BASIC BOOKS IN SCIENCE

– a Series of books that start *at the beginning*

Book 1

Number and symbols

— from counting to abstract algebras

Roy McWeeny

Professore Emerito di Chimica Teorica, Università di Pisa, Pisa (Italy)

Books in the Series will be available –*free of charge*–from the websites

<www.paricenter.com> (see 'Basic Books in Science')

<www.learndev.org> (see 'For the Love of Science')

(Last updated 10 May 2007)

BASIC BOOKS IN SCIENCE

– a Series of books that start *at the beginning*

About this Series

All human progress depends on **education**: to get it we need books and schools. Science Education is one of the great keys to progress.

Unfortunately, books and schools are not always easy to find. But now, with modern technology, all the world's knowledge can be made available to everyone through the enormous data banks available on the Internet – the 'Web' that connects all the world's computers.

The Series "Basic Books in Science" aims to exploit this new technology by bringing basic knowledge in all areas if science within the reach of everyone. Every Book will cover in some depth a well defined area, starting from the very beginning and leading up to university entrance level, and will be made freely available on the Internet *at no cost to the reader*. To obtain a copy it should be enough to make a single visit to any library or public office with a personal computer and a telephone line. Each book will serve as one of the 'building blocks' out of which Science is built; and together they will form a 'give-away' science library.

About this book

This book, like the others in the Series, is written in simple English – chosen only because it is the language most widely used in science and technology, industry, commerce, and international business and travel. Its subject "Number and symbols" is basic to the whole of science; but it introduces a *new language*, nothing like the one we use in everyday writing and speaking. The spoken word came first in the evolution of language; then the written word (starting about four thousand years ago). Mathematics began to develop somewhat later – in China, India, and the countries around the Mediterranean. But the symbols of mathematics, though still just marks on paper, are not connected in any way with speech, or sounds, or the written word: usually they stand for *operations* such as counting or moving things through space, operations which are often performed only in the mind. It is this *abstract* and *symbolic* nature of mathematics that makes it seem difficult to so many people and shuts them off from an increasingly large part of Science.

The aim of this first book in the Series is to open the door into Mathematics, ready for going on into Physics, Chemistry, and the other Sciences.

Looking ahead –

Your're starting on a long journey – a journey of discovery that takes you from ancient times, when people first invented languages and how to share their ideas with with each other by talking and writing, to the present day.

Science began to develop only a few thousand years ago, with the study of the stars in the sky (leading to **Astronomy**) and the measurement of distances in dividing out the land and sailing the seas (leading to **Mathematics**). With what we know now, the journey can be made in a very short time. But it's still the same journey – filled with surprises – and the further you go the more you will understand the world around you and the way it works. Along the way there are many important 'milestones':

- After the first two Chapters in Book 1, you will know how to work with numbers, going from the **operation** of counting, to various ways of **combining** numbers – like adding and multiplying. And you'll have learnt that other symbols (such as the letters of the alphabet) can be used to stand for *any* numbers, so that $a \times b = b \times a$ is a way of saying that one number (a) times another number (b) gives exactly the same answer as b times a – whatever numbers a and b may stand for. Mathematics is just *another language*.

- In Chapter 3 you'll find how questions that *seem* to have no answer can be given one by *inventing* new kinds of number – **negative numbers** and **fractions**.

- After Chapter 4 you will be able to use the **decimal system** and understand its meaning for all the **rational numbers**.

- In Chapter 5 you pass two more milestones: after the first you go from the rational numbers to the 'field' of *all* **real numbers**, including those that lie *between* the rational numbers and are called 'irrational'. The second breakthrough takes you into the field of **complex numbers** which can only be described if you define one completely new number, represented by the symbol i, with the strange property that $i \times i = -1$. There are no more numbers to find, as long as we stick to the rules set up so far.

- But we're not finished: human beings are very creative animals! The last Chapter shows how we can extend the use of symbols to include operations quite different from the ones we've used so far.

CONTENTS

Notes to the Reader.

When Chapters have several Sections they are numbered so that "Section 2.3" will mean "Chapter 2, Section 3". Similarly, "equation (2.3)" will mean "Chapter 2, equation 3". Important 'key' words are printed in **boldface**: they are collected in the Index at the end of the book, along with the numbers of the pages where you can find them.

Chapter 1

About numbers

1.1 Why do we need numbers?

In the 'exact' or 'physical' sciences we deal with **quantities** and the way different quantities are related: it is not enough to know what the quantities 'are' – we must be able to **measure** them. Measurement of a quantity means we must compare it with a certain 'unit' and decide *how many units* the quantity is equivalent to (in some agreed sense). If I am walking I might want to know *how far it is* to the school: that quantity is the **distance** and if I find it takes a thousand 'paces' or 'strides', then my stride is the unit of distance and a thousand of them is the distance to the school. As another example; if I say this room is three metres wide the unit is the metre and I mean that three one-metre rods, put end-to-end, will just reach from one side of the room to the other. The metre is the *standard unit of length* – that of a special 'measuring rod' kept in Paris since 1791. Standard measuring rods have been used for many thousands of years, long before history books were written: one of the most common was the 'cubit', which is about half a metre. Many one-cubit rods have been dug up in Egypt, where they were used in measuring the stones for building the great pyramids. But it was not until 1875 that the 'metric system', based on the metre, was accepted (by international agreement) as the standard system of measurements in science.

Any number of copies of the standard metre can be made using ordinary sticks: if the two ends of a stick can be put together with those of the standard metre, as in Fig.1a, then it also will have the same length as the standard unit. When three such sticks are put end-to-end, as in Fig.1b, they will just reach from one side of the room to the other and the width of the room is *equal* to that of three metres. Notice that width (or length) is the *distance between* two things, so width, length and distance are measured in the same units; and so is height (how high the room is).

The same ideas are used in measuring quantities of any kind: we must agree on the *unit*, on how we can *combine* two similar quantities, and on when they are *equivalent* (or *equal*). The ideas of **unit, combination** and **equivalence** (or *equality*) are always there. With length, the unit is the standard metre, the lengths of two objects are combined by putting them end-to-end as in Fig.1b, and two objects have equal length when they are 'able to

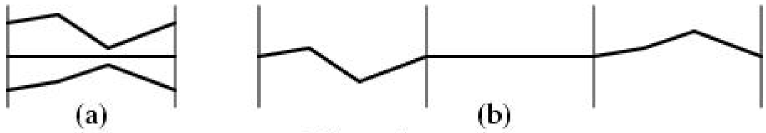


Figure 1

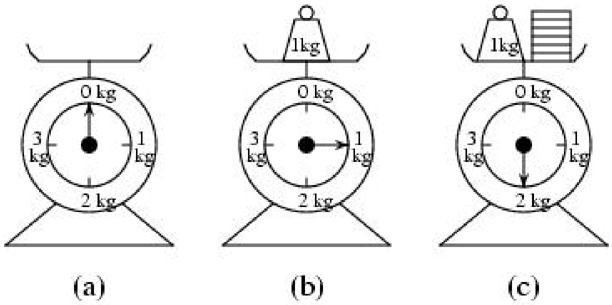


Figure 2

reach just as far' (e.g. three one-metre sticks or a single three-metre stick).

Another important quantity is the **mass** of an object. This can be measured by putting it in the 'scale pan' of a 'weighing machine', like the ones you find in the market (Fig.2a). The unit of mass is the **kilogram**, which is again kept in Paris. The shorter name for the unit is the 'kg'. The name for the international system of units, based on the metre (for length), kilogram (for mass), and second (for time), used to be called the 'mks system'. It has now become the Système International and is used in all the Sciences, throughout the world.

When a standard kilogram is put in the scale pan it makes the pointer move to a new position, marked with a "1 kg" in Fig.2b. If we take away the standard mass and put in its place a block of clay, then the clay will also have a mass of one kilogram when it makes the pointer move to exactly the same position.(Of course, the block of clay may not give exactly the same result, but by adding a bit more clay or taking a bit away we can get it just right.) Then we shall have *two* unit masses – the standard kilogram and our one-kilogram block of clay. Let's put them *both* in the scale pan as in Fig.2c and note the new position of the pointer: it will point to the "2 kg" on the scale. Putting the masses both together in the scale pan, and thinking of them as a single new object, is a way of combining masses. Then take both masses out and instead put a larger block of clay in the scale pan: if it moves the pointer to the same position ("2 kg"), then the larger block of clay will have a mass of *two* kilograms. Here, *combination* of masses means 'putting them in the same scale pan' and *equality* of two masses means 'moving the pointer to the same mark'. (Later we'll find a difference between mass and 'weight' – which is what the weighing machine measures – but more about that in Book 4. In everyday life we still compare the masses of two objects by weighing them.)

Finally, let's think about **time**. Time can be measured using a pendulum – which is a small heavy object, tied to a string and hanging from a support. Every (double-)swing of the pendulum, back and forth, takes a certain time; and as each swing seems to be the same as any other we can suppose they all take the same time – which can be used as a *unit*. The standard unit of time is the 'second' (measured using a standard pendulum) and the *combination* of two time units means 'waiting for one swing to follow another'. With a standard pendulum, sixty swings will mean the passing of one 'minute' – so by now it is clear that to measure things we must know about counting and numbers.

1.2 Counting: the natural numbers

Of course you can count – but did you ever really think about it? There are two ways of arriving at the idea of *number*. When I say "there are 'five' cows in that field" I may mean

> I have given names to the cows, one after another, from an 'ordered set' which I learnt by heart as a small child; and the last name I give out, when there are no more cows left, is the 'number' of cows in the field,

or I may mean

> I have pointed each finger of my right hand at a different cow, in turn, until no cows or fingers are left over; there is then the *same number* of cows as fingers and this is the number called 'five'.

The number we arrive at in the first way, by giving out names from the list 'one', 'two', 'three', 'four', 'five', 'six', ... (or, for short, 1, 2, 3, 4, 5, 6, ... , where the 'symbols' 1, 2, ... just stand for the words) is called the 'ordinal number', because the names must be given out in a standard *order*.

The number arrived at in the second way does not depend on the order of names in a list. Mathematicians call it the 'cardinal number': it depends on 'pairing' the objects in one set with those in another set – a 'standard set'. For example, I can pair the cows in the field with the members of a 'standard' set of objects, the fingers of my right hand, with one cow for each finger. But then we need to have a standard set for every number! So if I define 'two' as the number of hands I have, then I can say that I also have two eyes – because I can cover each of them with one of my hands. In the same way, by putting the fingers of my left hand tip-to-tip with those of my right hand (which was used as the standard set of '5' objects) I find there are 5 fingers also on my left hand. The kind of 'pairing' in these examples is described as 'making a one-to-one correspondence' between the members of two sets; and to remind us of what it means we can just as well use 'pairing number' as 'cardinal number.

Instead of using sets of hands or fingers to define numbers it is easier to use other objects such as heaps of pebbles or beads. Many people think this way of arriving at numbers was used thousands of years ago when men first started to keep animals: by having a heap of pebbles, one for each cow, it would be easy to tell if all the cows had come back at night by picking up a pebble for each cow and seeing if any pebbles were left in the

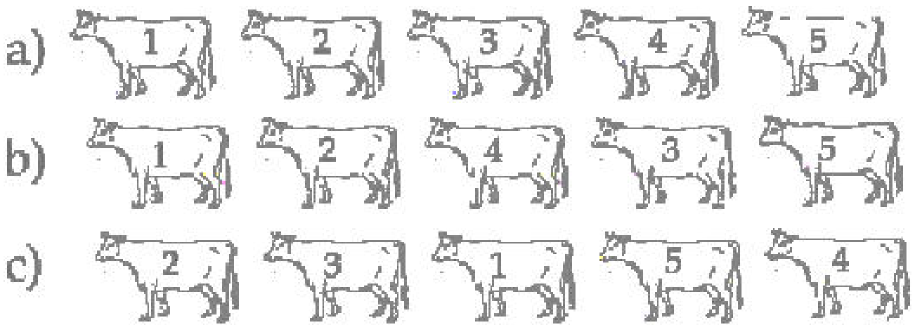


Figure 3

heap – if a pebble was left over, a cow was missing. Once we have learnt about numbers, we don't need to keep all those heaps of stones – as we'll now discover.

Does it matter how we find the number of objects in a set? The 5 cows are still 5 cows, whichever way we get the number. To make sure this is always true, let's write the symbols 1, 2, 3, 4, 5, 6, ... on pieces of sticky paper and stick one on each cow as in Fig.3a. Clearly 6 and the numbers that come after it will not be needed as 5 is the *ordinal* number given to the last cow. By sticking the labels on the cows we have made a one-to-one correspondence between the members in the sets of cows and pieces of paper; and so the two sets contain the same number of objects – the *cardinal* (or 'pairing') number 5. As long as we're talking about things we can count, *ordinal number and cardinal number must agree!* The reason for this is that the pairing between the members of two sets does not depend on the *way* in which the pairings are made: so if the labels were stuck on the cows in a different way, as in Fig.3b, their number would not be changed. Changing 3 into 4 and 4 into 3 is called an 'interchange'; and by making more and more interchanges we could mix up all the labels, as in Fig.3c, without changing the number of cows. In other words, *the number of objects in a set does not depend on the order of counting.* A famous mathematician of the last century has called this fact "one of the two most important principles in Mathematics" – the other one being that there is no way of telling left from right except by pointing. The arrangement of numbers in Fig.3c is called a 'permutation' of the numbers in their standard order (Fig.3a) and the number of possible permutations rises very fast when the number increases: 5 numbers can be arranged in more than a hundred different ways, but with 10 numbers there are *millions* of different permutations. To see how there can be so many let's think how the 5 cows might come into the field. The first one could be, perhaps, Number 4 – one out of 5 possibilities. The next to come in might be Number 2 – one out of the 4 cows left outside (because Number 4 is already in). So the first two cows can arrive in 20 ($= 5 \times 4$) different ways. Next comes another cow – which may be 1,3, or 5; so the first three cows to arrive could be labelled in 20×3 ($= 5 \times 4 \times 3$) different ways. Now you can guess the answer: there will be $5 \times 4 \times 3 \times 2 \times 1$ permutations of the cows. This number is called 'factorial 5' and is written as '5!'. If you want to count up to 10! you will see that the number certainly deserves an exclamation mark!

The numbers we arrive at by counting are called the ‘natural numbers’ or the ‘whole numbers’; they are **integers**.

1.3 The naming of numbers

There is no end to counting: we can go on forever – think of the stars in the sky. But we have to give names to the numbers and that becomes more and more difficult – and still more difficult to *remember* them all. The first few are (in words and in symbols)

‘one’	‘two’	‘three’	‘four’	‘five’	‘six’	‘seven’	‘eight’	‘nine’
1	2	3	4	5	6	7	8	9

but by the time we get to 9 we’re getting tired, so let’s stop there. And if there are more cows in the field what shall we call the next one?

To count beyond 9 we use a trick: we add just one more symbol to our set of 9, calling it ‘0’ or, as a word, ‘zero’ (or ‘nought’). The symbol set is then

$$1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8 \quad 9 \quad 0$$

and these particular symbols are called ‘digits’. The next number beyond 9 can now be called ‘10’ – which is a *new* name, neither 1 nor 0, known as ‘ten’. Our first set of ten integers is now

$$1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8 \quad 9 \quad 10$$

– matching the 10 members of the symbol set. If we want to count beyond 10 we can then change the zero into 1, 2, 3, ... to get more new names: they are (in symbols and in words)

11	12	13	14	15
‘eleven’	‘twelve’	‘thirteen’	‘fourteen’	‘fifteen’
16	17	18	19	20
‘sixteen’	‘seventeen’	‘eighteen’	‘nineteen’	‘twenty’

where the last number has been called ‘20’ (‘twenty’), ready for starting the next set of 10 numbers in which the zero will be replaced by 1, 2, 3, ... and so on. Of course, the names of the numbers, in words, will depend on what language you speak: in English ‘fourteen’ really means ‘four-and-ten’ (four places after the ‘ten’); but in German the same number is ‘vierzehn’ because ‘vier’ means ‘four’ and ‘zehn’ means ‘ten’. One of the nice things about Mathematics is that its language is the same for everyone – if you write ‘14’ nearly everyone in the world will know what you mean!

By going on in the same way we get the first ten sets of ten numbers. They can be set out as a ‘Table’:

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

where, in the last line, the '90' of the line before has been changed to '100' (the 9 changing to a 10) to show that it is the last number in Line 10. This number is called a 'hundred' and the one which will follow it (the first number on Line 11) will be called, in words, 'a hundred and one'. So we can go on counting as long as we like, giving every number a name by using these simple rules. Line 11 will end with '110'; the next one will start with '111', '112', ... and end with '120'; while Line 100 will start with '990', '991' and go up to '999' and then '1000' – the new number called a 'thousand'.

The natural numbers, counted in sets of ten using the symbols 1, 2, 3, ... 0 (the numbers 'to base 10'), were introduced into Europe by the Arabs – who brought them from India about 1300 years ago, *adding the zero*. Many things about them were not fully understood at the time, especially the meaning of the symbol 0, which (as we shall find later) can stand alone as a number in its own right, so the idea of number has had a long and difficult history. Indeed, it was not until the late 16th century that the symbols took the forms we use today.

We're lucky to be able to start with a 'ready-made' way of counting, invented so long ago and passed on from generation to generation – thanks to the discovery of writing.

Exercises

(1) Cut a unit stick (it doesn't have to be 1 m, it can be 1 'my-unit') and make several others of equal length. Then measure some distances. (You can also use a cubit as your standard length: the length plus the width of a page of this book give you almost exactly a cubit – and that is very nearly half a metre.)

(2) Make a simple weighing machine, using a small plastic bucket hanging from a piece of elastic. Fix a piece of wire to the bucket and put a card behind it, so you can see how far it moves when you put weights in the bucket. Then do the experiments with clay. (If you don't have a standard kilogram and clay, use a big pebble and plastic bags filled with sand.)

(3) Make a pendulum from a piece of string and a heavy bead. Then use it to compare times (for example, the times needed to fill different containers under a dripping water supply).

(4) Three friends (call them A, B, and C) come separately to your home: in how many ways can they arrive? Write out all the different possibilities (BAC, CBA, etc. – where BAC means first 'B', then 'A', and last 'C') and then count them.

(5) There are five cows (three white and two black) in a field. In how many ways can you take out two white and one black?

(6). Extend the Table (last page of this Section), whose last line is

$$91 \quad 92 \quad 93 \quad 94 \quad 95 \quad 96 \quad 97 \quad 98 \quad 99 \quad 100,$$

by adding 11 more lines.

(7). Write out all the numbers that Line 50 will contain.

Chapter 2

Combining numbers

2.1 Combining by addition

Now that we have numbers what are we going to do with them? First we need to know how they can be combined. The Table we just made shows two ways of doing this: let's look at the first two lines –

$$\begin{array}{cccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 & 19 & 20. \end{array}$$

the first kind of combination comes to mind when we read along any row and is called **addition**. From any integer, let's take 5 as an example, we can go to the next one by counting one more: this takes us to 6, the number following 5, and we write this result as $6 = 5 + 1$ where the 'plus' sign (+) stands for the act (or 'operation') of 'adding', the '1' stands for the 'one more', and the = stands for 'is the same as' or 'is equal to'.

The number we get by adding is the **sum** of the numbers on the two sides of the + sign. Of course, we can do the same thing again: after 6 comes 7, which means $7 = 6 + 1$, but $6 = 5 + 1$ so we can say $7 = 5 + 1 + 1$ and this can be written $7 = 5 + 2$ because $1 + 1 = 2$. And since $5 = 4 + 1$ we can say $7 = 4 + 1 + 2 = 4 + 3$, because two counts after 1 takes us to 3. Notice that two counts after 1 gives the same as one count after 2; and in symbols this means $1 + 2 = 2 + 1$ – the order in which the numbers are combined makes no difference. By using the symbols we can easily say things that would be long and complicated to say in words; and we can say things that are true for *any numbers whatever* – not just particular numbers like 6 and 7.

To make this clear, let's use the names a and b to stand for *any* two numbers: they could be 1 and 2, or 4 and 2, or 2 and 2 ... or whichever numbers we choose. When we add them we get a new number, which we can call c. If we have chosen – *just for the moment* – $a = 1$ and $b = 2$ then we know that $c = 1 + 2 = 3$. And, as we have noted above, $2 + 1$ gives the same result $c = 2 + 1$; so $a + b = c$ means also that $b + a = c$. In the same way, choosing for the moment $a = 4$ and $b = 2$, a glance at our Table shows that $4 + 2 = 6$ (two counts after 4) and that $2 + 4$ (4 counts after 2) gives the same result: again, with

a and b standing for the new pair of numbers (4 and 2) we see that

$$a + b = b + a \tag{2.1}$$

This is usually called a "law of combination". It is a law (or rule) for combining two numbers and it is true *whatever the numbers may be*. Mathematicians call it a **commutative law**, meaning that it doesn't matter which way the two things are combined – a added to b is the same as b added to a

We have also noted above that the way in which we *group* the numbers is not important, always giving the same result. For example, $7 = 4 + 1 + 2 = 4 + 3$ (in which we have added the last two numbers, before adding them to the 4) is the same as $7 = 5 + 2$ (in which we have added the first two numbers, before adding the 2). If we want to make the difference clear we can put the operations (additions) we do first within brackets, writing

$$7 = 4 + (1 + 2) = 4 + 3 \quad \text{and} \quad 7 = (4 + 1) + 2 = 5 + 2.$$

Again this is a *general* rule – it is true whatever numbers we are combining by addition. If, for the moment, we call the three numbers a, b, c then

$$(a + b) + c = a + (b + c) \tag{2.2}$$

This is the second important law for combining numbers by addition. Again mathematicians have a special name for it, calling it an **associative law**. Both (2.1) and (2.2) follow from the fact that number is independent of the order of counting.

When we first learnt about counting we probably played with *sets of objects* – not with numbers. If the objects are beads, on a string, the laws for combining them can be seen in pictures: for example

•••• •• and •• ••••

are the same set of beads, just arranged in two different ways. The pictures express the result (2.2), with $a = 4,\ b = 2$.

In the same way, the sum of three numbers, $a = 4,\ b = 1,\ c = 2$ gives us the picture

•••• • ••

(4 beads and 1 bead and 2 beads), which can be re-arranged as

••••• ••

(the picture for 5+2, 5 being 4+1) or, alternatively, as

•••• •••

(the one for 4+3, 3 being 1+2).

Now, however, instead of using sets of *objects*, we can play with *numbers* – which exist only in our *minds* and come from the *idea* of counting the members of a set. We have a symbol for every number and the 'rules of the game' are the laws of combination. There is no need always to count beads, or cows, ... – we can usually do it in our heads. Knowing that $3 + 4 = 7$ we can say at once that 3 cows together with 4 cows will make 7 cows – as also will 3 cows and 2 more and 2 more.

2.2 Combining by multiplication

Another way of combining numbers is called **multiplication**. Look at the first 3 lines of the Table in Section 1.3: they are

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30.

Each row contains 10 numbers and we see that 3 rows of 10 hold altogether 30 different numbers (30 being the last name if we count along the rows, one after another). The act of taking 3 sets of 10 objects, putting them together and thinking of them as *one new set*, lets us define multiplication: the numbers of elements in the three sets are related by saying "3 times 10 amounts to 30" and we write this in symbols as $3 \times 10 = 30$. The symbol $\times$ is the 'multiplication sign' and means 'multiply by' (just as the + means 'add to'); and the result is called the **product** (just as adding two numbers gives their 'sum').

By using a, b to stand for *any* two numbers, multiplication gives a number $c = a \times b$, such that, in words, a sets of b objects will contain altogether c objects. Again it does not matter what the objects are – they can be just the symbols that stand for numbers and it is not necessary that they be arranged in rows of ten, as in the Table. If we think of 3 rows of 4 objects we can picture them as $\begin{matrix} \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \end{matrix}$ and counting them tells us that there are altogether 12 objects: so $3 \times 4 = 12$. But suppose we turn the picture round, so as to get 4 rows of 3 objects. The set of objects has not been changed in any way – so the number it contains will not be changed. It will look like this:

$$\begin{matrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{matrix}$$

and will contain 4×3 objects. Counting along the 4 rows, one after the other, the result will be written $4 \times 3 = 12$; so the order in which the two numbers are written makes no difference. The law under which two numbers are combined by multiplication is thus

$$a \times b = b \times a \qquad (2.3)$$

– just as it was for addition, with the + sign in place of the $\times$.

Again, when we take three numbers, the operation of multiplication can be done in two different ways: $a \times b \times c$ can be obtained as $a \times (b \times c)$ or as $(a \times b) \times c$ – depending on which multiplication is done first. As in the second law (above) for addition. the two results are found to be equal:

$$a \times (b \times c) = (a \times b) \times c \qquad (2.4)$$

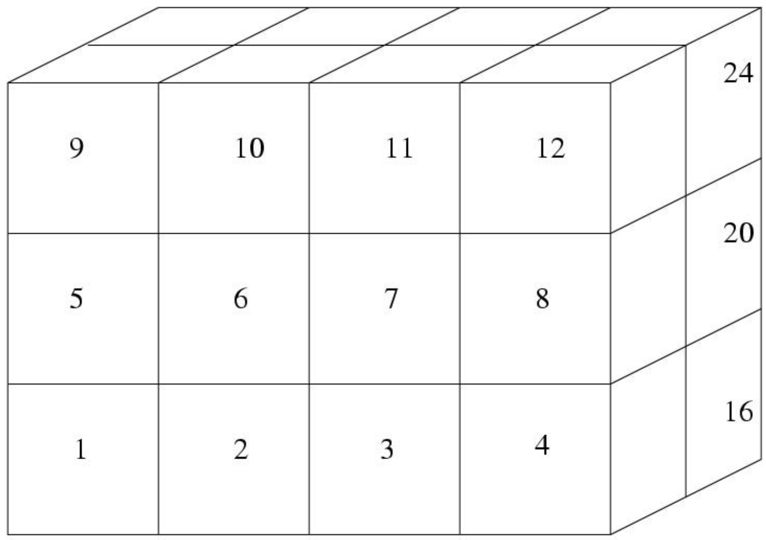


Figure 4

– it does not matter how we 'pair' the numbers in the product. To see why this is true, replace the numbers by numbered bricks and make them into a wall as in Fig.4, taking $a = 2,\ b = 4,\ c = 3$. The vertical side is 3 bricks high and the bottom layer contains $8 = 2 \times 4$ bricks; so does the next one and the 3 layers together hold $(2 \times 4) \times 3 = 24$ bricks. But we can also say there are 2 vertical slabs of bricks (back and front), each one holding $3 \times 4 = 12$ bricks, and so the number of bricks in the wall is $2 \times (3 \times 4) = 24$: whichever way we look at it, the wall contains 24 bricks. This is not a *proof* of what (2.4) says – it's only an example, showing that it is true when $a = 2,\ b = 3,\ c = 2$ and we are building a wall of bricks. To prove (2.4) for sets containing any numbers of members we must use only the properties of the numbers themselves – and that is more difficult.

Besides (2.1), (2.2), (2.3) and (2.4), we need only one more law for combining numbers: it reads

$$a \times (b + c) = (a \times b) + (a \times c), \tag{2.5}$$

where a, b, c are, as usual, any three integers. If the numbers refer to counted objects, (2.5) simply tells us that a rows of $b + c$ objects contain $a \times b$ objects (as if c were not there) plus $a \times c$ (when c objects are added to each row).

Exercises

1) Make up the following 'multiplication tables', using sets of dots and the counting method used in finding the result (2.3):

$1 \times 1 = 1$	$2 \times 1 = 2$	$3 \times 1 = 3$	$4 \times 1 = ?$	$5 \times 1 = ?$
$1 \times 2 = ?$	$2 \times 2 = ?$	$3 \times 2 = ?$	$4 \times 2 = ?$	$5 \times 2 = ?$
$1 \times 3 = ?$	$2 \times 3 = ?$	$3 \times 3 = ?$	$4 \times 3 = ?$	$5 \times 3 = ?$
$1 \times 4 = ?$	$2 \times 4 = ?$	$3 \times 4 = ?$	$4 \times 4 = ?$	$5 \times 4 = ?$
$1 \times 5 = ?$	$2 \times 5 = ?$	$3 \times 5 = ?$	$4 \times 5 = ?$	$5 \times 5 = ?$

Replace every question mark (?) by the right number.

2) Verify the laws (1.3), (1.4) and (1.5), taking first $a = 2,\ b = 3,\ c = 4$, and then with $a = 12,\ b = 3,\ c = 4$. (If you don't have a '12-times' table in your head, put $12 = 6+6$ and use the same laws with the smaller numbers.)

3) Make up a multiplication table going up to 9×9, starting from the '5-times' table in Exercise 1. (You'll need the numbers coming after 9, as listed in Section 1.3)

Chapter 3

Inventing new numbers – equations

3.1 Negative numbers and simple equations

We now know quite a lot about the 'whole' or 'natural' numbers, usually called **integers**. It has been said that God made the integers, leaving us to do the rest! And the integers can in fact be used to make new numbers, nothing like the ones we have used up to now.

To invent new numbers, we start from the **concept** of number, as developed so far in the rules of counting, comparing, and combining; and then notice that there are questions we can ask which seem to have no answers, leading only to nonsense. But instead of throwing away the 'nonsense' we look for ways of giving it a meaning, using only what we know already, and in this way we *extend* our concepts. This is the idea of *generalization* and it runs through the whole of mathematics.

Think first of the definition of addition: $a + b = c$. If someone tells us that $a = 2$ and $c = 5$ he may want to know what b must be. In this case b is an 'unknown'. An unknown number is often given the name 'x' – so we have to find x so that $2 + x = 5$, which is called an **equation**. The equation is 'satisfied' (found to be true) when $x = 3$ and this is called the **solution** of the equation.

The equation

$$a + x = c \tag{3.1}$$

has a solution when c is bigger than a because a set of a objects, after adding x objects, will contain *more* objects than before. We say there is a solution when $c > a$ (read as "c is greater than a"). But what if $c < a$ ("c is less than a")? As far as we know, there is no number which, when added to a, will make it smaller: if $a = 2$ and $c = 1$ there seems to be no answer to the question if $2 + x = 1$, what must x be? –for x cows added to 2 cows can never give 1!

Even if $a = 2,\ c = 2$ the set of numbers in the Table on p.7 does not contain one such that $2 + x = 2$. But we can *invent* one. We take the *zero*, in our symbol set (p.6), and think of it as our first new number: it has the property

$$a + 0 = a, \quad 0 + a = a, \tag{3.2}$$

whatever number a stands for! Adding zero to a number changes nothing.

Now let's go back to equation (3.1) and ask what the solution is when $c = 0$. Again, there is no number in our Table (p.7) of the integers such that $a + x = 0$. So let's again invent one, calling it 'a with a hat' or $\hat{a}$, such that any natural number a has a 'partner' $\hat{a}$ with the property

$$a + \hat{a} = 0. \tag{3.3}$$

As a result, the set of all our numbers is extended, becoming

$$\ldots \quad \hat{4}, \quad \hat{3}, \quad \hat{2}, \quad \hat{1}, \quad 0, \quad 1, \quad 2, \quad 3, \quad 4, \quad \ldots$$

where the three dots (...) mean that the list goes on forever both to left and to right.

We want all the new numbers to behave just like the ones we already know, obeying the same laws of combination (2.1) up to (2.5), so what must $\hat{a}$ mean? The natural numbers were first used in counting objects in a set: we have a simple picture of a set of 5 objects (e.g. the cows in a field); and even of an 'empty' set, containing 0 objects (e.g. a field with no cows); but how can we think of a set with $\hat{5}$ objects? At first sight it seems to be nonsense: but if we look at (3.3) it is clear that adding $\hat{a}$ objects must be the same as *taking away* a objects; so we can write $\hat{a} = -a$, where the 'minus' sign means 'take away' or 'subtract'.

It follows that (3.3) can be rewritten as $a + \hat{a} = a + (-a) = 0$, where the + sign and the brackets are not really needed, because the minus tells us to take away a from whatever goes before it. So we usually write (3.3) in the form

$$a + \hat{a} = a - a = 0 \tag{3.4}$$

– the number of objects in the empty set. In words, *adding $-a$ is the same as taking away a*. Remember also that, from (2.1), the order in which numbers are combined by addition is not important:

$$a + (-a) = (-a) + a \qquad \text{or} \qquad a - a = -a + a$$

The things separated by the = signs are all are equivalent to zero.

The act, or **operation**, of subtracting is said to be the **inverse** of addition: if b is any other number the effect of adding a to it is exactly undone by subtracting another a, for $b + a - a = b + 0 = b$ and b is left unchanged.

By using the minus sign, the complete list of whole numbers (above) becomes

$$\ldots \quad -4, \quad -3, \quad -2, \quad -1, \quad 0, \quad 1, \quad 2, \quad 3, \quad 4, \quad \ldots$$

where the numbers on the right of the zero are the **positive integers** ('natural numbers'), while those on the left are the **negative integers** ('invented numbers'). To use the full set of positive and negative integers we need only the rules (2.1) to (2.5). To find the number x, satisfying the equation (3.1), we note that (adding the same number $\hat{a}$ on both sides of the equal sign – so the two sides will stay equal) gives

$$x = c + \hat{a} = c - a.$$

Provided $c > a$ the answer is a positive integer (e.g. $x = 5 - 2 = 3$). But if $c < a$ we shall be taking away more than we have! So instead let's add $\hat{c}$ to both sides of (3.1), obtaining $a + x + \hat{c} = 0$, and then add $\hat{x}$ to both sides. The result is $\hat{x} = a + \hat{c} = a - c$ and this is a positive integer when $c < a$; so again there is no problem – the solution is the *negative* integer $\hat{x}$ or $-x$, which is the partner of x.

So now we have a set of all the integers, together with their negative partners and the zero, whose elements can be combined in two ways: (i) by the operation of addition, satisfying the laws (2.1) and (2.2); and (ii) by multiplication, satisfying (2.3), (2.4) and (2.5). The set also contains a 'unit under addition', namely the 0 (zero), which can be added to any number in the set without changing it ($a + 0 = a$); and for any integer a there is an 'inverse under addition', namely $-a$. Mathematicians call such a set a **ring**. The ring is 'closed' under the two operations, which always lead to another member of the set: the ring has the **closure** property.

There is one golden rule in playing with equations: if two things are equal and you put them on the two sides of an = sign, then the two sides will stay equal *provided you do exactly the same thing to both of them.*

Various simple rules follow from this principle. The first of them arises when we deal with combination by addition. If, for example,

$$a + b = c + d$$

then we can add $-b$ to both sides of the equation, getting $a + b - b = c + d - b$. But since $b - b = 0$, which can be added to any number without changing it, we are left with

$$a = c + d - b.$$

The b on the left, in the first equation, can be carried over to the right *provided* we replace it by $-b$ (its *inverse* under addition).

What happens, however, if b is already a negative integer ($-p$, say)? In that case the rule says we can move $-p$ over to the right if we replace it by $-(-p)$ - which we haven't met before! To find what the 'double-minus' means let's go back to the basic law (2.5), putting $c = -b$ and $a = -1$. The result is

$$-1 \times (b - b) = -1 \times b + -1 \times (-b)$$

or, since $b - b = 0$ and a factor 1 changes nothing, $0 = -b - (-b)$. But if we add b to both sides of the last equation we get (since $b - b = 0$) $b = -(-b)$. This is true for *any* number, so we can say

$$-(-a) = +a \qquad (3.5)$$

– the rule of 'two minus's make a plus'. So if you move a negative integer across the = sign you must make it positive. Whatever it is you move, you must reverse the sign!

As we know already, there is a second way of combining numbers, called multiplication, with somewhat similar properties; but we'll think about this in a later Section. For now, let's be content with what we have: we've already discovered a whole set of rules for combining various numbers $a, b, c, ...$ (it doesn't matter how many or what names we give them) and we need to test them carefully to be sure they always work.

The rules we have are given in 2.1), (2.2), (2.3), (2.4), and (2.5); they all apply to any three numbers – so it would seem – when they are combined using only the operations of addition and multiplication. Try them out by choosing many examples of what values the numbers might have (e.g. $a = 3,\ b = -5,\ c = 2$).

(Don't forget the properties of 0 and the fact that a negative integer is the inverse under addition of its positive partner.)

3.2 Representing numbers in a picture – vectors

Before going on, it is helpful to have a picture in which we can 'see' the properties of numbers. The number of objects or 'elements' in a set is always a positive integer. In the scale of numbers,

$$\ldots\quad -4,\quad -3,\quad -2,\quad -1,\quad 0,\quad 1,\quad 2,\quad 3,\quad 4,\quad \ldots,$$

such numbers always stand to the right of the zero 0 – the number in an empty set. But how can we picture the *negative* integers, which stand to the left of the 0? Even if we take away all the cows, to get an empty field with 0 cows, we cannot imagine a 'more than empty' field!

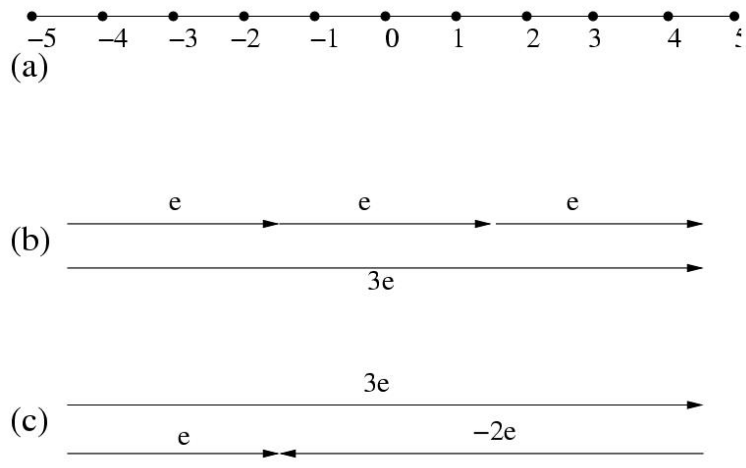


Figure 5

The way round this problem is to *represent* every number by a point on a line (the dots in Fig.5a), each with a label showing how many steps are needed to reach it, starting from the point labelled '0' – which is called the **origin** (often denoted by the capital letter 'O'), the line on which the points lie being called the **axis**. The points representing the positive integers are then reached by taking a whole number of steps from the origin to the right; those for the negative integers by taking a whole number of steps to the left. This reminds us of the way we measured distance (p.1) by counting 'strides' from one point to another. But now we are talking about steps 'to the right' or 'to the left', so both the length and the *direction* of the step are important. Let's use the symbol e to stand for one step to the right, $2\mathsf{e}$ for two, $3\mathsf{e}$ for three, and so on, noting that steps are *combined* by making them one after the other, starting always from the point already reached. A 'directed step' like e or $n\mathsf{e}$ (if you make n steps) is called a **vector**; it defines a *displacement* from one point to another. Vectors help us to make the connection between the ideas of *number* and *space* (the subject of Book 2). There is one vector for every dot on the right of the origin O in Fig.5a, the one for the dot labelled 3 being $3\mathsf{e} = \mathsf{e} + \mathsf{e} + \mathsf{e}$, shown in Fig.5b. In other words, there is a one-to-one correspondence (as explained in Chapter 2) between the numbers 1, 2, 3, ... and the vectors $1\mathsf{e}, 2\mathsf{e}, 3\mathsf{e}$, ... which can be used to point to them.

What about the numbers to the *left* of the origin in Fig.5a? If we use $\hat{\mathsf{e}}$ to mean one step *to the left* it is clear that

$$\mathsf{e} + \hat{\mathsf{e}} = \mathsf{0}, \tag{3.6}$$

where $\mathsf{0}$ is the 'step' from the origin that would leave you there (i.e. 'stand still!'). And it is clear that $-\mathsf{e}$ would also be a good name for 'one step to the left', for then we could say

$$\mathsf{e} + (-\mathsf{e}) = \mathsf{0}. \tag{3.7}$$

The negative integers $-1, -2, -3$, ... form a set whose elements are in one-to-one correspondence with the vectors $-1\mathsf{e}, -2\mathsf{e}, -3\mathsf{e}$, ... which point them out.

The laws which tell us how numbers are combined under addition also apply to the vectors representing displacements. In Fig.5c the displacements $3\mathsf{e}$ and $-2\mathsf{e}$ have a sum $3\mathsf{e} - 2\mathsf{e} = 1\mathsf{e} = (3-2)\mathsf{e}$; and in general if $\mathsf{a}, \mathsf{b}, \mathsf{c}$ are the three vectors for a steps, b steps, and the resultant c steps, then

$$\mathsf{a} + \mathsf{b} = \mathsf{c} \quad \text{means} \quad a\mathsf{e} + b\mathsf{e} = c\mathsf{e} = (a+b)\mathsf{e}. \tag{3.8}$$

The addition of displacements along the axis means *making them in succession* (i.e. one after another), while the addition of numbers refers only to the *counting* of objects in sets. But the rules of the game are the same: for example, $\mathsf{a} + \mathsf{b} = \mathsf{b} + \mathsf{a}$ – combining the displacements in a different order gives the same result – just as for the numbers a, b.

The pictorial way of representing numbers (Fig.5a) throws up still more questions. If the dots on a line in Fig.5a correspond to the negative and positive integers, then what about the points between them? Can they also be labelled by numbers? In the next Chapter we'll look for the answer.

3.3 More new numbers – fractions

First let's go back to the idea of multiplication of integers. We got the idea of *negative* integers by asking what x must be, given that $a + x = c$. Now let's ask a similar question when we are told that

$$a \times x = c. \tag{3.9}$$

If, say, $a = 5, c = 20$ then there *is* a solution, $x = 4$; because we know that $5 \times 4 = 20$. But what if $a = 5, c = 21$? There seems to be no solution, because we don't know of any integer x such that $5x = 21$.

Remember that in talking about addition we first invented a 'new number', 0 (zero), which could be added to any number (a, say) without changing it; and then we invented a new number $\hat{a}$ (for every a), as the solution of $a+x = 0$. That means $a+\hat{a} = 0$. Finally, we changed the name of $\hat{a}$ to $-a$ to get the set of *negative* integers, $-1, -2, -3, ...$ Can we do something similar for multiplication?

The zero, which can be added to any other number without changing it, has already been called an 'additive unit' or a 'unit under addition'. The name has nothing to do with the number 1 ('unity'); but 1 has a similar property under *multiplication* – it is a 'multiplicative unit', which can be used to multiply any other number without changing it.

Now let's use (3.8), putting $c = 1$ (the multiplicative unit), to define a new number $x = \bar{a}$ (called 'a-bar') with the property

$$a \times \bar{a} = 1 \tag{3.10}$$

– which looks like (3.3), but with 1 instead of 0 and $\times$ instead of $+$. The definition works fine for any number a, except $a = 0$ (which must be left out, as we'll see later).

To get a picture of what all this means, we go back to Fig.5a, looking at the part near the origin O. When a is a positive integer (3.10) tells us that, whatever number $\bar{a}$ may be,

$$(\bar{a} + \bar{a} + \bar{a} + ... + \bar{a}) \quad (a \text{ terms}) = 1,$$

so the number will be represented by a displacement which, repeated a times, will carry us from the origin to the point labelled 1. The number defined in this way is called a **fraction** and is usually written

$$\bar{a} = \frac{1}{a}$$

or, for short, $(1/a)$. For example, if $a = 2$ then $\bar{a} = \frac{1}{2}$ and is called 'half'. So two halves make one, $2 \times \frac{1}{2} = 1$. In Fig.5a, the number '$\frac{1}{2}$' would be represented by making a displacement of 'half a step to the right', going from the origin to the point labelled $\frac{1}{2}$; and two half steps, one after the other, would just reach the point labelled 1.

The number $\frac{1}{a}$ can also be written as $(1 \div a)$ or, in words, '1 *divided by a*', where the operation $\div$ is the inverse of $\times$: **division** is the inverse of multiplication. Thus for any number b,

$$b \times (a \times 1 \div a) = b \times 1 = b.$$

This is very similar to what we said about subtraction and addition, after equation (3.4).

Now we can go back to (3.9) and find a solution even when it is not a whole number. For we can multiply both sides of the equation $a \times x = b$ by the same number $1/a$, so the two sides will stay equal, to get the result

$$\text{If} \quad a \times x = b, \text{ then } \quad x = \frac{1}{a} \times b = b \times \frac{1}{a} = \frac{b}{a}. \tag{3.11}$$

The new number, defined in this way for *any* integers a, b (positive or negative) is called a **rational fraction**.

Now we have discovered rational fractions, which may be either positive or negative numbers (since the symbols a, b may each carry a $\pm$ sign), we have reached a milestone: the set of all rational fractions (which includes the integers whenever b divided by a gives exactly a whole number) is called the **rational number system**. Using the basic laws (2.1) to (2.5), and the rules that follow from them, to re-arrange or solve equations is called 'doing algebra'. The first known book on 'Algebra' was written by Al-Khwarizmi, of Baghdad, in the 9th century. It was translated from Arabic into Latin in about 1140 and was enormously important in the further development of Mathematics. The laying of these foundations was probably one of the Arab world's greatest gifts to humanity.

By introducing fractions we can solve quite complicated equations, extending the ideas we got in Section 3.1 as a result of introducing negative numbers. There we found a rule for getting an equation into a simpler form – but only when it involved addition and its inverse (subtraction). When $a + b = c + d$ we found that $a = c + d - b$ – so the b could be carried across the $=$ sign, provided we made it into $-b$. This was called "solving the equation for a". Now we need similar rules for dealing with multiplication and its inverse (division).

Suppose, for example, $a \times b = c \times d$. How can we separate a from the other numbers? Again, we can do the same thing to both sides of the equation and they will still be equal: so let's multiply both sides by $(1/b)$. The result is

$$a \times b \times \frac{1}{b} = c \times d \times \frac{1}{b}$$

or, since $b \times (1/b) = 1$ (which leaves unchanged anything that multiplies it)

$$a = c \times d \times \frac{1}{b} = c \times \frac{d}{b} = \frac{c \times d}{b}.$$

In other words, $\times b$ on the left of the $=$ sign can be carried over to the right, provided we make it into $(1/b)$: multiplying by something on one side of the $=$ corresponds to *dividing* by the 'something' on the other side.

As an example of using this rule let's find a simpler form for the sum of two rational fractions:

$$p = \frac{a}{b} + \frac{c}{d}.$$

We can multiply the first term on the right by (d/d) $(= 1)$ without changing it; and the second term by (b/b). And this gives

$$p = a \times \frac{1}{b} \times \frac{d}{d} + c \times \frac{1}{d} \times \frac{b}{b},$$

but, since the order of the factors in a product does not matter, this can be re-written as

$$p = \frac{a \times d}{b \times d} + \frac{c \times b}{b \times d} = \frac{a \times d + c \times b}{b \times d}.$$

Notice that the last step follows from the law (2.5), the common factor (the same in both terms) being $1/(b \times d)$. Usually, nowadays, the multiplication signs are not shown: when two symbols are shown side-by-side it is assumed they are multiplied. Thus ab means $a \times b$. With this shorter notation, which we use nearly always from now on, the last result becomes

$$p = \frac{a}{b} + \frac{c}{d} = \frac{ad + cb}{bd}.$$

The upper part of the fraction is called the **numerator**, the lower part the **denominator**; and the result was obtained by 'bringing the fractions (a/b and c/d) to a common denominator (bd)'.

Exercises

(1) Verify, by counting, that the basic rules (2.1), (2.2), (2.3), (2.4), (2.5), for combining numbers by addition and multiplication, are satisfied when $a = 5,\ b = 2\ c = 4$.

(2) Verify that this is still true when $a = -5,\ b = 3,\ c = -2$, using only the property of 0 and the definition of a negative integer as the inverse under addition of its positive partner.

(3) Think of the $-$ sign as an *instruction* – to reverse the direction of a vector, so it will point the opposite way. Then show, using pictures, that (3.5) is true also for vectors.

(4) Given that $\mathsf{a} - \mathsf{b} = \mathsf{c}$, what is a in terms of b and c? and how did you get your result?

(5) The vector $m\mathsf{v}$, with m an integer, is formed by taking m steps v one after the other. If $\mathsf{v} = v\mathsf{e}$ and $\mathsf{u} = u\mathsf{e}$, what is the vector $m(\mathsf{u} + \mathsf{v})$? What is the result when $v = 2,\ u = -3,\ m = 3$? Say *in words* what you have done.

(6) Answer the following questions, by bringing the fractions to a common denominator:

(a) $\frac{1}{2} + \frac{1}{3} =\ ?$ (b) $\frac{1}{4} + \frac{2}{3} =\ ?$ (c) $\frac{1}{2} - \frac{1}{3} =\ ?$

(d) $\frac{3}{4} - \frac{4}{7} =\ ?$ (e) $\frac{5}{4} - \frac{2}{3} =\ ?$ (f) $\frac{1}{2} - \frac{2}{3} =\ ?$

(g) $\frac{a}{2b} + \frac{a}{3b} =\ ?$ (h) $\frac{a}{b} + \frac{b}{a} =\ ?$ (i) $\frac{1}{b} - \frac{a}{3b} =\ ?$

Chapter 4

The decimal system

4.1 Rational fractions

In the *decimal* system, rational fractions such as

$$\frac{1}{10} \qquad \frac{1}{100} \qquad \frac{1}{1000} \; ....$$

have a special importance. To see why, let's return to pictures like Fig.5(a). We can think of $\frac{1}{10}$ as the length of a 'mini-step' $\frac{\mathsf{e}}{10}$ which would take us from the origin to the first point, labelled '0.1' and marked with a short vertical line (|) in Fig.6. By repeating this mini-step we arrive at the next | ; and so on. After ten such steps we arrive at the whole number 1; and this is the picture representing $10 \times \frac{1}{10} = 1$. This is all shown in Fig.6, which shows the interval 0 to 1 greatly magnified, along with vectors representing the 10 mini-steps, $\frac{\mathsf{e}}{10}$ which would take us to the point labelled '1'.

On going beyond 1, taking 11,12,13, ... mini-steps, we find first of all

$$\frac{11}{10} = 11 \times \frac{1}{10} = (10+1) \times \frac{1}{10} = 10 \times \frac{1}{10} + 1 \times \frac{1}{10} = 1 + \frac{1}{10},$$

and then the pattern repeats itself: every **interval** between two integers is marked off into 10 parts, shown by the short vertical lines – and every mark labels a new number.

There is no end to what we can do! If we define a 'mini-mini-step', such that there are ten of them to a mini-step and a hundred of them to a whole step, then we can put marks

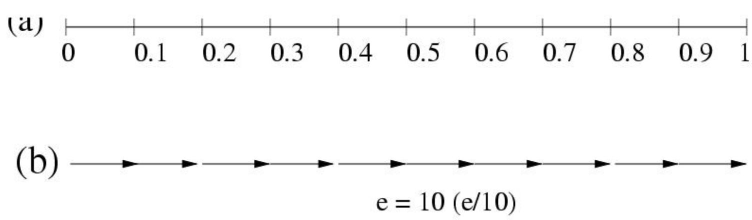


Figure 6

all the way along the axis (in both directions, left and right) and they get closer and closer together as we divide each step into smaller and smaller parts. At this point it seems possible that *all numbers* might be expressed as rational fractions. This, we shall find, is not so – even more new numbers remain to be discovered. But to understand the difficulty we must pause to look at decimal numbers in more detail.

The whole numbers (positive integers), which we talked about in Section 1.3, were constructed from the 10 symbols $1, 2, 3, 4, 5, 6, 7, 8, 9$ and 0. The number 10 is called the 'base' of the decimal system.

The numbers listed in Section 1.3 were collected in a Table, repeated below:

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

They were arrived at by starting from 1 and counting along the rows, left to right and one row after another, to reach any 'table entry'. For example, the 97th entry in the Table is the counted number 97: this means we have counted 9 complete rows of 10 (i.e. up to 90) and then 7 counts more, adding 1 for each count in the 10th row; thus we arrive at $97 = 9 \times 10 + 7 \times 1$. In the same way you will find

$$197 = 1 \times 100 + 9 \times 10 + 7 \times 1$$

where the first term is 100 ($= 10 \times 10$), reached after counting 10 complete rows of 10, and the remaining terms give the 97 already found. As a last example,

$$3107 = 3 \times 1000 + 1 \times 100 + 0 \times 10 + 7 \times 1$$

in terms of thousands ($1000 = 10 \times 10 \times 10$), hundreds ($100 = 10 \times 10$), tens (10) and units (1): this is a 'four-digit' number, the digits telling us the numbers of thousands, hundreds, tens, and ones it includes.

4.2 Powers and their properties

Terms like $(10 \times 10 \times 10)$ in the above representation of 3107 are called **powers** of 10. It is useful to write them all in the same way by agreeing that, for any number a, a^m will stand for $a \times a \times a \; ... \; \times a$ with m factors in the product (m being a positive integer). (Notice also that a can be a product, $a = bc$, and in that case $a^m = (bc)^m = b^m \times c^m$, as the order of the factors doesn't matter.) In representing numbers this way, a is called the 'base' and m the 'index'. It is easy to combine two powers of a number because, if a^n is a product with n factors, then

$$a^{(m+n)} = a^m \times a^n \tag{4.1}$$

(m factors of a, followed by n more). In other words, to multiply powers of a number we simply add the indices.

We can extend the 'law of indices' (4.1), making it apply also to *negative* powers, by noting that multiplication of a^m by $(1/a)$ has the effect of *taking away* a factor a – since $a \times (1/a) = 1$. So

$$a^m \times \frac{1}{a} = a^{(m-1)}.$$

This corresponds exactly to (4.1) with $n = -1$, which gives

$$a^{(m-1)} = a^m \times a^{(-1)},$$

provided we agree to *define* a negative power (until now meaningless!) by

$$a^{-1} = \frac{1}{a} \tag{4.2}$$

– so that a^{-1} becomes just another name for $(1/a)$. In the same way, a product of n such factors will be $a^{(-n)} = (1/a)^n$; and (4.1) will have a definite meaning for all integers m and n (positive or negative) – except when one of them is zero! What does a^0 mean?

If we multiply a $(= a^1)$ by a^{-1} the result is $a \times (1/a) = 1$. But, according to (4.1), with $m = 1$ and $n = -1$, this means

$$a^0 = 1. \tag{4.3}$$

In other words, *any number raised to the power zero gives 1.*

We need just one more rule about the powers of a number. When (later) we try to solve the equation $x^2 = a$, where a is any number and x is the unknown number we seek, we must ask for a new kind of *inverse* – the inverse of the *operation* of raising something to any given power. What is needed is another 'law of indices', not quite like the one given in (4.1).

The new rule will give the mth power of x^n; and when m and n are both positive integers it is clear that

$$(x^n)^m = x^{(mn)}, \tag{4.4}$$

because the left-hand side is a product of n xs *multiplied by itself* m times – which is a product $x \times x \times x \ldots \times x$ with mn factors. As usual, we insist that the rule should be true generally and try putting $m = 1/n$. In this case $(x^n)^{1/n} = x^1 = x$ and the rule tells us that:

$$\text{If } x^n = a \quad \text{then} \quad x = a^{1/n}. \tag{4.5}$$

In other words, the inverse of raising a number to the nth power is to raise it to the power $(1/n)$. In the special case $n = 2$, raising x to the power 2 is called 'squaring'; and the inverse operation of raising any x to the power $(1/n)$ is called 'taking the square root of

x' – often denoted by the symbol $\sqrt{x}$. Similarly, when the index n is greater than 2, we write $x^{(1/n)} = \sqrt[n]{x}$, which is called the nth root of x. These results are important in later Sections.

Now let's look back at the number 3107: we can write it as

$$3107 = 3 \times 10^3 + 1 \times 10^2 + 0 \times 10^1 + 7 \times 10^0$$

– the powers of 10 (from the greatest, 3, down to the smallest, 0), multiplied by the digits 3,1,0,7. And now that we know about negative powers we can extend the decimal representation to *all rational numbers*, including those with a fractional part – corresponding to points *between* the integers (as in Fig.6). Thus 3107.42 will be used to label a point 4 mini-steps and 2 mini-mini steps after the point with the integer label 3107: the 'decimal point' (.) simply separates the whole number on the left from the fractional part which follows it. In general, the six-digit number $rst.uvw$ will be

$$rst.uvw = r \times 10^2 + s \times 10^1 + t \times 10^0 + u \times 10^{-1} + v \times 10^{-2} + w \times 10^{-3},$$

where the fractional part is $u \times (1/10) + v \times (1/100) + w \times (1/1000)$.

4.3 Decimal numbers that never end

At the end of Section 3.2 we were wondering if *every* number, labelling a point on the axis in a picture such as Fig.5(a), was expressible as a rational fraction of the form p/q – with a big enough denominator q. Thus, for example, 1.414= 1414/1000 is the rational fraction corresponding to the decimal number in which the terms 1+(4/10)+(1/100)+(4/1000) have been brought to the common denominator 1000. In the decimal system the common denominator is always a power of 10, but this is not always the case. (A pocket calculator, for example, works by using electrical switches; and then the natural base is 2, not 10, corresponding to switch 'on' or switch 'off'.)

Let's ask another question: 1/9 is a very simple-looking rational number – how can we express it in the decimal system? We can write it as follows:

$$\frac{1}{9} = \frac{10}{9}\frac{1}{10} = \frac{9+1}{9} \times \frac{1}{10} = \left(1 + \frac{1}{9}\right) \times \frac{1}{10} = 0.1 + \frac{1}{9} \times 0.1.$$

So, very roughly, the answer is 0.1, but there is another term and this is just 0.1 times the original fraction; and if we do the division in this term we again find 0.1 (plus a 'remainder' of 1/9) but multiplied by 0.1, giving 0.01. When this term is added to the first, we get

$$\frac{1}{9} \approx 0.11$$

and if we continue the process (which is what you did a long time ago when learning to do 'long division sums') the result will be

$$\frac{1}{9} \approx 0.1111111111... ,$$

where the decimal number will go on and on forever; the last digit is always 1 and there is no way of stopping!

What, then, does it mean to say (1/9) = $0.11111^{\cdot}$ (where the upper dot means that the digits after the decimal point go on repeating themselves, or 'recurring')? Written as a sum of decimal fractions, we can say

$$\frac{1}{9} = \frac{1}{10} + \frac{1}{10^2} + \frac{1}{10^3} + \ ... \ + \frac{1}{10^n} + \ ...$$

which is the sum of an *infinite* number of smaller and smaller terms. A sum like this is called a **series** and if we stop after n terms, throwing away those that follow, we get an *approximation* to the rational fraction (1/9). The last term in the approximation then improves the previous approximation (with only $n-1$ terms) by 10^{-n} – a small correction. The = sign in the expression simply means that, as we take more and more terms we shall get closer and closer to the number (1/9) we are looking for. Looking at the approximations, for $n = 1, 2, 3, \ ... \ ,$ it is clear that $(1/9) > 0.1$ but < 0.2 (to one decimal place); but it is > 0.11 and < 0.12 (to two decimal places); and > 0.111 but < 0.112 (to three decimal places); and so on. The numbers like 0.111 and 0.112 are called **bounds**, the first being a 'lower' bound and the second an 'upper' bound. The upper and lower bounds mark the boundaries of an *interval* in which we know the number must lie.

But looking for the number (1/9) is like hunting a very small, and very slippery animal; even if we know it can be represented by a point on a line between, for example, the bounds 0.111111 and 0.111112, we still can't get hold of it! By going further we can say it must be somewhere between 0.1111111 and 0.1111112, but that still doesn't pin it down *exactly*. To be found in that last interval it must be very small, less than 0.0000001 of a step wide (for otherwise it couldn't get in!). In fact, the number we are looking for is represented by a *point*, which has zero width, so however small we make the interval there is room in it for millions and millions of numbers besides the one we want! The set of intervals, every one enclosing all those that follow it, is called a **nest** of intervals; and any number is thus defined, as accurately as we please, by giving a recipe for finding the nest in which it lives! Most numbers can't be expressed as either integers or rational fractions; they are said to be **irrational** and such a number is generally expressed as the sum of an infinite set of terms. Simple examples of the irrational numbers can arise as the 'roots' obtained by solving the equation (4.5). When $x^n = a$ the solution $x = a^{1/n}$ is called the 'nth root' of a. Such numbers are in general called **surds**.

Square roots are often needed and are quite hard to calculate by the usual 'schoolbook' methods (much worse than long division!) but if you can multiply and divide you can easily get very good approximations. Take $x = \sqrt{2}$, for example, where you expect the root to be greater than 1 (since $1^2 = 1$ and is too small), but less than 2 (since $2^2 = 4$ and is too big). Try 1.5, half way between. If you divide 2 by 1.5 you get $1.333^{\cdot}$; so $1.5 \times 1.333^{\cdot} = 2$ and the number you want (x, say, such that $x \times x = 2$) will be between $1.333^{\cdot}$ and 1.5. Half way between is $x \approx 1.416$; and if you go on you soon find $x \approx 1.41421356$. This is the number the Greeks were looking for and couldn't find!

The number set which includes all integers (including the sign ± and the Zero), the rational fractions, and all the irrational numbers, is called the **Real Number System**.

When scientists talk about 'ordinary numbers' this is the system they have in mind. We say just a bit more about it in the next Chapter.

Exercises

(1) Express the rational fractions 1/3, 2/3, 1/5, 1/7 in decimal form

(2) Use the method given on the last page of this Chapter to find $\sqrt{3}$, $\sqrt{5}$ and $\sqrt{7}$, accurate to 3 decimal places (i.e. to three figures after the point).

(3) Use the roots you have already found to get $\sqrt{6}, \sqrt{21}, \sqrt{63}$, again to 3 decimal places. (Use the fact that $(ab)^m = a^m \times b^m$.)

(4) Try to find a method (similar to the one used above) for getting the *cube* root of a number and use it to get $\sqrt[3]{25}$. (The number you want must be a bit less than 3, because $3^3 = 27$)

(5) Use the results in Section 4.2 to get simpler forms of the following:

$$a^2a^{-3}, \qquad a^3/a^{-4}, \qquad a^3/a^4, \qquad (a^2)^{1/2},$$

$$(a^3)^{1/2}, \qquad \sqrt{(a^3)}, \qquad \sqrt[3]{(a^2)}, \qquad \sqrt[3]{(a^3)}$$

(6) Write each of the numbers 295 and 3106 as a sum of powers of 10. Then multiply them together, using the laws of indices, and write the result in the usual (decimal) form.

(7) Do the same as in the last Exercise, but with the numbers 26.32 and 3.156.

(8) Suppose you have only *two* different symbols, 1 and 0 (instead of the usual ten digits (1,2,... 9,0). Try to express the numbers 7 and 17 in terms of powers of 2. (Remember that $2^0 = 1$, $2^1 = 2$, $2^2 = 4$, $2^3 = 8$, etc. and do as you did in Exercise (2).) Numbers written in this way are said to be expressed in **binary** form. For example, the binary number 10110 means $1 \times 2^5 + 0 \times 2^4 + 1 \times 2^3 + 1 \times 2^2 + 0 \times 2^10 \times 2^0$, which is $32 + 8 + 4$ or 48 in the decimal system.

(9) What is your age (in years) as a binary number?

(10) What does the binary number $abc.rst$ (where a, b, c, r, s, t can take values only 0 or 1) mean in the decimal system?

(11) Multiply the numbers 7 and 17, found in binary form in Exercise (8), by using the laws of indices. Put the result in decimal form.

(12) Multiply the binary numbers 101.11 and 110.01 and express the result as a decimal number. (Look back at your answers to Exercises (7) and (8).)

Chapter 5

Real and complex numbers

5.1 Real numbers and series

The set of all the numbers found so far (negative and positive integers, including zero, and both rational and irrational numbers) is called the real number **field**. Its elements, the numbers themselves, can be combined by addition (according to the laws (2.1) and (2.2)) and by multiplication (according to (2.3) to (2.5) – which make it a ring (see p.12) – and it becomes a *field* when we include along with every number a (apart from zero) an *inverse* $(1/a)$ such that $a \times (1/a) = 1$.

It's taken us a long time to get to the idea of the real number field – even if we knew how to count at the beginning. But we should remember that it took people thousands of years to get so far. Along the way, there were many hold-ups which seemed to put a stop to all progress. The Greeks, for example, more than 2000 years ago laid the foundations of geometry, using *pictures* rather than numbers to express their ideas; and when they couldn't find any number x such that $x^2 = 2$ (which measures the *area* of a square with sides of length x) they stopped looking – thinking that geometrical quantities could not be exactly represented by numbers. It was for this reason that algebra (the science of number) and geometry (the science of space) were developed by the Greeks in entirely different ways. And it was almost 2000 years later that the idea of an *irrational* number, defined in the last Section as the limit of a set of better and better approximations, broke down the wall between the two.

There are many interesting stories about the school of Pythagoras: The Pythagoreans almost certainly *did* invent 'algebraic geometry'; but their religious belief, which told them that Nature could only be understood in terms of the numbers given by God, was so strong that they swore to keep their discovery secret! We'll come back to algebraic geometry in Book 2, where we talk about Space. But for now, so as to understand real numbers better, let's look again at the difference between rational fractions such as p/q (p, q both integers) and an irrational number. The rational fraction (1/9) was written in the last Chapter as a sum of terms:

$$\frac{1}{9} = a_1 + a_2 + a_3 + \ldots + a_n + \ldots$$

where the nth term is $a_n = 10^{-n}$, the **subscript** 'n' simply showing *which* term has this value. The terms form a **sequence** $a_1, a_2, a_3, ...$ and their sum is called a **series**. If we keep only the first n terms, the series gives an n-term approximation to (1/9); to get the exact value, in decimal form, we must go on forever, getting the 'sum to infinity'. The series is said to **converge** to the **limit** (1/9) when the sum to n terms gets as close to the limit as we please for a large enough value of n. This is clearly the case for a number represented as a recurring decimal.

Another series, which may or may not converge, follows when we take $a_n = x^{n-1}$, x being any chosen number. The sequence is then

$$a_1 = x^0 = 1,\ a_2 = x^1 = x,\ a_3 = x^2,\ a_4 = x^3,\ ...$$

and the series is the sum

$$S = 1 + x + x^2 + x^3 + \dots\ .$$

Thus, if we put $x = 1/2$, we get

$$S = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \dots\ .$$

To see if the series converges let's look at the sum to n terms. This will be

$$S_n = 1 + x + x^2 + x^3 + \dots + x^{n-1},$$

which can easily be found exactly. We multiply S_n by x, getting

$$xS_n = x + x^2 + x^3 + \dots + x^n,$$

and subtract this result from S_n in the line before it. In this way we get

$$S_n - xS_n = 1 + (x - x) + (x^2 - x^2) + \dots - x^n = 1 - x^n,$$

as all the terms between the first and the last cancel in pairs to give zero. So $S_n(1-x) = 1 - x^n$ and, dividing both sides of this equation by $(1 - x)$,

$$S_n = \frac{1 - x^n}{1 - x}. \tag{5.1}$$

If x is any rational number, S_n is also a rational number: with $x = (1/2)$, for example, $S_n = 2(1 - 2^{-n})$. The interesting question now is: Does the series converge when we let $n \to \infty$ (read as 'n tends to infinity' i.e. becomes as large as we please)? And the answer is clearly yes – because the term 2^{-n} in the numerator then becomes vanishingly small. There is thus a limit

$$S_\infty = \frac{1}{1 - x}, \tag{5.2}$$

which has the integer value 2 when $x = 1/2$. Even an integer can be represented as an infinite series! In fact

$$2 = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \dots\ .$$

On the other hand, for $x > 1$ the series does *not* converge: if you put $x = 2$, for example, S_n given in (5.1) goes towards larger and larger values, without limit, as n becomes large.

Last of all, let's look at a series which does not correspond to anything simple – neither an integer nor a rational fraction – even though the series itself does not look complicated. The series is an expression for a very important *irrational* number, denoted by the symbol e^x, for any value of x:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots , \tag{5.3}$$

where $n!$ is the number defined in Section 1.2 – the product of all the integers from 1 up to n. The number e which (raised to the power x) appears in (5.3) is obtained by putting $x = 1$. The result is

$$e = 1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \ \dots \quad = 2.718281828\dots \tag{5.4}$$

and this is one of the most important numbers in the whole of mathematics!

Much more on the use of series will be found in other Books of the Series. But here we'll finish with one last generalization of the number system.

5.2 The field of complex numbers

The real number field is fine for describing anything we may want to measure: every physical quantity is a certain *number* of units (such as the 'paces' used on p.1 to measure distances) and that number, which is the *measure* of the quantity, belongs to the real number field. But mathematics goes a long way beyond that: we invented new kinds of number by trying to answer questions that seemed to have no answer – finding first the negative numbers, then the rational fractions, and finally the irrational numbers. And there is still one question we want to ask: What is the number x whose square, x^2, has a *negative* value such as -1? In other words, what value of x makes $x^2 = -1$? This question was first asked, it seems, by an Italian mathematician, Cardano, in a book he wrote in 1545; but it was much later (1777) that the famous Euler gave the solution the name i, which is still used. Thus $x = i$ is *defined* by its property

$$i^2 = -1. \tag{5.5}$$

This number is called the 'imaginary' unit, a term apparently first used by Descartes to distinguish it from the real numbers $+1$ and -1 which both have the square $(+1)^2 = (-1)^2 = 1$. Here, in contrast, $(+i)^2 = (-i)^2 = -1$.

Now we have the new number i we treat it like any other number, supposing it satisfies the same conditions (2.1), (2.2), etc: so any real number a will have a new companion ai and by (2.3) this will be the same as ia. Numbers which contain the symbol i are called **complex numbers**. Thus, $z = x + iy$, in which x and y are ordinary *real* numbers, is a *complex* number; and all numbers of this form belong to the *complex number field.* This field includes also the real numbers, which are of the form $x + iy$ with $y = 0$, and it is *closed* under the operations of both addition and multiplication: by combining any two complex numbers (let's call them z_1 and z_2, with subscripts giving them their full names) we find, on addition,

$$z_1 + z_2 = x_1 + iy_1 + x_2 + iy_2 = (x_1 + x_2) + i(y_1 + y_2), \tag{5.6}$$

which is just another complex number. And on multiplication, using the rules for re-arranging the result (as in Section 1.7), we find

$$z_1 \times z_2 = (x_1 + iy_1) \times (x_2 + iy_2) = (x_1x_2 - y_1y_2) + i(x_1y_2 + y_1x_2). \tag{5.7}$$

This is again simply another complex number, of the form $a + ib$, with a 'real part' $a = (x_1x_2 - y_1y_2)$ and an 'imaginary part' ib with $b = (x_1y_2 + y_1x_2)$. In short, combining complex numbers always leads us to other numbers of the same complex number field – never to anything *new*. Note that there are two special complex numbers, $0 = 0 + i0$ and $1 = 1 + i0$, which are the 'unit under addition' and the 'unit under multiplication' (as in earlier Sections, where there was no imaginary part); and that every complex number $z = x + iy$ has a 'partner' $z^* = x - iy$, obtained by putting $-i$ in place of i, which is called the *complex conjugate* of z and is indicated by the 'star'. From every complex number we also find a *real* number

$$zz^* = (x + iy)(x - iy) = x^2 - iyx + iyx - i^2y^2 = x^2 + y^2, \tag{5.8}$$

which is called the **square modulus** of z and is written $|z|^2$: thus $|z| = |z^*| = \sqrt{x^2 + y^2}$. With the invention of i, we can stop looking for new numbers – there seem to be no more questions we can ask that can't be answered in terms of numbers belonging to the complex field. Remember – we had to invent negative numbers to get an answer to a question:

> If $x + a = 0$ (a being a positive integer), then what is x? The solution of this equation is the negative integer $x = -a$. And since 0 is a 'unit under addition' (it can be added to any number without changing it) this showed that $-a$ was an 'inverse under addition' of a.

Similarly, the number 1 is the 'unit under multiplication' and to invent fractions we started from another question:

> If $xa = 1$ (a being a positive integer), then what is x? The solution is $x = 1/a$, which is the inverse under multiplication of a. On using the notation in (4.2) this can also be written as $x = a^{-1}$; and, in the same way, $xa^n = 1$ has the solution $x = a^{-n}$, which is the inverse of a^n since $a^na^{-n} = a^0 = 1$.

The solutions to these questions were first obtained for the real number field. But, by *defining* the number i as a solution of the equation $x^2 = -1$ (which we didn't know how to solve), we can now get answers for all numbers in the *complex* number field. For example, if $z = x + iy$ is any complex number it will have an inverse z^{-1}, with real and complex parts a and ib, provided we can satisfy the equation

$$z\,z^{-1} = (x + iy)(a + ib) = 1 + i0.$$

But two complex numbers will be equal only if their real and complex parts are *separately* equal; and (doing the multiplication) that means

$$xa - yb = 1 \quad \text{(real parts)} \quad \text{and} \quad xb + ya = 0 \quad \text{(complex parts)}.$$

The second equation tells us that we must choose $b = -ya/x$; and we can put this in the first equation to get $xa - (-y^2a/x) = 1$ or, multiplying both sides by x, $(x^2 + y^2)a = x$. So we find the solution

$$z^{-1} = a + ib, \quad \text{with} \quad a = \frac{x}{x^2 + y^2}, \quad b = \frac{-y}{x^2 + y^2}.$$

What we have shown is that an equation which involves only one unknown x $(= x^1)$ (with no powers such as x^2) has a solution provided we let x take *complex* values. Such equations are said to be of the 'first degree' (or 'linear'): those which involve x^2 are of the 'second degree' (or 'quadratic'). A more general equation is

$$a_0 + a_1x + a_2x^2 + a_3x^3 + \ldots + a_nx^n = 0, \tag{5.9}$$

in which the highest power of x that appears is the positive integer n and the numbers $a_0, a_1, \ldots a_n$ are called the **coefficients** of the powers of x: (5.9) is called an algebraic equation of the 'nth degree'. The expression on the left of the = is an nth degree **polynomial** ('poly' means 'many' and 'nomial' means 'term'). Thus, x^2+1 is a second degree polynomial and $x^2 + 1 = 0$ is a quadratic equation, with solution $x = \pm i = \pm(-1)^{1/2} = \pm\sqrt{-1}$. Let's now turn to something more general.

5.3 Equations with complex solutions

To see how other simple equations may have solutions which contain i, it's enough to look at the algebraic equation of degree 2, writing it

$$ax^2 + bx + c = 0, \tag{5.10}$$

where the coefficients a, b, c are real numbers and we want to find the unknown number x. The first step in getting a solution is to divide all terms by a and then add $-c/a$ on both sides of the = sign to obtain

$$x^2 + \frac{b}{a}x = -\frac{c}{a}. \tag{5.11}$$

If we had (something)^2 on the left we could say that (something)=$\sqrt{(-c/a)}$, as we did in solving $x^2 = -1$; but our 'something' is more complicated. However, noting that

$$(x + r)^2 = (x + r)(x + r) = x^2 + xr + rx + r^2 = x^2 + 2rx + r^2,$$

we can choose the number r as $r = \frac{1}{2}(b/a)$, which gives

$$(x + r)^2 = x^2 + \frac{b}{a}x + \left(\frac{b}{2a}\right)^2.$$

And now we can re-write (5.11), adding $(b/2a)^2$ to both sides, in the form

$$(x+r)^2 = \left(\frac{b}{2a}\right)^2 - \frac{c}{a} = \frac{b^2-4ac}{4a^2}$$

– so we can get the solution (for $x+r$, instead of x) just by taking the square root of the quantity on the right. Thus, remembering that $r = b/(2a)$ and bringing the terms on the right to a common denominator,

$$x + b/(2a) = \pm\sqrt{b^2-4ac}/(2a).$$

The solution of (5.10) follows at once and is usually written in the form

$$x = \frac{-b \pm \sqrt{b^2-4ac}}{2a}. \qquad (5.12)$$

This is a very important result: it shows that any algebraic equation of degree 2, whatever values we give to the coefficients a, b, c, must have just 2 solutions, $x = r_1$ and $x = r_2$. The two values, r_1 and r_2, arise on taking the upper and lower signs in (5.12) and are called the **roots** of equation (5.10).

Now let's go back to the general algebraic equation (5.9). The last question we are going to ask is: for what values of x is it satisfied? Remember that we are now working in the field of *complex* numbers, so x and all the coefficients $a_1, a_2, \dots, a_n$ may be complex. This is the 'last question' because the great German mathematician Carl Friedrich Gauss was able to prove that, when complex numbers are admitted, (5.9) has exactly n solutions $x = r_1,\ r_2\ \dots, r_n$: any of these n numbers, the roots, will satisfy the equation. This result is usually called the Fundamental Theorem of Algebra.

We are at the end of the road! All algebraic equations have solutions in the complex number field, so there is no *need* to look for new kinds of number. In the next (and last) Section, however, we'll find other things – *not numbers* – which can be described by using symbols of a new kind, leading us to new games with new rules. When we go *outside* the number field we find the road still goes on!

Exercises

(1) Look at the sequence of terms

$$a, \quad a+d, \quad a+2d, \quad a+3d, ... \ a+nd,$$

where the first term is a and every other term arises by adding the 'difference' d to the one before it. What is the series S_n formed from the first n terms?

(2) Show, by writing the terms in reverse order, and then adding the two series together, term by term, that

$$S_n = \frac{n}{2}[2a + (n-2)d].$$

(This is an n-term 'arithmetic series')

(3) What is the sum of the first n natural numbers, $(1+2+3+...+n)$? From your formula, get the sum of the numbers from 1 to 1000.

(4) The n-term geometric series

$$A_n = 1 + x + x^2 + x^3 + \dots + x^{n-1},$$

has a sum given in (5.1) and converges to a finite value when x is between 0 and 1. If you are given another series,

$$B_n = 1 + \frac{x}{2} + \frac{x^2}{3} + \dots + \frac{x^{n-1}}{n-1},$$

what can you say about it? Does it converge for $n \to \infty$? (Compare corresponding terms to see if one is always less than the other.)

(5) Combine by addition the following pairs of complex numbers:

$(2 + 3i)$ and $(3 - 5i)$; $(2 + 3i)$ and $(3 - 2i)$; $(2 - 3i)$ and $-(2 + 3i)$.

(6) Combine by multiplication the same pairs as in Exercise (1).

(7) Get a formula for finding the product of $z_1 = a + ib$ and $z_2 = c + id$, a, b, c, d being any real numbers.

(8) Write down the complex conjugate (z^*) of each of the numbers

$z = 3 + 2i$, $\quad z = 2 - 3i$, $\quad z = 3/(2 + i)$.

(9) Find the results of the following divisions

$$\frac{2+3i}{3-5i}, \qquad \frac{2+3i}{3-2i}, \qquad \frac{2-3i}{2+3i}.$$

(Note that any of the denominators multiplied by its complex conjugate gives a *real* number.)

(10) Use the formula (5.12) to find the roots (there are two in each case) of the equations

$$x^2 + 3x - 1 = 0; \qquad x^2 + 4x - 4 = 0; \qquad 2x^2 + 5x + 2 = 0; \qquad (x^2/4) - 3x + 25 = 0.$$

Call the two roots r_1 and r_2 in each case and show that the quadratic equations can all be written in the *factorized form* $(x - r_1)(x - r_2) = 0$.

Chapter 6

Beyond numbers – operators

6.1 Symmetry and groups

All around us we see objects which have **shapes**: sometimes the shapes please us and we say the object is beautiful; and very often the beauty is related to **symmetry**. Think of a human face where every feature on the right exactly matches one on the left; or a royal palace where for every window on the right there is one exactly like it on the left; or a flower with five petals, each one exactly like the others. At first sight, shapes seem to have little to do with counting or with symbols of any kind. But by going beyond numbers, and inventing new symbols, we find even shapes can sometimes be described in mathematical ways.

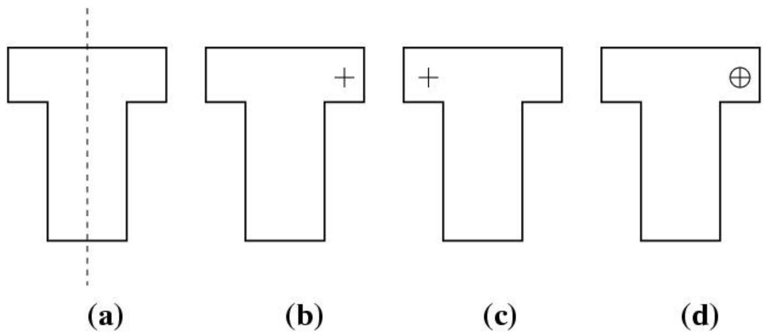


Figure 7

Let's take a very simple example: the capital letter T in Fig.7(a), which you can cut out of a piece of stiff white card, so you can turn it around and see what happens to it. The symmetry of the card depends on the way it responds to certain **operations**. If we could reflect it across the vertical line in Fig.7(a), the left- and right-hand 'arms' would simply change places and the card *after the operation* would look exactly as it did before. Think of the card as being in two different 'states', before and after the operation, calling them S (before) and S' (after) – noting that S and S' are *not numbers* but just the names of the two states. What makes the operation a *symmetry operation* is the fact that, as far as we can tell (supposing the card has no marks on it to spoil its symmetry) nothing has changed.

In words, we might say

> The new state of the card, after reflecting it across the vertical line, is no different (as far as we can see) from the state we started from.

But in symbols we can say instead simply

$$S' = \mathsf{R}S = S$$

where R denotes the **operator** that causes "reflection across a plane cutting the T into two equal halves " – interchanging left and right: the special 'typeface' (R) reminds us that the operator is, again, *not a number*. The = sign, as usual, means there is no difference (as far as we can tell) between the quantities it separates: they are 'equal' or *equivalent*.

The operation just described is not something you can actually *do*, for that would mean taking every bit of card from one side to the other and then putting them all together again! The important thing is that we can *imagine* the operation of reflection; and the card would seem to be unchanged by it.

Now, if a symmetry operation seems to make no change, how do we know what's happening? We can use a trick: let's put a small 'secret' mark, + say, at the top right-hand corner of the card as in Fig.7(b), so as to see what's going on. This will not be counted as part of the card and will not spoil its symmetry, but now we can 'keep track' of all the operations we make. The reflection operation R, for example, will send the mark to the top *left-hand* corner – as in Fig.7(c); and if we now use S^+ and ${}^+S$ (instead of S and S') for the states before and after the reflection we can say

$$\mathsf{R}S^+ = {}^+S = S^+.$$

Another symmetry operation will be reflection across the plane of the card – which interchanges back and front: let's call this operation R'. This will send our mark to the *underneath* side of the card, directly below the original mark + on the top side. In Fig.7(d) the moved mark is shown in a circle, like this $\oplus$, to mean it's really underneath. We'll need a new symbol for this new state; and if we use $S^\oplus$ then we can say

$$\mathsf{R}'S^+ = S^\oplus = S^+,$$

where the state after the reflection is again shown as equal to the one before (not counting the mark), showing that R' is also a *symmetry* operation.

The results of the two operations, R and R', are thus

$$\mathsf{R}S^+ = \ {}^+S \quad (A), \qquad \mathsf{R}'S^+ = S^\oplus \quad (B).$$

Now let's try *combining* operations, just as we did earlier in dealing with numbers, to find their 'laws of combination'. And remember that doing exactly the same thing to both sides of an equation is sure to leave them equal. For example, the two reflection operations, R and R', can each be applied twice; and using (A) and (B) above we find:

$$\mathsf{RR}S^{+} = \mathsf{R}\ {}^{+}S = S^{+},$$

$$\mathsf{R'R'}S^{+} = \mathsf{R'}S^{\oplus} = S^{+}.$$

What does all this mean? It simply means that a reflection, followed again by the same reflection, leaves the card exactly as it was in the beginning. Doing either of them twice is the same as doing nothing! In other words, they are equivalent to the **identity** or **unit** operator, usually called E when we're talking about symmetry, which is an instruction to leave the card alone. In this example a symbol such as S^{+} stands for the state of the card, but it doesn't matter which state we start from (as you can easily check): the operators have a life of their own! We then drop the state symbol and say

$$\mathsf{RR} = \mathsf{E}, \quad \mathsf{R'R'} = \mathsf{E}. \tag{6.1}$$

This is all beginning to look like what we did in inventing new numbers in Chapter 3; so we'll use the same kind of language. If the 'product' of two operators is equivalent to the unit operator, each is said to be the 'inverse' of the other: so here both operators are said to be 'self-inverse'.

Most of this book has been about 'ordinary numbers' (even when we use letters to stand for them) and belongs to what is nowadays called "elementary algebra". But now we're moving on to 'higher' algebra, which includes various kinds of "abstract algebra".

There are important differences between the algebra of ordinary numbers and the algebra of operators. For example, there is only *one* number (1) which is its own inverse ($1\times1 = 1$); but here we have *two* different reflection operations. Another big difference is that the *order* in which the operations are made can be important, as you will find in some of the Exercises.

Next, let's apply R and R' *one after the other.* This will give us another operation,

$$\mathsf{RR'}S^{+} = \mathsf{R}S^{\oplus} = {}^{\oplus}S$$

and $\mathsf{R'R}S^{+}$ gives exactly the same result! But the new state ${}^{\oplus}S$ could have been obtained by a *single operation*: if you *rotate* the card (in state S^{+}), through half a turn around the centre line of the T, you'll find the + has moved over from the right-hand arm to the left – but *behind* the card. The operator for rotation through half a turn is usually denoted by C_2, where the subscript 2 means the operator has to be applied *twice* to get a rotation through a full circle – which clearly returns the card to its original state. So, like R and R', C_2 is a self-inverse operator: $\mathsf{C}_2\mathsf{C}_2 = \mathsf{E}$.

Now we know that two different reflections, one after the other, can give the same result as a single rotation let's try combining one of each – a reflection and a rotation. If we apply first C_2 and then R, starting from state S^{+} and noting that the 'first' operation is the one nearest the state symbol (so you must read them in right-to-left order), then we get

$$\mathsf{RC}_2S^{+} = \mathsf{R}\ {}^{\oplus}S = S^{\oplus}.$$

And if we use R first and then C_2, we find the same result,

$$\mathsf{C}_2\mathsf{R}_v S^+ = \mathsf{C}_2\ {}^+S = S^{\oplus}.$$

What have we done so far? We have found four **symmetry operators**

$$\mathsf{E} \qquad \mathsf{C}_2 \qquad \mathsf{R} \qquad \mathsf{R}', \tag{6.2}$$

which, when applied to the letter T, leave it unchanged or to use the technical term 'invariant'. They describe the symmetry of the object; but they are related not only to one particular object (the card in the shape of the letter T) – just as the operation of counting is not related to what is being counted. The properties of the four operators can be studied without talking about what they are 'working on': C_2 applied twice over is equivalent to E, i.e. "do nothing" – *whatever it is that is being rotated.* The four operators together form a **group** $\{\mathsf{E}\ \mathsf{C}_2\ \mathsf{R}\ \mathsf{R}'\}$ in which each symbol stands for a certain kind of geometrical operation on an object. This particular group is called a **point group** because all its operations leave *at least one point* of the object unmoved.

Mathematicians go even further: they don't even want the pictures. They can think of the four symbols as the 'elements' of an **abstract group** – just a set of elements with certain 'laws of combination' – and that's just about as abstract as you can get. They insist only that any group must have the following properties: any two elements may be combined (in the symmetry group they were operations, applied one after another) to give a third, which is 'equivalent' to the 'product' of the first two. The group must be 'closed' (however we combine the four symbols we can never get anything new); it must contain an identity or unit element E; and for every element (R, say) there must be an *inverse* (often written R^{-1}, just as in 'ordinary' algebra), such that $\mathsf{R}\mathsf{R}^{-1} = \mathsf{E}$.

It seems that in going from elementary algebra to abstract algebras we've left ordinary numbers far behind. But the field of numbers, both real and complex, almost always comes in. For example, we can even put the elements of an abstract group together, with coefficients taken from the complex number field, to get the elements of an abstract **group algebra**: for a group $\{\mathsf{A}\ \mathsf{B}\ \mathsf{C}\ \mathsf{D}\}$, the combination

$$\mathsf{X} = a\mathsf{A} + b\mathsf{B} + c\mathsf{C} + d\mathsf{D},$$

where a, b, c, d are any complex numbers, is an element of the algebra of the group.

We won't worry about the details: it's enough to know that what we have done leads into one of the most important areas of mathematics – the theory of groups – with hundreds of applications in many parts of science.

6.2 Sorting things into categories

Suppose we go to a market where they sell animals, for example, asses (for carrying sacks of vegetables), bullocks (for pulling ploughs or heavy carts), cows (for milking), and donkeys (for children to ride). Each kind of animal belongs to a 'category'; and we can use a letter to stand for an animal in any category, a for an ass (it doesn't matter which one), b for a bullock (no matter which), and so on.

In words, we might want to say "The market contains 6 asses, 4 bullocks, 8 cows, and 5 donkeys". This is the 'state' of the market on that particular day (let's call it S, just as we did in talking about the T-shaped card in the last Section). Then we can say all this in symbols by writing

$$S = 6\mathsf{a} + 4\mathsf{b} + 8\mathsf{c} + 5\mathsf{d}.$$

The idea of sorting the animals into categories can now be put in terms of symbols by using, let's say, the capital letter A to stand for the *operation* of keeping only the asses, moving the other animals somewhere else – 'off limits' or 'not for sale'. The result of the operation is to get a new state of the market, S', in which the only animals on sale are asses. If we write

$$S' = \mathsf{A}S = 6\mathsf{a},$$

then S' is a 'sample' containing only asses. If I'm looking only for asses, that's fine – there are six of them.

Similarly, we could use B for keeping only the bullocks, and if I'm looking for bullocks then $\mathsf{B}S = 4\mathsf{b}$ – and again I'm happy, there are four of them.

But each of the samples is *pure*, containing only the animals in one category: there are no asses in a sample of bullocks. In symbols this reads

$$\mathsf{A}(\mathsf{B}S) = \mathsf{A}4\mathsf{b} = 0$$

and in the same way we won't find bullocks among the asses:

$$\mathsf{B}(\mathsf{A}S) = \mathsf{B}6\mathsf{a} = 0.$$

However, operating once again with A on the sample $\mathsf{A}S = 6\mathsf{a}$ just confirms that it contains six asses:

$$\mathsf{A}(\mathsf{A}S) = \mathsf{A}(6\mathsf{a}) = 6\mathsf{a} = \mathsf{A}S.$$

Putting it all together,

$\mathsf{AA}S = \mathsf{A}S \quad (= 6\mathsf{a})$ (still 6 asses in a sample of 6 asses)

$\mathsf{AB}S = 0$ (no asses in a sample of bullocks)

$\mathsf{BA}S = 0$ (no bullocks in a sample of asses)

– and so on for all the animals!

The important things to notice are that

$$\mathsf{AA}S = \mathsf{A}S, \quad \mathsf{BB}S = \mathsf{B}S, \quad \mathsf{CC}S = \mathsf{C}S \; ....$$

and that

$$\mathsf{AB}S = \mathsf{BA}S = \mathsf{AC}S = \mathsf{CA}S, ... = \mathsf{BC}S = \mathsf{CB}S, ... = 0$$

where 0 stands for the market in the state with *no* animals,

$$0 = 0\mathsf{a} + 0\mathsf{b} + 0\mathsf{c} + 0\mathsf{d}.$$

Also, by adding all the samples (the animals in the four categories) we put the market together again, in its original state – with 4 asses, 6 bullocks, 8 cows, and 5 donkeys:

$$\mathsf{A}S + \mathsf{B}S + \mathsf{C}S + \mathsf{D}S = (\mathsf{A} + \mathsf{B} + \mathsf{C} + \mathsf{D})S = S.$$

None of these results depend in any way on the 'state' of the market S, which may contain any numbers of animals in the various categories: $\mathsf{AA}S = \mathsf{A}S$ means $\mathsf{AA} = \mathsf{A}$ – whatever the state symbol S on which the operators act. In the same way $\mathsf{AB}S = 0 = \mathsf{0}S$, where $\mathsf{0}$ stands for the *zero operator*, which destroys any state symbol on which it acts. And the last result shows that $\mathsf{A} + \mathsf{B} + \mathsf{C} + \mathsf{D} = \mathsf{1}$, where $\mathsf{1}$ is the *unit operator*, which acts on any S without changing it.

It all begins to look a bit like what we found in the last Section, where we were dealing with *symmetry* operators. We've discovered yet another *operator algebra*, containing a $\mathsf{0}$ and a $\mathsf{1}$ and the four symbols $\mathsf{A}, \mathsf{B}, \mathsf{C}, \mathsf{D}$, whose properties can be summarized as

$$\mathsf{AA} = \mathsf{A}, \qquad \mathsf{BB} = \mathsf{B}, \qquad \mathsf{CC} = \mathsf{C}, \qquad \mathsf{DD} = \mathsf{D}, \tag{6.3}$$

$$\mathsf{AB} = \mathsf{BA} = \mathsf{0}, \quad \mathsf{AC} = \mathsf{CA} = \mathsf{0}, \quad ...., \quad \mathsf{CD} = \mathsf{DC} = \mathsf{0}, \tag{6.4}$$

and (what are the terms not shown in (6.4)?)

$$\mathsf{A} + \mathsf{B} + \mathsf{C} + \mathsf{D} = \mathsf{1}. \tag{6.5}$$

These properties define what mathematicians call a **spectral set** – 'spectral' because sorting the animals into their different kinds is like making a rainbow (a *spectrum*) of all the different colours that make up 'white' light.

It may seem that we're only playing games, which can't really be of any use to anyone. But it's hard to think of any parts of mathematics that can't be useful in science. In Physics, for example, there were two big 'revolutions' in the last century – Einstein's theory of **relativity** (1905 - 1915), which changed our ideas about space, time, and the whole universe; and the **quantum theory** (starting in the 1920s), which changed the way we think about the smallest particles of matter (such as **electrons**, which carry electricity, and **photons**, which carry light) – and led to so many spectacular advances in science and technology here on earth. Both depend on such 'games'.

6.3 Arguing with symbols – logic

At the beginning of this book we started from the idea of counting the number of objects in a 'set' – the set being a collection of *any* objects and the number being one particular 'property' of the set, the same for all sets whose objects can be 'paired' or 'put in one-to-one correspondence'.

We've come a long way since then: the 'objects' don't have to be things you can pick up or touch – they can be just ideas or names (like the numbers 1,2,3,... themselves), or 'operations' (like the ones we talked about in the last Section); or even 'qualities' (such as good and bad). They can all be represented by *symbols* with certain laws of combination. And over the years a very important branch of mathematics has taken shape: it's called **set theory** and can be applied in solving problems where you'd never believe mathematics could be useful – problems about the way people think and reason and try to make logical judgements.

Georg Cantor(1845-1918), a German mathematician, is usually given the credit for inventing set theory, during the last quarter of the nineteenth century; but the foundations were laid much earlier by George Boole (1815-1864). Boole never had a university education: he started his working life as an 'assistant schoolteacher' (age 16) and later started a school of his own. In his spare time he worked on mathematics: his work was published and he was invited to Cork (Ireland) in 1849 as Professor of Mathematics. And in 1854 he wrote a remarkable book called "The Laws of Thought", making use of his own theory of sets.

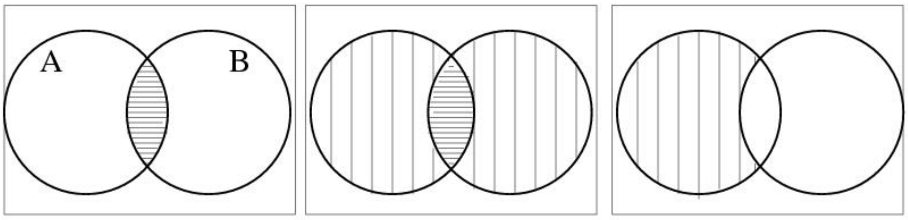
A **set** is any collection of distinct members or 'elements' with some well-defined meaning: for example, they might be the different letters used in writing "set theory" (s,e,t,h,o,r,y – each letter appearing only once, even though 'e' and 't' are used twice). The set is given a name, L say, and can be defined by two different methods: we may *list* the elements, showing them (or their names) in curly brackets (the order of the elements being immaterial); or we may call the general element x, say, and give a *recipe* for deciding if x belongs to the set. The set of letters in our example can thus be defined as $L = e, h, o, r, s, t, y$, in the 'list' method, or as $L = \{x|x =$ a letter needed in writing the words "set theory"$\}$, in the 'recipe' method. The vertical line $|$ simply separates x from the recipe.

To say that an element x "belongs to" or is "contained in" the set L we use a symbol $\in$ (like the Greek letter 'epsilon'); so $h \in L$ says that 'h' is one of the letters in the set L. And to say that an element does *not* belong to the set we use the symbol $\notin$; so clearly $z \notin L$.

There are two sets with special names: there is always an 'empty' set, denoted by $\emptyset$, which contains *no* members; and there is always a 'universal set', denoted by U, which contains *all* the elements we may want to talk about. U can contain all the elements of all sets! – but usually we're dealing with more limited sets. In the example above, we might take U to be the set A of all letters of the English alphabet, which is 'universal' if we're talking about writing in English; and in this case all the elements in L are also elements of A. When all the elements of one set, B, are contained in another set, A, we write $B \subset A$ – in words, B is a 'subset' of A. So in our example $L \subset A$.

We don't need many more 'fancy' signs, like $\in$, $\subset$; for example, if all the elements in B are contained in A *and vice versa* (i.e. all the elements of A are also contained in B) then the two sets must be exactly the same. In symbols

$$\text{If} \quad B \subset A \quad \text{and} \quad A \subset B \quad \text{then} \quad A = B,$$



(a) (b) (c)

Figure 8

where the = sign has its usual meaning.

Now let's think about *combining* sets, just as we did in thinking about the properties of ordinary numbers. In doing this it's a great help to use a pictorial way of indicating sets, due to John Venn (1834-1883), the inventor of the **Venn diagram**. We draw a big box to represent the universal set U and, within it, a number of circles to indicate the sets $A,\ B, ...$ we want to talk about. A member of any set A can then be represented by putting a labelled point anywhere inside the corresponding circle.

The two main methods of combining two sets are shown in Fig.8: in both cases the circles labelled A and B must overlap – otherwise the sets are quite separate and are said to be 'disjoint'.

In case Fig.8(a) the shaded area represents the **intersection** of sets A and B and any point in this area will stand for an element belonging to *both* A and B. To say this in symbols, we use $\cap$ for 'intersection' and (using the language introduced already) write

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

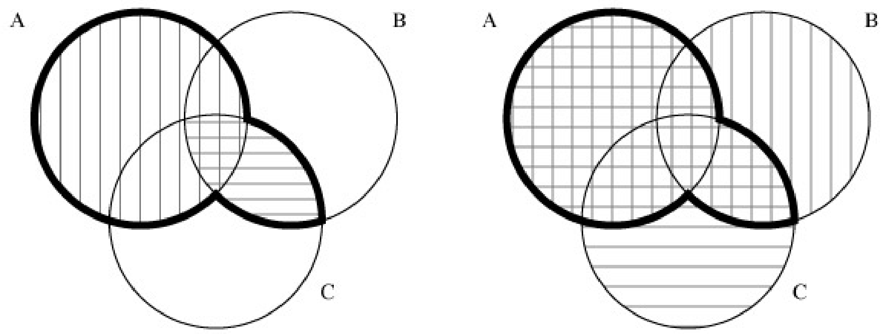
In words, "The intersection of sets A and B is the set which contains elements x, such that x belongs to set A *and* x belongs to set B".

In case Fig.8(b) the shaded area contains the points that stand for elements belonging to A, *or* B, or *both* A and B (the latter shown with heavier shading). The symbol used to indicate this **union** of two sets is $\cup$ and the definition becomes

$$A \cup B = \{x | x \in A \text{ and } x \in B \textit{ or in both}\}.$$

The *difference* of two sets is indicated in Fig.8(c). $A - B$ is the set whose elements belong to A, but *not* to B, the minus sign meaning they are excluded from B. Again, the shaded area shows the result.

The 'laws' for doing algebra with sets are similar to those used for numbers in Section 1.4. Here we'll just list them to show what they look like, but in illustrating them we'll find it easier to use Venn diagrams. There is a **commutative law** for each kind of combination, like that given in (2.1) for the integers, and also an **associative law**, similar to (2.2). These first four laws are



(a) (b)

- $A \cap B = B \cap A$
- $A \cup B = B \cup A$ (the order doesn't matter in either case)
- $A \cap (B \cap C) = (A \cap B) \cap C$
- $A \cup (B \cup C) = (A \cup B) \cup C$ (nor does the way of 'grouping')

When both $\cap$ and $\cup$ appear in a combination of three sets, there is a new kind of **distributive law**:

- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

– where the first equality reminds us of $a(b + c) = ab + ac$ for numbers.

All the laws follow from the definitions; but instead of trying to prove them it's easier simply to check them by using the diagrams. Fig.9 , for example, shows the pictures corresponding to the two sides of the last statement: the area bounded by the heavy line in (a) corresponds to the union of A and $B \cap C$ and thus represents $A \cup (B \cap C)$. Fig.9(b) shows the intersection of $(A \cup B)$ (shaded with vertical lines) and $(A \cup C)$ (shaded with horizontal lines): the region with both horizontal and vertical shading indicates their intersection. The two areas marked out in this way, in (a) and in (b), are clearly identical. This confirms that $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

It's often useful to introduce the **complement** of a set. If A is any set, containing elements $a_1, a_2, ...$, it will be a subset of the universal set U, represented as a big box with A (and any other sets) inside it. We define the complement of A, written A', as the subset of elements that are *not* members of A. This subset is represented by the part of the box lying *outside* A. And from this definition it is obvious that

$$A \cap A' = \emptyset, \qquad A \cup A' = U.$$

Other results involving $\emptyset$, U and the complement of a set, are easily confirmed from the corresponding diagrams. They are collected below:

- $A \cap A = A, \qquad A \cup A = A, \qquad A \cup \emptyset = A$
- $A \cap U = A, \qquad A \cup A' = U, \qquad A \cap A' = \emptyset.$

The results are all what you would expect if you think of the diagrams. Two more results (a bit less obvious) are given below for completeness, in case you ever need them. They are called "De Morgan's laws" and are:

- $(A \cap B)' = A' \cup B', \qquad (A \cup B)' = A' \cap B'$

Applications of the algebra of sets

In case you think this is getting too far away from real life to be useful, let's look at a problem that might come up any day – but which is hard to solve without using symbols or diagrams.

Example

A contractor has put in a tender for building a road. He has a construction team and all the heavy equipment he needs – cement mixers, bulldozers and cranes, etc.. He says he has to pay the 70 men in his team in line with what they can do: 30 can drive a mixer truck, 35 can work a bulldozer, and 22 can operate a crane; but some can do more than one job – 14 can use both a mixer and a bulldozer (and have to be paid more), 10 can use mixer and crane (and will also be paid more), 10 can work bulldozer and crane (again being paid more), and 4 can work all three (and must be paid most of all).

Is he telling the truth? Or is he cheating? – to get more money.

To get the answer let's use B for the set of men who can use a bulldozer; C for those who can use a crane; and M for those who can drive a mixer. These are the three basic sets and we represent them by the three circles in Fig.10: they must overlap, because some men can do more than one thing – they are members of more than one set. The areas labelled b, c, and m will refer to men who can do only one of the three things; and we'll also use b for the *number* who can only work a bulldozer, and so on. The areas labelled p, r, and s each refer to the overlap of two sets, and the the corresponding numbers of men who can do *two* jobs; while q will be the number who can do all three.

We now start to put in the numbers given to us by the contractor. Only 4 men can do all three jobs, and so belong to all three sets: they belong to the area where all three circles overlap; this area corresponds to $B \cap C \cap M$ and is a subset with q members; and so we can say $q = 4$.

The area of overlap of B and C corresponds to $B \cap C$ and this subset has $p + q$ elements; but the contractor says $p + q = 10$. Since we know $q = 4$ this means $p = 6$. In the same way (check it yourself) $B \cap M$ has $r + q$ elements, so $r = 10$; and $C \cap M$ has $s + q$ elements, so $s = 6$. The number of men who can do *more than* one job is therefore $p + q + r + s = 26$.

Now we're told that 33 men can operate a bulldozer; so $b + p + q + r$ and this means $b = 33 - 20 = 13$. Similarly, we find $c = 6$ and $m = 15$. The number of men who can do *only* one job is $b + c + m = 34$.

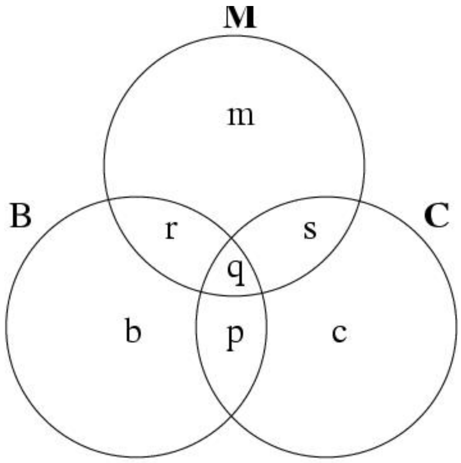


Figure 10

The total number of men working on the project is therefore $34 + 26 = 60$. But the contractor told us he had to pay wages to *seventy* men, not sixty – so he's trying to cheat!

Set theory has many applications nowadays – in all fields where logical decisions have to be made, whether in business and management or in the design of electrical circuits (where Boolean algebra – George Boole's great invention – is widely used). But some of its most difficult and important applications lie at the roots of Mathematics itself. For example, in talking about number we started from counting and the 'natural numbers' (the positive integers, without the zero. But the 'laws of combination' came to us from thinking about *finite* sets, sets of beads or bricks which could easily be counted. And then we went ahead, to the field of all real numbers, without saying any more about the foundations, even though the numbers – represented in earlier chapters by points on a line – were so crowded that they could never be counted! To deal with sets containing so many elements that they can't be counted raises problems so deep that only professional mathematicians talk about them. And that's where we stop.

Exercises

(1) Does the set of integers (all positive and negative numbers, including zero), with addition as the law of combination, form a group? What happens if we take away the zero? or if we use only the even numbers, 2, 4, 6, ... ?

(2) Remember the 'multiplication tables' in arithmetic (Section 2.3): we can set up similar tables for symmetry groups. If the group elements are called $\mathsf{A}, \mathsf{B}, \mathsf{C}, ...$ the *group* multiplication table will be

	A	B	C	D
A	AA	AB	AC	AD
B	BA	BB	BC	BD
C	CA	CB	CC	CD
D	DA	DB	DC	DD

where, for example, BC stands for the entry on Row-B and in Column-C.

Make up the multiplication table, for the symmetry group studied in this Section, by working out all the products.

(3) Cut out a square piece of card, with the underneath side painted black, and study the symmetry operations which leave it looking just the same (put a small + on the front, in the top right-hand corner, so you can see what's happening). How many reflection operations are there? and how many rotations? (You can use symbols C_2, C_4 to stand for rotations through a half, or a quarter turn, around an axis; and R_1, R_2, R'_1, R'_2 for the various reflections.) Say clearly what every operation does.

Then set up the multiplication table, as you did in Exercise 3, by finding the single symmetry operation equivalent to XY for the product in Row-X and Column-Y – doing this for every product.

Can you find any rules for combining rotations with rotations, rotations with reflections, and so on? Are there any pairs for which XY and YX are not the same?

(4) Suppose there are 40 students in a class and that their heights are as follows:

4	are between	1m 5cm and 1m 10cm
8	are between	1m 10cm and 1m 15cm
13	are between	1m 15cm and 1m 20cm
12	are between	1m 20cm and 1m 25cm
3	are between	1m 20cm and 1m 25cm.

The numbers in these five categories show the 'state' of the class: how can you put this in symbols?

(5) Use A, B, ...E for the 'operators' which take students from the five categories. What operators (call them U and V) would select students Under, and oVer, 1m 20cm ? And what properties will they have? (e.g. UU = ?)

Looking back –

We started from almost nothing – just a few ideas about counting – and we've come quite a long way. Let's take stock:

- Chapters 1 and 2 must have reminded you of when you first learnt about numbers – as a small child, and the endless chanting of multiplication tables – learning without understanding. But now you know what it all *means*: the rules of arithmetic are made by *us*, to help us to understand and use what we see around us. You know that you can use *any* symbols in place of numbers and write down the rules of arithmetic using only the symbols – and then you are doing **algebra**

- In Chapter 3 you met **equations**, containing a number you don't know (call it 'x'), along with the integers (1,2,3,...), and were able to **solve** the equations, finding for x the new numbers (0, zero; and $-1, -2, ...$, negative numbers). You found how useful *pictures* could be in talking about numbers and their properties; and how a number could be given to any point on a line – bringing with it the idea of **fractions** as the numbers labelling points in between the integers.

- But between any two such numbers, 'rational fractions' of the form a/b, *however close*, there were still uncountable millions of numbers that you can't represent in that simple way; you can get as close as you wish – but without ever really getting there. An 'irrational' number is defined only by a *recipe* that tells you how to reach it – but that's enough. The set of all the numbers defined so far was called the **Field of Real Numbers**; and it's enough for all everyday needs – like measuring.

- The last big step was made in Chapter 5, when we admitted the last 'new' number, denoted by i and called the "imaginary unit", with the property $i \times i = -1$ (*not* 1). And when i was included, and allowed to mix with all the real numbers, our number field had to be extended to include both **Real Numbers and Complex Numbers**. All equations involving only such numbers could then be solved without inventing anything new: the field was *closed*.

- In the last Chapter, however, you saw that symbols could be used to stand for other things – not only numbers. They can be used for **operations**, like moving things in space; or sorting a mixture of objects into objects 'of the same kind'; or just for arguing about things!

Index