x)$ is never less than 2 in value;

(ii) there are *two* values of x for which $f(x)$ takes any given value greater than 2;

(iii) since the graph is symmetrical about the line $x = 1$, the values of $f(x+1)$ and $f(-x+1)$ are equal for all values of x.

6. The zeros of polynomials; equations. Given a polynomial $f(x)$, particular importance attaches to those values (if any) of x for which

$$f(x) = 0.$$

The polynomial is then *equated to zero*, or, alternatively, *vanishes*. The values of x for which this happens are the *roots* of the equation.

An equation as first derived will probably not appear in a tidy form like

$$ax+b = 0,$$

$$ax^2+bx+c = 0.$$

Considerable algebraic manipulation may first be required to reduce it. But, once the preliminary manipulation is past, the equation (provided that it is of the polynomial type) must ultimately come to such a form; the method of solution then follows standard procedure, of which an account is given later.

***Illustration* 2.** *To solve the equation*

$$x-\tfrac{2}{3}(3x+4) = \tfrac{5}{6}-\tfrac{1}{2}(x+1).$$

Multiply throughout by 6:

$$6x-4(3x+4) = 5-3(x+1).$$

Remove brackets: $$6x-12x-16 = 5-3x-3.$$

Collect terms: $$3x+18 = 0.$$

Hence $$x = -6.$$

Examples 1

Reduce the following equations to linear or quadratic form, and then solve them.

1. $\tfrac{1}{2}(1-x) = \tfrac{1}{3}(1+x)$.
2. $\dfrac{3}{1-x} = \dfrac{2}{1-3x}$.
3. $\dfrac{x-1}{2}+\dfrac{x-2}{3}+\dfrac{x-3}{4} = 0$.
4. $(x+3)(x-3) = 3x-11$.
5. $x(x-4) = -3$.
6. $\dfrac{x-\frac{1}{2}}{\frac{1}{2}}+\dfrac{x-\frac{2}{3}}{\frac{2}{3}} = 0$.
7. $\tfrac{1}{2}(x+2a) = \tfrac{1}{4}(x-2a)$.
8. $\dfrac{x+a-b}{a} = \dfrac{x-a+b}{b}$.
9. $x(x-4) = x-6$.
10. $\dfrac{x}{3} = \dfrac{2}{7-x}$.
11. $\dfrac{x+3}{\frac{1}{2}}-\dfrac{x+1}{\frac{1}{3}} = \dfrac{x-1}{\frac{1}{4}}$.
12. $x-\tfrac{1}{5}(3x-5) = 2(\tfrac{1}{6}-\tfrac{1}{3}x)$.
13. $\tfrac{5}{3}(x-3)-2\{\tfrac{7}{2}(x-2)+3\} = 0$.
14. $x(x-1) = 6$.

7. Polynomials in several variables. The polynomials so far considered have all been functions of a single variable. The idea is easily extended. Expressions like

$$ax^2+bxy+cy^2+dx+ey+f,$$

$$x^3+y^3+x^2+y^2+x+y+1,$$

$$x^4y^2+3xy^3+5$$

are called *polynomials in* x, y; expressions like

$$x^3+y^2+z,$$

$$ax^2+by^2+cz^2+2fyz+2gzx+2hxy,$$

$$x^5y^2+y^4z^3+z^3x^4+14$$

are called *polynomials in* x, y, z. The characteristic feature of a polynomial is that it is a *sum of multiples of products of powers of the variables*; thus, for variables x, y, z, powers might be x^3, y, z^2, the product would be x^3yz^2, a multiple might be $5x^3yz^2$, and the polynomial is a sum of terms derived in this way.

The *order* of any term in the polynomial is the sum of the exponents of the powers of $x, y, z, \ldots$ in it; for example, the orders of terms like $3x^3yz^2$ and $7xy^5z^4$ are $(3+1+2) = 6$ and $(1+5+4) = 10$. The *degree* of a polynomial is the order of the term of highest order in it; for example, the three polynomials in x, y at the beginning of this paragraph have orders 2, 3, 6; and the polynomials in x, y, z have orders 3, 2, 7.

A polynomial is said to be *homogeneous* when all of its terms are of the same order. Typical examples are

$$2x+3y+7z,$$

$$5x^2+4xy+y^2-3xy,$$

$$x^3+y^3+z^3-3xyz.$$

8. The method of mathematical induction. This important method of investigation can be used for many purposes, and may be explained by two typical examples, of which the first is very elementary.

***Illustration* 3.** *To find a formula for the n-th odd number, in the sequence*

$$1, 3, 5, 7, 9, \ldots.$$

By inspection, the first, second, third odd numbers are

$$2.1-1, \quad 2.2-1, \quad 2.3-1$$

respectively, suggesting the formula

$$2n-1$$

for the nth. The method of induction proceeds by three stages:

(i) *Assume*, as a starting-point, that the formula $2n-1$ is valid up to some particular value k. Thus we *assume* that the kth odd number is

$$2k-1.$$

(ii) *Examine the* $(k+1)$*th term on the basis of this assumption*. The $(k+1)$th odd number, by definition, exceeds the kth by 2, so that it is

$$2k+1,$$

or $$2(k+1)-1.$$

(iii) *Observe* that the expression for the $(k+1)$th term is precisely the same as that for the kth with the number k replaced by $k+1$. Hence if the formula $2n-1$ is true, in the form $2k-1$, for $n = k$, it is necessarily true, in the form $2(k+1)-1$, for $n = k+1$.

But it is true, in the form $2.1-1$, for $n = 1$. Hence it is true for $n = 2$; hence it is true for $n = 3$; for $n = 4$; ...; and, successively, for all values of n. The formula $2n-1$ is therefore established.

***Illustration* 4.** *To find a formula for the sum*

$$1+3+5+7+\ldots+(2n-1)$$

of the first n odd numbers.

Write $S_1, S_2, S_3, \ldots$ to denote the sums of the first 1, 2, 3, ... odd numbers. Then

$$S_1 = 1,$$
$$S_2 = 1+3 = 4 = 2^2,$$
$$S_3 = 1+3+5 = 9 = 3^2,$$
$$S_4 = 1+3+5+7 = 16 = 4^2,$$

suggesting the formula $$S_n = n^2.$$

(i) Assume, as a starting-point, that the formula $S_n = n^2$ is valid up to some particular value k. Thus we *assume* that

$$S_k = k^2.$$

(ii) Examine S_{k+1} on the basis of this assumption. The sum S_{k+1} is found by adding to S_k the $(k+1)$th odd integer, so that (as above)

$$S_{k+1} = k^2+(2k+1)$$
$$= (k+1)^2.$$

(iii) The expression for S_{k+1} is precisely that for S_k with k replaced by $k+1$. Hence if the formula is true for $n = k$, it is necessarily true for $k+1$.

But it is true for $n = 1$, since $S_1 = 1^2$. Hence it is true for S_2; hence it is true for S_3; for S_4; ...; and, successively, for all values of n.

The formula $S_n = n^2$ is therefore established.

Examples 2

Use the method of mathematical induction to establish the following results:

1. The sum of the first n integers is $\frac{1}{2}n(n+1)$.

2. The sum of the first n powers of 2,

$$1+2+4+8+\ldots+2^{n-1},$$

is 2^n-1.

3. The sum of the first n even integers

$$2+4+6+8+\ldots+2n$$

is $n(n+1)$.

4. The number of straight lines joining all possible pairs of n points (in general position) is $\frac{1}{2}n(n-1)$.

[For example, the points A, B, C, D are joined in pairs by the six lines BC, CA, AB, AD, BD, CD.]

5. The number $3.5^{2n+1}+2^{3n+1}$ is divisible by 17.

6. The number $3^{2n+2}-8n-9$ is divisible by 64.

7. The number $3n^5+7n$ is a multiple of 5; hence $n^2(n^2+1)(n^2+4)$ is also a multiple of 5.

8. The sum

$$1^5+2^5+3^5+\ldots+n^5$$

is

$$\tfrac{1}{12}n^2(n+1)^2(2n^2+2n-1).$$